

(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2004-524815

(P2004-524815A)

(43) 公表日 平成16年8月19日(2004.8.19)

(51) Int. Cl. ⁷	F I	テーマコード (参考)
C 1 2 N 15/09	C 1 2 N 15/00	Z N A A
A 6 1 K 38/00	A 6 1 K 45/00	2 G O 4 5
A 6 1 K 45/00	A 6 1 K 49/00	4 B O 2 4
A 6 1 K 49/00	A 6 1 P 11/00	Z
A 6 1 P 11/00	A 6 1 P 25/00	4 B O 2 9
		4 B O 5 0
		4 B O 6 3
	審査請求 未請求 予備審査請求 有	(全 197 頁) 最終頁に続く

(21) 出願番号	特願2002-530761 (P2002-530761)	(71) 出願人	301005050
(86) (22) 出願日	平成13年9月28日 (2001.9.28)		インサイト・ゲノミックス・インコーポレイテッド
(85) 翻訳文提出日	平成15年3月25日 (2003.3.25)		アメリカ合衆国カリフォルニア州94304・パロアルト・ポータードライブ 3160
(86) 国際出願番号	PCT/US2001/030310	(74) 代理人	100089266
(87) 国際公開番号	W02002/026998		弁理士 大島 陽一
(87) 国際公開日	平成14年4月4日 (2002.4.4)	(72) 発明者	ユエ、ヘンリー
(31) 優先権主張番号	60/237,093		アメリカ合衆国カリフォルニア州94087・サニーバイル・ルイスアベニュー 826
(32) 優先日	平成12年9月29日 (2000.9.29)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/238,370		
(32) 優先日	平成12年10月6日 (2000.10.6)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/241,284		
(32) 優先日	平成12年10月17日 (2000.10.17)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 ヒドロラーゼ

(57) 【要約】

本発明はヒトのヒドロラーゼ (HYDR) およびHYDRを同定し、コードするポリヌクレオチドを提供する。本発明はまた、発現ベクター、宿主細胞、抗体、アゴニストおよびアンタゴニストをも提供する。本発明はまた、HYDRの異常発現に関連する疾患を診断、治療または予防する方法をも提供する。

【特許請求の範囲】

【請求項 1】

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

(a) SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列を含むポリペプチド

(b) SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列と少なくとも 90% が同一であるような天然のアミノ酸配列を含むポリペプチド

(c) SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片

(d) SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片 10

【請求項 2】

SEQ ID NO: 1 - 4 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: 5 - 8 からなる群から選択した請求項 5 に記載の単離されたポリヌクレオチド。 20

【請求項 6】

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

【請求項 7】

請求項 6 に記載の組換えポリヌクレオチドを用いて形質転換した細胞。

【請求項 8】

請求項 6 に記載の組換えポリヌクレオチドを含む遺伝形質転換体。

【請求項 9】

請求項 1 のポリペプチドを生産する方法であって、 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連結されたプロモーター配列を含む組換えポリヌクレオチドで形質転換された細胞を培養するステップと、

(b) そのように発現した前記ポリペプチドを回収する過程とからなり、前記組換えポリヌクレオチドが、請求項 1 に記載の前記ポリペプチドをコードするポリヌクレオチドに機能的に連結したプロモーター配列を有することを特徴とする方法。

【請求項 10】

前記ポリペプチドが、SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列を含むことを特徴とする請求項 9 に記載の方法。

【請求項 11】

請求項 1 に記載のポリペプチドと特異結合するような単離された抗体。 40

【請求項 12】

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

(a) SEQ ID NO: 5 - 8 からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド

(b) SEQ ID NO: 5 - 8 からなる群から選択したポリヌクレオチド配列と少なくとも 90% が同一であるような天然のポリヌクレオチド配列を含むポリヌクレオチド

(c) (a) のポリヌクレオチドに相補的なポリヌクレオチド

(d) (b) のポリヌクレオチドに相補的なポリヌクレオチド

(e) (a) ~ (d) の RNA 等価物 50

【請求項 13】

請求項 12 に記載のポリヌクレオチドの少なくとも 60 の連続したヌクレオチドを含む単離されたポリヌクレオチド。

14. 請求項 12 に記載のポリヌクレオチドの配列を有する標的ポリヌクレオチドをサンプル中から検出する方法であって、

(a) 前記サンプル中の前記標的ポリヌクレオチドに相補的な配列を有する少なくとも 20 の連続したヌクレオチドを含むプローブを用いて前記サンプルをハイブリダイズする過程と、

(b) 前記ハイブリダイゼーション複合体の存在・不存在を検出し、該複合体が存在する場合にはオプションでその量を検出する過程からなり、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の間でハイブリダイゼーション複合体が形成されるような条件下で、前記プローブが前記標的ポリヌクレオチドに特異的にハイブリダイズすることを特徴とする方法。

10

【請求項 14】

前記プローブが少なくとも 60 の連続したヌクレオチドを含むことを特徴とする請求項 14 に記載の方法。

【請求項 15】

請求項 12 に記載のポリヌクレオチドの配列を有する標的ポリヌクレオチドをサンプル中から検出する方法であって、

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

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

20

【請求項 16】

有効量の請求項 1 のポリペプチドと、薬剤として許容できる賦形剤とを有することを特徴とする成分。

【請求項 17】

前記ポリペプチドが、SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列を有することを特徴とする請求項 17 に記載の成分。

30

【請求項 18】

機能性 HYDR の発現低下に関連する疾患又は病状を治療する方法であって、そのような治療を必要とする患者に対して請求項 17 に記載の成分を投与する過程を含むことを特徴とする方法。

【請求項 19】

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

(a) 請求項 1 のポリペプチドを含むサンプルを化合物に曝すステップと、

(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。

40

【請求項 20】

請求項 20 に記載の方法によって同定したアゴニスト化合物と、薬剤として許容できる賦形剤とを含むことを特徴とする成分。

【請求項 21】

機能性 HYDR の発現低下に関連する疾患又は病状を治療する方法であって、そのような治療を必要とする患者に対して請求項 21 に記載の成分を投与する過程を含むことを特徴とする方法。

【請求項 22】

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

50

- (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すステップと、
- (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 2 3】

請求項 2 3 に記載の方法によって同定したアンタゴニスト化合物と、薬剤として許容できる賦形剤とを含むことを特徴とする成分。

【請求項 2 4】

機能性 H Y D R の過剰発現に関連する疾患又は病状の治療方法であって、そのような治療を必要とする患者に対して請求項 2 4 に記載の成分を投与する過程を含むことを特徴とする方法。

10

【請求項 2 5】

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

- (a) 適切な条件下で請求項 1 に記載のポリペプチドを少なくとも 1 つの試験化合物に混合させる過程と、
- (b) 請求項 1 に記載のポリペプチドの試験化合物との結合を検出し、それによって請求項 1 に記載のポリペプチドに特異結合する化合物を同定するステップを含むことを特徴とする方法。

【請求項 2 6】

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

20

- (a) 請求項 1 に記載のポリペプチドの活性が許容された条件下で、請求項 1 に記載のポリペプチドを少なくとも 1 つの試験化合物に混合させる過程と、
- (b) 請求項 1 に記載のポリペプチドの活性を試験化合物の存在下で算定する過程と、
- (c) 試験化合物の存在下での請求項 1 に記載のポリペプチドの活性を、試験化合物の不存在下での請求項 1 に記載のポリペプチドの活性と比較する過程を含み、試験化合物の存在下での請求項 1 に記載のポリペプチドの活性の変化が、請求項 1 に記載のポリペプチドの活性を調節する化合物を標示することを特徴とする方法。

【請求項 2 7】

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

30

- (a) 前記標的ポリヌクレオチドの発現に好適な条件下で、該標的ポリヌクレオチドを含むサンプルを化合物に曝露する過程と、
- (b) 前記標的ポリヌクレオチドの変異発現を検出する過程と、
- (c) 可変量の前記化合物の存在下と前記化合物の不存在下で、前記標的ポリヌクレオチドの発現を比較する過程とを含むことを特徴とする方法。

【請求項 2 8】

試験化合物の毒性を算定する方法であって、

- (a) 核酸を含む生物学的サンプルを前記試験化合物で処理する過程と、
- (b) 処理した前記生物学的サンプルの核酸と、請求項 1 2 に記載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過程であって、このハイブリダイゼーションが、前記プローブと前記生物学的サンプルの標的ポリヌクレオチドとの間で特異的なハイブリダイゼーション複合体が形成される条件下で行われ、前記標的ポリヌクレオチドが、請求項 1 2 に記載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、前記過程と、
- (c) ハイブリダイゼーション複合体の収量を定量する過程と、
- (d) 前記処理された生物学的サンプル中の前記ハイブリタイゼーション複合体の量を、処理されていない生物学的サンプル中の前記ハイブリタイゼーション複合体の量と比較する過程とを含み、前記処理された生物学的サンプル中の前記ハイブリタイゼーション複合体の量の差が、前記試験化合物の毒性を標示することを特徴とする方法。

40

50

【請求項 29】

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

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

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

【請求項 30】

前記抗体が、

(a) キメラ抗体

(b) 単鎖抗体

(c) Fab断片

(d) F(ab')₂断片

(e) ヒト化抗体のいずれかであることを特徴とする請求項11に記載の抗体。

10

【請求項 31】

請求項11に記載の抗体と、許容できる賦形剤とを含む化合物。

【請求項 32】

被検者のHYDRの発現に関連する病状又は疾患の診断方法であって、請求項32に記載の化合物の有効量を前記被検者に投与する過程を含むことを特徴とする方法。

20

【請求項 33】

前記抗体が標識されることを特徴とする請求項32に記載の化合物。

【請求項 34】

被検者のHYDRの発現に関連する病状又は疾患の診断方法であって、請求項34に記載の化合物の有効量を前記被検者に投与する過程を含むことを特徴とする方法。

【請求項 35】

請求項11に記載の抗体の特異性を有するポリクローナル抗体を調製する方法であって、

(a) 抗体反応を誘発する条件下で、SEQ ID NO: 1 - 4からなる群から選択したアミノ酸配列またはその免疫原性断片を含むポリペプチドを用いて動物を免疫化する過程と、

(b) 前記動物から抗体を単離する過程と、

(c) 前記単離された抗体をポリペプチドでスクリーニングし、それによって、SEQ ID NO: 1 - 4からなる群から選択したアミノ酸配列を有するポリペプチドに特異結合するようなポリクローナル抗体を同定する過程とを含むことを特徴とする方法。

30

【請求項 36】

請求項36に記載の方法で産出したポリクローナル抗体。

【請求項 37】

請求項37に記載のポリクローナル抗体及び適切なキャリアを含む組成物。

【請求項 38】

請求項11に記載の抗体の特異性を有するモノクローナル抗体を作製する方法であって、

(a) 抗体反応を誘発する条件下で、SEQ ID NO: 1 - 4からなる群から選択したアミノ酸配列またはその免疫原性断片を含むポリペプチドを用いて動物を免疫化する過程と、

(b) 前記動物から抗体産出細胞を単離する過程と、

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

(d) 前記ハイブリドーマ細胞を培養する過程と、

(e) SEQ ID NO: 1 - 4からなる群から選択したアミノ酸配列を有するポリペプチドに特異結合するような前記培養モノクローナル抗体から単離する過程とを含むことを特徴とする方法。

40

50

【請求項 39】

請求項 39 に記載の方法で産出したモノクローナル抗体。

【請求項 40】

請求項 40 に記載のモノクローナル抗体及び適切なキャリアを含む組成物。

【請求項 41】

F a b 発現ライブラリのスクリーニングにより前記抗体を産出することを特徴とする請求項 11 に記載の抗体。

【請求項 42】

組換え免疫グロブリンライブラリのスクリーニングにより前記抗体を産出することを特徴とする請求項 11 に記載の抗体。

10

【請求項 43】

S E Q I D N O : 1 - 4 からなる群から選択したアミノ酸配列を有するポリペプチドを検出する方法であって、

(a) 前記抗体と前記ポリペプチドの特異結合を許容する条件下で、サンプルを用いて請求項 11 に記載の抗体をインキュベートする過程と、

(b) 特異結合を検出する過程とを含み、該特異結合が、S E Q I D N O : 1 - 4 を有する群から選択したアミノ酸配列を有するポリペプチドがサンプル中に存在することを示唆することを特徴とする方法。

【請求項 44】

S E Q I D N O : 1 - 4 からなる群から選択したアミノ酸配列を有するポリペプチドを精製する方法であって、

20

(a) 前記抗体と前記ポリペプチドの特異結合を許容する条件下で、サンプルを用いて請求項 11 に記載の抗体をインキュベートする過程と、

(b) 前記サンプルから前記抗体を分離し、S E Q I D N O : 1 - 4 からなる群から選択したアミノ酸配列を有する精製ポリペプチドを得る過程とを含むことを特徴とする方法。

【請求項 45】

マイクロアレイの少なくとも 1 つが請求項 13 に記載のポリヌクレオチドであることを特徴とするマイクロアレイ。

【請求項 46】

ポリヌクレオチドを含むサンプルの転写イメージを作製する方法であって、

30

(a) サンプル中のポリヌクレオチドを標識化する過程

(b) ハイブリダイゼーション複合体が形成されるのに適した条件下で請求項 46 のマイクロアレイの要素とサンプル中の標識化ポリヌクレオチドを接触させる過程、および

(c) サンプル中のポリヌクレオチドの発現量を定量する過程を含む。

【請求項 47】

固体基板上的固有の物理的位置に付着された種々のヌクレオチド分子を含むアレイであって、少なくとも 1 つの前記ヌクレオチド分子が、標的ポリヌクレオチドの少なくとも 30 の連続したヌクレオチドと特異的にハイブリダイズ可能な最初のオリゴヌクレオチドまたはポリヌクレオチドを有し、前記の標的ポリヌクレオチドが請求項 12 に記載のポリヌクレオチドを特徴とするアレイ。

40

【請求項 48】

請求項 48 に記載のアレイで、前記のオリゴヌクレオチドまたはポリヌクレオチドの最初の配列が前記の標的ポリヌクレオチドの少なくとも 30 の連続したヌクレオチドに完全に相補的であることを特徴とするアレイ。

【請求項 49】

請求項 48 に記載のアレイで、前記のオリゴヌクレオチドまたはポリヌクレオチドの最初の配列が前記の標的ポリヌクレオチドの少なくとも 60 の連続したヌクレオチドに完全に相補的であることを特徴とするアレイ。

50

【請求項 5 0】

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

【請求項 5 1】

請求項 4 8 に記載のアレイで、マイクロアレイであることを特徴とするアレイ。

【請求項 5 2】

請求項 4 8 に記載のアレイで、前記のオリゴヌクレオチドまたはポリヌクレオチドの最初の配列を有するヌクレオチド分子にハイブリダイズした前記の標的ポリヌクレオチドを有することを特徴とするアレイ。

【請求項 5 3】

請求項 4 8 に記載のアレイで、リンカーが少なくとも 1 つの前記のヌクレオチド分子と前記の固体基板を連結していることを特徴とするアレイ。

【請求項 5 4】

請求項 4 8 に記載のアレイで、基板上的固有の物理的位置の各々が複数のヌクレオチド分子を含み、任意の単一の固有の物理的位置でのその複数のヌクレオチド分子は同一の配列を有し、基板上的固有の物理的位置の各々は、基板上的別の固有の物理的位置でのヌクレオチド分子の配列とは異なる配列を有するヌクレオチド分子を含むことを特徴とするアレイ。

【請求項 5 5】

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 5 6】

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 5 7】

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 5 8】

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 5 9】

SEQ ID NO : 5 のポリヌクレオチド配列を含む請求項 1 2 に記載のポリヌクレオチド。

【請求項 6 0】

SEQ ID NO : 6 のポリヌクレオチド配列を含む請求項 1 2 に記載のポリヌクレオチド。

【請求項 6 1】

SEQ ID NO : 7 のポリヌクレオチド配列を含む請求項 1 2 に記載のポリヌクレオチド。

【請求項 6 2】

SEQ ID NO : 8 のポリヌクレオチド配列を含む請求項 1 2 に記載のポリヌクレオチド。

【発明の詳細な説明】

【0 0 0 1】

(技術分野)

本発明は、ヒドロラーゼの核酸配列及びアミノ酸配列に関する。本発明はまた、これらの配列を利用した免疫系の疾患、神経の疾患、免疫不全、肺の疾患、および癌を含む細胞増殖異常の診断・治療・予防に関する。本発明はさらに、ヒドロラーゼ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影響についての評価に関する。

【0 0 0 2】

(発明の背景)

ヒドロラーゼは水分子の導入によって基質内のさまざまな共有結合の切断を触媒する酵素クラスである。この反応においては、水分子の酸素原子が基質内の標的結合に求核攻撃を加える。水分子が標的結合のところで分かれて結合を破り、二つの分子が生成される。ヒ

10

20

30

40

50

ドロラーゼは細胞成分の合成と分解などの機能および細胞シグナル伝達、細胞増殖、炎症、アポトーシス、分泌および排泄を含む細胞機能の調節に必須な反応にかかわっている。ヒドロラーゼはこれらの機能にかかわる病気プロセスの重要な過程に関連している。加水分解酵素、すなわちヒドロラーゼは基質特異性によってホスファターゼ、ペプチダーゼ、リゾホスホリパーゼ、ホスホジエステラーゼ、グリコシダーゼおよびグリオキサラーゼなどに分類される。

ホスファターゼは加水分解によりタンパク質からリン酸基を除去する。これは、多くの細胞内シグナル伝達経路を調節し、ひいては細胞成長と分化、細胞間接触、細胞周期および発癌を制御するエネルギー供給過程である。

【0003】

ペプチダーゼはプロテアーゼとも呼ばれ、ペプチドやタンパク質の鎖のバックボーンを形成するペプチド結合を切断する。タンパク分解は細胞増殖、分化、再構築や恒常性ならびに炎症および免疫反応において重要である。典型的なタンパク質の半減期は数時間から数日の範囲であるため、ペプチダーゼはつねにタンパク質前駆体を切断して活性型に変え、標的タンパク質からシグナル配列を除去し、古いタンパク質や欠陥のあるタンパク質を分解している。ペプチダーゼはバクテリア、寄生生物およびウィルスの侵入と宿主内での複製の際に機能する。ペプチダーゼの例としてはトリプシンおよびキモトリプシン、補体カスケードと血液凝固カスケードの成分、リソソームのカテプシン、カルパイン、ペプシン、レニンおよびキモシンなどが挙げられる (Beynon, R. J. および J. S. Bond (1994) Proteolytic Enzymes: A Practical Approach, Oxford University Press, New York, NY, 1-5頁)。

【0004】

リゾホスホリパーゼ (LPL) はエステル結合の加水分解でアシル基を除去するという、脂質分解の重要な過程を触媒することによって細胞内脂質を調節する。約 15 ~ 30 kD の小さい LPL アイソフォームはヒドロラーゼとして機能し、大きなアイソフォームはヒドロラーゼとアシル基転移酵素の両方の機能をする。LPL の特定の基質であるリゾホスファチジルコリンは細胞膜の溶解を起こす。LPL の活性は、炎症反応を含む数多くの経路で重要なシグナル伝達分子によって調節される。

【0005】

ホスホジエステラーゼはホスホジエステル化合物内の 2 つのエステル結合のうちの 1 つの加水分解を触媒する。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 DNA および RNA のエンドヌクレアーゼおよびエキソヌクレアーゼが含まれるが、これらは細胞の成長と複製およびタンパク質合成に必須である。エンドヌクレアーゼ V (デオキシノシン 3' エンドヌクレアーゼ) はタイプ II の部位特異性デオキシリボヌクレアーゼの一例であり、このタイプのデオキシリボヌクレアーゼは、ヒポキサンチン、ウラシルまたはミスマッチ塩基を含む DNA を切断する、推定上の DNA 修復酵素である。大腸菌エンドヌクレアーゼ V は、デオキシキサンチンへの第 2 のホスホジエステル結合 3' においてデオキシキサンチンを含む DNA を切断し、切断部位で 3' ヒドロキシル基および 5' ホスホリル基を生成することが示されている (He, B. 他 (2000) Mutat. Res. 459: 109-114)。大腸菌エンドヌクレアーゼ V は脱アミノ化されたグアニンすなわちキサンチンを DNA から除去する際に役割を果たし、ひいてはニトロソ基の脱アミノ化による突然変異誘発効果から細胞を保護すると示唆されている (Schouten KA および Weiss B (1999) Mutat. Res. 435: 245-254)。真核生物においては、tRNA スプライシング過程において、アンチコドンループ 1 塩基 3' からアンチコドン妨害する小さな tRNA イントロンを除去する必要がある。この過程にはエンドヌクレアーゼ、リガーゼおよびホスホトランスフェラーゼの段階的な作用が必要とされる (Hong, L. 他 (1998) Science 280: 279-284)。リボヌクレアーゼ P (Rnase P) は、tRNA スプライシング過程に成熟した tRNA 5' 末端

10

20

30

40

50

を生成するために必要な遍在性のRNAプロセシングエンドヌクレアーゼである。これは tRNA 前駆体上の特異的な部位で P-3' O 結合の切断を触媒することによって 5' リン酸および 3' ヒドロキシル末端基を生成することで達成される。RNase P による触媒反応は Mg^{2+} または Mn^{2+} などの 2 価のカチオンに絶対的に依存している (Kuruz, J. C. 他 (2000) *Curr. Opin. Chem. Biol.* 4: 553 - 558)。RNase P の基質認識機構は真核生物、古細菌 (Archaea) およびバクテリアの間でよく保存されていることが実証されている (Fabbri, S. 他 (1998) *Science* 280: 284 - 286)。S. cerevisiae において POP1 (RNA 前駆体の処理) と名づけられている遺伝子が、RNase P および RNase MRP (もう一つの RNA 処理タンパク質) の両方のタンパク質成分をコードしている。酵母の POP1 の突然変異は致死的であることが示されている (Lygerou, Z. 他 (1994) *Genes Dev.* 8: 1423 - 1433)。もう一つのホスホジエステラーゼである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。ホスホリルコリンはホスファチジルコリンの合成に使用されるが、ホスファチジルコリンは数多くの細胞内シグナル伝達経路に関与している。セラミドは、神経組織に高濃度に存在する膜脂質であるガングリオシドの生成に必須な前駆体である。酸性スフィンゴミエリナーゼが欠損すると、リソソームにおいてスフィンゴミエリン分子が蓄積され、それによってニーマン-ピック病が引き起こされる。

10

20

【0006】

グリコシダーゼはグリコシドのヘミアセチル結合を切断する。グリコシドは 1 つ以上の糖を含む化合物である。たとえば、哺乳動物のラクターゼ・フロリジンヒドロラーゼは乳糖を分割する腸酵素である。哺乳動物のベータガラクトシダーゼはガングリオシド、糖タンパク質およびグリコサミノグリカンから末端ガラクトースを除去し、この酵素の欠損はモルキオ病タイプ B として知られるガングリオシド蓄積症と関連している。脊椎動物のリソソームのアルファグルコシダーゼ (グリコーゲン、マルトースおよびイソマルトースを加水分解) および脊椎動物の腸スクラーゼ・イソマルターゼ (スクロース、マルトースおよびイソマルトースを加水分解) はこのファミリーの広範囲に分布するメンバーであり、活性部位に高度に保存された配列を持っている。

グリオキシラーゼ系は糖新生、すなわち体内に保存された化合物からグルコースを産生することに関与している。この系はグリオキシラーゼ I (トリオースリン酸エネルギー代謝の副産物であるメチルグリオキサールから S-D ラクトイルグルタチオンを形成する反応を触媒) およびグリオキシラーゼ II (S-D ラクトイルグルタチオンを D 乳酸と還元グルタチオンに加水分解) からなっている。グリオキシラーゼは高血糖症、非インシュリン依存糖尿病、バクテリア毒素の解毒、および細胞増殖と微小管アセンブリの制御に関与している。

30

【0007】

エーテル結合に作用するヒドロラーゼの小さなサブクラスとして、チオエーテルヒドロラーゼなどが挙げられる。S アデノシル L ホモシステインヒドロラーゼは AdoHcyase または SAHH (PROSITE PDOC00603; EC 3.3.1.1) としても知られているが、S アデノシル L ホモシステイン (AdoHcy) を可逆的にアデノシンとホモシステインに加水分解する活性を持つものとしてラットの肝臓抽出物中で初めて記述されたチオエーテルヒドロラーゼである (Sganga, M. W. 他 (1992) *PNAS* 89: 6328 - 6332)。SAHH はサイトゾル酵素であって、大腸菌およびある種の関連バクテリアを除いて、試験されたすべての細胞で見ついている (Walker, R. D. 他 (1975) *Can. J. Biochem.* 53: 312 - 319; Shimizu, S. 他 (1988) *FEMS Microbiol. Lett.* 51: 177 - 180; Shimizu, S. 他 (1984) *Eur. J. Biochem.* 141: 385 - 392)。SAHH は補助因子として NAD^+ に依存している。SAHH の欠損は高メチオニン血症 (Online Mend 50

40

50

elian Inheritance in Man (OMIM) #180960 Hypermethioninemia)に関連しているが、これは新生児胆汁鬱滞、成長障害、知的および運動能力の遅進、異常な毛および歯を含む顔面異形症、および心筋症によって特徴付けられる病理学的状態である(Labruno, P. 他(1990) J. Pediat. 117:220-226)。

ヒドロラーゼのもう1つのサブクラスにはペプチド結合以外の炭素窒素(C-N)結合に作用する酵素が含まれる。このサブクラスにはアミド、アミジンおよびその他のC-N結合を加水分解する酵素が属している。このサブクラスはさらにリニアアミド、サイクリックアミド、リニアアミジン、サイクリックアミジン、ニトリルおよびその他の化合物を含む基質の特異性によって分けられている。この酵素サブクラスに属するヒドロラーゼでサイクリックアミジンに作用するのがアデノシンジアミナーゼ(ADA)である。ADAはアデノシンをイノシンに分解する反応の触媒である。ADAは胎盤、筋肉、肺、胃、消化性憩室、脾臓、赤血球、胸腺、精漿、甲状腺、T細胞、骨髄幹細胞および肝臓を含む多くの哺乳動物組織に存在する。ADAのサブクラスであるADARはRNAに作用し、RNAエディターゼ(editase)として分類されている。ショウジョウバエからのADARであるdADARは発生中の神経系で発現することが示されているため、中枢神経系のパラ電圧依存性Na⁺チャネル転写物に作用するエディターゼの候補になっている(Palladino, M. J. 他(2000) RNA 6:1004-1018)。ADAの欠損は重症複合型免疫不全(SCID)を伴う重大なリンパ球減少症を引き起こす。ADA欠損を示す患者の細胞はADA触媒活性およびADARタンパク質の量が正常より少ないか、ときには検出不能な場合もある。ADA欠損はADA遺伝子の遺伝的な突然変異に起因し、SCIDを引き起こすことが示されている(Hershfield, M. S. (1998) Semin. Hematol. 4:291-298)。マウスにおけるADA欠損の代謝への影響は肺胞生成(alveogenesis)の欠陥、肺の炎症および気道閉塞と関連していることが示されている(Blackburn, M. R. 他(2000) J. Exp. Med. 192:159-170)。

【0008】

膵臓リボヌクレアーゼ(RNase)は、ある種の哺乳動物分類群および一部の爬虫類の膵臓に大量に見られるピリミジン特異的なエンドヌクレアーゼである(Beintema, J. J. 他(1988) Prog. Biophys. Mol. Biol. 51:165-192)。哺乳類の膵臓RNaseスーパーファミリーは、2段階リン酸転移加水分解反応を通じてRNAを分解する非サイトゾルのエンドヌクレアーゼである(Beintema, J. J. 他(1986) Mol. Biol. Evol. 3:262-275)。特に、これらの酵素は3'ホスホモノヌクレオチドと3'ホスホオリゴヌクレオチドの内ヌクレオチド結合を切断してC-PまたはU-Pと2'、3'サイクリックリン酸中間体にする過程に関与している。リボヌクレアーゼは一本鎖のDNAと複合体を作ってDNAヘリックスを巻き戻すことができる。この複合体は酵素のリジンおよびアルギニン残基とヌクレオチドのリン酸基の間の拡張多部位カチオン・アニオン相互作用によって生じる。このファミリーに属する酵素の一部は純粋な消化機能を演じているようであるが、その他のものは強力な異例な生物学的活性を示す(D'Alessio, G. (1993) Trends Cell Biol. 3:106-109)。膵臓RNaseファミリーに属するタンパク質としては、ウシの精嚢および脳リボヌクレアーゼ、腎臓非分泌型リボヌクレアーゼ(Beintema, J. J. 他(1986) FEBS Lett. 194:338-343)、肝臓タイプリボヌクレアーゼ(Rosenberg, H. F. et al. (1989) PNAS U.S.A. 86:4460-4464)、正常および悪性組織の血管新生を誘発するアンジオジェニン、リボヌクレアーゼ活性を持つ細胞毒素および寄生虫毒素である好酸球カチオン性タンパク質(Hofstenge, J. et al. (1989) Biochemistry 28:9806-9813)、およびかえる肝臓リボヌクレアーゼとカエルのシアル酸結合レクチンが含まれる。膵臓RNaseの配列には4つの保存されたジスルフィド結合と触媒活性に関与

10

20

30

40

50

する3つのアミノ酸残基が含まれている。

セリンヒドロラーゼは活性部位にセリン残基を含む加水分解酵素の機能クラスである。このクラスの酵素としては、さまざまな基質を加水分解するプロテイナーゼ、エステラーゼおよびリパーゼが含まれ、生物学的な役割もさまざまである。。このスーパーファミリのタンパク質は基質特異性またはアミノ酸類似性によってさらにサブファミリに分けることができる (P u e n t e , X . S . および L o p e z - O n t , C . (1 9 9 5) J . B i o l . C h e m . 2 7 0 (2 1) : 1 2 9 2 6 - 1 2 9 3 2) 。

【0009】

カルボキシルエステラーゼはカルボキシルエステルを加水分解するタンパク質であり、A、BおよびCの3つのカテゴリーに分類される。ほとんどのタイプBカルボキシルエステラーゼは進化的に関連しており、1つのタンパク質ファミリを構成すると考えられている。タイプBカルボキシルエステラーゼファミリのタンパク質としては、脊椎動物のアセチルコリンエステラーゼ、哺乳動物の肝臓ミクロソームカルボキシルエステラーゼ、哺乳動物の胆汁塩活性化リパーゼおよびアヒルの脂肪酸CoAヒドロラーゼが含まれる。このタンパク質ファミリの一部は触媒活性がないが、他のタイプBカルボキシルエステラーゼと進化的に関係するドメインを含んでおり、例としてはサイログロブリンとショウジョウバエのタンパク質である *neuractin* が挙げられる。

10

【0010】

ヌクレオチダーゼはヌクレオチドから遊離ヌクレオシドを生成する反応を触媒する。ハトの心臓からクローン化されたサイトゾルヌクレオチダーゼ *cN-I* (5'ヌクレオチダーゼI) はATPの加水分解中に生成されるAMPからアデノシンを生成する反応を触媒する (S a l a - N e w b y , G . B . 他 (1 9 9 9) J . B i o l . C h e m . 2 7 4 : 1 7 7 8 9 - 1 7 7 9 3) 。アデノシン濃度の増加は代謝ストレスのシグナルであると考えられており、アデノシン受容体は血管拡張、興奮性ニューロン発火の減少および心臓内の虚血性プレコンディショニングを含む効果を仲介する (S c h r a d e r , J . (1 9 9 0) C i r c u l a t i o n 8 1 : 3 8 9 - 3 9 1 ; R u b i n o , A . 他 (1 9 9 2) E u r . J . P h a r m a c o l . 2 2 0 : 9 5 9 8 ; d e J o n g , J . W . 他 (2 0 0 0) P h a r m a c o l . T h e r . 8 7 : 1 4 1 1 4 9) 。ピリミジン5'ヌクレオチダーゼの欠損は遺伝的な溶血性貧血 (O M I M E n t r y 2 6 6 1 2 0) を引き起こすことがある。

20

30

【0011】

ADPリボシル化は、ADPリボース部分が - NADからアルギニンまたはシステインなどの標的アミノ酸に移される可逆的な翻訳後タンパク質修飾である。ADPリボシルアルギニンヒドロラーゼはタンパク質からADPリボースを除去してアルギニンを再生することによりADPリボシル化サイクルを完了する (M o s s , J . 他 (1 9 9 7) A d v . E x p . M e d . B i o l . 4 1 9 : 2 5 - 3 3) 。ADPリボシル化はバクテリア毒素でよく知られた反応である。たとえば、コレラ毒素は興奮性Gタンパク質のサブユニットをADPリボシル化することによりアデニルシクラーゼ系を乱して細胞内cAMPの増加を引き起こす (M o s s , J . および V a u g h a n , M . (e d s) (1 9 9 0) A D P - r i b o s y l a t i n g T o x i n s a n d G - P r o t e i n s : I n s i g h t s i n t o S i g n a l T r a n s d u c t i o n , A m e r i c a n S o c i e t y f o r M i c r o b i o l o g y , W a s h i n g t o n , D . C .) 。また、ADPリボシル化は、真核生物において調節機能を持ち、細胞毒性T細胞の細胞増殖 (W a n g , J . 他 (1 9 9 6) J . I m m u n o l . 1 5 6 : 2 8 1 9 - 2 8 2 7) および細胞骨格アセンブリ (Z h o u , H . 他 (1 9 9 6) A r c h . B i o c h e m . B i o p h y s . 3 3 4 : 2 1 4 - 2 2 2) に影響している可能性がある。

40

【0012】

新規のヒドロラーゼ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。

50

この新規の組成物は、免疫系の疾患、免疫不全、神経の疾患、肺の疾患、および癌を含む細胞増殖異常の診断・治療・予防において有用であり、また、ヒドロラーゼ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影響についての評価にも有用である。

【0013】

(発明の概要)

本発明は、総称して「HYDR」、個別にはそれぞれ「HYDR-1」、「HYDR-2」、「HYDR-3」、および「HYDR-4」と呼ぶヒドロラーゼである精製されたポリペプチドを提供する。或る実施態様において本発明は、(a) SEQ ID NO: 1-4 からなる群から選択したアミノ酸配列からなるポリペプチド、(b) SEQ ID NO: 1-4 からなる群から選択したアミノ酸配列と少なくとも90%が同一である天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1-4 からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または(d) SEQ ID NO: 1-4 からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片を含む群から選択した単離されたポリペプチドを提供する。

10

一実施態様では、SEQ ID NO: 1-4 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。

【0014】

また、本発明は(a) SEQ ID NO: 1-4 からなる群から選択したアミノ酸配列からなるポリペプチド、(b) SEQ ID NO: 1-4 からなる群から選択したアミノ酸配列と少なくとも90%が同一である天然のアミノ酸配列からなるポリペプチド、(c) SEQ ID NO: 1-4 からなる群から選択したアミノ酸配列からなるポリペプチドの生物学的活性断片、または(d) SEQ ID NO: 1-4 からなる群から選択したアミノ酸配列からなるポリペプチドの免疫原性断片、からなる群から選択されたポリペプチドをコードする、単離されたポリヌクレオチドを提供する。

20

一実施態様では、ポリヌクレオチドはSEQ ID NO: 1-4 からなる群から選択したポリペプチドをコードする。別の実施態様では、ポリヌクレオチドはSEQ ID NO: 5-8 からなる群から選択される。

【0015】

本発明は更に、(a) SEQ ID NO: 1-4 からなる群から選択したアミノ酸配列を含むポリペプチド、(b) SEQ ID NO: 1-4 からなる群から選択したアミノ酸配列と少なくとも90%の相同性のある天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1-4 からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または(d) SEQ ID NO: 1-4 からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片、からなる群から選択したポリペプチドをコードするようなポリヌクレオチドと機能的に連結したプロモーター配列を有する組換えポリヌクレオチドを提供する。一実施態様では、本発明は組換えポリヌクレオチドを用いて形質転換した細胞を提供する。別の実施態様では、本発明はこの組換えポリヌクレオチドを含む遺伝形質転換体を提供する。

30

【0016】

また、本発明は、(a) SEQ ID NO: 1-4 からなる群から選択したアミノ酸配列を有するポリペプチド、(b) SEQ ID NO: 1-4 からなる群から選択したアミノ酸配列と少なくとも90%が同一である天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1-4 からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、および(d) SEQ ID NO: 1-4 からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片から成る群から選択されたポリペプチドを製造する方法を提供する。

40

製造方法は、(a) 組換えポリヌクレオチドを用いて形質転換した細胞をポリペプチドの発現に適した条件下で培養する過程と、(b) そのように発現したポリペプチドを回収する過程とを有し、組換えポリヌクレオチドはポリペプチドをコードするポリヌクレオチドに機能的に連結したプロモーター配列を有する。

50

【0017】

本発明は更に、(a) SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列を有するポリペプチド、(b) SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列と少なくとも90%同一である天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、および(d) SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片から構成される群から選択されたポリペプチドに特異結合するような単離された抗体を提供する。

【0018】

本発明は更に、(a) SEQ ID NO: 5 - 8 からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド、(b) SEQ ID NO: 5 - 8 からなる群から選択したポリヌクレオチド配列と少なくとも90%の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、(c) (a) に相補的なポリヌクレオチド、(d) (b) に相補的なポリヌクレオチド、および(e) (a) ~ (d) のRNA等価物からなる群から選択された単離されたポリヌクレオチドを提供する。一実施態様では、ポリヌクレオチドは少なくとも60の連続したヌクレオチドを含む。 10

【0019】

本発明は更に、サンプル中の標的ポリヌクレオチドを検出する方法を提供する。ここで、標的ポリヌクレオチドは(a) SEQ ID NO: 5 - 8 からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド、(b) SEQ ID NO: 5 - 8 からなる群から選択したポリヌクレオチド配列と少なくとも90%の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、(c) (a) に相補的なポリヌクレオチド、(d) (b) に相補的なポリヌクレオチド、および(e) (a) ~ (d) のRNA等価物からなる群から選択されたポリヌクレオチドを含む。検出方法は、(a) サンプル中の標的ポリヌクレオチドに相補的な配列を含む少なくとも20の連続したヌクレオチドを含むプローブを用いて該サンプルをハイブリダイズする過程と、(b) ハイブリダイゼーション複合体の存在・不存在を検出し、複合体が存在する場合にはオプシオンでその量を検出する過程からなり、プローブと標的ポリヌクレオチドあるいはその断片の間でハイブリダイゼーション複合体が形成されるような条件下で、プローブは標的ポリヌクレオチドに特異的にハイブリダイズする。一実施態様では、プローブは少なくとも60の連続したヌクレオチドを含む。 20 30

【0020】

本発明はまた、サンプル中の標的ポリヌクレオチドを検出する方法を提供する。ここで、標的ポリヌクレオチドは(a) SEQ ID NO: 5 - 8 からなる群から選択したポリヌクレオチド配列を有するポリヌクレオチド、(b) SEQ ID NO: 5 - 8 からなる群から選択したポリヌクレオチド配列と少なくとも90%の相同性を有する天然のポリヌクレオチド配列を有するポリヌクレオチド、(c) (a) に相補的なポリヌクレオチド、(d) (b) に相補的なポリヌクレオチド、または(e) (a) ~ (d) のRNA等価物を含む群から選択されたポリヌクレオチドの配列を有する。検出方法は、(a) ポリメラーゼ連鎖反応増幅を用いて標的ポリヌクレオチドまたはその断片を増幅する過程と、(b) 標的ポリヌクレオチドまたはその断片の存在・不存在を検出し、該標的ポリヌクレオチドまたはその断片が存在する場合にはオプシオンでその量を検出する過程を含む。 40

【0021】

本発明は更に、有効量のポリペプチドと薬剤として許容できる賦形剤とを含む成分を提供する。有効量のポリペプチドは、(a) SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列を有するポリペプチド、(b) SEQ ID NO: 1 - 4 を有する群から選択したアミノ酸配列と少なくとも90%の相同性を有する天然のアミノ酸配列を有するポリペプチド、(c) SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または(d) SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片を含む。一実施例で 50

は、SEQ ID NO: 1 - 4 からなる一群から選択されたアミノ酸配列を含む組成物を提供する。更に、本発明は、患者にこの組成物を投与することを含む、機能的ヒドロラーゼ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。

【0022】

本発明はまた、(a) SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列、(b) SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列と少なくとも90%の同一性を有する天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、および(d) SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片からなる群から選択されたポリペプチドのアゴニストとしての有効性を確認するために化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a) ポリペプチドを有するサンプルを化合物に曝す過程と、(b) サンプル中のアゴニスト活性を検出する過程とを含む。別法では、本発明は、この方法によって同定されたアゴニスト化合物と好適な医薬用賦形剤とを含む組成物を提供する。更なる別法では、本発明は、この組成物の患者への投与を含む、機能的HYDR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

【0023】

本発明は更に、(a) SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列を含むポリペプチド、(b) SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列と少なくとも90%の同一性を有する天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、および(d) SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片からなる群から選択されたポリペプチドのアンタゴニストとしての有効性を確認するために化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a) ポリペプチドを含むサンプルを化合物に曝す過程と、(b) サンプル中のアンタゴニスト活性を検出する過程とを含む。一実施態様で本発明は、この方法によって同定したアンタゴニスト化合物と薬剤として許容できる賦形剤とを含む成分を提供する。更なる別法では、本発明は、このような治療を必要とする患者へのこの組成物の投与を含む、機能的HYDR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

30

【0024】

本発明は更に、(a) SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列を含むポリペプチド、(b) SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列と少なくとも90%の同一性を有する天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または(d) SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片、からなる群から選択されたポリペプチドに特異結合する化合物をスクリーニングする方法を提供する。

スクリーニング方法は、(a) ポリペプチドを適切な条件下で少なくとも1つの試験化合物に混合させる過程と、(b) 試験化合物とのポリペプチドの結合を検出し、それによってポリペプチドに特異結合する化合物を同定する過程とを含む。

40

【0025】

本発明は更に、(a) SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列を含むポリペプチド、(b) SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列と少なくとも90%の相同性を有する天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列を含むポリペプチドの生物学的活性断片、および(d) SEQ ID NO: 1 - 4 からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片からなる群から選択されたポリペプチドの活性を調節する化合物をスクリーニングする方法を提供する。

スクリーニング方法は、(a) ポリペプチドの活性が許容された条件下で、ポリペプチド

50

を少なくとも1つの試験化合物と混合させる過程と、(b)ポリペプチドの活性を試験化合物の存在下で算定する過程と、(c)試験化合物の存在下でのポリペプチドの活性を試験化合物の不存在下でのポリペプチドの活性と比較する過程とを含み、試験化合物の存在下でのポリペプチドの活性の変化は、ポリペプチドの活性を調節する化合物であることを意味する。

【0026】

本発明は更に、標的ポリヌクレオチドの変異発現の有効性を確認するために化合物をスクリーニングする方法を提供する。標的ポリヌクレオチドは、SEQ ID NO: 5 - 8を有する群から選択した配列を含む。スクリーニング方法は、(a)標的ポリヌクレオチドを含むサンプルを化合物に曝す過程と、(b)標的ポリヌクレオチドの変異発現を検出する過程とを含む。

10

【0027】

本発明は更に、(a)核酸を含む生物学的サンプルを試験化合物で処理する過程と、(b)(i)SEQ ID NO: 5 - 8からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド、(ii)SEQ ID NO: 5 - 8からなる群から選択したポリヌクレオチド配列と少なくとも90%の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、(iii)(i)に相補的な配列を有するポリヌクレオチド、(iv)(ii)のポリヌクレオチドに相補的なポリヌクレオチド、(v)(i)~(iv)のRNA等価物からなる群から選択したポリヌクレオチドの少なくとも20の連続したヌクレオチドを含むプローブを用いて、処理した生物学的サンプルの核酸をハイブリダイズする過程とを含む試験化合物の毒性の算定方法を提供する。ハイブリダイゼーションは、上記プローブと生物学的サンプル中の標的ポリヌクレオチドの間に特定のハイブリダイゼーション複合体が形成されるような条件下で発生し、上記標的ポリヌクレオチドは、(i)SEQ ID NO: 5 - 8からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド、(ii)SEQ ID NO: 5 - 8からなる群から選択したポリヌクレオチド配列と少なくとも90%の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、(iii)(i)のポリヌクレオチドに相補的な配列を有するポリヌクレオチド、(iv)(ii)のポリヌクレオチドに相補的なポリヌクレオチド、および(v)(i)~(iv)のRNA等価物からなる群から選択する。或いは、標的ポリヌクレオチドは、上記(i)~(v)からなる群から選択したポリヌクレオチド配列の断片と、(c)ハイブリダイゼーション複合体の量を定量する過程と、(d)処理した生物学的サンプルのハイブリダイゼーション複合体の量を、非処理の生物学的サンプルのハイブリダイゼーション複合体の量と比較する過程を含み、処理した生物学的サンプルのハイブリダイゼーション複合体の量の差は、試験化合物の毒性を示す。

20

30

(発明を実施するための形態)

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

40

【0028】

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

【0029】

本明細書中で用いる全ての専門用語及び科学用語は、特に定義されている場合を除き、当業者に一般に理解されている意味と同じ意味を有する。本明細書で説明するものと類似あるいは同等の任意の装置、材料及び方法を用いて本発明の実施または試験を行うことがで

50

きるが、ここでは好適な装置、材料、方法について説明する。本発明で言及する全ての刊行物は、刊行物中で報告されていて且つ本発明に関係があるであろう細胞株、プロトコル、試薬及びベクターについて説明及び開示する目的で引用しているものである。本明細書のいかなる開示内容も、本発明が先行技術の効力によってこのような開示に対して先行する権利を与えられていないことを認めるものではない。

定義

用語「HYDR」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HYDRのアミノ酸配列を指す。

【0030】

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

【0031】

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

【0032】

HYDRをコードする「変異(altered)」核酸配列は、種々のヌクレオチドを欠失、挿入または置換する核酸配列を有し、HYDRと同一またはHYDRの機能的特徴を少なくとも1つ有するポリペプチドを産出する。この定義には、HYDRをコードするポリヌクレオチド配列の正常な染色体の遺伝子座ではない位置での対立遺伝子変異配列との不適當或いは予期しないハイブリダイゼーション、並びにHYDRをコードするポリヌクレオチドの特定のオリゴヌクレオチドプローブを用いて容易に検出可能な或いは検出困難な多型性を含む。コードされたタンパク質も変異され得り、サイレント変化を生じHYDRと機能的に等価となるような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HYDRの活性が保持される範囲で、残基の極性、電荷、溶解度、疎水性、親水性、及び/または両親媒性についての類似性に基づいて成され得る。例えば、負に帯電したアミノ酸にはアスパラギン酸及びグルタミン酸があり、正に帯電したアミノ酸にはリジン及びアルギニンがある。親水性値が近似している非荷電極性側鎖を有するアミノ酸には、アスパラギンとグルタミン、セリンとスレオニンがある。親水性値が近似している非荷電側鎖を有するアミノ酸には、ロイシンとイソロイシンとバリン、グリシンとアラニン、フェニルアラニンとチロシンがある。

【0033】

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

【0034】

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

【0035】

用語「アンタゴニスト」は、HYDR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、HYDRに直接相互作用するか、或いはHYDRが関与する生物学的

10

20

30

40

50

経路の成分と作用して、HYDR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。

【0036】

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

10

【0037】

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

【0038】

「アプタマー」の語は特定の分子標的に結合する核酸またはオリゴヌクレオチド分子を指す。アプタマーは*in vitro*の進化プロセスから誘導される(たとえば米国特許番号第5,270,163号に記載のSELEX(Systematic Evolution of Ligands by EXponential Enrichment))。このプロセスでは大きな組み合わせライブラリから標的に特異的なアプタマーの配列が選択される。アプタマーの組成は二本鎖または一本鎖で、デオキシリボヌクレオチド、リボヌクレオチド、ヌクレオチド誘導體またはその他のヌクレオチド様分子を含むことがある。アプタマーのヌクレオチド成分の糖基を修飾することにより(たとえばリボヌクレオチドの2'-OH基を2'-Fまたは2'-NH₂で置換)、核酸分解酵素への耐性を改善したり、血液中における寿命を延ばすなど、望ましい改良を加えることができる。循環系からアプタマーが除去される速度を遅くするために、アプタマーを高分子量担体等の分子に抱合させることができる。アプタマーは、たとえば光活性化または架橋剤によって各々のリガンドと特異的に架橋させることができる。(Brody, E.N. および L. Gold (2000) J. Biotechnol. 74:5-13等を参照。)「intramer」の語は*in vivo*で発現されるアプタマーを意味する。たとえば、ワクシニアウイルスに基づくRNA発現系を使って、白血球の細胞質で特定のRNAアプタマーが高レベルで発現されている(Blind, M. 他(1999) Proc. Natl Acad. Sci. USA 96:3606-3610)。

20

30

【0039】

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

40

【0040】

「アンチセンス」の語は、特定の核酸配列の「センス」(コーディング)鎖と塩基対を形成することが可能な任意の成分を指す。アンチセンス成分には、DNAや、RNAや、ペプチド核酸(PNA)や、ホスホロチオ酸、メチルホスホン酸またはベンジルホスホン酸等の修飾されたバックボーン連鎖を有するオリゴヌクレオチドや、2'-メトキシエチル糖または2'-メトキシエトキシ糖等の修飾された糖類を有するオリゴヌクレオチドや、或いは5-メチルシトシン、2-デオキシウラシルまたは7-デアザ-2'-デオキシグアノシン等の修飾された塩基を有するオリゴヌクレオチドが含まれうる。アンチセンス分

50

子は、化学合成または転写を含む任意の方法で製造することができる。相補的アンチセンス分子は、ひとたび細胞に導入されたら、細胞が形成した天然の核酸配列と塩基対を形成し、転写または翻訳を妨害する二重鎖を形成する。「負」若しくは「マイナス(-)」の語が参照DNA分子のアンチセンス鎖を、「正」若しくは「プラス(+)」がセンス鎖を指すことがある。

【0041】

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

10

【0042】

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

【0043】

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

20

【0044】

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

30

【0045】

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

元の残基	保存的な置換	
Ala	Gly, Ser	
Arg	His, Lys	
Asn	Asp, Gln, His	
Asp	Asn, Glu	
Cys	Ala, Ser	10
Gln	Asn, Glu, His	
Glu	Asp, Gln, His	
Gly	Ala	
His	Asn, Arg, Gln, Glu	
Ile	Leu, Val	
Leu	Ile, Val	20
Lys	Arg, Gln, Glu	
Met	Leu, Ile	
Phe	His, Met, Leu, Trp, Tyr	
Ser	Cys, Thr	
Thr	Ser, Val	
Trp	Phe, Tyr	30
Tyr	His, Phe, Trp	
Val	Ile, Leu, Thr	

【 0 0 4 6 】

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

【 0 0 4 7 】

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

【 0 0 4 8 】

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

。

【0049】

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

【0050】

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

【0051】

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

【0052】

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

【0053】

SEQ ID NO: 5 - 8の断片には、固有のポリヌクレオチド配列領域が含まれる。この領域は、SEQ ID NO: 5 - 8を特異的に同定するものであり、例えばこの断片を得たゲノム中のその他の配列とは異なるものである。SEQ ID NO: 5 - 8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: 5 - 8を関連ポリヌクレオチド配列から区別する類似の方法に有用である。SEQ ID NO: 5 - 8の断片の正確な長さ及び断片に対応するSEQ ID NO: 5 - 8の領域は、断片に対する意図した目的に基づき当業者が慣例的に決定することが可能である。

【0054】

SEQ ID NO: 1 - 4の断片は、SEQ ID NO: 5 - 8の断片によってコードされる。SEQ ID NO: 1 - 4の断片には、SEQ ID NO: 1 - 4を特異的に同定する固有のアミノ酸配列領域が含まれている。例えば、SEQ ID NO: 1 - 4の断片は、SEQ ID NO: 1 - 4を特異認識する抗体を産出するための免疫原性ペプチドとして有用である。SEQ ID NO: 1 - 4の断片及び断片に対応するSEQ ID NO: 1 - 4の領域の正確な長さは、断片に対する意図した目的に基づき当業者が慣例的に決定することが可能である。

【0055】

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

【0056】

「相同性」の語は、2つ以上のポリヌクレオチド配列または2つ以上のポリペプチド配列

10

20

30

40

50

の配列間での配列類似性、または配列同一性を意味する。

【0057】

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

【0058】

ポリヌクレオチド配列間の一致率は、MEGALIGN version 3.12e 配列アラインメントプログラムに組込まれているようなCLUSTAL Vアルゴリズムのデフォルトのパラメータを用いて決定できる。このプログラムは、LASERGENE ソフトウェアパッケージ(一組の分子生物学的分析プログラム)(DNASTAR, Madison WI)の一部である。このCLUSTAL Vは、Higgins, D.G. 及び P.M. Sharp (1989) CABIOS 5:151-153、Higgins, D.G. 他 (1992) CABIOS 8:189-191に記載されている。ポリヌクレオチド配列を2つ1組でアラインメントする際のデフォルトパラメータは、Ktuple=2、gap penalty=5、window=4、「diagonals saved」=4と設定する。「重み付けされた」残基重み付け表が、デフォルトとして選択された。CLUSTAL Vは、アラインメントされたポリヌクレオチド配列対間の「類似率」として一致率を報告する。

【0059】

或いは、米国国立バイオテクノロジー情報センター(NCBI)のBasic Local Alignment Search Tool (BLAST)が一般的に用いられ、且つ、無料で入手可能な配列比較アルゴリズム一式を提供している(Altschul, S.F. ら (1990) J. Mol. Biol. 215:403-410)。このアルゴリズムは、幾つかの情報源から入手可能であり、メリーランド州ベセスダにあるNCBI及びインターネット(<http://www.ncbi.nlm.nih.gov/BLAST/>)からも入手可能である。BLASTソフトウェア一式には様々な配列分析プログラムが含まれており、既知のポリヌクレオチド配列を種々のデータベースから得た別のポリヌクレオチド配列とアラインメントする「blastn」もその1つである。その他にも、2つのヌクレオチド配列を対で直接比較するために用いる「BLAST 2 Sequences」と称されるツールも利用可能である。「BLAST 2 Sequences」は、<http://www.ncbi.nlm.nih.gov/gorf/bl2.html>にアクセスして対話形式で利用することが可能である。「BLAST 2 Sequences」ツールは、blastn と blastp (後述)の両方に用いることができる。BLASTプログラムは、一般的には、ギャップ及びデフォルト設定に設定された他のパラメータと共に用いる。例えば、2つのヌクレオチド配列を比較するために、デフォルトパラメータとして設定された「BLAST 2 Sequences」ツールVersion 2.0.12(2000年4月21日)を用いてblastnを実行してもよい。デフォルトパラメータの設定例を以下に示す。

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

Open Gap: 5 及び Extension Gap: ペナルティ 2

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

10

20

30

40

50

【 0 0 6 0 】

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

【 0 0 6 1 】

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

【 0 0 6 2 】

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

【 0 0 6 3 】

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

【 0 0 6 4 】

或いは、NCBI BLASTソフトウェア一式を用いてもよい。例えば、2つのポリペプチド配列を対で比較をする場合、ある者は、デフォルトパラメータで設定された「BLAST 2 Sequences」ツールVersion 2.0.12（Apr - 21 - 2000）でblastpを使用するであろう。デフォルトパラメータの設定例を以下に示す。

Matrix: BLOSUM62

Open Gap: 11 及び Extension Gap : ペナルティ 1

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

【 0 0 6 5 】

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

10

20

30

40

50

対して一致率を測定し得る長さを記述し得ることを理解されたい。

【0066】

「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、染色体の複製、分離及び維持に必要な全てのエレメントを含む直鎖状の小染色体である。

【0067】

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

【0068】

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

【0069】

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

【0070】

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

【0071】

「ハイブリダイゼーション複合体」の語は、相補的塩基対間の水素結合の形成力によって2つの核酸配列間に形成された複合体を指す。ハイブリダイゼーション複合体は、溶解状態で形成し得る (C_0 または R_0 解析等)。或いは、一方の核酸配列が溶解状態で

10

20

30

40

50

存在し、もう一方の核酸配列が固体支持体（例えば紙、膜、フィルタ、チップ、ピンまたはガラススライド、或いは他の適切な基板であって細胞若しくはその核酸が固定される基板）に固定されているような2つの核酸配列間に形成され得る。

【0072】

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

【0073】

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

10

【0074】

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HYDR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HYDRのポリペプチド断片またはオリゴペプチド断片を含む。

【0075】

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

【0076】

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

20

【0077】

用語「調節」は、HYDRの活性の変化を指す。例えば、調節によって、HYDR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。

【0078】

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

30

【0079】

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

【0080】

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

40

【0081】

HYDRの「翻訳後修飾」は、脂質化、グリコシル化、リン酸化、アセチル化、ラセミ化、タンパク分解性分割及びその他の当分野で既知の修飾を含み得る。これらのプロセスは、合成的または生化学的に発生し得る。生化学的修飾は、HYDRの酵素環境に依存し

50

、細胞の種類によって異なり得る。

【0082】

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

10

【0083】

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

【0084】

プローブ及びプライマーの調製及び使用方法については、Sambrook, J. 他(1989) Molecular Cloning: A Laboratory Manual, 第2版, 1-3巻, Cold Spring Harbor Press, Plainview NY、Ausubel, F.M. 他, (1987) Current Protocols in Molecular Biology, Greene Publi. Assoc. & Wiley-Intersciences, New York NY、Innis他(1990) PCR Protocols, A Guide to Methods and Applications Academic Press, San Diego CA等を参照されたい。PCRプライマー対は、その目的のためのコンピュータプログラム、例えばPrimer(Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA)を用いるなどして既知の配列から得ることができる。

20

30

【0085】

プライマーとして用いるオリゴヌクレオチドの選択は、そのような目的のために本技術分野でよく知られているソフトウェアを用いて行う。例えばOLIGO 4.06ソフトウェアは、各100ヌクレオチドまでのPCRプライマー対の選択に有用であり、オリゴヌクレオチド及び最大5,000ヌクレオチドまでの大きめのポリヌクレオチドであって32キロベースまでのインプットポリヌクレオチド配列から得たものを分析するのにも有用である。類似のプライマー選択プログラムには、拡張能力のための追加機能が組み込まれている。例えば、PrimOUプライマー選択プログラム(テキサス州ダラスにあるテキサス大学南西部医療センターのゲノムセンターから一般向けに入手可能)は、メガベース配列から特定のプライマーを選択することが可能であり、従ってゲノム全体の範囲でプライマーを設計するのに有用である。Primer3プライマー選択プログラム(マサチューセッツ州ケンブリッジのWhitehead Institute/MITゲノム研究センターから一般向けに入手可能)ではユーザーが「ミスプライミング・ライブラリ」をインプットすることができ、ここでプライマー結合部位として避けたい配列はユーザーが指定する。Primer3は特に、マイクロアレイのためのオリゴヌクレオチドの選択に有用である。(後二者のプライマー選択プログラムのソースコードは、各自のソースから得てユーザー固有のニーズを満たすように変更してもよい。)PrimerGenプログラム(英国ケンブリッジ市の英国ヒトゲノムマッピングプロジェクト・リソースセンターか

40

50

ら一般向けに入手可能)は、多数の配列アラインメントに基づいてプライマーを設計し、それによって、アラインメントされた核酸配列の最大保存領域または最小保存領域の何れかとハイブリダイズするようなプライマーの選択を可能にする。従って、このプログラムは、固有な、および保存されたオリゴヌクレオチド及びポリヌクレオチドの断片の同定に有用である。上記選択方法のいずれかによって同定したオリゴヌクレオチド及びポリヌクレオチドの断片は、ハイブリダイゼーション技術において、例えばPCRまたはシーケンシングプライマーとして、マイクロアレイエレメントとして、或いは核酸のサンプルにおいて完全または部分的相補的ポリヌクレオチドを同定する特異プローブとして有用である。オリゴヌクレオチドの選択方法は、上記の方法に限定されるものではない。

【0086】

10

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

【0087】

20

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

【0088】

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

【0089】

30

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

【0090】

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

【0091】

40

「サンプル」の語は、その最も広い意味で用いられる。HYDR、HYDR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。

【0092】

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

50

【0093】

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

【0094】

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

【0095】

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

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

【0096】

「形質転換」は、外来性のDNAが受入細胞に入り込むプロセスを表す。形質転換は、本技術分野で知られている種々の方法に従って自然条件または人工条件下で生じ得るものであり、外来性の核酸配列を原核または真核宿主細胞に挿入する任意の既知の方法を基にし得る。形質転換の方法は、形質転換する宿主細胞の種類によって選択する。限定するものではないが形質転換方法には、バクテリオファージまたはウイルス感染、電気穿孔法（エレクトロポレーション）、熱ショック、リポフェクション及び微粒子銃を用いる方法がある。「形質転換された」細胞には、導入されたDNAが自律的に複製するプラスミドとして或いは宿主染色体の一部として複製可能である安定的に形質転換された細胞が含まれる。さらに、限られた時間に一過的に導入DNA若しくは導入RNAを発現する細胞も含まれる。

【0097】

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

【0098】

特定の核酸配列の「変異体」は、核酸配列1本全部の長さに対して特定の核酸配列と少なくとも40%の相同性を有する核酸配列であると定義する。その際、デフォルトパラメータに設定した「BLAST 2 Sequences」ツールVersion 2.0.9（1999年5月7日）を用いてblastnを実行する。このような核酸対は、所定の長さに対して、例えば少なくとも50%、60%、70%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%またはそれ以上の相同性を示し得る。或る変異体は、例えば「対立遺伝子」変異体（前述）、「スプライス」変異体、「種」変異体または「多型性」変異体として説明し得る。スプライス変異体は参照分子とかなりの相同性を有し得るが、mRNAプロセッシング中のエキソンの

交互スプライシングによって通常、ポリヌクレオチドがより多くまたはより少数の塩基を有することになる。対応するポリペプチドは、追加機能ドメインを有するか或いは参照分子に存在するドメインが欠落していることがある。種変異体は、種相互に異なるポリヌクレオチド配列である。結果的に生じるポリペプチドは通常、相互にかなりのアミノ酸相同性を有する。多型性変異体は、与えられた種の個体間で特定の遺伝子のポリヌクレオチド配列が異なる。また、多型性変異体は、1つのヌクレオチド塩基によってポリヌクレオチド配列が変化する「1塩基多型性」(SNP)を含み得る。SNPの存在は、例えば特定の個体群、病状または病状性向を示し得る。

【0099】

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

発明

本発明は、新規のヒトヒドロラーゼ(HYDR)及びHYDR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免疫系の疾患、免疫不全、神経の疾患、肺の疾患、および癌を含む細胞増殖異常の診断、治療、及び予防に関する。

【0100】

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

【0101】

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

【0102】

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

10

20

30

40

50

【0103】

表2及び3は共に、本発明の各々のポリペプチドの特性を要約しており、それら特性が請求の範囲に記載されたポリペプチドがヒドロラーゼであることを確立している。例えば、SEQ ID NO: 1はシロイヌナズナの推定上のアデノシンデアミナーゼ (GenBank ID g7267246) と38%の同一性を有することがBasic Local Alignment Search Tool (BLAST) によって示された。(表2参照) BLAST確率スコアは 2.6×10^{-59} であり、これは観察されたポリペプチド配列アラインメントが偶然に得られる確率を示している。SEQ ID NO: 1もまた、アデノシン/AMPデアミナーゼドメインを有するが、これは、隠れマルコフモデル (HMM) を基にした保存されたタンパク質ファミリドメインのPFAMデータベースにおいて、統計的に有意な一致を検索して決定された。(表3参照) BLIMP5解析よりのデータは、SEQ ID NO: 1がデアミナーゼである、さらに実証的な証拠を提供する。別の例において、SEQ ID NO: 2は Xylella fastidiosa のエンドヌクレアーゼV (デオキシノシン3'エンドヌクレアーゼ) (GenBank ID g9106985) に40%の同一性を有するが、これはBasic Local Alignment Search Tool (BLAST) によって決定された。(表2参照) BLAST確率スコアは 2.0×10^{-28} であり、これは観察されたポリペプチド配列アラインメントが偶然に得られる確率を示している。BLIMP5解析よりのデータは、SEQ ID NO: 2がエンドヌクレアーゼである、さらに実証的な証拠を提供する。別の例として、SEQ ID NO: 3はヒトのADPリボシルアルギニンヒドロラーゼ (GenBank ID g402478) と46%の同一性を有することがBasic Local Alignment Search Tool (BLAST) によって示された。(表2参照) BLAST確率スコアは 1.2×10^{-79} であり、これは観察されたポリペプチド配列アラインメントが偶然に得られる確率を示している。SEQ ID NO: 3はADPリボシルアルギニンヒドロラーゼに特徴的なドメインを含んでいるが、これはRPODOMデータベースで見出された。(表3参照) 同様に、SEQ ID NO: 4はハトの5'ヌクレオチダーゼ (GenBank ID g4902474) と85%の同一性を持っている。BLAST確率スコアは 3.1×10^{-161} である。これらのデータはSEQ ID NO: 3およびSEQ ID NO: 4がヒドロラーゼであるという証拠を提供する。SEQ ID NO: 1-4の解析のためのアルゴリズム及びパラメータが表7で記述されている。

10

20

30

【0104】

表4に示すように、本発明の完全長ポリヌクレオチド配列は、cDNA配列またはゲノムDNA由来のコード (エキソン) 配列を用いて、或いはこれら2種類の配列を任意に組み合わせ構築した。列1および列2はそれぞれ、本発明の各ポリヌクレオチドのポリヌクレオチド配列識別番号 (ポリヌクレオチド SEQ ID NO:) およびそれに対応するIncyteポリヌクレオチドコンセンサス配列番号 (Incyteポリヌクレオチド ID) を示す。列3は、各ポリヌクレオチド配列の長さ (塩基対単位) を示す。列4は、例えば、SEQ ID NO: 5-8を同定するため、或いはSEQ ID NO: 5-8と関連するポリヌクレオチド配列とを区別するためのハイブリダイゼーションまたは増幅技術に有用なポリヌクレオチド配列の断片を示す。列5はcDNA配列、ゲノムDNAから予想されたコード配列 (エキソン) 及び/またはcDNA及びゲノムDNAを共に有する配列集合に対応する識別番号を示している。これらの配列は、本発明の完全長ポリヌクレオチド配列を構築するのに用いた。表4の列6および列7はそれぞれ、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。

40

【0105】

表4の列5の識別番号は、特に例えばIncyte cDNAとそれに対応するcDNAライブラリを指す場合もある。例えば、3941090F8はIncyte cDNA配列の識別番号であり、SCORNOT04はそれが由来するcDNAライブラリの識別番

50

号である。cDNAライブラリが示されていないIncyte cDNAは、プールされているcDNAライブラリ(例えば、71742519V1)に由来する。または、列5の識別番号は、完全長の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、g2165516)の識別番号の場合もある。また、列5の識別番号は、ENSEMBL(The Sanger Centre, Cambridge, UK)データベースに由来する配列(すなわち、名前に「ENST」を含む配列)の識別番号である場合もある。また、列5の識別番号は、NCBI RefSeqヌクレオチド配列記録データベースに由来する配列(すなわち、名前に「NM」または「NT」を含む配列)あるいはNCBI RefSeqタンパク質配列記録に由来する配列(すなわち、名前に「NP」を含む配列)の識別番号である場合もある。または列5の識別番号は、「エキソンスティック(exon-stitching)」アルゴリズムにより結び合わせたcDNA及びGenscan予想エキソンの両方の集合を指す場合もある。例えば、FL_XXXXXX_N1_N2_YYYYY_N3_N4はアルゴリズムが適用される配列のクラスターの識別番号がXXXXXXXXであり、アルゴリズムにより生成される予測の番号がYYYYYであり、(もし存在すれば)N₁, N₂, N₃...が解析中に手動で編集された可能性のある特定のエキソンであるような「縫合された」配列である(実施例5参照)。または列5の識別番号は、「エキソストレッチ(exon-stretching)」アルゴリズムにより結び合わせたエキソンの集合を指す場合もある。例えば、FLXXXXXXXX_gAAAAA_gBBBBB_1_Nは「ストレッチ」配列の識別番号である。ここでXXXXXXXXはIncyteプロジェクト識別番号、gAAAAAは「エキソストレッチ」アルゴリズムを適用したヒトゲノム配列のGenBank識別番号、gBBBBBは一番近いGenBankタンパク質相同体のGenBank識別番号またはNCBI RefSeq識別番号である。(実施例5を参照。)RefSeq配列が「エキソストレッチ」アルゴリズム用のタンパク質相同体として津筒されているときは、RefSeq識別子(「NM」、「NP」、または「NT」で表示)をGenBank識別し(たとえばgBBBBB)の代わりに使うことができる。

10

20

【0106】

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

30

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

10

20

30

【0107】

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

【0108】

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

40

【0109】

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

【0110】

本発明はまた、HYDR をコードするポリヌクレオチドを提供する。或る例では、HYDR をコードする SEQ ID NO: 5 - 8 からなる群から選択された配列を有するポリヌクレオチド配列が本発明に含まれている。SEQ NO ID: 5 - 8 のポリヌクレオチド配列は、配列表に示されているように等価 RNA 配列と同等の価値を有しているが、窒素

50

塩基チミンの出現はウラシルに置換され、糖のバックボーンはデオキシリボースではなくシリボースから構成されている。

【0111】

本発明はまた、HYDR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、HYDR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: 5 - 8 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: 5 - 8 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、HYDR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

【0112】

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

20

【0113】

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

【0114】

本発明はまた、HYDR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、当分野でよく知られている試薬を用いて、この合成配列を任意の様々な入手可能な発現ベクター及び細胞系中に挿入し得る。

更に、合成化学を用いてHYDRまたはその任意の断片をコードする配列に突然変異を誘導し得る。

【0115】

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: 5 - 8 及びそれらの断片とハイブリダイズ可能なポリヌクレオチド配列が含まれる(例えば、Wahl, G. M. 及びS. L. Berger (1987) *Methods Enzymol.* 152: 399 - 407、Kimmel, A. R. (1987) *Methods Enzymol.* 152: 507 - 511等を参照)。アニーリング及び洗浄条件を含むハイブリダイゼーションの条件は、「定義」に記載されている。

40

【0116】

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

50

(Applied Biosystems)、熱安定性T7ポリメラーゼ(Amersham, Pharmacia Biotech, Piscataway NJ)を用いることができる。或いは、例えばELONGASE増幅システム(Life Technologies, Gaithersburg MD)において見られるように、ポリメラーゼと校正エキソヌクレアーゼを併用することができる。好適には、MICROLAB 2200液体転移システム(Hamilton, Reno, NV)、PTC 200サーマルサイクラー(MJ Research, Watertown MA)及びABI CATALYST 800サーマルサイクラー(Applied Biosystems)等の装置を用いて配列の調製を自動化する。次に、ABI 373 或いは 377 DNAシーケンシングシステム(Applied Biosystems)、MEGABACE 1000 DNAシーケンシングシステム(Molecular Dynamics, Sunnyvale CA)または当分野でよく知られている他の方法を用いてシーケンシングを行う。結果として得られた配列を当分野でよく知られている種々のアルゴリズムを用いて分析する。(Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7、Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, 856-853ページ.等を参照)。

10

【0117】

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

20

30

40

【0118】

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

50

【0119】

市販されているキャピラリー電気泳動システムを用いて、シーケンシングまたはPCR産物のサイズを分析し、またはそのヌクレオチド配列を確認することができる。具体的には、キャピラリーシーケンシングは、電気泳動による分離のための流動性ポリマーと、4つの異なるヌクレオチドに特異的であるような、レーザーで活性化される蛍光色素と、放出された波長の検出に利用するCCDカメラとを有し得る。出力/光の強度は、適切なソフトウェア(Applied Biosystems社のGENOTYPER、SEQUENCE NAVIGATOR等)を用いて電気信号に変換し得る。サンプルのロードからコンピュータ分析及び電子データ表示までの全プロセスがコンピュータ制御可能である。キャピラリー電気泳動法は、特定のサンプルに少量しか存在しないようなDNA小断片のシーケンシングに特に適している。 10

【0120】

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

【0121】

種々の目的でHYDRをコードする配列を変えるために、当分野で一般的に知られている方法を用いて、本発明のヌクレオチド配列を組換えることができる。この目的には、遺伝子産物のクローン化、プロセッシング及び/または発現の調節が含まれるが、これらに限定されるものではない。遺伝子断片及び合成オリゴヌクレオチドのランダムなフラグメンテーション及びPCR再アセンブリによるDNAシャッフリングを用い、ヌクレオチド配列を組み換えることが可能である。例えば、オリゴヌクレオチド仲介特定部位突然変異誘導を利用して、新規な制限部位の生成、グリコシル化パターンの変更、コドン優先の変更、スプライス変異体の生成等を行う突然変異を導入し得る。 20

【0122】

本発明のヌクレオチドを、MOLECULAR BREEDING (Maxygen Inc., Santa Clara CA; 米国特許第5,837,458号; Chang, C.-C. 他(1999) Nat. Biotechnol. 17:793-797; Christians, F.C. 他(1999) Nat. Biotechnol. 17:259-264; Cramer, A. 他(1996) Nat. Biotechnol. 14:315-319)などのDNAシャッフリング技術を用いてシャッフリングして、HYDR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HYDRの生物学的特性を変更或いは改良することができる。DNAシャッフリングは、遺伝子断片のPCR仲介再組換えを用いて遺伝子変異体のライブラリを生成するプロセスである。ライブラリはその後、所望の特性を有する遺伝子変異体を同定するための選択またはスクリーニングにかける。次にこれらの好適な変異体をプールし、更に反復してDNAシャッフリング及び選択/スクリーニングを行ってもよい。このように、遺伝の多様性は「人為的」品種改良及び急速な分子の進化を経て創生される。例えば、ランダムポイント突然変異を有する単一の遺伝子の断片を組み換えて、スクリーニングし、その後所望の特性が最適化されるまでシャッフリングすることができる。或いは、所定の遺伝子の断片を同種または異種のいずれかから得た同一遺伝子ファミリーの相同遺伝子と組み換えて、それによって天然に存在する複数の遺伝子の遺伝多様性を、指図された制御可能な方法で最大化させることができる。 30 40

【0123】

別の実施例によれば、SECP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. 50

225 - 232を参照)。別法として、化学的方法を用いてHYDR自体またはその断片を合成することが可能である。例えば、種々の液相または固相技術を用いてペプチド合成を行うことができる(Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, 55 - 60ページ、Roberge, J. Y.ら(1995) Science 269:202 - 204等を参照)。自動合成はABI 431Aペプチドシンセサイザ(Applied Biosystems)を用いて達成し得る。更にHYDRのアミノ酸配列または任意のその一部は、直接的な合成の際の変更、及び/または化学的方法を用いた他のタンパク質または任意のその一部からの配列との組み合わせにより、天然のポリペプチド配列を有するポリペプチドまたは変異体ポリペプチドを作製することが可能である。 10

【0124】

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

【0125】

生物学的に活性なHYDRを発現させるために、HYDR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HYDRをコードするポリヌクレオチド配列におけるエンハンサー、構成型及び発現誘導型のプロモーター、5'及び3'の非翻訳領域などの調節配列が含まれる。このような要素は、長さ及び特異性が様々である。特定の開始シグナルによって、HYDRをコードする配列のより効果的な翻訳を達成することが可能である。このようなシグナルには、ATG開始コドンと、コザック配列などの近傍の配列が含まれる。HYDRをコードする配列及びその開始コドン、上流の調節配列が好適な発現ベクターに挿入された場合は、更なる転写調節シグナルや翻訳調節シグナルは必要なくなるであろう。しかしながら、コーディング配列或いはその断片のみが挿入された場合は、インフレームのATG開始コドンを含む外来性の翻訳調節シグナルが発現ベクターに含まれるようにすべきである。外来性の翻訳要素及び開始コドンは、様々な天然物及び合成物を起源とし得る。用いられる特定の宿主細胞系に好適なエンハンサーを含めることで発現の効率を高めることが可能である。(Scharf, D. ら(1994) Results Probl. Cell Differ. 20:125 - 162.等を参照)。 20 30

【0126】

当業者に周知の方法を用いて、HYDRをコードする配列、好適な転写及び翻訳調節エレメントを含む発現ベクターを作製することが可能である。これらの方法には、in vitro組換えDNA技術、合成技術、及びin vivo遺伝子組換え技術が含まれる。(例えば、Sambrook, J. 他.(1989) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, 4章及び8章, 及び16 - 17章; 及びAusubel, F. M. 他.(1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, 9章、13章および16章を参照)。 40

【0127】

種々の発現ベクター/宿主系を利用して、HYDRをコードする配列の保持及び発現が可能である。限定するものではないがこのような発現ベクター/宿主系には、組換えバクテリオファージ、プラスミドまたはコスミドDNA発現ベクターで形質転換させた細菌や、酵母菌発現ベクターで形質転換させた酵母菌や、ウイルス発現ベクター(例えばバキュロウイルス)に感染した昆虫細胞系や、ウイルス発現ベクター(例えばカリフラワーモザイ 50

クウイルス、CaMVまたはタバコモザイクウイルスTMV)または細菌発現ベクター(例えばTiまたはpBR322プラスミド)で形質転換させた植物細胞系、動物細胞系などの微生物等がある。(前出のSambrook、前出のAusubel、Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509、; Engelhard, E.K. ら (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227、Sandig, V. ら (1996) Hum. Gene Ther. 7:1937-1945、Takamatsu, N. (1987) EMBOJ. 6:307-311、; 『マグローヒル科学技術年鑑』(The McGraw Hill Yearbook of Science and Technology) (1992) McGraw Hill New York NY, 191-196頁、Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659、Harrington, J.J. ら (1997) Nat. Genet. 15:345-355等を参照)レトロウイルス、アデノウイルス、ヘルペスウイルスまたはワクシニアウイルス由来の発現ベクター、または種々の細菌性プラスミド由来の発現ベクターを用いて、ヌクレオチド配列を標的器官、組織または細胞集団へ輸送することができる(Di Nicola, M. ら (1998) Cancer Gen. Ther. 5(6):350-356、Yu, M. 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。

【0128】

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。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、HYDR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、HYDR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(St rat agene, 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HYDRをコードする配列をライゲーションするとlacZ遺伝子が破壊され、組換え分子を含む形質転換された細菌の同定のための比色スクリーニング法が可能となる。更にこれらのベクターは、クローニングされた配列における*in vitro*転写、ジデオキシのシーケンシング、ヘルパーファージによる一本鎖の救出、入れ子状態の欠失の生成にも有用であろう(例えば、Van Heeke, G. および S.M. Schuster (1989) J. Biol. Chem. 264:55035509を参照)。例えば、抗体の産生のためなどに多量のHYDRが必要な場合は、HYDRの発現をハイレベルで誘導するベクターが使用できる。例えば、強力な誘導SP6バクテリオファージプロモーターまたは誘導T7バクテリオファージプロモーターを含むベクターが使用できる。

【0129】

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

【0130】

植物系もHYDRの発現に使用可能である。HYDRをコードする配列の転写は、ウイルスプロモーター、例えば単独或いはTMV (Takamatsu, N. (1987) EMBO J 6:307-311)由来のオメガリーダー配列と組み合わせて用いられるようなCaMV由来の35S及び19Sプロモーターによって促進される。或いは、RUBISCOの小サブユニット等の植物プロモーターまたは熱ショックプロモーターを用いてもよい(例えば、Coruzzi, G. 他(1984) EMBO J. 3:1671-1680; Broglie, R. 他(1984) Science 224:838-843; および Winter, J. 他(1991) Results Probl. Cell Differ. 17:85-105を参照)。これらの構成物は、直接DNA形質転換または病原体を媒介とする形質移入によって、植物細胞内に導入可能である。(『マグローヒル科学技術年鑑』(The McGraw Hill Yearbook of Science and Technology) (1992) McGraw Hill New York NY, 191-196ページ等を参照)。

10

【0131】

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

20

【0132】

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

30

【0133】

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

40

【0134】

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

50

ランスフェラーゼに対する耐性を各々与える (Wigler, M. ら (1980) Proc. Natl. Acad. Sci. USA 77: 3567-3570、Colbere-Garapin, F. ら (1981) J. Mol. Biol. 150: 1-14 等を参照)。この他の選択可能な遺伝子、例えば、代謝のための細胞の必要条件を変える *trpB* 及び *hisD* は、文献に記載されている (Hartman, S. C. および R. C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85: 80478051 を参照。) 可視マーカー、例えばアントシアニン、緑色蛍光タンパク質 (GFP; Clontech)、グルクロニダーゼ及びその基質 グルクロニド、またはルシフェラーゼ及びその基質ルシフェリン等を用いてもよい。これらのマーカーを用いて、トランスフォーマントを特定するだけでなく、特定のベクター系に起因する一過性或いは安定したタンパク質発現を定量することが可能である (Rhodes, C. A. (1995) Methods Mol. Biol. 55: 121131 等を参照)。

10

【0135】

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

20

【0136】

一般に、HYDR をコードする核酸配列を含み且つ HYDR を発現する宿主細胞は、当業者によく知られている種々の方法を用いて同定することが可能である。限定するものではないが当業者によく知られている方法には、DNA-DNA 或いは DNA-RNA ハイブリダイゼーション、PCR 法、核酸或いはタンパク質の検出、定量、或いはその両方を行うための膜系、溶液ベース或いはチップベースの技術を含むタンパク質バイオアッセイまたはイムノアッセイ技術がある。

【0137】

特異的ポリクローナル抗体または特異的モノクローナル抗体を用いて HYDR の発現の検出及び計測を行うための免疫学的方法は、当分野で公知である。このような技術の例としては、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)、蛍光活性化細胞選別 (FACS) などが挙げられる。HYDR 上の2つの非干渉エピトープに反応するモノクローナル抗体を用いた、2部位のモノクローナルベースイムノアッセイ (two-site, monoclonal-based immunoassay) が好ましいが、競合の結合アッセイも用いることもできる。これらのアッセイ及びこれ以外のアッセイは、当分野で公知である (Hampton, R. ら (1990) Serological Methods, a Laboratory Manual, APS Press, St Paul, MN, Sect. IV、Coligan, J. E. ら (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY、Pound, J. D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ 等を参照)。

30

40

【0138】

多岐にわたる標識方法及び結合方法が、当業者に知られており、様々な核酸アッセイおよびアミノ酸アッセイにこれらの方法を用い得る。HYDR 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、HYD

50

Rをコードする配列、またはその任意の断片をmRNAプローブを生成するためのベクターにクローニングすることも可能である。このようなベクターは、当分野において知られており、市販もされており、T7、T3またはSP6等の好適なRNAポリメラーゼ及び標識されたヌクレオチドを加えて、in vitroでRNAプローブの合成に用いることができる。このような方法は、例えばAmersham Pharmacia Biotech、Promega (Madison WI)、U.S. Biochemical等から市販されている種々のキットを用いて実行することができる。検出を容易にするために用い得る好適なレポーター分子或いは標識には、基質、補助因子、インヒビター、磁気粒子のほか、放射性核種、酵素、蛍光剤、化学発光剤、発色剤等がある。

【0139】

HYDRをコードするヌクレオチド配列で形質転換された宿主細胞は、細胞培地でのこのタンパク質の発現及び回収に好適な条件下で培養される。形質転換細胞から製造されたタンパク質が分泌されるか細胞内に留まるかは、使用される配列、ベクター、或いはその両者に依存する。HYDR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HYDR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。

【0140】

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

【0141】

本発明の別の実施例では、HYDR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HYDRタンパク質が、HYDR活性のインヒビターに対するペプチドライブラリのスクリーニングを促進し得る。また、異種タンパク質部分及び異種ペプチド部分も、市販されている親和性基質を用いて融合タンパク質の精製を促進し得る。限定されるものではないがこのような部分には、グルタチオンSトランスフェラーゼ(GST)、マルトース結合タンパク質(MBP)、チオレドキシン(Trx)、カルモジュリン結合ペプチド(CBP)、6-His、FLAG、c-myc、赤血球凝集素(HA)がある。GSTは固定化グルタチオン上で、MBPはマルトース上で、Trxはフェニルアルシンオキシド上で、CBPはカルモジュリン上で、そして6-Hisは金属キレート樹脂上で、同族の融合タンパク質の精製を可能にする。FLAG、c-myc及び赤血球凝集素(HA)は、これらのエピトープ標識を特異的に認識する市販されているモノクローナル抗体及びポリクローナル抗体を用いて、融合タンパク質の免疫親和性精製を可能にする。また、HYDR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、HYDRが精製の後に異種部分から切断され得る。融合タンパク質の発現及び精製方法は、前出のAusubel(1995)10章に記載されている。市販されている種々のキットを用いて融合タンパク質の発現及び精製を促進することもできる。

【0142】

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HYDRの合成が可能である。これらの系は、T7、T3またはSP6プロモーターと機能的に連結したタンパク質コ

10

20

30

40

50

ード配列の転写及び翻訳を結合する。翻訳は、例えば³⁵Sメチオニンのような放射能標識したアミノ酸前駆体の存在下で起こる。

【0143】

本発明のHYDRまたはその断片を用いて、HYDR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、HYDR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。

【0144】

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HYDRの天然のリガンド、または天然の基質、構造的または機能的な擬態性または自然結合パートナーに密接に関連している（Coligan, J. E. 他（1991）Current Protocols in Immunology 1（2）の5章等を参照）。同様に、化合物は、HYDR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HYDRを発現する好適な細胞の作製が含まれる。好適な細胞には、哺乳動物、酵母、ショウジョウバエ、大腸菌からの細胞が含まれる。HYDRを発現する細胞またはHYDRを含有する細胞膜断片を試験化合物と接触させて、HYDRまたは化合物の何れかの結合、刺激または活性の阻害を分析する。

10

20

【0145】

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

【0146】

本発明のHYDRまたはその断片を用いて、HYDR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、HYDRが少なくとも1つの試験化合物と結合する、HYDR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HYDRの活性が試験化合物不在下でのHYDRの活性と比較する。試験化合物の存在下でのHYDRの活性の変化は、HYDR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HYDRの活性に適した条件下でHYDRを含む*in vitro*または無細胞再構成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、HYDR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

40

【0147】

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、HYDRまたはその哺乳動物相同体をコードするポリヌクレオチドを「ロックアウト」する。このような技術は当技術分野において周知であり、ヒト疾患動物モデルの作製に有用である（米国特許第5,175,383号及び第5,767,337号等を参照）。例えば129/SvJ細胞株等のマウスES細胞は初期のマウス胚に由来し、培地で増殖させることができる。このES細胞は、ネオマイシンホスホトランスフェラーゼ遺伝子（neo: Capecchi, M. R. (1989) Science 244: 1288-1292）等のマーカー遺伝子で破壊した目的の遺伝子を含むベクターで形質転換する

50

。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(Martha, J. D. (1996) Clin. Invest. 97:1999-2002; Wagner, K. U. 他(1997) Nucleic Acids Res. 25:4323-4330)。形質転換したES細胞を同定し、例えばC57BL/6マウス系等から採取したマウス細胞胚盤胞に微量注入する。胚盤胞を偽妊娠メスに外科的に導入し、得られるキメラ子孫の遺伝形質を決め、これを交配させてヘテロ接合性系またはホモ接合性系を作製する。このようにして作製した遺伝子組換え動物は、可能性のある治療薬や毒性薬剤で検査することができる。

【0148】

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

【0149】

HYDR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、HYDRをコードするポリヌクレオチドの或る領域を動物ES細胞に注入し、注入した配列を動物細胞ゲノムに組み込ませる。形質転換細胞を胞胚に注入し、胞胚を上記のように移植する。遺伝子組換え子孫または近交系について研究し、可能性のある医薬品を用いて処理し、ヒトの疾患の治療に関する情報を得る。別法では、例えばHYDRを乳汁内に分泌するなどHYDRを過剰に発現する哺乳動物近交系は、便利なタンパク質源となり得る(Janne, J. 他(1998) Biotechnol. Annu. Rev. 4:55-74)。

【0150】

治療

HYDRのある領域とヒドロラーゼ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、HYDRの発現は、腎臓、肺腫瘍および心臓血管組織に密接に関連する。従って、HYDRは、免疫系の疾患、免疫不全、神経の疾患、肺の疾患、および癌を含む細胞増殖異常においてある役割を果たすと考えられる。HYDR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、HYDRの発現または活性を低下させることが望ましい。また、HYDR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、HYDRの発現または活性を亢進させることが望ましい。

【0151】

従って、一実施例において、HYDR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HYDR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免疫系の疾患、免疫不全、神経の疾患、肺の疾患、および癌を含む細胞増殖異常が含まれ、免疫系の疾患の中には、後天性免疫不全症候群(AIDS)及び副腎機能不全、成人呼吸窮迫症候群、アレルギー、強直性脊椎炎、アミロイド症、貧血、喘息、アテローム性動脈硬化症、自己免疫性溶血性貧血、自己免疫性甲状腺炎、自己免疫性多腺性内分泌カンジダ性外胚葉ジストロフィ(APECED)、気管支炎、胆嚢炎、接触皮膚炎、クローン病、アトピー性皮膚炎、皮膚筋炎、糖尿病、肺気腫、リンパ球毒素性一時性リンパ球減少症、赤芽球症、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、過好酸球増加症、過敏性大腸症候群、多発性硬化症、重症筋無力症、心筋または心膜炎、骨関節炎、骨粗しょう症、腓炎、乾癬、ライター症候群、リウマチ様関節炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性エリテマトーデス、全身性硬化症、原発性血小板血症、血小板減少症、潰瘍性大腸炎、ウェルナー症候群、癌合併症、血液透析、体外循環、ウイルス感染症、細菌感染症、真菌感染症、寄生虫感染症、原虫感染症、蠕虫感染症

10

20

30

40

50

、外傷が含まれ、免疫不全の中には後天性免疫不全症候群（AIDS）、ブルトン型伴性無ガンマロブリン血症、後天性免疫グロブリン血症（CVI）、ディジョージ症候群（胸腺形成不全症）、胸腺異型性、IgA単独欠損症、重症複合型免疫不全（SCID）、血小板減少および湿疹を伴う免疫不全症（ウイスコット アルドリッチ症候群）、チェディアック 東症候群、慢性肉芽腫症、遺伝性血管神経性浮腫、クッシング病が含まれ、神経の疾患の中には、癲癇、虚血性脳血管障害、脳卒中、大脳新生物、アルツハイマー病、ピック病、ハンチントン病、痴呆、パーキンソン病及びその他の錐体外路障害、筋萎縮性側索硬化及びその他の運動ニューロン障害、進行性神経性筋萎縮症、色素性網膜炎、遺伝性運動失調、多発性硬化症及び他の脱髄疾患、細菌性及びウイルス性髄膜炎、脳膿瘍、硬膜下蓄膿症、硬膜外膿瘍、化膿性頭蓋内血栓性静脈炎、脊髄炎及び神経根炎、ウイルス性中枢神経系疾患、プリオン病（クールー、クロイツフェルト ヤコブ病、ゲルストマン症候群、及びGerstmann - Straussler - Scheinker症候群を含む）、致死性家族性不眠症、神経系性栄養病及び代謝病、神経線維腫症、結節硬化症、小脳網膜血管芽腫（cerebelloretinal hemangioblastomatosis）、脳3叉神経血管症候群、ダウン症候群を含む中枢神経系の精神薄弱及び他の発生障害、脳性麻痺、神経骨格異常症、自律神経系障害、脳神経障害、脊髄障害、筋ジストロフィー及びその他の神経筋障害、末梢神経疾患、皮膚筋炎及び多発性筋炎、遺伝性、代謝性、内分泌性、及び中毒性の筋疾患、重症筋無力症、周期性四肢麻痺、精神病（気分性、不安性の障害、分裂病性疾患）、季節性障害（SAD）、静座不能、健忘症、緊張病、糖尿病性ニューロパシー、錐体外路性終末欠陥症候群、ジストニー、分裂病性精神障害、帯状疱疹後神経痛、及びトゥレット病と、進行性核上麻痺、大脳皮質基底核部変性症、及び家族性の前頭側頭性健忘症とが含まれ、肺の疾患の中には先天性肺異常、肺拡張不全、肺うっ血及び肺水腫、肺動脈塞栓症、肺出血、肺梗塞、肺高血圧症、血管硬化症、閉塞性肺疾患、拘束型肺疾患、慢性閉塞性肺疾患、肺気腫、慢性気管支炎、気管支喘息、細気管支拡張症、細菌性肺炎、ウイルス性肺炎及びマイコプラズマ肺炎、肺膿瘍、肺結核、びまん性間質性疾患、塵肺症、サルコイド症、特発性肺線維症、剥離性間質性肺炎、過敏症肺炎、肺好酸球増加閉塞性細気管支炎 器質性肺炎（pulmonary eosinophilia bronchiolitis obliterans - organizing pneumonia）、びまん性肺出血症候群、グッドパスチャー症候群、特発性肺血鉄症、肺併発膠原血管病併発性肺疾患、肺胞たんぱく症、肺腫瘍、炎症性及び非炎症性胸水、気胸症、胸膜腫瘍、薬物性肺疾患、放射線による肺疾患、及び肺移植の合併症などが含まれ、細胞増殖異常には日光性角化症、動脈硬化、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、及び奇形癌を含む癌、具体的には、副腎、膀胱、骨、骨髄、脳、乳房、頸部、胆嚢、神経節、消化管、心臓、腎臓、肝臓、肺、筋肉、卵巣、膵臓、副甲状腺、陰茎、前立腺、唾液腺、皮膚、脾臓、精巣、胸腺、甲状腺、子宮の癌が含まれる。

10

20

30

【0152】

別の実施例では、HYDR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して、限定するものではないが上記した疾患を含むHYDRの発現または活性の低下に関連した疾患を治療または予防することも可能である。

40

【0153】

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HYDR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HYDR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。

【0154】

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HYDR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、HYDR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。

【0155】

50

更なる実施例では、HYDR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HYDR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、上記した免疫系の疾患、免疫不全、神経の疾患、肺の疾患、および癌を含む細胞増殖異常が含まれる。一実施態様では、HYDR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HYDRを発現する細胞または組織に薬剤を運ぶターゲティング或いは運搬機構として間接的に用いられ得る。

【0156】

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HYDR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、HYDR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。

10

【0157】

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

【0158】

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

20

【0159】

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

30

【0160】

HYDRに対する抗体を誘発するために用いられるオリゴペプチド、ペプチド、または断片は、少なくとも約5個のアミノ酸からなり、一般的には約10個以上のアミノ酸からなるものが好ましい。これらのオリゴペプチド、ペプチドまたは断片は、天然のタンパク質のアミノ酸配列の一部と同一であり且つ小さな天然の分子の全アミノ酸配列を含むことが望ましい。HYDR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。

40

【0161】

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

50

l. Acad. Sci. USA 80:2026-2030、Cole, S.P. 他 (1984) Mol. Cell Biol. 62:109-120等を参照)。

【0162】

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. 81:6851-6855; Neuberger, M.S. 他 (1984) Nature 312:604-608; Takeda, S. 他 (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、HYDR特異的一本鎖抗体を生成する。関連特異性を有するガイディオタイプ組成が異なるような抗体を、ランダムな組合せの免疫グロブリンライブラリからチェーンシャッフリングによって産生することもできる(Burton D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137等を参照)。

【0163】

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

【0164】

HYDRに対する特異的な結合部位を含む抗体も得ることができる。例えば、限定するものではないが、このような断片には、抗体分子のペプシン消化によって作製される $F(ab')_2$ 断片と、 $F(ab')_2$ 断片のジスルフィド架橋を還元することによって作製される Fab 断片とがある。或いは、Fab発現ライブラリを作製することによって、モノクローナル Fab 断片を所望の特異性と迅速且つ容易に同定することが可能となる(Huse, W.D. 他 (1989) Science 256:1275-1281等を参照)。

【0165】

種々のイムノアッセイを用いてスクリーニングし、所望の特異性を有する抗体を同定することができる。確立された特異性を有するポリクローナル抗体またはモノクローナル抗体の何れかを用いる免疫放射線アッセイまたは競合結合アッセイに対する数々のプロトコルは、本技術分野において公知である。このようなイムノアッセイは通常、HYDRとその特異性抗体間の複合体形成の計測に關与している。二つの非干渉性HYDRエピトープに対して反応性のモノクローナル抗体を用いる、2部位モノクローナルベースのイムノアッセイが一般に利用されるが、競合的結合アッセイも利用することができる(Pound、前出)。

【0166】

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、HYDR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HYDR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HYDR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、HYDR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HYDRエピトープに単一特異的なモノクローナル抗体医薬の K_a は、親和性の真の測定値を表す。 K_a 値が $10^9 \sim 10^{12}$ L/mol の高親和性抗体医薬は、HYDR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 $10^6 \sim 10^7$ liter/mol の低親和性抗体医薬は、HYDR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) Antibo

dies, 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)。

【0167】

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

【0168】

本発明の別の実施例では、HYDRをコードするポリヌクレオチド、またはその任意の断片や相補配列が、治療目的で使用することができる。ある実施態様では、HYDR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、HYDRをコードする配列の制御領域から、またはコード領域に沿ったさまざまな位置から設計可能である。(Agrawal, S.他(1996) Antisense Therapeutics, Humana Press Inc., Totawa NJを参照)。

【0169】

治療に用いる場合、アンチセンス配列を好適な標的細胞に導入するのに好適な任意の遺伝子送達系を用いることができる。アンチセンス配列は、転写時に標的タンパク質をコードする細胞配列の少なくとも一部に相補的な配列を発現する発現プラスミドの形で細胞内に輸送することが可能である。(Slater, J. E. 他(1998) J. Allergy Clin. Immunol. 102(3): 469-475 及び Scanlon, K. J. 他(1995) 9(13): 1288-1296等を参照。)アンチセンス配列はまた、例えばレトロウイルスやアデノ関連ウイルスベクター等のウイルスベクターを用いて細胞内に導入することもできる(Miller, A. D. (1990) Blood 76: 271、前出のAusubel、Uckert, W. and W. Walther (1994) Pharmacol. Ther. 63(3): 323-347等を参照。)その他の遺伝子送達機構には、リポソーム系、人工的なウイルスエンベロープ及び当分野で公知のその他の系が含まれる(Rossi, J. J. (1995) Br. Med. Bull. 51(1): 217-225; Boado, R. J. 他(1998) J. Pharm. Sci. 87(11): 1308-1315、Morris, M. C. 他(1997) Nucleic Acids Res. 25(14): 2730-2736等を参照)。

【0170】

本発明の別の実施例では、HYDRをコードするポリヌクレオチドを、体細胞若しくは生殖細胞の遺伝子治療に用いることが可能である。遺伝子治療を行うことにより、(i)遺伝子欠損症(例えばX染色体鎖遺伝(Cavazzana-Calvo, M. 他(2000) Science 288: 669-672)により特徴付けられる重度の複合型免疫欠損(SCID)-X1の場合)、先天性アデノシンデアミナーゼ(ADA)欠損症に関連する重度の複合型免疫欠損(Blaese, R. M. ら(1995) Science 270: 475-480、Bordignon, C. 他(1995) Science 270: 470-475)、嚢胞性繊維症(Zabner, J. ら(1993) Cell 75: 207-216; Crystal, R. G. 他(1995) H 40 40

um. 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因子欠損に起因する血友病 (Crystal, R.G. (1995) Science 270:404-410、Verma, I.M. および N. Somia (1997) Nature 389:239-242) を治療し、(ii) 条件的致死性遺伝子産物を発現させ (例えば制御不能な細胞増殖に起因する癌の場合)、(iii) 細胞内の寄生虫 (例えばヒト免疫不全ウイルス (HIV) (Baltimore, D. (1988) Nature 335:395-396、Poeschl, E. 他 (1996) Proc. Natl. Acad. Sci. USA. 93:11395-11399)、B型若しくはC型肝炎ウイルス (HBV、HCV)、Candida albicans 及び Paracoccidioides brasiliensis 等の真菌寄生虫、並びに Plasmodium falciparum 及び Trypanosoma cruzi 等の原虫寄生体に対する防御機能を有するタンパク質を発現させることができる。HYDR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HYDR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。

【0171】

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

【0172】

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

【0173】

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

【0174】

本発明の別の実施例では、HYDR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HYDRをコードするポリヌクレオチドと、(ii)好適なRNAパッケージングシグナルと、(iii)追加のレトロウイルス・シス作用性RNA配列及び効率的なベクターの増殖に必要なコーディング配列を伴うRev応答性エレメント(RRE)とからなるレトロウイルスベクターを作製して治療することができる。レトロウイルスベクター（例えばPFB及びPFBNEO）はStratagene社から市販されており、刊行データ(Riviere, I. 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。ベクターは、好適なベクター産生細胞系(VPCL)において増殖され、VPCLは、標的細胞上の受容体に対する親和性を有するエンベロープ遺伝子またはVSVg等の汎親和性エンベロープタンパク質を発現する(Armentano, D. ら (1987) *J. Virol.* 61: 1647 - 1650、Bender, M. A. ら (1987) *J. Virol.* 61: 1639 - 1646、Adam, M. A. および A. D. Miller (1988) *J. Virol.* 62: 3802 - 3806、Dull, T. 他 (1998) *J. Virol.* 72: 8463 - 8471、Zufferey, R. 他 (1998) *J. Virol.* 72: 9873 - 9880)。Riggに付与された米国特許第5,910,434号(「Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant」)において、レトロウイルスパッケージング細胞系を得るための方法が開示されており、引用することをもって本明細書の一部とする。レトロウイルスベクターの増殖、細胞集団（例えばCD4⁺T細胞）の形質導入、及び形質導入した細胞の患者への戻しは、遺伝子治療の分野では当業者に公知の方法であり、多数の文献に記載されている(Ranga, U. 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)。

【0175】

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、HYDRの発現に関連する1つ或いは複数の遺伝子異常を有する細胞にHYDRをコードするポリヌクレオチドを送達する。アデノウイルス系ベクターの作製及びパッケージングについては、当業者に公知である。複製欠損型アデノウイルスベクターは、免疫調節タンパク質をコードする遺伝子を脾臓の無損傷の脾島内に導入するために可変性であることが証明された(Csete, M. E. ら (1995) *Transplantation* 27: 263 - 268)。使用できる可能性のあるアデノウイルスベクターは、Armentanoに付与された米国特許第5,707,618号(「Adenovirus vectors for gene therapy」)に記載されており、引用することをもって本明細書の一部

とする。アデノウイルスベクターについては、Antinozzi, P. A. ら (1999) *Annu. Rev. Nutr.* 19: 511-544 及び Verma, I. M. および N. Somia (1997) *Nature* 18: 389: 239-242 も参照されたい。両文献は、引用することをもって本明細書の一部とする。

【0176】

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

【0177】

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

【0178】

転写開始部位由来のオリゴヌクレオチドを用いて遺伝子発現を阻害することも可能である

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

10

【0179】

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

【0180】

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

20

【0181】

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

30

【0182】

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

40

【0183】

本発明の更なる実施例は、HYDR をコードするポリヌクレオチドの発現の変化に有効な化合物をスクリーニングする方法を含む。限定するものではないが特異ポリヌクレオチドの変異発現に有効な化合物には、オリゴヌクレオチド、アンチセンスオリゴヌクレオチド、三重らせん形成オリゴヌクレオチド、転写因子その他のポリペプチド転写制御因子、及び特異ポリヌクレオチド配列と相互作用し得る非高分子化学的実体がある。有効な化合物

50

は、ポリヌクレオチド発現のインヒビターまたはエンハンサーのいずれかとして作用することによりポリヌクレオチド発現を変異し得る。従って、HYDR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、HYDR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、HYDR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、HYDR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。

【0184】

特異ポリヌクレオチドの変異発現における有効性に対して、少なくとも1個から複数個の試験化合物をスクリーニングし得る。試験化合物は、当分野で通常知られている任意の方法により得られる。このような方法には、ポリヌクレオチドの発現を変異させる場合と、既存の、市販のまたは専売の、天然または非天然の化合物ライブラリから選択する場合と、標的ポリヌクレオチドの化学的及び/または構造的特性に基づく化合物を合理的にデザインする場合と、組合せ的にまたは無作為に生成した化合物のライブラリから選択する場合に有効であることが知られているような化合物の化学修飾がある。HYDRをコードするポリヌクレオチドを含むサンプルは、少なくとも1つの試験化合物に曝露して得る。サンプルには例えば、無傷細胞、透過処理した細胞、*in vitro* 無細胞系または再構成生化学系があり得る。HYDR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、HYDRをコードするポリヌクレオチドの配列に相補的なヌクレオチド配列を有するプローブを用いたハイブリダイゼーションにより、特定のヌクレオチドの発現を検出する。ハイブリダイゼーション量を定量し、それによって1つ以上の試験化合物に曝露される及び曝露されないポリヌクレオチドの発現の比較に対する基礎を形成し得る。試験化合物に曝露されるポリヌクレオチドの発現変化の検出は、ポリヌクレオチドの発現を変異する際に試験化合物が有効であることを示している。特異ポリヌクレオチドの変異発現に有効な化合物に対して、例えば *Schizosaccharomyces pombe* 遺伝子発現系 (Atkins, D. ら (1999) 米国特許第5,932,435号、Arndt, G.M. ら (2000) *Nucleic Acids Res.* 28:E15) または HeLa 細胞等のヒト細胞系 (Clarke, M.L. ら (2000) *Biochem. Biophys. Res. Commun.* 268:8-13) を用いてスクリーニングを実行する。本発明の特定の実施例は、特異的ポリヌクレオチド配列に対するアンチセンス活性のためのオリゴヌクレオチド (デオキシリボヌクレオチド、リボヌクレオチド、ペプチド核酸、修飾オリゴヌクレオチド) の組み合わせライブラリをスクリーニングすることに関与している (Bruce, T.W. ら (1997) 米国特許第5,686,242号、Bruce, T.W. ら (2000) 米国特許第6,022,691号)。

【0185】

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

【0186】

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

【0187】

本発明の追加実施例は、通常薬剤として許容できる賦形剤で処方される活性成分を有する成分の投与に関連する。賦形剤には例えば、糖、でんぷん、セルロース、ゴム及びタンパク質がある。様々な処方が通常知られており、詳細は *Remington's Pharmaceutical Sciences* (Maack Publishing, Eas

10

20

30

40

50

t o n P A) の最新版に記載されている。このような組成物は、HYDR、HYDRの抗体、擬態、アゴニスト、アンタゴニスト、またはHYDRのインヒビターなどからなる。

【0188】

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

【0189】

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

【0190】

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

【0191】

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

【0192】

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

【0193】

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

【0194】

正確な投与量は、治療が必要な被験者に関する要素を考慮して、現場の医者が決定することになる。効果的なレベルの活性成分を与え、或いは所望の効果を維持するべく、投与量及び投与を調節する。被験者に関する要素としては、疾患の重症度、患者の通常健康状態、患者の年齢、体重及び性別、投与の時間及び頻度、薬剤の配合、反応感受性及び治療

10

20

30

40

50

に対する応答等を考慮する。作用期間が長い成分は、特定の製剤の半減期及びクリアランス率によって3～4日毎に1度、1週間に1度、或いは2週間に1度の間隔で投与し得る。

【0195】

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

10

【0196】

(診断)

別の実施例では、HYDRの発現によって特徴付けられる疾患の診断のために、或いはHYDRやHYDRのアゴニスト、アンタゴニストまたは阻害剤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おいて、HYDRを特異的に結合する抗体が用いられることがある。診断目的に有用な抗体は、上記の治療の箇所で記載した方法と同じ方法で調合される。HYDR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HYDRを検出する方法が含まれる。この抗体は修飾されたものもされていないものも可能であり、レポーター分子との共有結合または非共有結合で標識化できる。レポーター分子としては広くさまざまな種類が本分野で知られており、また使用可能であるが、そのうちのいくつかは上記で説明されている。

20

【0197】

HYDR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HYDR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HYDRの発現の値は、複合体の形成に適した条件下で、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HYDRに対する抗体とを結合させることによって決定する。標準複合体形成量は、種々の方法、例えば測光法で定量できる。被験者のHYDRの発現の量、対照及び疾患、生検組織からのサンプルが基準値と比較される。基準値と被験者との偏差が疾患を診断するパラメータとなる。

30

【0198】

別の実施例によれば、HYDRをコードするポリヌクレオチドを診断目的に用いることもできる。用いることができるポリヌクレオチドには、オリゴヌクレオチド配列、相補的RNA及びDNA分子、そしてPNAが含まれる。このポリヌクレオチドを用いて、疾患と相関し得るHYDR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、HYDRの存在の有無、更に過剰な発現を調べ、治療中のHYDR値の調節を監視する。

【0199】

一実施形態では、HYDR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、HYDR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HYDRをコードする自然界の配列のみを同定するかどうか、或いは対立遺伝子や関連配列を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。

40

【0200】

プローブはまた、関連する配列の検出に利用され、HYDR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。本発明のハイブリダイゼーションプローブはDNAまたはRNAとすることができ、SEQ ID NO: 5-8の配列、或いはHYD

50

R 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。

【0201】

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

【0202】

HYDRをコードするポリヌクレオチド配列を用いて、HYDR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免疫系の疾患、免疫不全、神経の疾患、肺の疾患、および癌を含む細胞増殖異常が含まれ、免疫系の疾患の中には、後天性免疫不全症候群(AIDS)及び副腎機能不全、成人呼吸窮迫症候群、アレルギー、強直性脊椎炎、アミロイド症、貧血、喘息、アテローム性動脈硬化症、自己免疫性溶血性貧血、自己免疫性甲状腺炎、自己免疫性多腺性内分泌カンジダ性外胚葉ジストロフィ(APECED)、気管支炎、胆嚢炎、接触皮膚炎、クローン病、アトピー性皮膚炎、皮膚筋炎、糖尿病、肺気腫、リンパ球毒素性一時性リンパ球減少症、赤芽球症、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、過好酸球増加症、過敏性大腸症候群、多発性硬化症、重症筋無力症、心筋または心膜炎、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、免疫不全の中には後天性免疫不全症候群(AIDS)、ブルトン型伴性無ガンマグロブリン血症、後天性免疫グロブリン血症(CVI)、ディジョージ症候群(胸腺形成不全症)、胸腺異型性、IgA単独欠損症、重症複合型免疫不全(SCID)、血小板減少および湿疹を伴う免疫不全症(ウイスコットアルドリッチ症候群)、チェディアック東症候群、慢性肉芽腫症、遺伝性血管神経性浮腫、クッシング病が含まれ、神経の疾患の中には、癲癇、虚血性脳血管障害、脳卒中、大脳新生物、アルツハイマー病、ピック病、ハンチントン病、痴呆、パーキンソン病及びその他の錐体外路障害、筋萎縮性側索硬化及びその他の運動ニューロン障害、進行性神経性筋萎縮症、色素性網膜炎、遺伝性運動失調、多発性硬化症及び他の脱髄疾患、細菌性及びウイルス性髄膜炎、脳膿瘍、硬膜下蓄膿症、硬膜外膿瘍、化膿性頭蓋内血栓性静脈炎、脊髄炎及び神経根炎、ウイルス性中枢神経系疾患、プリオン病(クールー、クロイツフェルトヤコブ病、ゲルストマン症候群、及びGerstmann-Straussler-Scheinker症候群を含む)、致死性家族性不眠症、神経系性栄養病及び代謝病、神経線維腫症、結節硬化症、小脳網膜血管芽腫(cerebelloretinal hemangioblastomatosis)、脳3叉神経血管症候群、ダウン症候群を含む中枢神経系の精神薄弱及び他の発生障害、脳性麻痺、神経骨格異常症、自律神経系障害、脳神経障害、脊髄障害、筋ジストロフィー及びその他の神経筋障害、末梢神経疾患、皮膚筋炎及び多発性筋炎、遺伝性、代謝性、内分泌性、及び中毒性の筋疾患、重症筋無力症、周期性四肢麻痺、精神病(気分性、不安性の障害、分裂病性疾患)、季節性障害(SAD)、静座不能、健忘症、緊張病、糖尿病性ニューロパシー、錐体外路性終末欠陥症候群、ジストニー、分裂病性精神障害、帯状疱疹後神経痛、及びトゥレット病と、進行性核上麻痺、大脳皮質基底核部変性症、及び家族性の前頭側頭性健忘症とが含まれ、肺の疾患の中には先天性肺異常、肺拡張不全、肺うっ血及び肺水腫、肺動脈塞栓症、肺出血、肺梗塞、肺高血圧症、血管硬化症、閉塞性肺疾患、拘束型肺疾患、慢性閉塞性肺疾患、肺気腫、

慢性気管支炎、気管支喘息、細気管支拡張症、細菌性肺炎、ウイルス性肺炎及びマイコプラズマ肺炎、肺膿瘍、肺結核、びまん性間質性疾患、塵肺症、サルコイド症、特発性肺線維症、剥離性間質性肺炎、過敏症肺炎、肺好酸球増加閉塞性細気管支炎 器質性肺炎 (pulmonary eosinophilia bronchiolitis obliterans - organizing pneumonia)、びまん性肺出血症候群、グッドパスチャー症候群、特発性肺血鉄症、肺併発膠原血管病併発性肺疾患、肺胞たんぱく症、肺腫瘍、炎症性及び非炎症性胸水、気胸症、胸膜腫瘍、薬物性肺疾患、放射線による肺疾患、及び肺移植の合併症などが含まれ、細胞増殖異常には日光性角化症、動脈硬化、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)、骨髓線維症、発作性夜間ヘモグロビン尿症、真性多血症、乾癬、原発性血小板血症、並びに腺癌及び白血病、リンパ腫、黒色腫、骨髄腫、肉腫、及び奇形癌を含む癌、具体的には、副腎、膀胱、骨、骨髄、脳、乳房、頸部、胆嚢、神経節、消化管、心臓、腎臓、肝臓、肺、筋肉、卵巣、膵臓、副甲状腺、陰茎、前立腺、唾液腺、皮膚、脾臓、精巣、胸腺、甲状腺、子宮の癌が含まれる。HYDR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 (dipstick)、ピン (pin)、ELISA式アッセイ、及び変異HYDRの発現を検出するために患者から採取した体液或いは組織を利用するマイクロアレイに使用することが可能である。このような定性方法または定量方法は、当分野で公知である。

10

【0203】

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

20

【0204】

HYDR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、HYDRをコードする配列或いはその断片とを結合させることにより達成され得る。実質的に精製されたポリヌクレオチドを既知量用いて行った実験から得た値を正常な対象から得た値と比較することにより、標準ハイブリダイゼーションを定量することができる。このようにして得た標準値は、疾患の徴候を示す患者から得たサンプルから得た値と比較することができる。標準値からの偏差を用いて疾患の存在を確定する。

30

【0205】

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

40

【0206】

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

【0207】

50

H Y D R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。これらのオリゴマーは、化学的に合成するか、酵素により生産するか、或いは*in vitro*で産出し得る。オリゴマーは、好ましくはH Y D Rをコードするポリヌクレオチドの断片、或いはH Y D R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、最適な条件下で、特定の遺伝子や条件を識別するために利用される。また、オリゴマーは、やや緩いストリンジент条件下で、近縁のD N A或いはR N A配列の検出、定量、或いはその両方のため用いることが可能である。

【0208】

或る実施態様において、H Y D R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、多くの場合にヒトの先天性または後天性遺伝病の原因となるような置換、挿入及び欠失である。限定するものではないがS N Pの検出方法には、S S C P (single-stranded conformation polymorphism) 及び蛍光S S C P (f S S C P) 法がある。S S C Pでは、H Y D R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。D N Aは例えば、病変組織または正常組織、生検サンプル、体液その他に由来し得る。D N A内のS N Pは、一本鎖形状のP C R生成物の2次及び3次構造に差異を生じさせる。差異は非変性ゲル中でのゲル電気泳動法を用いて検出可能である。f S S C Pでは、オリゴヌクレオチドプライマーを蛍光性に標識する。それによってD N Aシーケンシング機などの高処理機器でアンプリマー(amplimer)の検出が可能になる。更に、インシリコS N P (in silico S N P, i s S N P) と呼ばれる配列データベース分析法は、一般的なコンセンサス配列に配列されるような個々の重畳するD N A断片の配列を比較することにより、多型性を同定し得る。これらのコンピュータベースの方法は、D N Aの実験室での調整及び統計モデル及びD N A配列クロマトグラムの自動分析を用いたシーケンシングのエラーに起因する配列の変異をフィルタリングして除去する。別の態様では、例えば高処理M A S S A R R A Yシステム(Sequenom, Inc., San Diego CA)を用いた質量分析によりS N Pを検出し、特徴付ける。

【0209】

H Y D Rの発現を定量するために用い得る方法には、ヌクレオチドの放射標識またはビオチン標識、調節核酸の相互増幅(coamplification) 及び標準曲線から得た結果の補間もある(例えば、Melby, P. C.ら(1993) J. Immunol. Methods, 159: 235-244; Duplaa, C.ら(1993) Anal. Biochem. 212: 229-236を参照)。目的のオリゴマーが種々の希釈液中に存在し、分光光度法または比色反応によって定量が迅速になるような高処理フォーマットのアッセイを行うことによって、複数のサンプルの定量速度を加速することができる。

【0210】

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

【0211】

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

【0212】

或る実施例は、或る組織または細胞タイプの転写イメージを生成するような本発明のポリヌクレオチドの使用に関連する。転写イメージは、特定の組織または細胞タイプにより遺伝子発現の包括的パターンを表す。包括的遺伝子発現パターンは、所定の条件下で所定の時間に発現した遺伝子の数及び相対存在量を定量することにより分析される (Seilliamer 他、米国特許第5,840,484号の「Comparative Gene Transcript Analysis」を参照。この特許に言及することを以って本明細書の一部とする)。従って、特定の組織または細胞タイプの転写または逆転写全体に本発明のポリヌクレオチドまたはその補体をハイブリダイズすることにより、転写イメージを生成し得る。或る実施例では、本発明のポリヌクレオチドまたはその補体がマイクロアレイ上のエレメントのサブセットを複数含むような高処理フォーマットでハイブリダイゼーションを発生させる。結果として得られる転写イメージは、遺伝子活性のプロファイルを提供し得る。

10

【0213】

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

20

【0214】

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

30

40

【0215】

或る実施例では、核酸を含有する生物学的サンプルを試験化合物で処理することにより試験化合物の毒性を算定する。処理した生物学的サンプル中で発現した核酸は、本発明のポリヌクレオチドに特異的な1つ若しくは複数のプローブでハイブリダイズし、それによって本発明のポリヌクレオチドに対応する転写レベルを定量し得る。処理した生物学的サン

50

ブル中の転写レベルを、非処理生物学的サンプル中のレベルと比較する。両サンプルの転写レベルの差は、処理されたサンプル中で試験化合物が引き起こす毒性反応を示す。

【0216】

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

10

20

【0217】

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

30

【0218】

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

40

【0219】

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

50

【0220】

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

【0221】

マイクロアレイは、本技術分野でよく知られている方法を用いて調製し、使用し、そして分析する (Brennan, T. M. ら (1995) の米国特許第 5, 474, 796 号、Schena, M. ら (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619、Baldeschweiler らの (1995) PCT 出願第 WO95/251116 号、Shalon, D. らの (1995) PCT 出願第 WO95/35505 号、Heller, R. A. ら (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155、Heller, M. J. らの (1997) 米国特許第 5, 605, 662 号等を参照)。様々なタイプのマイクロアレイが公知であり、詳細については、DNA Microarrays: A Practical Approach, M. Schena, 編集。(1999) Oxford University Press, London に記載されている。該文献は、特に引用することを以って本明細書の一部となす。

【0222】

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

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

【0223】

確定した染色体マーカーを用いた結合分析等の物理的マッピング技術及び染色体標本原位ハイブリッド形成法を用いて、遺伝地図を拡張することができる。例えばマウスなど別の哺乳動物の染色体上に遺伝子を配置することにより、正確な染色体の遺伝子座がわかっていない場合でも関連するマーカーを明らかにし得る。この情報は、位置クローニングその他の遺伝子発見技術を用いて疾患遺伝子を探す研究者にとって価値がある。疾患または

症候群に關与する遺伝子が、血管拡張性失調症の11q22-23領域等、特定の遺伝子領域への遺伝的結合によって大まかに位置決めがなされると、該領域に対するいかなる配列マッピングも、更なる調査のための関連遺伝子或いは調節遺伝子を表すことができる。(Gattì, R. A. 他 (1988) Nature 336: 577-580等を参照) 転座、反転等に起因する、健常者、保有者、感染者の三者間における染色体位置の相違を發見するために本發明のヌクレオチド配列を用いてもよい。

【0224】

本發明の別の実施例では、HYDR、その触媒作用断片或いは免疫原断片またはそのオリゴペプチドを、種々の任意の薬剤スクリーニング技術における化合物のライブラリのスクリーニングに用いることができる。薬剤スクリーニングに用いる断片は、溶液中に遊離しているか、固体支持物に固定されるか、細胞表面上に保持されるか、細胞内に位置することになる。HYDR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。

10

【0225】

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

20

【0226】

別の実施例では、HYDRと結合可能な中和抗体がHYDR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、HYDR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。

【0227】

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

30

【0228】

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

【0229】

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60/237,093号、同第60/238/370号、および同第60/241,284号に言及することをもって本明細書の一部とする。

【0230】

(実施例)

40

1 cDNAライブラリの作製

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

50

ずれか一方、或いは別の方法を用いて、溶解物からRNAを沈殿させた。

【0231】

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

【0232】

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

【0233】

2 cDNAクローンの単離

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

【0234】

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

10

20

30

40

50

ene OR)及びFluoroskan II蛍光スキャナ(Labsystems Oy, Helsinki, Finland)を用いて蛍光分析的に定量した。

【0235】

3 シークエンシング及び分析

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

【0236】

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

10

20

30

40

50

メントした配列と配列の一致率も計算するMEGALIGNマルチシーケンスアライメントプログラム(DNASTAR)に組み込まれているようなCLUSTALアルゴリズムによって特定されるデフォルトパラメータを用いて生成する。

【0237】

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

10

【0238】

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: 5 - 8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。

【0239】

4 ゲノムDNAからのコード配列の同定及び編集

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

20

30

40

【0240】

5 cDNA配列データを使ったゲノム配列データの構築

ステッチ配列(Stitched Sequence)

部分cDNA配列は、実施例4に記載のGenscan遺伝子同定プログラムにより予測されたエキソンを用いて伸長させた。実施例3に記載されたように構築された部分cDNAは、ゲノムDNAにマッピングし、関連するcDNA及び1つ若しくは複数のゲノム配列から予測されたGenscanエキソンを含むクラスタに分解した。cDNA及びゲノム情報を統合するべくグラフ理論及び動的プログラミングに基づくアルゴリズムを用いて

50

各クラスタを分析し、引き続いて確認、編集または伸長して完全長配列を産出するような潜在的スプライス変異体を生成した。間隔全体の長さがクラスタ中の2以上の配列に存在するような配列を同定し、そのように同定された間隔は推移により等しいと考えられた。例えば、1つのcDNA及び2つのゲノム配列に間隔が存在する場合、3つの間隔は全て等しいと考えられる。このプロセスは、無関係であるが連続したゲノム配列をcDNA配列により結び合わせて架橋し得る。このようにして同定された区間を、親配列 (parent sequence) に沿って現われるようにステッチアルゴリズムで縫い合わせ、可能な最も長い配列および変異配列を作製する。1種類の親配列に沿って発生した間隔と間隔との連鎖 (cDNA - cDNA または ゲノム配列 - ゲノム配列) は、親の種類を変える連鎖 (cDNA - ゲノム配列) に優先した。結果として得られるステッチ配列は、BLAST分析により公共データベース genpept 及び gbpr1 に翻訳されて比較された。Genscanにより予測された不正確なエキソンは、genpept からヒットしたトップのBLASTと比較することにより修正した。必要な場合には、追加cDNA配列を用いるかゲノムDNAの検査により配列を更に伸長させた。

10

【0241】

ストレッチ配列 (Stretched Sequence)

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

20

【0242】

6 HYDR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: 5 - 8 を構築するために用いた配列を、BLAST及びSmith-Watermanアルゴリズムを用いて、Incyte LIFESEQデータベース及び公共のドメインデータベースの配列と比較した。SEQ ID NO: 5 - 8 と一致するこれらのデータベースの配列を、Phrap (表7) などの構築アルゴリズムを使用して、連続及び重複した配列のクラスターに組み入れた。スタンフォード・ヒトゲノムセンター (SHGC)、ホワイトヘッド・ゲノム研究所 (WIGR)、Genethon等の公的な情報源から入手可能な放射線ハイブリッド及び遺伝地図データを用いて、クラスタ化された配列が前もってマッピングされたかを決定した。マッピングされた配列がクラスタに含まれている場合は、個々の配列番号を含めてそのクラスタの全配列が地図上の位置に割り当てられた。

30

40

【0243】

地図上の位置は、ヒト染色体の範囲または間隔として表される。センチモルガン間隔の地図上の位置は、染色体のpアームの末端に関連して測定する。(センチモルガン (cM) は、染色体マーカー間の組換え頻度に基づく計測単位である。平均して、1cMは、ヒト中のDNAの1メガベース (Mb) にほぼ等しい。尤も、この値は、組換えのホットスポット及びコールドスポットに起因して広範囲に変化する。) cM距離は、配列が各クラスタ内に含まれるような放射線ハイブリッドマーカーに対して境界を提供するようなGenethonによってマッピングされた遺伝マーカーに基づく。NCBI「GeneMap99」 (<http://www.ncbi.nlm.nih.gov/genemap>)

50

)などの一般個人が入手可能なヒト遺伝子マップおよびその他の情報源を用いて、上記した区間が既に同定されている疾患遺伝子マップ内若しくは近傍に位置するかを決定できる。

【0244】

7 ポリヌクレオチド発現の分析

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

【0245】

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

【0246】

【数1】

(BLASTスコア×配列一致率)

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

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

【0247】

或いは、HYDRをコードするポリヌクレオチド配列は、由来する組織に対して分析する。例えば或る完全長配列は、Incyte cDNA配列(実施例3を参照)と少なくとも一部は重畳するように構築される。各cDNA配列は、ヒト組織から作製されたcDNAライブラリに由来する。各ヒト組織は、以下の生物/組織カテゴリー即ち心血管系、結合組織、消化器系、胎芽構造、内分泌系、外分泌腺、女性生殖器、男性生殖器、生殖細胞、血液及び免疫系、肝、筋骨格系、神経系、膵臓、呼吸器系、感覚器、皮膚、顎口腔系、非分類性/混合性または尿路の1つに分類される。各カテゴリーのライブラリ数を数えて、全カテゴリーの総ライブラリ数で除する。同様に、各ヒト組織は、以下の疾患/病状カテゴリー即ち癌、細胞株、発達、炎症、神経性、外傷、心血管、鬱血、その他の1つに分類される。

各カテゴリーのライブラリ数を数えて、全カテゴリーの総ライブラリ数で除する。得られるパーセンテージは、HYDR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。

【0248】

8 HYDRをコードするポリヌクレオチドの伸長

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

【0249】

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

【0250】

高忠実度の増幅が、当業者によく知られている方法を利用したPCR法によって得られた。PCRは、PTC-200サーマルサイクラー(MJ Research, Inc.)を用いて96穴プレート内で行った。反応混合液は、鋳型DNA及び200nmol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度で1分間 ステップ4: 68で2分間 ステップ5: ステップ2、3、4を20回繰り返す ステップ6: 68で5分間 ステップ7: 4で保存。プライマー対T7、SK+に対しては、上記パラメータに代えて以下のパラメータを用いた。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、4を20回繰り返す ステップ6: 68で5分間 ステップ7: 4で保存

【0251】

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

【0252】

伸長させたヌクレオチドは、脱塩及び濃縮して384穴プレートに移し、CviJIコラウイルスエンドヌクラーゼ(Molecular Biology Research, Madison WI)を用いて消化し、pUC18ベクター(Amersham Pharmacia Biotech)への再連結反応前に音波処理またはせん断した。ショットガン・シー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で分離し、断片を切除し、寒天をAgar ACE(Promega)で消化した。伸長させたクローンをT4リガーゼ(New England Biolabs, Beverly MA)を用いてpUC18ベクター(Amersham Pharmacia Biotech)に再連結し、Pfu DNAポリメラーゼ(Str

10

20

30

40

50

a t a g e n e) で処理して制限部位のオーバーハングを満たし、大腸菌細胞に形質移入した。形質移入した細胞を選択して抗生物質を含む培地に移し、それぞれのコロニーを切りとってLB / 2 X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。

【0253】

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

【0254】

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

【0255】

9 個々のハイブリダイゼーションプローブの標識及び使用

SEQ ID NO: 5 - 8 から導き出されたハイブリダイゼーションプローブを用いて、cDNA、mRNA、またはゲノムDNAをスクリーニングする。約20塩基対からなるオリゴヌクレオチドの標識について特に記載するが、より大きなヌクレオチド断片に対しても事実上同一の手順が用いられる。オリゴヌクレオチドは、OLIGO 4.06ソフトウェア (National Biosciences) 等の最新ソフトウェアを用いて設計し、各オリゴマー50 pmolと、[γ -³²P]アデノシン3リン酸 (Amersham Pharmacia Biotech) 250 μ Ciと、T4ポリヌクレオチドキナーゼ (DuPont NEN, Boston MA) を混合することにより標識する。標識したオリゴヌクレオチドは、SEPHADEX G-25超細繊維分子サイズ排除デキストラン ビードカラム (Amersham Pharmacia Biotech) を用いて十分に精製する。Ase I、Bgl II、Eco RI、Pst I、Xba I またはPvu II (DuPont NEN) のいずれか1つのエンドヌクレアーゼで消化されたヒトゲノムDNAの典型的な膜ベースのハイブリダイゼーション解析において、毎分10⁷ カウントの標識されたプローブを含むアリコットを用いる。

【0256】

各消化物から得たDNAは、0.7%アガロースゲル上で分画してナイロン膜 (Nytran Plus, Schleicher & Schuell, Durham NH) に移す。ハイブリダイゼーションは、40 で16時間行う。非特異的シグナルを除去するため、例えば0.1x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に一致する条件下で、プロットを室温で順次洗浄する。オートラジオグラフィまたはそれに代わるイメージング手段を用いてハイブリダイゼーションパターンを視覚化し、比較する。

【0257】

10 マイクロアレイ

マイクロアレイの表面上でアレイエレメントの結合または合成は、フォトリソグラフィ、圧電印刷 (インクジェット印刷、前出のBaldeschweiler等を参照)、機械

的マイクロスポッティング技術及びこれらから派生したものを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、均一かつ非多孔性の固体とするべきである (S c h e n a (1 9 9 9) 前出)。推奨する基板には、シリコン、シリカ、スライドガラス、ガラスチップ及びシリコンウエハがある。或いは、ドットプロット法またはスロットプロット法に類似のアレイを利用して、熱的、紫外線的、化学的または機械的結合手順を用いて基板の表面にエレメントを配置及び結合させてもよい。通常のアレイは、手作業で、または利用可能な方法や機械を用いて作製でき、任意の適正数のエレメントを有し得る (S c h e n a , M . ら (1 9 9 5) S c i e n c e 2 7 0 : 4 6 7 - 4 7 0 , S h a l l o n . D . ら (1 9 9 6) G e n o m e R e s . 6 : 6 3 9 - 6 4 5 , M a r s h a l l , A . および J . H o d g s o n (1 9 9 8) N a t . B i o t e c h n o l . 1 6 : 2 7 - 3 1 . を参照)。

10

【0258】

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

20

【0259】

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)セルソース法を用いてポリ(A)⁺RNAを精製する。各ポリ(A)⁺RNAサンプルを、MMLV逆転写酵素、0.05 pg/μlのオリゴ(dT)プライマー(21mer)、1×第一鎖合成バッファー、0.03 unit/μlのRNアーゼ阻害因子、500 μMのdATP、500 μMのdGTP、500 μMのdTTP、40 μMのdCTP、40 μMのdCTP-Cy3(BDS)またはdCTP-Cy5(Amersham Pharmacia Biotech)を用いて逆転写する。逆転写反応は、GEMBRIGHTキット(Incyte)を用いてポリ(A)⁺RNA含有の25体積ml内で行う。特異的対照ポリ(A)⁺RNAは、非コード酵母ゲノムDNAから*in vitro*転写により合成する。各反応サンプル(1つはCy3、もう1つはCy5標識)は、2.5 mlの0.5 M水酸化ナトリウムで処理し、85 °Cで20分間インキュベートし、反応を停止させてRNAを分解させる。サンプルは、2つの連続するCHROMA SPIN 30ゲル濾過スピンカラム(CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA)を用いて精製する。混合後、2つの反応サンプルは、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析出させる。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の5×SSC/0.2% SDS中で再懸濁する。

30

40

【0260】

マイクロアレイの調製

本発明の配列を用いて、アレイエレメントを生成する。各アレイエレメントは、クローン化cDNAインサートによりベクター含有細菌性細胞から増幅する。PCR増幅は、cDNAインサートの側面に位置するベクター配列に相補的なプライマーを用いる。30サイ

50

クルのPCRで1~2 ngの初期量から5 µgより大きい最終量までアレイエレメントを増幅する。増幅されたアレイエレメントは、SEPHACRYL-400 (Amersham Pharmacia Biotech)を用いて精製される。

【0261】

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

10

【0262】

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

【0263】

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

20

【0264】

ハイブリダイゼーション

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

30

【0265】

検出

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

40

【0266】

2つの異なるスキャンのうち、混合ガスマルチラインレーザは2つの蛍光色素を連続的に励起する。発光された光は、波長に基づき分離され、2つの蛍光色素に対応する2つの光

50

電子増倍管検出器 (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) に送られる。アレイと光電子増倍管間に設置された好適なフィルターを用いて、シグナルをフィルターする。用いる蛍光色素の最大発光の波長は、Cy3では565nm、Cy5では650nmである。装置は両方の蛍光色素からのスペクトルを同時に記録し得るが、レーザー源において好適なフィルターを用いて、蛍光色素1つにつき1度スキャンし、各アレイを通常2度スキャンする。

【0267】

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

10

【0268】

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

20

【0269】

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

30

【0270】

1.1 相補的ポリヌクレオチド

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

40

【0271】

1.2 HYDRの発現

HYDRの発現及び精製は、細菌またはウイルスをベースにした発現系を用いて行うことができる。細菌でHYDRを発現するために、抗生物質耐性及びcDNAの転写レベルを高める誘導性のプロモーターを含む好適なベクターにcDNAをサブクローニングする。このようなプロモーターには、lacオペレーター調節エレメントに関連するT5またはT7バクテリオファージプロモーター及びtrp-lac(tac)ハイブリッドプロモーターが含まれるが、これらに限定するものではない。組換えベクターを、BL21(DE3)等の好適な細菌宿主に形質転換する。抗生物質耐性をもつ細菌が、イソプロピル-Dチオガラクトピラノシド(IPTG)で誘発されるとHYDRを発現する。真核細胞

50

でのHYDR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* 核多角体病ウイルス (AcMNPV) を感染させて行う。バキュロウイルスの可欠ポリヘドリン遺伝子を、相同的組換え、或いは転移プラスミドの仲介に關与する細菌仲介遺伝子転移のどちらかによって、HYDRをコードするcDNAと置換する。ウイルスの感染力は維持され、強い多角体プロモーターによって高いレベルのcDNAの転写が行われる。組換えバキュロウイルスは、多くの場合は *Spodoptera frugiperda* (Sf9) 昆虫細胞に感染に用いられるが、ヒト肝細胞の感染にも用いられることもある。後者の感染の場合は、バキュロウイルスの更なる遺伝的変更が必要になる。(Engelhard, E. K. ら (1994) Proc. Natl. Acad. Sci. USA 91: 3224-3227, Sandig, V. ら (1996) Hum. Gene Ther. 7: 1937-1945. 等を参照)。

【0272】

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

【0273】

13 機能的アッセイ

HYDR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HYDRをコードする配列の発現によって評価する。cDNAを、cDNAを高いレベルで発現する強いプロモーターを含む哺乳動物発現ベクターにサブクローニングする。選択されるベクターには、PCMV SPORTプラスミド (Life Technologies) 及びPCR 3.1 (Invitrogen) が含まれ、どちらもサイトメガロウイルスプロモーターを有する。リポソーム製剤或いは電気穿孔法を用いて、5~10µgの組換えベクターをヒト細胞株、例えば内皮由来または造血由来の細胞株に一過的に形質移入する。更に、標識タンパク質をコードする配列を含む1~2µgのプラスミドを同時に形質移入する。標識タンパク質の発現により、形質移入細胞と非形質移入細胞を区別する手段が与えられる。また、標識タンパク質の発現によって、cDNAの組換えベクターからの発現を正確に予想できる。標識タンパク質は、例えば緑色蛍光タンパク質 (GFP; Clontech)、CD64またはCD64-GFP融合タンパク質から選択できる。自動化された、レーザー光学に基づく技術であるフローサイトメトリー (FCM) を用いて、GFPまたはCD64-GFPを発現する形質移入された細胞を同定し、その細胞のアポトーシス状態や他の細胞特性を評価する。FCMは、細胞死に先行するか或いは同時に発生する現象を診断する蛍光分子の取込を検出して計量する。このような現象として挙げられるのは、プロピジウムヨウ化物によるDNA染色によって計測される核DNA内容物の変化、前方光散乱と90°側方光散乱によって計測される細胞サイズと顆粒状性の変化、プロモデオキシウリジンの取込量の低下によって計測されるDNA合成の下方調節、特異抗体との反応性によって計測される細胞表面及び細胞内におけるタンパク質の発現の変化、及び蛍光複合アネキシンVタンパク質の細胞表面への結合によって計測される原形質膜組成の変化とがある。フローサイトメトリー法については、Ormerod, M

. G. (1994) Flow Cytometry Oxford, New York, NY. に記述がある。

【0274】

遺伝子発現におけるHYDRの影響は、HYDRをコードする配列とCD64またはCD64-GFPのどちらかが形質移入された高度に精製された細胞集団を用いて評価することができる。CD64またはCD64-GFPは、形質転換された細胞表面で発現し、ヒト免疫グロブリンG(IgG)の保存領域と結合する。形質転換された細胞と形質転換されない細胞とは、ヒトIgGが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。HYDR及び対象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

【0275】

1.4 HYDR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.* 182: 488-495を参照) または他の精製技術で実質的に精製されたHYDR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。

【0276】

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

20

【0277】

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

【0278】

1.5 特異的な抗体を用いる天然HYDRの精製

天然HYDR或いは組換えHYDRを、HYDRに特異的な抗体を用いるイムノアフィニティークロマトグラフィにより実質的に精製する。イムノアフィニティークラムは、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HYDR抗体とを共有結合させることにより形成する。結合後に、製造者の使用説明書に従って樹脂をブロックし、洗浄する。

40

【0279】

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

【0280】

1.6 HYDRと相互作用する分子の同定

50

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

【0281】

別法では、HYDR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。 10

【0282】

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

【0283】

17 HYDR 活性の実証

HYDRの活性はさまざまな特定の酵素アッセイによって実証される。その一部を以下に概説する。 20

【0284】

タンパク質ホスファターゼ (PP) 活性は、P-ニトロフェニルリン酸 (PNPP) の加水分解によって測定する。HYDRを、0.1%のメルカプトエタノールの存在下で、pH 7.5のHEPESバッファー中にPNPPと共に37°Cで60分インキュベートする。この反応は6mlの10N NaOHを加えて止める (Diamond, R. H. 他 (1994) Mol. Cell. Biol. 14:375262)。或るいは、HYDRの酸性ホスファターゼ活性は、pH 4.5の0.1Mクエン酸ナトリウム、及び500µlの40mM塩化ナトリウム中にHYDRを含む抽出液と100µlの10mM PNPPを37°Cで20分インキュベートして実証する。この反応は0.5mlの0.4Mグリシン/NaOH (pH 10.4) を加えて止める (Safitig, P. 他 (1997) J. Biol. Chem. 272:1862818635)。PNPPの加水分解によって起こる410nmでの光吸収の増加を分光光度計で測定する。このアッセイでは、光吸収の増加がHYDRの活性に比例する。 30

【0285】

別法では、HYDR活性は、リン酸化したタンパク質基質から除去されたリン酸の量を測定して決定する。反応は、60mM Tris、pH 7.6、1mM EDTA、1mM EGTA、0.1% 2-メルカプトエタノールおよび10µM基質、適量の ^{32}P 標識したセリン/トレオニン/またはチロシンを含む最終容量30µlの2nM或いは4nMのHYDRにおいて行なう。反応は基質で開始し、37°Cで10~15分インキュベートする。0.6M HCl、90mM $\text{Na}_4\text{P}_2\text{O}_7$ 、および2mM NaH_2PO_4 中の4% (w/v) 活性化木炭450µlで停止させ、次に12,000xgで5分間遠心分離する。酸可溶性 ^{32}P iを液体シンチレーションカウンターで定量する (Sinclair, C. 他 (1999) J. Biol. Chem. 274:2366623672)。 40

【0286】

HYDRのアデノシンデアミナーゼ活性は、アデノシン基質をHYDRとインキュベートしたときに起こるデアミノ化の速度を測定することにより決定される。反応は、最終的に 50

53.3 mMのリン酸カリウムと0.045 mMのアデノシンを含む最終容積3.0 mlに所定量のHYDRを含むサンプルで行われる。HYDRを除くアッセイ試薬を石英のキュベット内で混合し、25℃で平衡化する。反応はHYDRを加えてただちに逆さまにして混合することで開始する。アデノシンのイノシンへの加水分解によって起こる265 nmでの光吸収の増加を分光光度計で測定する。A_{265 nm}の減少を約5分間記録する。このアッセイでは、光吸収の減少がHYDRの活性に比例する。

HYDRの活性は、前出のSala-Newby他に記述されているように、AMPの加水分解によって生成される遊離アデノシンの量を決定することによって測定できる。短時間の間、37℃の適切なバッファ内でHYDRとAMPを10分間インキュベートする。遊離アデノシンをAMPから分離して逆相HPLCで測定する。

10

【0287】

別方法として、ADPリボシルアルギニンの加水分解によってHYDR活性を測定することもできる(Konczalik, P. および J. Moss (1999) J. Biol. Chem. 274:16736-16740)。50 mMリン酸カリウム(pH 7.5)、5 mMジチオスレイトール、10 mM MgCl₂の中で50 ngのHYDRを100 μMのADPリボシル[¹⁴C]アルギニン(78,000 cpm)と共にインキュベートする(最終容積100 μl)。37℃で1時間経過した後、サンプルの90 μlをAffiGel 601(boronate)カラム(0.5 × 4 cm)に注入する。このカラムを5倍の1 mlの0.1 Mグリシン(pH 9.0)、0.1 M NaClおよび10 mM MgCl₂の溶液で平衡化し、溶出する。全溶出液中の遊離¹⁴C-Argを液体シンチレーション計数により測定する。

20

【0288】

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

(表の簡単な説明)

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

30

【0289】

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

【0290】

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

【0291】

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

40

【0292】

表5は、本発明のポリヌクレオチドの代表的なcDNAライブラリを示す。

【0293】

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

【0294】

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

50

【表 1】

表 1

Incyte パート ID	Incyte パート ID	Incyte パート ID	Incyte パート ID	Incyte パート ID
3731660	3731660CD1	3731660CD1	3731660CD1	3731660CD1
2656987	2656987CD1	2656987CD1	2656987CD1	2656987CD1
4009684	4009684CD1	4009684CD1	4009684CD1	4009684CD1
7473183	7473183CD1	7473183CD1	7473183CD1	7473183CD1

10

20

30

【表 2】

表 2

ポリペプチド SEQ ID NO.	Incyte ポリペプチド ID	GenBank ID NO.	確率スコア	GenBank 相同体
1	3731660CD1	g7267246	2.6e-59	[シロイヌナズナ] 推定上のアデノシンデアミナーゼ
2	2656987CD1	g9106985	2.0e-28	[Xylella fastidiosa] エンドヌクレアーゼ V (デオキシイノシン3' エンドヌクレアーゼ) Silvestri, M.L. 他 (2000) Nature 405:151-157
3	4009684CD1	g402478	1.2e-79	(ヒト) ADP リボシルアルギニルヒドロラーゼ (Takada, T. 他 (1993) J. Biol. Chem. 268:17837-17843)
4	7473183CD1	g1360379 g4902474	0 3.1e-161	[ヒト] 5'ヌクレオチダーゼ I [ハト] 5'ヌクレオチダーゼ (Sala-Newby, G.B. 他 (1999) J. Biol. Chem. 274:17789-17793)

10

20

30

【表 3】

表 3

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチン配列、ドメインおよびモチーフ	分析方法及びデータベース
1	3731660CD1	355	S124 S132 S336 S78 S79 T107 T117 T120 T159 T219 T337 T35 T53 T62	N112 N26 N314	デアミンナーゼ ヒドロラーゼ スクレスチド 代謝 アデノシン AMP アミノヒドロラーゼ アイソフォーム多重遺伝子 アラミリ PD005960:E65-W345 アデノシンおよびVAMPデアミンナーゼタンパク質 BL00485:T269-L323	BLAST_PRODOM
2	2656987CD1	308	S203 S261 S60 S71 T178 T258 T300		アデノシン/AMP デアミンナーゼ A_deaminase:P9-K344 タンパク質 エンドスクレアーゼ ヒドロラーゼ スクレアーゼ DNA 修繕 推定上 V デオキシイノシン 3' エントヌクレアーゼ PD009651:V51-E245	HMMER_PPFAM BLAST_PRODOM
3	4009684CD1	354	S12 S238 S264 S290 S350 S49 T151 T210 T259	N26	ADP リボシルアルギニン ヒドロラーゼ ADP リボシルアルギニン 切断 酵素 マグネシウム PD024785:M1-E353	BLAST_PRODOM
4	7473183CD1	366	S219 S55 T279 T293 T352 Y69			MOTIFS

【 表 4 】

10

20

30

表 4

ホムレオチド SEQ ID NO:	Incyte ホムレオチド ID	配列長	選択された断片	配列断片	5'位置	3'位置
5	3731660CB1	2269	1-47, 433-640, 1650-2269	3941090F8 (SCORNOT04)	1	738
				71742519V1	758	1429
				71743734V1	632	1130
				71744365V1	1301	1826
6	2656987CB1	1150	233-363	71743665V1	1652	2269
				6740376H1 (BRADIT02)	1	489
7	4009684CB1	2000	1-248	92165516	672	1150
				2656987T6 (LUNGUT09)	510	1122
				2656987F6 (LUNGUT09)	324	859
				70791352V1	483	1036
				4009684F6 (MUSCNOT10)	1092	1632
				70788530V1	1282	1768
				5277542F7 (MUSLNOT01)	1	507
8	7473183CB1	1101	1-822, 993-1101	70792322V1	1431	2000
				189560T6 (CARDNOT01)	661	1289
				GNN.95262300_000003_006.edit	1	1101

【表 5】

10

20

30

表 5

ポリヌクレオチド SEQ ID NO:	inocyte プロジェクト ID	代表的ライブラリ
5	37311660CB1	LUNCAST01
6	2656987CB1	LUNGTUT09
7	4009684CB1	LVENNCT03

【表 6】

10

20

表 6

ライブラリ	ベクター	ライブラリの説明
LUNGAST01	PSPORT1	ライブラリは、脳外傷で死亡した17才白人男子の肺組織から単離されたRNAを用いて作製された。患者の病歴には、喘息がある。
LUNGTUT09	P1N1CY	ライブラリは、肺の部分除去中に68才白人男性から取り除かれた肺腫瘍組織から単離されたRNAを用いて作製された。病歴字後5で は、侵襲性のグレード3扁平上皮細胞癌と転移性腫瘍が見られた。患者の病歴はタイプIIの糖尿病、甲状腺疾患、抑鬱 性障害、高脂血症、高血圧、およびタバコの使用等である。
LYENNOT03	PSPORT1	ライブラリは31才の男性の左心室組織から単離したRNAを用いて作製した。

【表 7】

10

20

30

表 7-1

プログラム	説明	引用文献	パラメータ関連値
ABI FACTURA	核酸配列においてペクター配列を除去して、不定の塩基をマスクするプログラム	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	Fast Data Finderは、アミノ酸または核酸配列の比較、注釈付けに有用である。	Applied Biosystems, Foster City, CA. Paracel Inc., Pasadena, CA.	不一致<50%
ABI AutoAssembler	核酸配列を構築するプログラム。	Applied Biosystems, Foster City, CA.	
BLAST	Basic Local Alignment Search Tool はアミノ酸および核酸配列の配列類似性検索に有用であり、BLAST には5つの機能がある：blastp、blastn、blastx、tblastnおよびtblastx。	Altschul, S.F. 他 (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. 他 (1997) (1997) Nucleic Acids Res. 25:3389-3402.	ESTs: 確率値=1.0E-8 以下 完全長配列: 確率値= 1.0E-10以下
FASTA	問合せ配列と同種の配列群配との類似性を検索する Pearson および Lipman アルゴリズム。FASTAは最低5つの機能からなる：fasta、tfasta、fastx、tfastxおよび Ussearch。	Pearson, W.R. 及び Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1980) Methods Enzymol. 183:63-98; および Smith, T.F. および M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E 値=1.0E-6 構築された ESTs: fasta 同一性=95%以上 一致した長さ=200塩基以上 fastx E 値=1.0E-8 以下 完全長配列: fastx スコア=100以上
BLIMPS	配列をBLOCKS、PRINTS、DOMO、PRODOMおよびPFAM データベースの配列と対応させて遺伝子ファミリー、配列相同性および構造的フィンガープリント領域を検索するBLocks IMProved サーチャー。	Henikoff, S. 及び J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. 及び U.S. Henikoff (1996) Methods Enzymol. 266:88-105; Attwood, T.K. 他. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	確率値= 1.0E-3以下
HMMER	PFAMのようなタンパク質ファミリーコンセンサス配列の隠れマルコフモデル(HMM)に基づいたデータベースに対して問合せ配列を検索するアルゴリズム。	Krogh, A. 他. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. 他 (1988) Nucleic Acids Res. 26:320-322; Durbin, R. 他. (1998) Our World View, in a Nutshell, Cambridge Univ. Press, 1-350ページの. 1-350.	PFAM ヒット: 確率値=1.0E-3 以下 シングルレベル: ヒット: スコア<0以上

10

20

30

【表 8】

表 7-2

プログラム	説明	引用文献	パラメータ関連
ProfileScan	Prositeで定義された配列パターンと一致するタンパク質配列の構造的モチーフおよび配列モチーフを検索するアルゴリズム。	Gribskov, M. 他 (1988) CABIOS 4(6):66; Gribskov, M. 他 (1989) Methods Enzymol.183:146-159; Baroch, A. 他 (1997) Nucleic Acids Res. 25:217-221.	標準化された質のスコアとその特定のPrositeモチーフに対するCCG指定” HIGH” 値 通常、スコア=1.4-2.1.
Phred	高い感度と確率で自動配列決定機のトレースを調べる塩基誤出しアルゴリズム。	Ewing, B. 他 (1998) Genome Res. 8:175-185; Ewing, B. 及び P. Green (1998) Genome Res. 8:186-194.	
Phrap	Smith-Waterman アルゴリズムの効率的なインプリメンテーションに基づくSWATやCrossMatchを含むPhils Revised Assembly プログラムで、配列相同性の検索やDNA配列の構築に有用である。	Smith, T.F. および M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. および M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	スコア=120以上 一致した長さ=56以上
Consed	Phrap で構築したものの表示、編集用グラフィカルツール。	Gordon, D. 他 (1988) Genome Res. 8:195-202.	
SPScan	タンパク質配列をスキヤンして、分泌シグナルペプチドの存在を調べる重み付けマトリックス解析プログラム。	Nielson, H. 他 (1997) Protein Engineering 10:1-6; Claverie, J.M. 及び S. Audic (1997) CABIOS 12:431-439.	スコア=3.5以上
TMAP	加重マトリックスを用いて蛋白配列での膜貫通セグメントの描写および配向を決定するプログラム。	Persson, B. および P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. および P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	隠れMarkov モデル(HMM)を使って蛋白配列の膜貫通セグメントを構写し、配向を決定するプログラム。	Sonhammer, E.L. 他 (1998) Proc.Sixth Intl. Conf. on Intelligent Systems for Mol.Biol., Glasgow 他, eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, 175-182ページ。175-182.	
Motifs	Prositeで定義された配列と一致したパターンのあるミノ酸配列を検索するプログラム。	Bairoch, A. 他 (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual 第9版, M51-59頁, Genetics Computer Group, Madison, WI	

10

20

30

【国際公開パンフレット】

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
4 April 2002 (04.04.2002)

PCT

(10) International Publication Number
WO 02/26998 A2(51) International Patent Classification: **C12N 15/55**,
9/14, C07K 16/40, G01N 33/573, 33/577, C12Q 1/34,
1/68, A61K 38/46, 39/395, C12N 1/21, 1/15, 5/10, A01K
67/027, A01H 5/00Road #102, Glenview, IL 60025 (US). **YAO, Monique**,
G. [US/US]; 111 Frederick Court, Mountain View, CA
94043 (US). **LAL, Preeti** [IN/US]; P.O. Box 5142, Santa
Clara, CA 95056 (US). **THORNTON, Michael** [US/US];
9 Medway Road, Woodside, CA 94062 (US).

(21) International Application Number: PCT/US01/30310

(74) Agents: **HAMLET-COX, Diana** et al.; Incyte Genomics,
Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).(22) International Filing Date:
28 September 2001 (28.09.2001)

(25) Filing Language: English

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

(26) Publication Language: English

(30) Priority Data:
60/237,093 29 September 2000 (29.09.2000) US
60/238,370 6 October 2000 (06.10.2000) US
60/241,284 17 October 2000 (17.10.2000) US(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD,
TG).(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): **YUE, Henry**
[US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US).
BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road,
San Leandro, CA 94577 (US). **WARREN, Bridget, A.**
[US/US]; 10130 Parkwood Drive #2, Cupertino, CA 95014
(US). **TRIBOULEY, Catherine, M.** [US/US]; 1121 Ten-
nessee Street, #5, San Francisco, CA 94107 (US). **TANG,
Y., Tom** [US/US]; 4230 Ranwick Court, San Jose, CA
95118 (US). **KHAN, Farrah, A.** [IN/US]; 3617 Central

Published:

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

WO 02/26998 A2

(54) Title: HYDROLASES

(57) Abstract: The invention provides human hydrolases (HYDR) and polynucleotides which identify and encode HYDR. 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 HYDR.

WO 02/26998

PCT/US01/30310

HYDROLASES**TECHNICAL FIELD**

This invention relates to nucleic acid and amino acid sequences of hydrolases and to the use of these sequences in the diagnosis, treatment, and prevention of immune system disorders, immune deficiencies, neurological disorders, pulmonary disorders, and cell proliferative disorders, including cancer, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of hydrolases.

BACKGROUND OF THE INVENTION

Hydrolases are a class of enzymes that catalyze the cleavage of various covalent bonds in a substrate by the introduction of a molecule of water. The reaction involves a nucleophilic attack by the water molecule's oxygen atom on a target bond in the substrate. The water molecule is split across the target bond, breaking the bond and generating two product molecules. Hydrolases participate in reactions essential to such functions as synthesis and degradation of cell components, and for regulation of cell functions including cell signaling, cell proliferation, inflammation, apoptosis, secretion and excretion. Hydrolases are involved in key steps in disease processes involving these functions. Hydrolytic enzymes, or hydrolases, may be grouped by substrate specificity into classes including phosphatases, peptidases, lysophospholipases, phosphodiesterases, glycosidases, and glyoxalases.

Phosphatases hydrolytically remove phosphate groups from proteins, an energy-providing step that regulates many cellular processes, including intracellular signaling pathways that in turn control cell growth and differentiation, cell-cell contact, the cell cycle, and oncogenesis.

Peptidases, also called proteases, cleave peptide bonds that form the backbone of peptide or protein chains. Proteolytic processing is essential to cell growth, differentiation, remodeling, and homeostasis as well as inflammation and the immune response. Since typical protein half-lives range from hours to a few days, peptidases are continually cleaving precursor proteins to their active form, removing signal sequences from targeted proteins, and degrading aged or defective proteins. Peptidases function in bacterial, parasitic, and viral invasion and replication within a host. Examples of peptidases include trypsin and chymotrypsin, components of the complement cascade and the blood-clotting cascade, lysosomal cathepsins, calpains, pepsin, renin, and chymosin (Beynon, R.J. and J.S. Bond (1994) Proteolytic Enzymes: A Practical Approach, Oxford University Press, New York, NY, pp. 1-5).

Lysophospholipases (LPLs) regulate intracellular lipids by catalyzing the hydrolysis of ester

WO 02/26998

PCT/US01/30310

bonds to remove an acyl group, a key step in lipid degradation. Small LPL isoforms, approximately 15-30 kD, function as hydrolases; larger isoforms function both as hydrolases and transacylases. A particular substrate for LPLs, lysophosphatidylcholine, causes lysis of cell membranes. LPL activity is regulated by signaling molecules important in numerous pathways, including the inflammatory response.

The phosphodiesterases 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 endo- and exo-nucleases, which are essential to cell growth and replication as well as protein synthesis. Endonuclease V (deoxyinosine 3'-endonuclease) is an example of a type II site-specific deoxyribonuclease, a putative DNA repair enzyme that cleaves DNAs containing hypoxanthine, uracil, or mismatched bases. *Escherichia coli* endonuclease V has been shown to cleave DNA containing deoxyxanthosine at the second phosphodiester bond 3' to deoxyxanthosine, generating a 3'-hydroxyl and a 5'-phosphoryl group at the nick site (He, B. et al. (2000) Mutat. Res. 459:109-114). It has been suggested that *Escherichia coli* endonuclease V plays a role in the removal of deaminated guanine, i.e., xanthine, from DNA, thus helping to protect the cell against the mutagenic effects of nitrosative deamination (Schouten KA and Weiss B (1999) Mutat. Res. 435:245-254). In eukaryotes, the process of tRNA splicing requires the removal of small tRNA introns that interrupt the anticodon loop 1 base 3' to the anticodon. This process requires the stepwise action of an endonuclease, a ligase, and a phosphotransferase (Hong, L. et al. (1998) Science 280:279-284). Ribonuclease P (RNase P) is a ubiquitous RNA processing endonuclease that is required for generating the mature tRNA 5'-end during the tRNA splicing process. This is accomplished through the catalysis of the cleavage of P-3'O bonds to produce 5'-phosphate and 3'-hydroxyl end groups at a specific site on pre-tRNA. Catalysis by RNase P is absolutely dependent on divalent cations such as Mg²⁺ or Mn²⁺ (Kurz, J.C. et al. (2000) Curr. Opin. Chem. Biol. 4:553-558). Substrate recognition mechanisms of RNase P have been demonstrated to be well conserved among the Eucarya, the Archaea, and the Bacteria (Fabbri, S. et al. (1998) Science 280:284-286). In *S. cerevisiae*, a gene designated POP1 for 'processing of precursor RNAs', encodes a protein component of both RNase P and RNase MRP, another RNA processing protein. Mutations in yeast POP1 have been shown to be lethal (Lygerou, Z. et al. (1994) Genes Dev. 8:1423-1433). Another phosphodiesterase is acid sphingomyelinase, which hydrolyzes the membrane phospholipid sphingomyelin to ceramide and phosphorylcholine. Phosphorylcholine is used in the synthesis of phosphatidylcholine, which is involved in numerous intracellular signaling pathways. Ceramide is an essential precursor for the generation of gangliosides, membrane lipids found in high concentration in neural tissue. Defective acid sphingomyelinase phosphodiesterase leads to a build-up of sphingomyelin molecules in lysosomes,

WO 02/26998

PCT/US01/30310

resulting in Niemann-Pick disease.

Glycosidases catalyze the cleavage of hemiacetyl bonds of glycosides, which are compounds that contain one or more sugar. Mammalian lactase-phlorizin hydrolase, for example, is an intestinal enzyme that splits lactose. Mammalian beta-galactosidase removes the terminal galactose from
5 gangliosides, glycoproteins, and glycosaminoglycans, and deficiency of this enzyme is associated with a gangliosidosis known as Morquio disease type B. Vertebrate lysosomal alpha-glucosidase, which hydrolyzes glycogen, maltose, and isomaltose, and vertebrate intestinal sucrase-isomaltase, which hydrolyzes sucrose, maltose, and isomaltose, are widely distributed members of this family with highly conserved sequences at their active sites.

10 The glyoxylase system is involved in gluconeogenesis, the production of glucose from storage compounds in the body. It consists of glyoxylase I, which catalyzes the formation of S-D-lactoylglutathione from methylglyoxal, a side product of triose-phosphate energy metabolism, and glyoxylase II, which hydrolyzes S-D-lactoylglutathione to D-lactic acid and reduced glutathione. Glyoxylases are involved in hyperglycemia, non-insulin-dependent diabetes mellitus, the detoxification
15 of bacterial toxins, and in the control of cell proliferation and microtubule assembly.

A small subclass of hydrolases acting on ether bonds includes the thioether hydrolases. S-adenosyl-L-homocysteine hydrolase, also known as AdoHcyase or SAHH (PROSITE PDOC00603; EC 3.3.1.1), is a thioether hydrolase first described in rat liver extracts as the activity responsible for the reversible hydrolysis of S-adenosyl-L-homocysteine (AdoHcy) to adenosine and homocysteine
20 (Sganga, M.W. et al. (1992) PNAS 89:6328-6332). SAHH is a cytosolic enzyme that has been found in all cells that have been tested, with the exception of *Escherichia coli* and certain related bacteria (Walker, R.D. et al. (1975) Can. J. Biochem. 53:312-319; Shimizu, S. et al. (1988) FEMS Microbiol. Lett. 51:177-180; Shimizu, S. et al. (1984) Eur. J. Biochem. 141:385-392). SAHH activity is dependent on NAD⁺ as a cofactor. Deficiency of SAHH is associated with hypermethioninemia
25 (Online Mendelian Inheritance in Man (OMIM) #180960 Hypermethioninemia), a pathologic condition characterized by neonatal cholestasis, failure to thrive, mental and motor retardation, facial dysmorphism with abnormal hair and teeth, and myocardiopathy (Labrune, P. et al. (1990) J. Pediat. 117:220-226).

Another subclass of hydrolases includes those enzymes which act on carbon-nitrogen (C-N)
30 bonds other than peptide bonds. To this subclass belong those enzymes hydrolyzing amides, amidines, and other C-N bonds. This subclass is further subdivided on the basis of the substrate specificity such as linear amides, cyclic amides, linear amidines, cyclic amidines, nitriles and other compounds. A hydrolase belonging to the sub-subclass of enzymes acting on the cyclic amidines is adenosine deaminase (ADA). ADA catalyzes the breakdown of adenosine to inosine. ADA is present in many

WO 02/26998

PCT/US01/30310

mammalian tissues, including placenta, muscle, lung, stomach, digestive diverticulum, spleen, erythrocytes, thymus, seminal plasma, thyroid, T-cells, bone marrow stem cells, and liver. A subclass of ADAs, ADAR, act on RNA and are classified as RNA editases. An ADAR from *Drosophila*, dADAR, has been shown to be expressed in the developing nervous system, making it a candidate for the editase that acts on para voltage-gated Na⁺ channel transcripts in the central nervous system (Palladino, M.J. et al. (2000) RNA 6:1004-1018). A deficiency of ADA causes profound lymphopenia with severe combined immunodeficiency (SCID). Cells from patients with ADA deficiency contain less than normal, and sometimes undetectable, amounts of ADA catalytic activity and ADA protein. It has been shown that ADA deficiency stems from genetic mutations in the ADA gene, resulting in SCID (Hershfield, M.S. (1998) Semin. Hematol. 4:291-298). Metabolic consequences of ADA deficiency in mice have been found to be associated with defects in alveogenesis, pulmonary inflammation, and airway obstruction (Blackburn, M.R. et al. (2000) J. Exp. Med. 192:159-170).

Pancreatic ribonucleases (RNase) are pyrimidine-specific endonucleases found in high quantity in the pancreas of certain mammalian taxa and of some reptiles (Beintema, J.J. et al (1988) Prog. Biophys. Mol. Biol. 51:165-192). Proteins in the mammalian pancreatic RNase superfamily are noncytosolic endonucleases that degrade RNA through a two-step transphosphorolytic-hydrolytic reaction (Beintema, J.J. et al. (1986) Mol. Biol. Evol. 3:262-275). Specifically, the enzymes are involved in endonucleolytic cleavage of 3'-phosphomononucleotides and 3'-phosphooligonucleotides ending in C-P or U-P with 2',3'-cyclic phosphate intermediates. Ribonucleases can unwind the DNA helix by complexing with single-stranded DNA; the complex arises by an extended multi-site cation-anion interaction between lysine and arginine residues of the enzyme and phosphate groups of the nucleotides. Some of the enzymes belonging to this family appear to play a purely digestive role, whereas others exhibit potent and unusual biological activities (D'Alessio, G. (1993) Trends Cell Biol. 3:106-109). Proteins belonging to the pancreatic RNase family include: bovine seminal vesicle and brain ribonucleases; kidney non-secretory ribonucleases (Beintema, J.J. et al (1986) FEBS Lett. 194:338-343); liver-type ribonucleases (Rosenberg, H.F. et al. (1989) PNAS U.S.A. 86:4460-4464); angiogenin, which induces vascularisation of normal and malignant tissues; eosinophil cationic protein (Hofsteenge, J. et al. (1989) Biochemistry 28:9806-9813), a cytotoxin and helminthotoxin with ribonuclease activity; and frog liver ribonuclease and frog sialic acid-binding lectin. The sequences of pancreatic RNases contain 4 conserved disulphide bonds and 3 amino acid residues involved in the catalytic activity.

Serine hydrolases are a functional class of hydrolytic enzymes that contain a serine residue in their active site. This class of enzymes contains proteinases, esterases, and lipases which hydrolyze a variety of substrates and, therefore, have different biological roles. Proteins in this superfamily can be

WO 02/26998

PCT/US01/30310

further grouped into subfamilies based on substrate specificity or amino acid similarities (Puentes, X.S. and Lopez-Ont, C. (1995) *J. Biol. Chem.* 270(21): 12926-12932).

Carboxylesterases are proteins that hydrolyze carboxylic esters and are classified into three categories- A, B, and C. Most type-B carboxylesterases are evolutionarily related and are considered
5 to comprise a family of proteins. The type-B carboxylesterase family of proteins includes vertebrate acetylcholinesterase, mammalian liver microsomal carboxylesterase, mammalian bile-salt-activated lipase, and duck fatty acyl-CoA hydrolase. Some members of this protein family are not catalytically active but contain a domain related evolutionarily to other type-B carboxylesterases, such as thyroglobulin and *Drosophila* protein neuractin.

10 Nucleotidases catalyze the formation of free nucleosides from nucleotides. The cytosolic nucleotidase cN-I (5' nucleotidase-I) cloned from pigeon heart catalyzes the formation of adenosine from AMP generated during ATP hydrolysis (Sala-Newby, G.B. et al. (1999) *J. Biol. Chem.* 274:17789-17793). Increased adenosine concentration is thought to be a signal of metabolic stress, and adenosine receptors mediate effects including vasodilation, decreased stimulatory neuron firing
15 and ischemic preconditioning in the heart (Schrader, J. (1990) *Circulation* 81:389-391; Rubino, A. et al. (1992) *Eur. J. Pharmacol.* 220:95-98; de Jong, J.W. et al. (2000) *Pharmacol. Ther.* 87:141-149). Deficiency of pyrimidine 5'-nucleotidase can result in hereditary hemolytic anemia (OMIM Entry 266120).

ADP-ribosylation is a reversible post-translational protein modification in which an ADP-
20 ribose moiety is transferred from β -NAD to a target amino acid such as arginine or cysteine. ADP-ribosylarginine hydrolases regenerate arginine by removing ADP-ribose from the protein, completing the ADP-ribosylation cycle (Moss, J. et al. (1997) *Adv. Exp. Med. Biol.* 419:25-33). ADP-ribosylation is a well-known reaction among bacterial toxins. Cholera toxin, for example, disrupts the adenylyl cyclase system by ADP-ribosylating the α -subunit of the stimulatory G-protein, causing an
25 increase in intracellular cAMP (Moss, J. and Vaughan, M. (eds) (1990) *ADP-ribosylating Toxins and G-Proteins: Insights into Signal Transduction*, American Society for Microbiology, Washington, D.C.). ADP-ribosylation may also have a regulatory function in eukaryotes, affecting such processes as cytoskeletal assembly (Zhou, H. et al. (1996) *Arch. Biochem. Biophys.* 334:214-222) and cell proliferation in cytotoxic T-cells (Wang, J. et al. (1996) *J. Immunol.* 156:2819-2827).

30 The discovery of new hydrolases, 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 immune system disorders, immune deficiencies, neurological disorders, pulmonary disorders, and cell proliferative disorders, including cancer, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of hydrolases.

WO 02/26998

PCT/US01/30310

SUMMARY OF THE INVENTION

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

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:1-4, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-4, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-4, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-4. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-4. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:5-8.

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:1-4, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-4, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-4, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-4. 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:1-4, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-4, c) a

WO 02/26998

PCT/US01/30310

biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-4, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-4. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with
5 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:1-4, b) a polypeptide comprising a naturally
10 occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-4, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-4, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-4.

The invention further provides an isolated polynucleotide selected from the group consisting of
15 a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:5-8, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:5-8, 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
20 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
25 a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:5-8, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:5-8, 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
30 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.

WO 02/26998

PCT/US01/30310

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:5-8, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:5-8, 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:1-4, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-4, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-4, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-4, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-4. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional HYDR, comprising administering to a patient in need of such treatment the composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-4, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-4, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-4, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-4. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional HYDR, comprising administering to a patient in need of such treatment the composition.

WO 02/26998

PCT/US01/30310

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:1-4, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid
5 sequence selected from the group consisting of SEQ ID NO:1-4, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-4, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-4. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the
10 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 HYDR, comprising administering to a patient in need of such treatment the composition.

The invention further provides a method of screening for a compound that specifically binds to
15 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:1-4, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-4, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-4, and d) an
20 immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-4. 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
25 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:1-4, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-4, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-4, and d) an
30 immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-4. 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

WO 02/26998

PCT/US01/30310

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

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a polynucleotide sequence selected from the group consisting of SEQ ID NO:5-8, 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:5-8, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:5-8, iii) a polynucleotide having a sequence complementary to i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:5-8, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:5-8, 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

WO 02/26998

PCT/US01/30310

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.

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

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

10 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

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

20 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.

25 Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the
30 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

"HYDR" refers to the amino acid sequences of substantially purified HYDR obtained from

WO 02/26998

PCT/US01/30310

any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

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

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

20 The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent HYDR. 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 HYDR is retained. For example, negatively charged amino acids may
25 include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

30 The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to a sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

WO 02/26998

PCT/US01/30310

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

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

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

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

WO 02/26998

PCT/US01/30310

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

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

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA;
10 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
15 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
20 functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic" refers to the capability of the natural, recombinant, or synthetic HYDR, 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
25 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.
30 Compositions comprising polynucleotide sequences encoding HYDR or fragments of HYDR 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 02/26998

PCT/US01/30310

"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 an

WO 02/26998

PCT/US01/30310

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.

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

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

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

A fragment of SEQ ID NO:5-8 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:5-8, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:5-8 is useful, for example, in
30 hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:5-8 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:5-8 and the region of SEQ ID NO:5-8 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-4 is encoded by a fragment of SEQ ID NO:5-8. A fragment of

WO 02/26998

PCT/US01/30310

SEQ ID NO:1-4 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-4. For example, a fragment of SEQ ID NO:1-4 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-4. The precise length of a fragment of SEQ ID NO:1-4 and the region of SEQ ID NO:1-4 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

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

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

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

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

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/b2.html>. The

WO 02/26998

PCT/US01/30310

"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:

5 *Matrix: BLOSUM62*

Reward for match: 1

Penalty for mismatch: -2

Open Gap: 5 and Extension Gap: 2 penalties

Gap x drop-off: 50

10 *Expect: 10*

Word Size: 11

Filter: on

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

20 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.

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

WO 02/26998

PCT/US01/30310

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

10 *Expect: 10*

Word Size: 3

Filter: on

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

20 "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.

25 "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

WO 02/26998

PCT/US01/30310

varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml sheared, denatured salmon sperm DNA.

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

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

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

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

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

WO 02/26998

PCT/US01/30310

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of HYDR 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 HYDR which is useful in any of the antibody production methods disclosed herein or known in the art.

5 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.

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

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
15 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
20 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
25 may be pegylated to extend their lifespan in the cell.

"Post-translational modification" of an HYDR 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 HYDR.

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

WO 02/26998

PCT/US01/30310

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

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

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

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

WO 02/26998

PCT/US01/30310

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

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

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

A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions (UTRs). Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

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

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

The term "sample" is used in its broadest sense. A sample suspected of containing HYDR, nucleic acids encoding HYDR, 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

WO 02/26998

PCT/US01/30310

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
5 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.

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

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

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

"Transformation" describes a process by which exogenous DNA is introduced into a recipient cell. Transformation may occur under natural or artificial conditions according to various methods
20 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
25 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
30 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,

WO 02/26998

PCT/US01/30310

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*.

5 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 10 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding 15 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 20 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 25 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 30 identity over a certain defined length of one of the polypeptides.

THE INVENTION

The invention is based on the discovery of new human hydrolases (HYDR), the polynucleotides encoding HYDR, and the use of these compositions for the diagnosis, treatment, or prevention of immune system disorders, immune deficiencies, neurological disorders, pulmonary

WO 02/26998

PCT/US01/30310

disorders, and cell proliferative disorders, including cancer.

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
5 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
10 BLAST analysis against the GenBank protein (genpept) database. Columns 1 and 2 show the polypeptide sequence identification number (Polypeptide SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for polypeptides of the invention. Column 3 shows the GenBank identification number (Genbank ID NO:) of the nearest GenBank homolog. Column 4 shows the probability score for the match between each polypeptide and its GenBank
15 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
20 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
25 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 hydrolases. For example, SEQ ID NO:1 is 38% identical to *Arabidopsis thaliana* putative adenosine deaminase (GenBank ID g7267246) as determined
30 by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 2.6e-59, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:1 also contains an adenosine/AMP deaminase domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS analysis provides further corroborative evidence that SEQ ID NO:1 is an adenosine deaminase. In an alternative

WO 02/26998

PCT/US01/30310

example, SEQ ID NO:2 is 40% identical to *Xylella fastidiosa* endonuclease V (deoxyinosine 3'-
endonuclease) (GenBank ID g9106985) as determined by the Basic Local Alignment Search Tool
(BLAST). (See Table 2.) The BLAST probability score is 2.0e-28, which indicates the probability of
obtaining the observed polypeptide sequence alignment by chance. Data from PRODOM analysis
5 provides further corroborative evidence that SEQ ID NO:2 is an endonuclease. In an alternative
example, SEQ ID NO:3 is 46% identical to human ADP-ribosylarginine hydrolase (GenBank ID
g402478) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The
BLAST probability score is 1.2e-79, which indicates the probability of obtaining the observed
polypeptide sequence alignment by chance. SEQ ID NO:3 also contains a domain characteristic for
10 ADP-ribosylarginine hydrolases as found in the PRODOM database. (See Table 3.) Similarly, SEQ
ID NO:4 is 85% identical to *Columba livia* 5' nucleotidase (GenBank ID g4902474). The BLAST
probability score is 3.1e-161. These data provide evidence that SEQ ID NO:3 and SEQ ID NO:4 are
hydrolases. The algorithms and parameters for the analysis of SEQ ID NO:1-4 are described in Table
7.

15 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.
20 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:5-8 or that distinguish between SEQ ID NO:5-8 and related
polynucleotide sequences. Column 5 shows identification numbers corresponding to cDNA
sequences, coding sequences (exons) predicted from genomic DNA, and/or sequence assemblages
25 comprised of both cDNA and genomic DNA. These sequences were used to assemble the full length
polynucleotide sequences of the invention. Columns 6 and 7 of Table 4 show the nucleotide start (5')
and stop (3') positions of the cDNA and/or genomic sequences in column 5 relative to their respective
full length sequences.

The identification numbers in Column 5 of Table 4 may refer specifically, for example, to
30 Incyte cDNAs along with their corresponding cDNA libraries. For example, 3941090F8 is the
identification number of an Incyte cDNA sequence, and SCORNOT04 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., 71742519V1). Alternatively, the identification numbers in column 5 may
refer to GenBank cDNAs or ESTs (e.g., g2165516) which contributed to the assembly of the full

WO 02/26998

PCT/US01/30310

length polynucleotide sequences. In addition, the identification numbers in column 5 may identify sequences derived from the ENSEMBL (The Sanger Centre, Cambridge, UK) database (*i.e.*, those sequences including the designation "ENST"). Alternatively, the identification numbers in column 5 may be derived from the NCBI RefSeq Nucleotide Sequence Records Database (*i.e.*, those sequences including the designation "NM" or "NT") or the NCBI RefSeq Protein Sequence Records (*i.e.*, those sequences including the designation "NP"). Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. For example, FL_XXXXX_N₁_N₂_YYYY_N₃_N₄ represents a "stitched" sequence in which XXXXX is the identification number of the cluster of sequences to which the algorithm was applied, and YYYY is the number of the prediction generated by the algorithm, and N_{1,2,3,4}, if present, represent specific exons that may have been manually edited during analysis (See Example V). Alternatively, the identification numbers in column 5 may refer to assemblages of exons brought together by an "exon-stretching" algorithm. For example, FLXXXXX_gAAAAA_gBBBBB_1_N is the identification number of a "stretched" sequence, with XXXXX being the Incyte project identification number, gAAAAA being the GenBank identification number of the human genomic sequence to which the "exon-stretching" algorithm was applied, gBBBBB being the GenBank identification number or NCBI RefSeq identification number of the nearest GenBank protein homolog, and N referring to specific exons (See Example V). In instances where a RefSeq sequence was used as a protein homolog for the "exon-stretching" algorithm, a RefSeq identifier (denoted by "NM," "NP," or "NT") may be used in place of the GenBank identifier (*i.e.*, gBBBBB).

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

Prefix	Type of analysis and/or examples of programs
GNN, GFG, ENST	Exon prediction from genomic sequences using, for example, GENSCAN (Stanford University, CA, USA) or FGENES (Computer Genomics Group, The Sanger Centre, Cambridge, UK).
GBI	Hand-edited analysis of genomic sequences.
FL	Stitched or stretched genomic sequences (see Example V).

WO 02/26998

PCT/US01/30310

INCY	Full length transcript and exon prediction from mapping of EST sequences to the genome. Genomic location and EST composition data are combined to predict the exons and resulting transcript.
------	---

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

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

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

The invention also encompasses polynucleotides which encode HYDR. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:5-8, which encodes HYDR. The polynucleotide sequences of SEQ ID NO:5-8, 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 HYDR. 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 HYDR. 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:5-8 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:5-8. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of HYDR.

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 HYDR, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be

WO 02/26998

PCT/US01/30310

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

Although nucleotide sequences which encode HYDR and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring HYDR under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding HYDR or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally
10 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 HYDR 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,
15 than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode HYDR and HYDR 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
20 mutations into a sequence encoding HYDR 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:5-8 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-
25 511.) 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
30 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 02/26998

PCT/US01/30310

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.)

5 The nucleic acid sequences encoding HYDR 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
10 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
15 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.)
20 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,
25 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)
30 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-

WO 02/26998

PCT/US01/30310

specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Applied Biosystems), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

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

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

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

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

WO 02/26998

PCT/US01/30310

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, HYDR itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques. (See, e.g.,

5 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 HYDR, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a
10 sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chicz, 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 HYDR, the nucleotide sequences encoding HYDR or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences
20 encoding HYDR. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding HYDR. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding HYDR and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may
25 be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) *Results Probl.*
30 *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 HYDR and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. (See, e.g., Sambrook, J. et al. (1989) *Molecular Cloning, A Laboratory*

WO 02/26998

PCT/US01/30310

Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding HYDR. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. (See, e.g., Sambrook, supra; Ausubel, supra; Van Heeke, G. and S.M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509; Engelhard, E.K. et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:3224-3227; Sandig, V. et al. (1996) *Hum. Gene Ther.* 7:1937-1945; Takamatsu, N. (1987) *EMBO J.* 6:307-311; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659; and Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) *Cancer Gen. Ther.* 5(6):350-356; Yu, M. et al. (1993) *Proc. Natl. Acad. Sci. USA* 90(13):6340-6344; Baller, R.M. et al. (1985) *Nature* 317(6040):813-815; McGregor, D.P. et al. (1994) *Mol. Immunol.* 31(3):219-226; and Verma, I.M. and N. Somia (1997) *Nature* 389:239-242.) The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding HYDR. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding HYDR can be achieved using a multifunctional *E. coli* vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding HYDR 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 HYDR are needed, e.g. for the production of antibodies, vectors which direct high level expression of HYDR 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 HYDR. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH

WO 02/26998

PCT/US01/30310

promoters, may be used in the yeast *Saccharomyces cerevisiae* or *Pichia pastoris*. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, *supra*; Bitter, G.A. et al. (1987) *Methods Enzymol.* 153:516-544; and Scorer, C.A. et al. (1994)

5 Bio/Technology 12:181-184.)

Plant systems may also be used for expression of HYDR. Transcription of sequences encoding HYDR 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

10 promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. (1984) *Science* 224:838-843; and Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., *The McGraw Hill Yearbook of Science and Technology* (1992) McGraw Hill, New York NY, pp. 191-196.)

15 In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding HYDR 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 HYDR in host cells. (See, e.g., Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

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

For long term production of recombinant proteins in mammalian systems, stable expression of HYDR in cell lines is preferred. For example, sequences encoding HYDR can be transformed into

30 cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the

WO 02/26998

PCT/US01/30310

introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* and *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., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), β glucuronidase and its substrate β -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

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

WO 02/26998

PCT/US01/30310

competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding HYDR include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide.

Alternatively, the sequences encoding HYDR, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding HYDR may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode HYDR may be designed to contain signal sequences which direct secretion of HYDR through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding HYDR may be ligated to a heterologous sequence resulting in translation of a

WO 02/26998

PCT/US01/30310

fusion protein in any of the aforementioned host systems. For example, a chimeric HYDR protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of HYDR 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 HYDR encoding sequence and the heterologous protein sequence, so that HYDR 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 HYDR 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.

HYDR of the present invention or fragments thereof may be used to screen for compounds that specifically bind to HYDR. At least one and up to a plurality of test compounds may be screened for specific binding to HYDR. 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 HYDR, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner. (See, e.g., Coligan, J.E. et al. (1991) *Current Protocols in Immunology* 1(2): Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which HYDR 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 HYDR, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, *Drosophila*, or *E. coli*. Cells expressing HYDR or cell membrane fractions which contain HYDR are then contacted with a test compound and binding, stimulation, or inhibition of activity of either HYDR or the

WO 02/26998

PCT/US01/30310

compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with HYDR, either in solution
5 or affixed to a solid support, and detecting the binding of HYDR 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.

HYDR of the present invention or fragments thereof may be used to screen for compounds
10 that modulate the activity of HYDR. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for HYDR activity, wherein HYDR is combined with at least one test compound, and the activity of HYDR in the presence of a test compound is compared with the activity of HYDR in the absence of the test compound. A change in the activity of HYDR in the presence of the test compound is indicative of a
15 compound that modulates the activity of HYDR. Alternatively, a test compound is combined with an *in vitro* or cell-free system comprising HYDR under conditions suitable for HYDR activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of HYDR 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 HYDR or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES)
20 cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent No. 5,175,383 and U.S. Patent No. 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and
25 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.
30 (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.

WO 02/26998

PCT/US01/30310

Polynucleotides encoding HYDR may also be manipulated *in vitro* in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. 5 (1998) Science 282:1145-1147).

Polynucleotides encoding HYDR 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 HYDR 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 10 are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress HYDR, e.g., by secreting HYDR 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

15 Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of HYDR and hydrolases. In addition, the expression of HYDR is closely associated with kidney, lung tumor, and cardiovascular tissues. Therefore, HYDR appears to play a role in immune system disorders, immune deficiencies, neurological disorders, pulmonary disorders, and cell proliferative disorders, including cancer. In the treatment of disorders associated with increased 20 HYDR expression or activity, it is desirable to decrease the expression or activity of HYDR. In the treatment of disorders associated with decreased HYDR expression or activity, it is desirable to increase the expression or activity of HYDR.

Therefore, in one embodiment, HYDR 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 HYDR. Examples of such disorders include, but are not limited to, an immune system disorder, such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic 30 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, hyper eosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid

WO 02/26998

PCT/US01/30310

arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; an immune deficiency, such as acquired

5 immunodeficiency syndrome (AIDS), X-linked agammaglobinemia of Bruton, common variable immunodeficiency (CVID), DiGeorge's syndrome (thymic hypoplasia), thymic dysplasia, isolated IgA deficiency, severe combined immunodeficiency disease (SCID), immunodeficiency with thrombocytopenia and eczema (Wiskott-Aldrich syndrome), Chediak-Higashi syndrome, chronic granulomatous diseases, hereditary angioneurotic edema, and immunodeficiency associated with

10 Cushing's disease; a neurological disorder, such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural

15 empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central

20 nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD),

25 akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a pulmonary disorder, such as congenital lung anomalies, atelectasis, pulmonary congestion and edema, pulmonary embolism, pulmonary hemorrhage, pulmonary infarction, pulmonary hypertension, vascular sclerosis, obstructive pulmonary disease, restrictive

30 pulmonary disease, chronic obstructive pulmonary disease, emphysema, chronic bronchitis, bronchial asthma, bronchiectasis, bacterial pneumonia, viral and mycoplasmal pneumonia, lung abscess, pulmonary tuberculosis, diffuse interstitial diseases, pneumoconioses, sarcoidosis, idiopathic pulmonary fibrosis, desquamative interstitial pneumonitis, hypersensitivity pneumonitis, pulmonary eosinophilia bronchiolitis obliterans-organizing pneumonia, diffuse pulmonary hemorrhage syndromes,

WO 02/26998

PCT/US01/30310

Goodpasture's syndromes, idiopathic pulmonary hemosiderosis, pulmonary involvement in collagen-vascular disorders, pulmonary alveolar proteinosis, lung tumors, inflammatory and noninflammatory pleural effusions, pneumothorax, pleural tumors, drug-induced lung disease, radiation-induced lung disease, and complications of lung transplantation; and a cell proliferative disorder, such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus.

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

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

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

In a further embodiment, an antagonist of HYDR may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of HYDR. Examples of such disorders include, but are not limited to, those immune system disorders, immune deficiencies, neurological disorders, pulmonary disorders, and cell proliferative disorders, including cancer, described above. In one aspect, an antibody which specifically binds HYDR 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 HYDR.

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

WO 02/26998

PCT/US01/30310

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.

5 An antagonist of HYDR may be produced using methods which are generally known in the art. In particular, purified HYDR may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind HYDR. Antibodies to HYDR 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
10 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 HYDR or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to
15 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
20 HYDR 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 HYDR 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 HYDR may be prepared using any technique which provides for the
25 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
30 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,

WO 02/26998

PCT/US01/30310

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 HYDR-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for HYDR may also be generated. For example, such fragments include, but are not limited to, F(ab)₂ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab)₂ fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between HYDR and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering HYDR 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 HYDR. Affinity is expressed as an association constant, K_a , which is defined as the molar concentration of HYDR-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 HYDR epitopes, represents the average affinity, or avidity, of the antibodies for HYDR. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular HYDR 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 HYDR-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 HYDR, preferably in active form, from the

WO 02/26998

PCT/US01/30310

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).

5 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 HYDR-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
10 Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding HYDR, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene expression can be achieved by designing complementary sequences or antisense molecules (DNA, RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding
15 HYDR. 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 HYDR. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ.)

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

30 In another embodiment of the invention, polynucleotides encoding HYDR may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) Science 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency

WO 02/26998

PCT/US01/30310

(Blase, R.M. et al. (1995) *Science* 270:475-480; Bordignon, C. et al. (1995) *Science* 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) *Cell* 75:207-216; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:643-666; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:667-703), thalassemias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) *Science* 270:404-410; Verma, I.M. and N. Somia (1997) *Nature* 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) *Nature* 335:395-396; Poeschla, E. et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as Candida albicans and Paracoccidioides brasiliensis; and protozoan parasites such as Plasmodium falciparum and Trypanosoma cruzi). In the case where a genetic deficiency in HYDR expression or regulation causes disease, the expression of HYDR from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

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

Expression vectors that may be effective for the expression of HYDR include, but are not limited to, the pCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX, PCR2-TOPOTA vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). HYDR 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

WO 02/26998

PCT/US01/30310

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

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding HYDR to cells which have one or more genetic abnormalities with respect to the expression of HYDR. 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 (Csote, M.E. et al. (1995) *Transplantation* 27:263-268). Potentially useful adenoviral vectors are

WO 02/26998

PCT/US01/30310

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

5 In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding HYDR to target cells which have one or more genetic abnormalities with respect to the expression of HYDR. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing HYDR 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
10 ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) *Exp. Eye Res.* 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent No. 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent No. 5,804,413 teaches the use of recombinant HSV d92
15 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) *J. Virol.* 73:519-532 and Xu, H. et al. (1994) *Dev. Biol.* 163:152-161, hereby incorporated by reference. The manipulation of cloned
20 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
25 deliver polynucleotides encoding HYDR 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,
30 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 HYDR into the alphavirus genome in place of the capsid-coding region results in the production of a large number of HYDR-coding RNAs and the synthesis of high levels of HYDR in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a

WO 02/26998

PCT/US01/30310

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 HYDR into a variety of cell types. The specific transduction of a subset of
5 cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr,
15 Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme
20 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 HYDR.

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,
25 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
30 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 HYDR. Such DNA sequences may be incorporated into a wide variety of

WO 02/26998

PCT/US01/30310

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 HYDR. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased HYDR expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding HYDR may be therapeutically useful, and in the treatment of disorders associated with decreased HYDR expression or activity, a compound which specifically promotes expression of the polynucleotide encoding HYDR 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 HYDR 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 HYDR 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

WO 02/26998

PCT/US01/30310

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

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

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

An additional embodiment of the invention relates to the administration of a 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 HYDR, antibodies to HYDR, and mimetics, agonists, antagonists, or inhibitors of HYDR.

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-

WO 02/26998

PCT/US01/30310

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
5 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
10 macromolecules comprising HYDR or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, HYDR 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
15 al. (1999) Science 285:1569-1572).

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

A therapeutically effective dose refers to that amount of active ingredient, for example HYDR or fragments thereof, antibodies of HYDR, and agonists, antagonists or inhibitors of HYDR, 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
25 calculating the ED₅₀ (the dose therapeutically effective in 50% of the population) or LD₅₀ (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD₅₀/ED₅₀ ratio. Compositions which exhibit large 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
30 preferably within a range of circulating concentrations that includes the ED₅₀ with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the 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

WO 02/26998

PCT/US01/30310

active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting compositions may be administered every 3 to 4 days, every week, or
5 biweekly depending on the half-life and clearance rate of the particular formulation.

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

DIAGNOSTICS

In another embodiment, antibodies which specifically bind HYDR may be used for the diagnosis of disorders characterized by expression of HYDR, or in assays to monitor patients being
15 treated with HYDR or agonists, antagonists, or inhibitors of HYDR. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for HYDR include methods which utilize the antibody and a label to detect HYDR in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of
20 reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring HYDR, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of HYDR expression. Normal or standard values for HYDR expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibodies to HYDR under
25 conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of HYDR 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 HYDR may be used for
30 diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of HYDR may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of HYDR, and to monitor regulation of HYDR levels during therapeutic intervention.

WO 02/26998

PCT/US01/30310

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding HYDR or closely related molecules may be used to identify nucleic acid sequences which encode HYDR. 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 HYDR, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the HYDR 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:5-8 or from genomic sequences including promoters, enhancers, and introns of the HYDR gene.

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

Polynucleotide sequences encoding HYDR may be used for the diagnosis of disorders associated with expression of HYDR. Examples of such disorders include, but are not limited to, an immune system disorder, such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; an immune deficiency, such as acquired immunodeficiency syndrome (AIDS), X-linked agammaglobinemia of Bruton, common

WO 02/26998

PCT/US01/30310

variable immunodeficiency (CVI), DiGeorge's syndrome (thymic hypoplasia), thymic dysplasia, isolated IgA deficiency, severe combined immunodeficiency disease (SCID), immunodeficiency with thrombocytopenia and eczema (Wiskott-Aldrich syndrome), Chediak-Higashi syndrome, chronic granulomatous diseases, hereditary angioneurotic edema, and immunodeficiency associated with

5 Cushing's disease; a neurological disorder, such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural

10 empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central

15 nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD),

20 akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a pulmonary disorder, such as congenital lung anomalies, atelectasis, pulmonary congestion and edema, pulmonary embolism, pulmonary hemorrhage, pulmonary infarction, pulmonary hypertension, vascular sclerosis, obstructive pulmonary disease, restrictive

25 pulmonary disease, chronic obstructive pulmonary disease, emphysema, chronic bronchitis, bronchial asthma, bronchiectasis, bacterial pneumonia, viral and mycoplasma pneumoniae pneumonia, lung abscess, pulmonary tuberculosis, diffuse interstitial diseases, pneumoconioses, sarcoidosis, idiopathic pulmonary fibrosis, desquamative interstitial pneumonitis, hypersensitivity pneumonitis, pulmonary eosinophilia, bronchiolitis obliterans-organizing pneumonia, diffuse pulmonary hemorrhage syndromes,

30 Goodpasture's syndromes, idiopathic pulmonary hemosiderosis, pulmonary involvement in collagen-vascular disorders, pulmonary alveolar proteinosis, lung tumors, inflammatory and noninflammatory pleural effusions, pneumothorax, pleural tumors, drug-induced lung disease, radiation-induced lung disease, and complications of lung transplantation; and a cell proliferative disorder, such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective

WO 02/26998

PCT/US01/30310

tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus. The polynucleotide sequences encoding HYDR 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 HYDR expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding HYDR may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding HYDR 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 HYDR 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 HYDR, 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 HYDR, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

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

WO 02/26998

PCT/US01/30310

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

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

In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding HYDR 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 HYDR are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSSCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed *in silico* SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

Methods which may also be used to quantify the expression of HYDR 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.

WO 02/26998

PCT/US01/30310

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.

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

In another embodiment, HYDR, fragments of HYDR, or antibodies specific for HYDR 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.

20 A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time. (See Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent No.
25 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
30 resultant transcript image would provide a profile of gene activity.

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

Transcript images which profile the expression of the polynucleotides of the present invention

WO 02/26998

PCT/US01/30310

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

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

Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by

WO 02/26998

PCT/US01/30310

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 HYDR to quantify the levels of HYDR 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 (Lusking, A. et al. (1999) *Anal. Biochem.* 270:103-111; Mendozze, L.G. et al. (1999) *Biotechniques* 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

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

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

WO 02/26998

PCT/US01/30310

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

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

WO 02/26998

PCT/US01/30310

region of DNA associated with that disorder and thus may further positional cloning efforts.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps.

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

In another embodiment of the invention, HYDR, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between HYDR 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 HYDR, or fragments thereof, and washed. Bound HYDR is then detected by methods well known in the art. Purified HYDR can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

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

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

Without further elaboration, it is believed that one skilled in the art can, using the preceding

WO 02/26998

PCT/US01/30310

description, utilize the present invention to its fullest extent. The following embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications and publications, mentioned above and below, including U.S. Ser. No. 60/237,093, U.S. Ser. No. 60/238/370, and U.S. Ser. No. 60/241,284 are expressly incorporated by reference herein.

EXAMPLES

I. Construction of cDNA Libraries

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

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

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSOFT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, *supra*, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESOFT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), PCR2-TOPOTA plasmid (Invitrogen),

WO 02/26998

PCT/US01/30310

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

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 Plasmid, 10 QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4 °C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal 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

20 Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows. Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Applied Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared 25 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 polymucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI 30 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

WO 02/26998

PCT/US01/30310

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

5 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.

10 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

15 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

20 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.

25 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

30 applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score or the lower the probability value, the greater the identity between two sequences).

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

WO 02/26998

PCT/US01/30310

NO-5-8. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 4.

IV. Identification and Editing of Coding Sequences from Genomic DNA

Putative hydrolases were initially identified by running the Genscan gene identification program against public genomic sequence databases (e.g., gbpri and gblatg). 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 hydrolases, the encoded polypeptides were analyzed by querying against PFAM models for hydrolases. Potential hydrolases were also identified by homology to Incyte cDNA sequences that had been annotated as hydrolases. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbpri public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from genpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide sequences were obtained by assembling Genscan-predicted coding sequences with Incyte cDNA sequences and/or public cDNA sequences using the assembly process described in Example III. Alternatively, full length polynucleotide sequences were derived entirely from edited or unedited Genscan-predicted coding sequences.

V. Assembly of Genomic Sequence Data with cDNA Sequence Data

"Stitched" Sequences

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

WO 02/26998

PCT/US01/30310

be equivalent by transitivity. For example, if an interval was present on a cDNA and two genomic sequences, then all three intervals were considered to be equivalent. This process allows unrelated but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together by the stitching algorithm in the order that they appear along their parent sequences to generate the longest possible sequence, as well as sequence variants. Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or genomic sequence to genomic sequence) were given preference over linkages which change parent type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbprl 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 HYDR Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:5-8 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:5-8 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.

WO 02/26998

PCT/US01/30310

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

VII. Analysis of Polynucleotide Expression

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

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

The basis of the search is the product score, which is defined as:

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

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

WO 02/26998

PCT/US01/30310

either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

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

VIII. Extension of HYDR 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 Mg²⁺, (NH₄)₂SO₄, and 2-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters

WO 02/26998

PCT/US01/30310

for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

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

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

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

In like manner, full length polynucleotide sequences are verified using the above procedure or are used to obtain 5' regulatory sequences using the above procedure along with oligonucleotides

WO 02/26998

PCT/US01/30310

designed for such extension, and an appropriate genomic library.

IX. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:5-8 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μ Ci of [γ - 32 P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

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

20 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 (Sचना (1999), supra). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Sचना, M. et al. (1995) *Science* 270:467-470; Shalon, D. et al. (1996) *Genome Res.* 6:639-645; Marshall, A. and J. Hodgson (1998) *Nat. Biotechnol.* 16:27-31.)

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

WO 02/26998

PCT/US01/30310

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

Tissue or Cell Sample Preparation

10 Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)⁺ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)⁺ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/ μ l oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/ μ l RNase inhibitor, 500 μ M dATP, 500 μ M dGTP, 500 μ M dTTP, 40 μ M dCTP, 40 μ M dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse

15 transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)⁺ RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)⁺ RNAs are synthesized by *in vitro* transcription from non-coding yeast genomic DNA. After incubation at 37°C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85°C to the stop the reaction and degrade the RNA. Samples are purified

20 using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 μ l 5X SSC/0.2% SDS.

25 Microarray Preparation

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 μ g.

30 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

WO 02/26998

PCT/US01/30310

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

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

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

Hybridization

Hybridization reactions contain 9 μ l of sample mixture consisting of 0.2 μ g each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65°C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm² coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140 μ l of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60°C. The arrays are washed for 10 min at 45°C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45°C in a second wash buffer (0.1X SSC), and dried.

Detection

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

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source,

WO 02/26998

PCT/US01/30310

although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

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

XI. Complementary Polynucleotides

Sequences complementary to the HYDR-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring HYDR. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of HYDR. 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 HYDR-encoding transcript.

XII. Expression of HYDR

Expression and purification of HYDR is achieved using bacterial or virus-based expression systems. For expression of HYDR 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

WO 02/26998

PCT/US01/30310

promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express HYDR upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of HYDR 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 HYDR 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, HYDR 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 HYDR 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 HYDR obtained by these methods can be used directly in the assays shown in Examples XVI and XVII where applicable.

XIII. Functional Assays

HYDR function is assessed by expressing the sequences encoding HYDR 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

WO 02/26998

PCT/US01/30310

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

The influence of HYDR on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding HYDR 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 HYDR and other genes of interest can be analyzed by northern analysis or microarray techniques.

XIV. Production of HYDR Specific Antibodies

HYDR 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 HYDR 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

WO 02/26998

PCT/US01/30310

oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-HYDR activity by, for example, binding the peptide or HYDR to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

5 **XV. Purification of Naturally Occurring HYDR Using Specific Antibodies**

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

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

XVI. Identification of Molecules Which Interact with HYDR

HYDR, or biologically active fragments thereof, are labeled with ¹²⁵I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) *Biochem. J.* 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled HYDR, washed,
20 and any wells with labeled HYDR complex are assayed. Data obtained using different concentrations of HYDR are used to calculate values for the number, affinity, and association of HYDR with the candidate molecules.

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

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

30 **XVII. Demonstration of HYDR Activity**

HYDR activity is demonstrated through a variety of specific enzyme assays some of which are outlined below.

Protein phosphatase (PP) activity can be measured by the hydrolysis of P-nitrophenyl phosphate (PNPP). HYDR is incubated together with PNPP in HEPES buffer pH 7.5, in the

WO 02/26998

PCT/US01/30310

presence of 0.1% β -mercaptoethanol at 37°C for 60 min. The reaction is stopped by the addition of 6 ml of 10 N NaOH (Diamond, R.H. et al. (1994) Mol. Cell. Biol. 14:3752-62). Alternatively, acid phosphatase activity of HYDR is demonstrated by incubating HYDR containing extract with 100 μ l of 10 mM PNPP in 0.1 M sodium citrate, pH 4.5, and 50 μ l of 40 mM NaCl at 37°C for 20 min. The reaction is stopped by the addition of 0.5 ml of 0.4 M glycine/NaOH, pH 10.4 (Saltig, P. et al. (1997) J. Biol. Chem. 272:18628-18635). The increase in light absorbance at 410 nm resulting from the hydrolysis of PNPP is measured using a spectrophotometer. The increase in light absorbance is proportional to the activity of HYDR in the assay.

In the alternative, HYDR activity is determined by measuring the amount of phosphate removed from a phosphorylated protein substrate. Reactions are performed with 2 or 4 nM HYDR in a final volume of 30 μ l containing 60 mM Tris, pH 7.6, 1 mM EDTA, 1 mM EGTA, 0.1% 2-mercaptoethanol and 10 μ M substrate, 32 P-labeled on serine/threonine or tyrosine, as appropriate. Reactions are initiated with substrate and incubated at 30°C for 10-15 min. Reactions are quenched with 450 μ l of 4% (w/v) activated charcoal in 0.6 M HCl, 90 mM $\text{Na}_4\text{P}_2\text{O}_7$, and 2 mM NaH_2PO_4 , then centrifuged at $12,000 \times g$ for 5 min. Acid-soluble 32 Pi is quantified by liquid scintillation counting (Sinclair, C. et al. (1999) J. Biol. Chem. 274:23666-23672).

The adenosine deaminase activity of HYDR is determined by measuring the rate of deamination that occurs when adenosine substrate is incubated with HYDR. Reactions are performed with a predetermined amount of HYDR in a final volume of 3.0 ml containing 53.3 mM potassium phosphate and 0.045 mM adenosine. Assay reagents excluding HYDR are mixed in a quartz cuvette and equilibrated to 25°C. Reactions are initiated by the addition of HYDR and are mixed immediately by inversion. The decrease in light absorbance at 265 nm resulting from the hydrolysis of adenosine to inosine is measured using a spectrophotometer. The decrease in the $A_{265 \text{ nm}}$ is recorded for approximately 5 minutes. The decrease in light absorbance is proportional to the activity of HYDR in the assay.

HYDR activity can be measured by determining the amount of free adenosine produced by the hydrolysis of AMP, as described by Sala-Newby, et al. *supra*. Briefly, HYDR is incubated with AMP in a suitable buffer for 10 minutes at 37°C. Free adenosine is separated from AMP and measured by reverse phase HPLC.

Alternatively, HYDR activity is measured by the hydrolysis of ADP-ribosylarginine (Konczalik, P. and J. Moss (1999) J. Biol. Chem. 274:16736-16740). 50 ng of HYDR are incubated with 100 μ M ADP-ribosyl-[14 C]arginine (78,000 cpm) in 50 mM potassium phosphate, pH 7.5, 5 mM dithiothreitol, 10 mM MgCl_2 in a final volume of 100 μ l. After 1 h at 37°C, 90 μ l of the sample is applied to a column (0.5 \times 4 cm) of Affi-Gel 601 (boronate) equilibrated and eluted with five 1-ml

WO 02/26998

PCT/US01/30310

portions of 0.1 M glycine, pH 9.0, 0.1 M NaCl, and 10 mM MgCl₂. Free ¹⁴C-Arg in the total eluate is measured by liquid scintillation counting.

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

WO 02/26998

PCT/US01/30310

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
375287	1	24569CD1	2	2460CE1
409364	2	409364CD1	7	409367CE1
409364	3	409364CD1	7	409367CE1
7473183	4	7473183CD1	8	7473183CE1

WO 02/26998

PCT/US01/30310

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO:	Probability score	GenBank Homolog
1	573160CD1	97267246	2.6e-59	[Arabidopsis thaliana] Putative adrenomedullin
2	2656987CD1	93106985	2.0e-28	[Xylella fastidiosa] endonuclease V (deoxyinosine 3' endonuclease) Silvestri, M.L. et al. (2000) Nature 406:151-157
3	4009684CD1	9402478	1.2e-79	[Homo sapiens] ADP-ribosylarginine hydrolase Takada, T. et al. (1993) J. Biol. Chem. 268:1757-1763
4	7473183CD1	513603379 94902474	0 3.1e-161	[Homo sapiens] 5'-nucleotidase I (Columbia livia) 5'-nucleotidase I (Sala-Newby, G.B. et al. (1999) J. Biol. Chem. 274:17789-17793)

WO 02/26998

PCT/US01/30310

Table 3

SEQ ID NO.	Incyte Polypeptide	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1	3731660CD1	355	S124, S132, S336, S78, S79, T107, T117, T120, T159, T219, T337, T35, T53, T62	N12, N28, N314	DEAMINASE HYDROLASE NUCLEOTIDE METABOLISM ADENOSINE AMP AMINOHYDROLASE ISOFORM MULTIGENE FAMILY PD005960:E65-W345 Adenosine and AMP deaminase proteins BL00485:T269-L323 Adenosine/AMP deaminase A, deaminase:P9-K344	BLAST_PRODOM BLIMFS_BLOCKS HMMER_PFM
2	2656987CD1	308	S203, S261, S60, S71, T178, T258, T300, S301, S302, S240, S235, S46, T151, T210, T259		PROTEIN ENDONUCLEASE HYDROLASE NUCLEASE DNA REPAIR POTATIVE Y DEOXYRIBOSINE 3' ENDONUCLEASE PD09651:Y51-E245 ADENOSINE AMP AMINOHYDROLASE NUCLEASE ENZYME MAGNESIUM PD024785:ML-E353	HMMER_PFM BLAST_PRODOM
3	4009684CD1	354	S203, S261, S60, S71, T178, T258, T300, S301, S302, S240, S235, S46, T151, T210, T259	N26	PROTEIN ENDONUCLEASE HYDROLASE NUCLEASE DNA REPAIR POTATIVE Y DEOXYRIBOSINE 3' ENDONUCLEASE PD09651:Y51-E245 ADENOSINE AMP AMINOHYDROLASE NUCLEASE ENZYME MAGNESIUM PD024785:ML-E353	BLAST_PRODOM
4	7473183CD1	366	S219, S55, T279, T293, T352, Y69			MOTIFS

WO 02/26998

PCT/US01/30310

Table 4

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
5	3731660CBL	2269	1-47, 433- 640, 1650- 2269	394109078 (SCORNF04)	1	738
				71742451901	758	1259
				7174373401	632	1130
				7174436501	1301	1826
				7174366501	1652	2269
6	2656987CBL	1150	233-363	6740976H1 (BRAFDT02)	1	489
				92165516	672	1150
				2656987F6 (LJNGTWT09)	510	1122
				2656987F6 (LJNGTWT09)	324	859
				7079113201	483	1036
7	4009684CBL	2000	1-248	4009684F6 (MISCN010)	1092	1632
				7078853001	1282	1768
				5277543F7 (MUSLAN010)	1	507
				7079233201	1431	2000
				18956006 (CARDN010)	661	1289
8	7473183CBL	1101	1-822, 993- 1101	GNN.g95262300_000003_006.edit	1	1101

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID	Representative Library
3	373L60031	LDNGAST01
6	265697CE1	LDNGST09
7	409664CE1	LV2NR0703

WO 02/26998

PCT/US01/30310

Table 6

Library	Vector	Library Description
LONGAST01	PSFORT1	Library was constructed using RNA isolated from the lung tissue of a 17-year-old Caucasian male, constructed from total RNA. Patient history included asthma, eczema, and a history of asthma. The patient was hospitalized from a 69-year-old Caucasian male during segmental lung resection. Pathology indicated invasive grade 3 squamous cell carcinoma and a metastatic tumor. Patient history included type II diabetes, thyroid disorder, depressive disorder, hyperlipidemia, esophageal ulcer, and tobacco use.
LUNGSTU09	PINCY	
LIVERH0T03	PSFORT1	Library was constructed using RNA isolated from the left ventricle tissue of a 31-year-old male.

WO 02/26998

PCT/US01/30310

Table 7

Program	Description	Reference	Parameter/Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Parned, Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	ESY: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, ffasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESY: fasta E value=1.0E-6 Assembled ESY: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less Full Length sequences: fastx score=100 or greater
BLIMPS	A BLOCKS IMPROVED Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Aronow, J.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; Samadpour, E.L.L. et al. (1988) Nucleic Acids Res. 26:3204-3222; Durbin, R. et al. (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-530.	PFAM hits: Probability value= 1.0E-5 or less Signal peptide hits: Scores= 0 or greater

WO 02/26998

PCT/US01/30310

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; Bairach, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality scores; GCG-specified "HIGR" value for that particular Prosite motif. Generally, scores=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. 8:175-183; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	Scores=120 or greater. Match length= 56 or greater
Phrap	A Phrap Revised Assembly Program including SWAT and CrossMatch programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1980) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	
Consed	A graphical tool for viewing and editing Phrap assemblies.	Grouton, D. et al. (1998) Genome Res. 8:195-202.	
SFScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Chaveris, I.M. and S. Autic (1997) CABIOS 12:431-439.	Scores=1.5 or greater
TMAP	A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation.	Pearson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Pearson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMEMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Sambharner, 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. 173-182.	
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairach, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

WO 02/26998

PCT/US01/30310

What is claimed is:

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

WO 02/26998

PCT/US01/30310

b) recovering the polypeptide so expressed.

10. A method of claim 9, wherein the polypeptide has an amino acid sequence selected from the group consisting of SEQ ID NO:1-4.

5

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

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

- 10 a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:5-8,
b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:5-8,
15 c) a polynucleotide complementary to a polynucleotide of a),
d) a polynucleotide complementary to a polynucleotide of b), and
e) an RNA equivalent of a)-d).

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

20

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

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

30

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

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

WO 02/26998

PCT/US01/30310

- 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.
- 5
17. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable excipient.
18. A composition of claim 17, wherein the polypeptide has an amino acid sequence selected from the group consisting of SEQ ID NO:1-4.
19. A method for treating a disease or condition associated with decreased expression of functional HYDR, comprising administering to a patient in need of such treatment the composition of claim 17.
- 15
20. A method of screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:
- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- b) detecting agonist activity in the sample.
- 20
21. A composition comprising an agonist compound identified by a method of claim 20 and a pharmaceutically acceptable excipient.
22. A method for treating a disease or condition associated with decreased expression of functional HYDR, comprising administering to a patient in need of such treatment a composition of claim 21.
- 25
23. A method of screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:
- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- b) detecting antagonist activity in the sample.
- 30
24. A composition comprising an antagonist compound identified by a method of claim 23 and a pharmaceutically acceptable excipient.

WO 02/26998

PCT/US01/30310

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

26. A method of screening for a compound that specifically binds to the polypeptide of claim 1, the method comprising:

- a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and
- b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.

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

- a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,
- b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and
- c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

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

comprising:

- a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,
- b) detecting altered expression of the target polynucleotide, and
- c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

29. A method of assessing toxicity of a test compound, the method comprising:

- a) treating a biological sample containing nucleic acids with the test compound,
- b) hybridizing the nucleic acids of the treated biological sample with a probe comprising

WO 02/26998

PCT/US01/30310

- at least 20 contiguous nucleotides of a polynucleotide of claim 12 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 12 or fragment thereof,
- 5 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.
- 10
30. A diagnostic test for a condition or disease associated with the expression of HYDR in a biological sample, the method comprising:
- a) combining the biological sample with an antibody of claim 11, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex,
- 15 and
- b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.
31. The antibody of claim 11, wherein the antibody is:
- 20 a) a chimeric antibody,
- b) a single chain antibody,
- c) a Fab fragment,
- d) a F(ab')₂ fragment, or
- e) a humanized antibody.
- 25
32. A composition comprising an antibody of claim 11 and an acceptable excipient.
33. A method of diagnosing a condition or disease associated with the expression of HYDR in a subject, comprising administering to said subject an effective amount of the composition of claim
- 30 32.
34. A composition of claim 32, wherein the antibody is labeled.
35. A method of diagnosing a condition or disease associated with the expression of HYDR

WO 02/26998

PCT/US01/30310

in a subject, comprising administering to said subject an effective amount of the composition of claim 34.

36. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 11, the method comprising:

- a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-4, 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:1-4.

37. A polyclonal antibody produced by a method of claim 36.

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

39. A method of making a monoclonal antibody with the specificity of the antibody of claim 11, the method comprising:

- a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-4, 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:1-4.

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

41. A composition comprising the monoclonal antibody of claim 40 and a suitable carrier.

WO 02/26998

PCT/US01/30310

42. The antibody of claim 11, wherein the antibody is produced by screening a Fab expression library.
43. The antibody of claim 11, wherein the antibody is produced by screening a recombinant immunoglobulin library.
44. A method of detecting a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-4 in a sample, the method comprising:
- a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and
 - b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-4 in the sample.
45. A method of purifying a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-4 from a sample, the method comprising:
- a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and
 - b) separating the antibody from the sample and obtaining the purified polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-4.
46. A microarray wherein at least one element of the microarray is a polynucleotide of claim 13.
47. A method of generating a transcript image of a sample which contains polynucleotides, the method comprising:
- a) labeling the polynucleotides of the sample,
 - b) contacting the elements of the microarray of claim 46 with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex, and
 - c) quantifying the expression of the polynucleotides in the sample.
48. An array comprising different nucleotide molecules affixed in distinct physical locations on a solid substrate, wherein at least one of said nucleotide molecules comprises a first oligonucleotide

WO 02/26998

PCT/US01/30310

or polynucleotide sequence specifically hybridizable with at least 30 contiguous nucleotides of a target polynucleotide, and wherein said target polynucleotide is a polynucleotide of claim 12.

49. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is
5 completely complementary to at least 30 contiguous nucleotides of said target polynucleotide.

50. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 60 contiguous nucleotides of said target polynucleotide.

10 51. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to said target polynucleotide.

52. An array of claim 48, which is a microarray.

15 53. An array of claim 48, further comprising said target polynucleotide hybridized to a nucleotide molecule comprising said first oligonucleotide or polynucleotide sequence.

54. An array of claim 48, wherein a linker joins at least one of said nucleotide molecules to said solid substrate.

20

55. An array of claim 48, wherein each distinct physical location on the substrate contains multiple nucleotide molecules, and the multiple nucleotide molecules at any single distinct physical location have the same sequence, and each distinct physical location on the substrate contains nucleotide molecules having a sequence which differs from the sequence of nucleotide molecules at
25 another distinct physical location on the substrate.

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

57. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:2.

30

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

59. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:4.

WO 02/26998

PCT/US01/30310

- 60. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:5.
- 61. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:6.
- 5 62. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:7.
- 63. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:8.

WO 02/26998

PCT/US01/30310

<110> INCYTE GENOMICS, INC.
 YUE, HENRY
 Baughn, Mariah R.
 WARREN, Bridget A.
 TRIDOLEY, Catherine M.
 TANG, Y. Tom
 KHAN, Farrah A.
 YAO, Monique G.
 LAL, Preeti
 THORNTON, Michael

<120> HYDROLASES

<130> FI-0243 PCT

<140> To Be Assigned
 <141> Herewith

<150> 60/237,093; 60/238,370; 60/241,284
 <151> 2000-09-29; 2000-10-06; 2000-10-17

<160> 8
 <170> PERL Program

<210> 1
 <211> 355
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3731660CD1

<400> 1
 Met Ile Glu Ala Glu Glu Gln Gln Pro Cys Lys Thr Asp Phe Tyr
 1 5 10 15
 Ser Glu Leu Pro Lys Val Glu Leu His Ala His Leu Asn Gly Ser
 20 25 30
 Ile Ser Ser His Thr Met Lys Lys Leu Ile Ala Gln Lys Pro Asp
 35 40 45
 Leu Lys Ile His Asp Gln Met Thr Val Ile Asp Lys Gly Lys Lys
 50 55 60
 Arg Thr Leu Glu Glu Cys Phe Gln Met Phe Gln Thr Ile His Gln
 65 70 75
 Leu Thr Ser Ser Pro Glu Asp Ile Leu Met Val Thr Lys Asp Val
 80 85 90
 Ile Lys Glu Phe Ala Asp Asp Gly Val Lys Tyr Leu Glu Leu Arg
 95 100 105
 Ser Thr Pro Arg Arg Glu Asn Ala Thr Gly Met Thr Lys Lys Thr
 110 115 120
 Tyr Val Glu Ser Ile Leu Glu Gly Ile Lys Gln Ser Lys Gln Glu
 125 130 135
 Asn Leu Asp Ile Asp Val Arg Tyr Leu Ile Ala Val Asp Arg Arg
 140 145 150
 Gly Gly Pro Leu Val Ala Lys Glu Thr Val Lys Leu Ala Glu Glu
 155 160 165
 Phe Phe Leu Ser Thr Glu Gly Thr Val Leu Gly Leu Asp Leu Ser
 170 175 180
 Gly Asp Pro Thr Val Gly Gln Ala Lys Asp Phe Leu Glu Pro Leu
 185 190 195
 Leu Glu Ala Lys Lys Ala Gly Leu Lys Leu Ala Leu His Leu Ser

WO 02/26998

PCT/US01/30310

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

<210> 2
 <211> 308
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2656987CD1

<400> 2
 Met Ala Leu Glu Ala Ala Gly Gly Pro Pro Glu Glu Thr Leu Ser
 1 5 10 15
 Leu Trp Lys Arg Glu Gln Ala Arg Leu Lys Ala His Val Val Asp
 20 25 30
 Arg Asp Thr Glu Ala Trp Gln Arg Asp Pro Ala Phe Ser Gly Leu
 35 40 45
 Gln Arg Val Gly Gly Val Asp Val Ser Phe Val Lys Gly Asp Ser
 50 55 60
 Val Arg Ala Cys Ala Ser Leu Val Val Leu Ser Phe Pro Glu Leu
 65 70 75
 Glu Val Val Tyr Glu Glu Ser Arg Met Val Ser Leu Thr Ala Pro
 80 85 90
 Tyr Val Ser Gly Phe Leu Ala Phe Arg Glu Val Pro Phe Leu Leu
 95 100 105
 Glu Leu Val Gln Gln Leu Arg Glu Lys Glu Pro Gly Leu Met Pro
 110 115 120
 Gln Val Leu Leu Val Asp Gly Asn Gly Val Leu His His Arg Gly
 125 130 135
 Phe Gly Val Ala Cys His Leu Gly Val Leu Thr Asp Leu Pro Cys
 140 145 150
 Val Gly Val Ala Lys Lys Leu Leu Gln Val Asp Gly Leu Glu Asn
 155 160 165
 Asn Ala Leu His Lys Glu Lys Ile Arg Leu Leu Gln Thr Arg Gly
 170 175 180
 Asp Ser Phe Pro Leu Leu Gly Asp Ser Gly Thr Val Leu Gly Met
 185 190 195
 Ala Leu Arg Ser His Asp Arg Ser Thr Arg Pro Leu Tyr Ile Ser
 200 205 210
 Val Gly His Arg Met Ser Leu Glu Ala Ala Val Arg Leu Thr Cys

WO 02/26998

PCT/US01/30310

```

215          220          225
Cys Cys Cys Arg Phe Arg Ile Pro Glu Pro Val Arg Gln Ala Asp
230          235          240
Ile Cys Ser Arg Glu His Ile Arg Lys Ser Leu Gly Leu Pro Gly
245          250          255
Pro Pro Thr Pro Arg Ser Pro Lys Ala Gln Arg Pro Val Ala Cys
260          265          270
Pro Lys Gly Asp Ser Gly Glu Ser Ser Gly Gly Ala Pro Ser Pro
275          280          285
Gln Arg Gln Ala Asp Arg Thr Thr Pro Gly Gly Arg Arg Ser Thr
290          295          300
Ala Gln His Gln Val Gly Gln Arg
305

<210> 3
<211> 354
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4009684CD1

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

```

WO 02/26998

PCT/US01/30310

```

275      280      285
Ala Ala Gly Asn Ser Trp Thr Glu Leu Cys His Arg Ala Met Phe
290      295      300
His Gly Gly Glu Ser Ala Ala Thr Gly Thr Ile Ala Gly Cys Leu
305      310      315
Phe Gly Leu Leu Tyr Gly Leu Asp Leu Val Pro Lys Gly Leu Tyr
320      325      330
Gln Asp Leu Glu Asp Lys Glu Lys Leu Glu Asp Leu Gly Ala Ala
335      340      345
Leu Tyr Arg Leu Ser Thr Glu Glu Lys
350

```

```

<210> 4
<211> 366
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7473183CD1

```

```

<400> 4
Met Glu Pro Gly Gln Pro Arg Glu Pro Gln Glu Pro Arg Glu Pro
1      5      10      15
Gly Pro Gly Ala Glu Thr Ala Ala Ala Pro Val Trp Glu Glu Ala
20      25      30
Lys Ile Phe Tyr Asp Asn Leu Ala Pro Lys Lys Lys Pro Lys Ser
35      40      45
Pro Gln Asn Ala Val Thr Ile Ala Val Ser Ser Arg Ala Leu Phe
50      55      60
Arg Met Asp Glu Glu Gln Gln Ile Tyr Thr Glu Gln Gly Val Glu
65      70      75
Glu Tyr Val Arg Tyr Gln Leu Glu His Glu Asn Glu Pro Phe Ser
80      85      90
Pro Gly Pro Ala Phe Pro Phe Val Lys Ala Leu Glu Ala Val Asn
95      100      105
Arg Arg Leu Arg Glu Leu Tyr Pro Asp Ser Glu Asp Val Phe Asp
110      115      120
Ile Val Leu Met Thr Asn Asn His Ala Gln Val Gly Val Arg Leu
125      130      135
Ile Asn Ser Ile Asn His Tyr Asp Leu Phe Ile Glu Arg Phe Cys
140      145      150
Met Thr Gly Gly Asn Ser Pro Ile Cys Tyr Leu Lys Ala Tyr His
155      160      165
Thr Asn Leu Tyr Leu Ser Ala Asp Ala Glu Lys Val Arg Glu Ala
170      175      180
Ile Asp Glu Gly Ile Ala Ala Ala Thr Ile Phe Ser Pro Ser Arg
185      190      195
Asp Val Val Val Ser Gln Ser Gln Leu Arg Val Ala Phe Asp Gly
200      205      210
Asp Ala Val Leu Phe Ser Asp Glu Ser Glu Arg Ile Val Lys Ala
215      220      225
His Gly Leu Asp Arg Phe Phe Glu His Glu Lys Ala His Glu Asn
230      235      240
Lys Pro Leu Ala Gln Gly Pro Leu Lys Gly Phe Leu Glu Ala Leu
245      250      255
Gly Arg Leu Gln Lys Lys Phe Tyr Ser Lys Gly Leu Arg Leu Glu
260      265      270
Cys Pro Ile Arg Thr Tyr Leu Val Thr Ala Arg Ser Ala Ala Ser
275      280      285
Ser Gly Ala Arg Ala Leu Lys Thr Leu Arg Ser Trp Gly Leu Glu

```

WO 02/26998

PCT/US01/30310

Thr	Asp	Glu	Ala	290	Leu	Phe	Leu	Ala	Gly	Ala	Pro	Lys	Gly	Pro	Leu	300
				305												315
Leu	Glu	Lys	Ile	Arg	Pro	His	Ile	Phe	Phe	Asp	Asp	Gln	Met	Phe		330
				320												345
His	Val	Ala	Gly	Ala	Gln	Glu	Met	Gly	Thr	Val	Ala	Ala	His	Val		355
				335												360
Pro	Tyr	Gly	Val	Ala	Gln	Thr	Pro	Arg	Arg	Thr	Ala	Pro	Ala	Lys		
				350												
Gln	Ala	Pro	Ser	Ala	Gln											
				365												

<210> 5
 <211> 2269
 <212> DNA
 <213> Homo sapiens
 <220>
 <221> misc_feature
 <223> Incyte ID No: 3731660CB1

<400> 5
 gtggagagg agaggttgaa ggaacagaaa aatgagaaga caactcactg gaggtggaag 60
 tgggagagatc gcttgagctt ggaagcttga gactgcagca agotgtgatc gtgacactgc 120
 actccagcct gggcaacaaa gtaagatcct gtctcaaaag aaaaaaaaaa aaaggtgaag 180
 ctgaaccocag atctacacat acagctaact ttaccaaaat gtgtggaagt aaagcaatc 240
 gaaggaaatt cagtaccata caatttactg actgaaatac atcatattgc tcatactaag 300
 aataagagtg gagaagaatc atttttttcc tgctaaaatg atagagggag aagagcaaca 360
 gcttgcagag acagacttct atcttgaatt gccaaaagtg gaacttcag cccacttgaa 420
 tggatccatt agttctcata ccabgaagcc cagaagccag atcttaaat 480
 ccacgatcag atgactgtga ttgacaaggg aaagaaaaga actttggaag aatgtttcca 540
 gatgtttcaa actattcatc agcttactag tagccctgaa gatattetaa tggtaacaaa 600
 agatgtcata aaagaatttg cagatgacgg cgtcaagtao ctggaactaa ggagcaacc 660
 cagaagagaa aatgctactg gaatgactaa aaagacttat gtggaacta tacttgagg 720
 tataaacacag tccaacaacg aaaacttga cattgatgtt agytattga tagcagttga 780
 cagaagaggc ggccttttag tagccaagga gactgtaaaa ottgccgagg agtttctct 840
 ttctactgag gttacagttc ttggccttga cctcagtgga gaccctactg taggacaacg 900
 aaaaagcttc ttggaactc ttttagaagc taagaagca ggtctgaagt tagcattgca 960
 tcttccagag attccaaccc aaaaaaaga aacacaata ctctggatc tgcctcctga 1020
 cagatccagg catggaacat ttctcaactc cggtagggga ggatccctg atctggtgga 1080
 ctttgtgagg caacatcgga taccactgga actctgtttg acctcaaacg tcaaaagta 1140
 gacagttcca tcttatgacc agcaccattt cggattctgg tacagcattg cccatcctc 1200
 tgtgatctgt actgatgata aggggtgttt tgcaacacac ctttctcaag agtaccagct 1260
 ggcagctgaa acatlttaatt tgaccocagtc tcaggtgtgg gatctgtctt atgaatccat 1320
 caactacatc ttgtctctg acagaccagc atctgaaact aggaagaat ggaatccct 1380
 gaagcccaga gtgttacata ttaagctat aatgaggtga actactctg agtatgtgtt 1440
 tcaatcaagt tcttgcacata tcccacttag taaaacagtc caaccctct ttgaagcata 1500
 gcaaccaagt tcttgggct ctatcaccag caccttacac atggcaggta ctcagtaaat 1560
 acgtgtcttc aactgactca caagctctca ggtgcttact gggtaggact tgactgtgt 1620
 tgcataatba atccocatc caccagtga atctctgtt gttctcatca cacagcaaaa 1680
 tgataacaat ggtaaacaaca gctgactgag tgctttttat tggccatgca ctgtgctgag 1740
 tbattagac aaatcaattc acttcatctt catgacagcc cttaaaaggg aacctattta 1800
 ggtcctaagg tccaggaaaa aatttacag atgtggaac caaggtgag agcggggctg 1860
 acttggcagc aagtggctga tcccagacct catctcagac tttctgact ctgaagocaa 1920
 cactcacagg acctgtcac actgaaactt tgccagtgag gagtcagacc acctgacttc 1980
 tgggctctac ctagcctgcc agtgcactct atgtgtaact ttgggctcat caagcattct 2040
 ctittgtgga cgaatgaaaa tgctaccaaa gcctacttcc tggggttagc ttgagcggagc 2100
 acatgaaact ggaatgaaatg taacaccacc aaacacaaga gggcgaccag acacactagt 2160
 gagccagatc aaccocgggga cataatcccg gaccggtact tgacagcggg ctgaggactc 2220
 gcatccagct aatcgaaacg tcacccaacg agccgaaccc attgccaaag 2269

WO 02/26998

PCT/US01/30310

<210> 6
<211> 1150
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2656987CB1

<400> 6
gcatggccct ggaggcggcg ggaggcggcg cggaggaaac gctgtcactg tggaaacggg 60
agcaagctcg gctgaaggcc cacgtcgttag accgggacac ctaggctgtg cagcggagacc 120
ccgectctct ggtctgpcag aggttcgggg gcgttgacgt gtctctctg aaaggggaca 180
gtgtccggcg tttgtgttcc ctggtgtgtg tcaagctccc ttagctcgag gtggtgtatg 240
aggaagoccg catgttcaag ctacacagccc octactgttc gggtctctcg gctctcogag 300
aggtgcccctt cttgtgtggag ctggtgcagc agctgcggga gaaggagccg gcctctatgc 360
cccaagctct tcttgtggat ggaacggggg tactccacca ccgaggtctt ggggtggcct 420
gccaccttgg cgtccttaca gacctgccgt gtgttggggt ggccaagaaa cttctgcagg 480
tggatgggct ggagaacaac gcctgcaca aggagaagat ccgactcctg cagactcgag 540
gagaotcatt ccctctgtgt ggagactctg ggactgtcct ggaatggcc ctgaggagcc 600
acgacggag caacagccc ctctacatct ccgtgggcca agagttagc ctgaggagcc 660
ctgtgcgctc gacttctgtc tctgtcaggt tccggatccc agagccctg gcccaaggctg 720
acatctgtct ccgagagcac atccgcaagt cgtgtgact ccccgggcca ccccaaccga 780
ggagcccga ggcgagagg ccagtgccat gccccaagg agactcogga gactctccag 840
gtggagcacc cagtcccaca agacaggctg accgcaccac cccaggggga cgcgcagca 900
cagcccagca ccagtgggg cagaggtgac caccgcccct cttgtctcg tcatggctg 960
gtcagctgtg gtcacgtgtc ctacagggac agatctctat gggggcaagt gccagatcct 1020
gagagcgcac gagagctttt cccggagccg acgaagggga ctggagctg cagcctgac 1080
gacccctgca gctctgtctt tgcccacccc tttcaataga tggaaactgc ttgtctctta 1140
aaaaaaaaaa 1150

<210> 7
<211> 2000
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4009684CB1

<400> 7
ctccgcagct gttggggaag aggagctgcc tcttggatg gagaaatta aggtctcgat 60
gttgcgtggg agcgtcggcg atgtctcttg ctacagaaat gtctgcaagg agaacagcac 120
tctagccatg aagatccagg aggagctgca acgttccggg ggctctggacc acctcgtact 180
ctcgcaggga gaatggcccg tgagtgcaca caacctatg cacatcgcaa ccgcccaggg 240
cctcaccaca gactactggt gcotggatga tctgtaccgg gagatggtga gatgctatgt 300
ggaatcgttt gagagcttc cagaaccccg gccagaccac gotaccattg aaggctgtgc 360
tcagctaaag ccaataact accttctcgc ctggcacaca ccgttcaatg aaaaaggctc 420
agggtttggg gcggccacca aggacctgtg catcgccctg ccgtactgga agcctgagcg 480
gctggagacc ctcatcgagg tcaagctgga gtgcggccgg atgaccacca acctccccac 540
aggcttctct ggtccctctg gcaaggccct gtttgtgtcg ttgcgccac aaggaaagcc 600
cctgtctcag tggggagag acatgtctgc gccggtgct ctggcagaag agtactgtag 660
gagacacato cggccaccgg cagaatacca ggagcactg tttbacttgg aagctaaatg 720
gcaattttat ttggagaga ggaataatcag taazgactca gaaaataaag ccatctccc 780
cgacaattat gatgcagaag agagggaaaa gacctacag aagtggagct cggaggtgc 840
agggggaaga cagaccacag atgcccocat gatagcctat gaocccctc ttgcagcagg 900
aaacagctgg actgagctgt gtcaccgggc catgtttcat ggaggggaga gcgcgcccac 960
ggggcaccat gcaggctgcc tgttccgggt gtgttacggc ctggacctcg tccccaaagg 1020
cttgaaccag gacctggagg acaaggagaa gctggaggac ctgggcccgg ctctctaccg 1080
cctgtccaca gaggagaagt aaagccattt ctgccccttt cccctctag agccgatccc 1140
acccccgggg ccgtagggcc ctctgcagc cctgggtgga ggggtctctc gtagggctcc 1200

WO 02/26998

PCT/US01/30310

```

actgoggtct gtgctgact ggccacatct aactctctgt ttccaatttc agaatcctaa 1260
ctgtttgcata aaatacattg ttgtctctgc gagaatatit tccgtctctc accatcaaca 1320
ttgacactgc gtgagattgc cgcacttgga cctccatgctc tggcactcac ccgcagtctc 1380
ctggacaggc gctgtatttt attctgtogc agagctaagt ctgtttactc actcaactoa 1440
acaacactaa ctgcggtgggt ggctctcagc agggccccc cctgcagacc ctctgtctctg 1500
cctctgctct caggcatgct ttccccctgt agggccaatg cactcccccc cccccccaa 1560
cctcccaggt ccccagtggt tctgtgtgtc ctggacagag aaacagtcac cactggggcc 1620
tgccggcaoc atatagcagc atattttgct cttaaccaca cccacctttt taatccactc 1680
agattttaag atcaatccot ttttgaaca actcacggag aaaaccagaa cataaatggc 1740
ctcctgccag ctccggcgtc tctctgtggt ctgccttagt gggccaagtc caaatgcaga 1800
gaagcccttt ccctcccgct cctgcccaat cgggctcgct gacgaggaag cgctgtccct 1860
gtgatggagt tctctctcag agagtcttgg aaaagagacc acttgctctt gtttaaaaa 1920
aatttggacc tgatttttcc atgcagcctc tgggttgaat aaaacagcag ttgactgatg 1980
tttaaaaaaa aaaaaaaaaa

```

```

<210> 8
<211> 1101
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7473183CB1

```

```

<400> 8
atggaacctg ggcagcccc ggagcccag gagccccgc agccccggc aggagcggag 60
accgtgcgg ccccggtctg ggaggaagcc aagattttct acgacaacct cgcgcccag 120
aagaaaccca aatcgctcca gaatgcagtc accatcgctg tctcctccc agccttggtt 180
cgaatggacg aggagcagca gatctacacg gagcagggcg tggaggagta cgtgcgctac 240
cagctggaac atgagaagc acccttcagt cccgggccc ccttcccctt tctgaaggt 300
ctggagggcc tgaacaggg gctgcgggag ctgtaccctg atagtggaga cgtctctgac 360
atgctctcca tgactaacaa ccatgctcaa gtgggtgtcc gctcatcaa cagtatcaac 420
cactatgacc tgttcatgca gaggttctgc atgaccgat ggaacagccc gatctgctac 480
ctcaaggcct atcacaccaa cctctacttg tcagccgatg cggaaaaagt gcgagaagcc 540
attgatgagg ggaatgcagc tgcccaccatc ttcagcccca gaagggatgt ggttgtgtcc 600
cagagtcagc tgcgcgtggc ctctgatggg gacgcctgct tcttctegga cagatcggag 660
cgaatcgtca agcccacag gctggaccga ttctctgagc atgagaaggc ccacgagaac 720
aaactctgg ctcaaggccc cttaagggc ttctctggag cactgggtag gttgcagaa 780
aagtctact ccaaggcct cgggctggag tgcacaatc gtacctact ggtgacagca 840
cgaatgcaag ccagttccgg gcccgggct ccaagacc tgcgcagctg gggcctggag 900
acagatgaa ccttgttctc tctgtggag cccaaggccc ctctctctga gaagatccc 960
ccacacatct tctttgatga ccagatgttc catgtggctg gggctcagga gatgggcact 1020
gtggccccc atgtgcctta tgggttgga cagacacccc ggcggactgc acctgcaaa 1080
caggcccact ctgcacagta g

```

【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
4 April 2002 (04.04.2002)

PCT

(10) International Publication Number
WO 02/026998 A3(51) International Patent Classification: C12N 15/55,
9/14, C07K 16/40, G01N 33/573, 33/577, C12Q 1/34,
1/68, A61K 38/46, 39/395, C12N 1/21, 1/15, 5/10, A01K
67/027, A01H 5/00G. [US/US]; 111 Frederick Court, Mountain View, CA
94043 (US); LAL, Preeti [IN/US]; P.O. Box 5142, Sema
Clara, CA 95056 (US); THORNTON, Michael [US/US];
9 Medway Road, Woodside, CA 94062 (US).

(21) International Application Number: PCT/US01/30310

(74) Agents: HAMLET-COX, Diana et al.; Inceyte Genomics,
Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).(22) International Filing Date:
28 September 2001 (28.09.2001)(81) Designated States (national): AI, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK,
SI, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA,
ZW.

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/237,093 29 September 2000 (29.09.2000) US
60/238,370 6 October 2000 (06.10.2000) US
60/241,284 17 October 2000 (17.10.2000) US(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SI, 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, PL, SE, TR), OAPI patent (BF, BJ, CI,
CG, CI, CM, GA, GN, GQ, GW, ML, MR, NI, SN, TD,
TG).(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): YUE, Henry
[US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US).
BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road,
San Leandro, CA 94577 (US). WARREN, Bridget, A.
[US/US]; 10130 Parkwood Drive #2, Cupertino, CA 95014
(US). TRIBOULEY, Catherine, M. [US/US]; 1121 Ten-
nessee Street, #5, San Francisco, CA 94107 (US). FANG,
Y., Tom [US/US]; 4230 Rutwick Court, San Jose, CA
95118 (US). KHAN, Farrah, A. [IN/US]; 3617 Central
Road #102, Glenview, IL 60025 (US). YAO, Monique,Published:
— with international search report(88) Date of publication of the international search report:
17 July 2003For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

WO 02/026998 A3

(54) Title: HYDROLASES

(57) Abstract: The invention provides human hydrolases (HYDR) and polynucleotides which identify and encode HYDR. 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 HYDR.

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International Application No. PCT/US 01/30310
A. CLASSIFICATION OF SUBJECT MATTER		
IPC 7	C12N15/55 C12N9/14 C12Q1/34 C12Q1/68 C12N1/15 C12N5/10	C07K16/40 601N33/573 A61K38/46 A61K39/395 A01K67/027 A01H5/00
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols) IPC 7 C12N		
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)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL 'Online! EBI Similarity to E. coli P22333 Adenosine deaminase, 3 March 2000 (2000-03-03) NCI-CGAP: "gf44c04.x1 Soares testis NHT Homo sapiens cDNA clone" Database accession no. AI150008 XP002210564	1-16, 20, 23, 26-29, 31, 36-56, 60
Y	the whole document --- -/-	17-19, 30, 32-35
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents:		
A document defining the general state of the art which is not considered to be of particular relevance		*I* 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
E earlier document but published on or after the international filing date		*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
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)		*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, each combination being obvious to a person skilled in the art
O document referring to an oral disclosure, use, exhibition or other means		*Z* document member of the same patent family
P document published prior to the international filing date but later than the priority date claimed		
Date of the actual completion of the international search	Date of mailing of the international search report	
3 September 2002	12. 03. 03	
Name and mailing address of the ISA European Patent Office, P.O. 5518 Patentlaan 2 NL - 2526 HP Rijswijk Tel: (+31-70) 340-2040, Tx: 31 651 epo nl, Fax: (+31-70) 340-2016	Authorized officer Welland, S	

INTERNATIONAL SEARCH REPORT		Inventor's Application No. PCT/US 01/30310
C/(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>DATABASE EMBL 'Online! EB18 June 2000 (2000-06-08) HEGDE P. ET AL.: "Assessment of gene expression patterns in a model of colon tumor metastasis using a 19200 element cDNA microarray" retrieved from EMBL Database accession no. AW964086 XP002210565</p>	1-16, 20, 23, 26-29, 31, 36-56, 60
Y	the whole document	17-19, 30, 32-35
T	<p>-& HEGDE, P. ET AL.: "Identification of tumor markers in models of human colorectal cancer using a 19200-element complementary DNA microarray" CANCER RESEARCH, vol. 61, no. 21, 1 November 2001 (2001-11-01), pages 7792-7797, XP002210615 the whole document</p>	1-20, 23, 26-56, 60
A	<p>WO 00 28045 A (INCYTE PHARMA INC ;AZIMZAI YALDA (US); CORLEY NEIL C (US); YUE HEN) 18 May 2000 (2000-05-18) abstract page 5, paragraph 2 page 20, paragraph 5 -page 23, paragraph 3 page 54, paragraph 2 claim 7</p>	1-20, 23, 26-56, 60
A	<p>CHANG Z ET AL: "DEDUCED AMINO ACID SEQUENCE OF ESCHERICHIA-COLI ADENOSINE DEAMINASE REVEALS EVOLUTIONARILY CONSERVED AMINO ACID RESIDUES IMPLICATIONS FOR CATALYTIC FUNCTIONS" BIOCHEMISTRY, vol. 30, no. 8, 1991, pages 2273-2280, XP002210562 ISSN: 0006-2960 abstract page 2277, right-hand column, paragraph 2 -page 2279, left-hand column, paragraph 2 figure 4B table 1</p>	1-20, 23, 26-56, 60

	-/--	

INTERNATIONAL SEARCH REPORT		International Application No. PCT/US 01/30310
C/(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MORISAKI T ET AL: "ADENYLATE DEAMINASE A MULTIGENE FAMILY IN HUMANS AND RATS" JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 265, no. 20, 1990, pages 11482-11486, XP002210563 ISSN: 0021-9258 abstract page 11482, left-hand column, paragraph 1 -right-hand column, paragraph 1 page 11485, right-hand column, paragraph 3 -page 11486, left-hand column, paragraph 2	1-20, 23, 26-56, 60
A	HERSHFIELD M S ET AL: "Clinical expression, genetics and therapy of adenosine deaminase (ADA) deficiency." JOURNAL OF INHERITED METABOLIC DISEASE, vol. 20, no. 2, 1997, pages 179-185, XP001097451 ISSN: 0141-8955 page 181, paragraph 3 -page 182, paragraph 2	1-20, 23, 26-56, 60
A	VALENZUELA A ET AL: "HIV-1 envelope gp120 and viral particles block adenosine deaminase binding to human CD26." ADVANCES IN EXPERIMENTAL MEDICINE AND BIOLOGY. UNITED STATES 1997, vol. 421, 1997, pages 185-192, XP001097454 ISSN: 0065-2598 page 185, paragraph 1 -page 186, paragraph 2 page 188, paragraph 3 -page 191, paragraph 1	1-20, 23, 26-56, 60

INTERNATIONAL SEARCH REPORT		International application No. PCT/US 01/30310
Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)		
This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:		
1.	<input checked="" type="checkbox"/>	Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely: Although claims 19, 33 and 35 are directed to methods of treatment of the human/animal body and to diagnostic methods practised on the human/ animal body, resp., the search has been carried out and based on the alleged effects of the compounds and compositions.
2.	<input checked="" type="checkbox"/>	Claims Nos.: 21, 22, 24 and 25 because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically: see FURTHER INFORMATION sheet PCT/ISA/210
3.	<input type="checkbox"/>	Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).
Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)		
This International Searching Authority found multiple inventions in this International application, as follows:		
1.	<input type="checkbox"/>	As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2.	<input type="checkbox"/>	As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3.	<input type="checkbox"/>	As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4.	<input checked="" type="checkbox"/>	No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims. It is covered by claims Nos.: 1-20, 23 and 26-55(all partial) and 56 and 60(complete)
Remark on Protest		
	<input type="checkbox"/>	The additional search fees were accompanied by the applicant's protest.
	<input type="checkbox"/>	No protest accompanied the payment of additional search fees.

International Application No. PCT/US 01 80310

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 21, 22, 24 and 25

Present claims 21, 22, 24 and 25 relate to compounds defined by reference to a desirable characteristic or property, namely to compositions comprising a compound effective as agonist or antagonist of a polypeptide according to SEQ ID NO. 1-4, and to methods defined by the administration of said compounds, resp.

The claims cover all compounds having this characteristic or property and methods employing said compounds, whereas the application provides no support within the meaning of Article 6 PCT and disclosure within the meaning of Article 5 PCT for such compounds. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the compounds by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search of the claimed scope impossible.

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

International Application No. PCT/US 01 /0310

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

1. Claims: 1-20,
23 and 26-
55 (all partial) and 56 and 60 (complete)

Polypeptides comprising an amino acid sequence according to SEQ ID NO.1, polynucleotides encoding said polypeptide, in particular polynucleotides comprising a polynucleotide sequence according to SEQ ID NO. 5, recombinant polynucleotides comprising said polynucleotide sequence, transformed cells and transgenic organisms comprising said recombinant polynucleotides, methods of preparing and/or purifying said polypeptides, pharmaceutical compositions comprising said polypeptides, antibodies binding to said polypeptides, said antibodies for diagnosing a disease associated with the expression of said polypeptides, methods for preparing monoclonal or polyclonal antibodies binding to said polypeptides and compositions comprising said antibodies, methods for detecting a polynucleotide according to SEQ ID NO. 5 by hybridization or amplification techniques, methods of screening for a compound effective as agonist or antagonist or compounds modulating the activity of the polypeptides according to SEQ ID NO. 1, method of assessing the toxicity of a test compound, method for detecting said polypeptides in a sample, microarrays with said polynucleotides and use thereof for generating a transcript image of said sample.

2. Claims: 1-20,
23 and 26-
55 (all partial) and 57 and 61 (complete)

Polypeptides comprising an amino acid sequence according to SEQ ID NO.2, polynucleotides encoding said polypeptide, in particular polynucleotides comprising a polynucleotide sequence according to SEQ ID NO. 6, recombinant polynucleotides comprising said polynucleotide sequence, transformed cells and transgenic organisms comprising said recombinant polynucleotides, methods of preparing and/or purifying said polypeptides, pharmaceutical compositions comprising said polypeptides, antibodies binding to said polypeptides, said antibodies for diagnosing a disease associated with the expression of said polypeptides, methods for preparing monoclonal or polyclonal antibodies binding to said polypeptides and compositions comprising said antibodies, methods for detecting a polynucleotide according to SEQ ID NO. 6 by hybridization or amplification techniques, methods of screening for a compound effective as agonist or antagonist or compounds modulating the activity of the polypeptides according to SEQ ID NO. 2, method of assessing the toxicity of a test compound, method for detecting said polypeptides in a sample, microarrays with said polynucleotides and use thereof for generating a

International Application No. PCT/US 01 80310

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

transcript image of said sample.

3. Claims: 1-20,
23 and 26-
55 (all partial) and 58 and 62 (complete)

Polypeptides comprising an amino acid sequence according to SEQ ID NO.3, polynucleotides encoding said polypeptide, in particular polynucleotides comprising a polynucleotide sequence according to SEQ ID NO. 7, recombinant polynucleotides comprising said polynucleotide sequence, transformed cells and transgenic organisms comprising said recombinant polynucleotides, methods of preparing and/or purifying said polypeptides, pharmaceutical compositions comprising said polypeptides, antibodies binding to said polypeptides, said antibodies for diagnosing a disease associated with the expression of said polypeptides, methods for preparing monoclonal or polyclonal antibodies binding to said polypeptides and compositions comprising said antibodies, methods for detecting a polynucleotide according to SEQ ID NO. 7 by hybridization or amplification techniques, methods of screening for a compound effective as agonist or antagonist or compounds modulating the activity of the polypeptides according to SEQ ID NO. 3, method of assessing the toxicity of a test compound, method for detecting said polypeptides in a sample, microarrays with said polynucleotides and use thereof for generating a transcript image of said sample.

4. Claims: 1-20,
23 and 26-
55 (all partial) and 59 and 63 (complete)

Polypeptides comprising an amino acid sequence according to SEQ ID NO.4, polynucleotides encoding said polypeptide, in particular polynucleotides comprising a polynucleotide sequence according to SEQ ID NO. 8, recombinant polynucleotides comprising said polynucleotide sequence, transformed cells and transgenic organisms comprising said recombinant polynucleotides, methods of preparing and/or purifying said polypeptides, pharmaceutical compositions comprising said polypeptides, antibodies binding to said polypeptides, said antibodies for diagnosing a disease associated with the expression of said polypeptides, methods for preparing monoclonal or polyclonal antibodies binding to said polypeptides and compositions comprising said antibodies, methods for detecting a polynucleotide according to SEQ ID NO. 8 by hybridization or amplification techniques, methods of screening for a compound effective as agonist or antagonist or compounds modulating the activity of the polypeptides according to SEQ ID NO. 4, method of

International Application No. PCT/US 01/80310

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

assessing the toxicity of a test compound, method for detecting said polypeptides in a sample, microarrays with said polynucleotides and use thereof for generating a transcript image of said sample.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International Application No
PCT/US 01/30310

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0028045 A	18-05-2000	AU 1914400 A EP 1131443 A	29-05-2000 12-09-2001

フロントページの続き

(51) Int.Cl. ⁷	F I	テーマコード(参考)
A 6 1 P 25/00	A 6 1 P 35/00	4 B 0 6 4
A 6 1 P 35/00	A 6 1 P 37/02	4 B 0 6 5
A 6 1 P 37/02	A 6 1 P 43/00	4 C 0 8 4
A 6 1 P 43/00	C 0 7 K 1/22	4 C 0 8 5
C 0 7 K 1/22	C 0 7 K 16/18	4 H 0 4 5
C 0 7 K 16/18	C 0 7 K 16/40	
C 0 7 K 16/40	C 0 7 K 16/46	
C 0 7 K 16/46	C 1 2 M 1/00	A
C 1 2 M 1/00	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 9/14	
C 1 2 N 5/10	C 1 2 P 21/08	
C 1 2 N 9/14	C 1 2 Q 1/02	
C 1 2 N 15/02	C 1 2 Q 1/34	
C 1 2 P 21/08	C 1 2 Q 1/68	A
C 1 2 Q 1/02	G 0 1 N 33/15	Z
C 1 2 Q 1/34	G 0 1 N 33/50	Z
C 1 2 Q 1/68	G 0 1 N 33/53	M
G 0 1 N 33/15	G 0 1 N 33/573	A
G 0 1 N 33/50	G 0 1 N 37/00	1 0 2
G 0 1 N 33/53	C 1 2 N 15/00	C
G 0 1 N 33/573	C 1 2 N 15/00	F
G 0 1 N 37/00	C 1 2 N 5/00	A
	A 6 1 K 37/02	

(81) 指定国 AP(GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW

- (72) 発明者 ボーゲン、マライア・アール
アメリカ合衆国カリフォルニア州 9 4 5 7 7 ・サンレアンドロ・サンティアゴロード 1 4 2 4 4
- (72) 発明者 ワレン、ブリジット、エイ
アメリカ合衆国カリフォルニア州 9 5 0 1 4 ・クーペルティノー・# 2 ・パーウッドドライブ 1 0 1 3 0
- (72) 発明者 トリボレー、キャサリーン・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 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 カーン、ファラ・エイ
アメリカ合衆国イリノイ州 6 0 0 2 5 ・グレンビュー・# 1 0 2 ・セントラルロード 3 6 1 7
- (72) 発明者 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 ラル、ブリーティ

アメリカ合衆国カリフォルニア州 9 5 0 5 6 ・ サンタクララ ・ ピーオーボックス 5 1 4 2

(72)発明者 ソートン、マイケル

アメリカ合衆国カリフォルニア州 9 4 0 6 2 ・ ウッドサイド ・ メッドウェイロード 9

F ターム(参考) 2G045 AA25 AA40 BB03 BB20 DA12 DA13 DA14 DA36 DA37 DA77
 FB01 FB02 FB03 FB07
 4B024 AA01 AA11 BA11 BA44 CA04 CA05 CA06 CA07 CA09 CA10
 CA11 DA02 DA06 EA02 EA04 FA02 FA10 GA05 GA11 GA18
 GA19 HA03 HA08 HA09 HA12 HA14
 4B029 AA07 AA21 AA23 BB20 CC03 FA12
 4B050 CC01 CC03 DD11 LL01 LL03
 4B063 QA01 QA18 QQ20 QQ42 QQ63 QQ89 QR08 QR10 QR13 QR14
 QR20 QR32 QR41 QR42 QR55 QR58 QR59 QR62 QR63 QR66
 QR69 QR77 QR82 QS12 QS24 QS25 QS28 QS34 QS39 QX01
 QX02 QX07
 4B064 AG27 CA10 CA20 CC24 CE12 DA01 DA13
 4B065 AA26X AA90X AA93Y AB01 AC14 BA01 BA25 CA31 CA44 CA46
 4C084 AA01 AA07 AA17 DC50 NA14 ZA011 ZA591 ZB011 ZB071 ZB261
 ZC19
 4C085 KA04 LL09 LL13 LL18
 4H045 AA10 AA11 AA20 AA30 BA10 BA41 CA40 DA76 EA20 EA50
 FA72 FA74 GA26

专利名称(译)	水解酶		
公开(公告)号	JP2004524815A	公开(公告)日	2004-08-19
申请号	JP2002530761	申请日	2001-09-28
[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	ユエヘンリー ボーグンマライアアール ワレンブリジットエイ トリボレーキャサリーンエム タングワイトム カーンファラエイ ヤオモニークジー ラルプリーティ ソーントンマイケル		
发明人	ユエ、ヘンリー ボーグン、マライア・アール ワレン、ブリジット、エイ トリボレー、キャサリーン・エム タング、ワイトム カーン、ファラ・エイ ヤオ、モニーク・ジー ラル、プリーティ ソーントン、マイケル		
IPC分类号	G01N33/50 A61K38/00 A61K45/00 A61K49/00 A61P11/00 A61P25/00 A61P35/00 A61P37/02 A61P43/00 C07K1/22 C07K16/18 C07K16/40 C07K16/46 C12M1/00 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N9/14 C12N15/02 C12N15/09 C12N15/55 C12P21/08 C12Q1/02 C12Q1/34 C12Q1/68 G01N33/15 G01N33/53 G01N33/573 G01N37/00		
CPC分类号	A01K2217/05 A61K38/00 A61K2039/505 A61P11/00 A61P25/00 C12N9/14		
FI分类号	C12N15/00.ZNA.A A61K45/00 A61K49/00.Z A61P11/00 A61P25/00 A61P35/00 A61P37/02 A61P43/00.111 C07K1/22 C07K16/18 C07K16/40 C07K16/46 C12M1/00.A C12N1/15 C12N1/19 C12N1/21 C12N9/14 C12P21/08 C12Q1/02 C12Q1/34 C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.M G01N33/573.A G01N37/00.102 C12N15/00.C C12N15/00.F C12N5/00.A A61K37/02		
F-TERM分类号	2G045/AA25 2G045/AA40 2G045/BB03 2G045/BB20 2G045/DA12 2G045/DA13 2G045/DA14 2G045/DA36 2G045/DA37 2G045/DA77 2G045/FB01 2G045/FB02 2G045/FB03 2G045/FB07 4B024/AA01 4B024/AA11 4B024/BA11 4B024/BA44 4B024/CA04 4B024/CA05 4B024/CA06 4B024/CA07 4B024/CA09 4B024/CA10 4B024/CA11 4B024/DA02 4B024/DA06 4B024/EA02 4B024/EA04 4B024/FA02 4B024/FA10 4B024/GA05 4B024/GA11 4B024/GA18 4B024/GA19 4B024/HA03 4B024/HA08 4B024/HA09 4B024/HA12 4B024/HA14 4B029/AA07 4B029/AA21 4B029/AA23 4B029/BB20 4B029/CC03 4B029/FA12 4B050/CC01 4B050/CC03 4B050/DD11 4B050/LL01 4B050/LL03 4B063/QA01 4B063/QA18 4B063/QQ20 4B063/QQ42 4B063/QQ63 4B063/QQ89 4B063/QR08 4B063/QR10 4B063/QR13 4B063/QR14 4B063/QR20 4B063/QR32 4B063/QR41 4B063/QR42 4B063/QR55 4B063/QR58 4B063/QR59 4B063/QR62 4B063/QR63 4B063/QR66 4B063/QR69 4B063/QR77 4B063/QR82 4B063/QS12 4B063/QS24 4B063/QS25 4B063/QS28 4B063/QS34 4B063/QS39 4B063/QX01 4B063/QX02 4B063/QX07 4B064/AG27 4B064/CA10 4B064/CA20 4B064/CC24 4B064/CE12 4B064/DA01 4B064/DA13 4B065/AA26X 4B065/AA90X 4B065/AA93Y 4B065/AB01 4B065/AC14 4B065/BA01 4B065/BA25 4B065/CA31 4B065/CA44 4B065/CA46 4C084/AA01 4C084/AA07 4C084/AA17 4C084/DC50 4C084/NA14		

4C084/ZA011 4C084/ZA591 4C084/ZB011 4C084/ZB071 4C084/ZB261 4C084/ZC19 4C085/KA04
 4C085/LL09 4C085/LL13 4C085/LL18 4H045/AA10 4H045/AA11 4H045/AA20 4H045/AA30 4H045
 /BA10 4H045/BA41 4H045/CA40 4H045/DA76 4H045/EA20 4H045/EA50 4H045/FA72 4H045/FA74
 4H045/GA26

優先権
 60/237093 2000-09-29 US
 60/238370 2000-10-06 US
 60/241284 2000-10-17 US

外部リンク Espacenet

摘要(译)

本发明提供了人类水解酶 (HYDR) 和鉴定和编码HYDR的多核苷酸。
 本发明还提供表达载体, 宿主细胞, 抗体, 激动剂和拮抗剂。本发明还
 提供了用于诊断, 治疗或预防与HYDR的异常表达有关的疾病的方法。

		特許公開番号		特許公開日	
		P2004-524815A		平成16年8月19日(2004.8.19)	
		(43) 公表日		平成16年8月19日(2004.8.19)	
(51) Int. Cl. ⁷	F I	C 1 2 N 15/00	Z N A A	テーマコード (参考)	
C 1 2 N 15/09	A 6 1 K 38/00	A 6 1 K 45/00		2 G 0 4 5	
A 6 1 K 45/00	A 6 1 K 49/00	A 6 1 P 11/00	Z	4 B 0 2 4	
A 6 1 K 49/00	A 6 1 P 25/00			4 B 0 2 9	
A 6 1 P 11/00				4 B 0 5 0	
A 6 1 P 11/00				4 B 0 6 3	
		審査請求	未請求	予備審査請求	有 (全 197 頁) 最終頁に続
(21) 出願番号	特願2002-530761 (P2002-530761)	(71) 出願人	301005050		
(86) (22) 出願日	平成13年9月28日 (2001.9.28)		インサイト・ゲノミクス・インコーポレ イテッド		
(85) 翻訳文提出日	平成15年3月25日 (2003.3.25)		アメリカ合衆国カリフォルニア州943C 4・パロアルト・ボータードライブ 31 60		
(86) 国際出願番号	PCT/US2001/030310	(74) 代理人	100089266		
(87) 国際公開番号	W02002/026998		弁理士 大島 陽一		
(87) 国際公開日	平成14年4月4日 (2002.4.4)	(72) 発明者	ユエ、ヘンリー		
(31) 優先権主張番号	60/237,093		アメリカ合衆国カリフォルニア州9408 7・サニーベイル・ルイスアベニュー 8 26		
(32) 優先日	平成12年9月29日 (2000.9.29)				
(33) 優先権主張国	米国 (US)				
(31) 優先権主張番号	60/238,370				
(32) 優先日	平成12年10月6日 (2000.10.6)				
(33) 優先権主張国	米国 (US)				
(31) 優先権主張番号	60/241,284				
(32) 優先日	平成12年10月17日 (2000.10.17)				
(33) 優先権主張国	米国 (US)				
最終頁に続く					