

(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2004-501632

(P2004-501632A)

(43) 公表日 平成16年1月22日(2004.1.22)

(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 38/43	A 6 1 K 45/00	2 G O 4 5
A 6 1 K 45/00	A 6 1 P 17/02	4 B O 2 4
A 6 1 P 17/02	A 6 1 P 29/00	4 B O 5 0
A 6 1 P 29/00	A 6 1 P 35/00	4 B O 6 3
		4 B O 6 4
審査請求 未請求 予備審査請求 未請求 (全 187 頁) 最終頁に続く		

(21) 出願番号	特願2002-505956 (P2002-505956)	(71) 出願人	301005050
(86) (22) 出願日	平成13年6月13日 (2001.6.13)		インサイト・ゲノミックス・インコーポレイテッド
(85) 翻訳文提出日	平成14年2月25日 (2002.2.25)		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/US2001/019166	(74) 代理人	100089266
(87) 国際公開番号	W02002/000840		弁理士 大島 陽一
(87) 国際公開日	平成14年1月3日 (2002.1.3)	(72) 発明者	ソーントン、マイケル
(31) 優先権主張番号	60/213,383		アメリカ合衆国カリフォルニア州94062・ウッドサイド・メッドウェイロード 9
(32) 優先日	平成12年6月23日 (2000.6.23)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/215,544		
(32) 優先日	平成12年6月30日 (2000.6.30)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/222,818		
(32) 優先日	平成12年8月4日 (2000.8.4)		
(33) 優先権主張国	米国 (US)		
最終頁に続く			

(54) 【発明の名称】 ヒトリアーゼ

(57) 【要約】

本発明は、ヒトリアーゼ (H L Y A) と、H L Y A を同定しコードするポリヌクレオチドとを提供する。本発明はまた、発現ベクター、宿主細胞、抗体、アゴニスト及びアンタゴニストを提供する。更に、本発明は、H L Y A の発現に関連する疾患を診断、治療または予防する方法も提供する。

【特許請求の範囲】

【請求項 1】

以下の (a) 乃至 (d) を有する群から選択した実質上単離されたポリペプチド。

(a) 配列番号 1 乃至 3 を有する群から選択したポリペプチド

(b) 配列番号 1 乃至 3 を有す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アミノ酸配列を有する天然のポリペプチド

(c) 配列番号 1 乃至 3 を有する群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片

(d) 配列番号 1 乃至 3 を有する群から選択したアミノ酸配列を有するポリペプチドの免疫抗原性断片

10

【請求項 2】

配列番号 1 乃至 3 を有する群から選択した請求項 1 に記載の単離されたポリペプチド。

【請求項 3】

請求項 1 のポリペプチドをコードする単離されたポリヌクレオチド。

【請求項 4】

請求項 2 のポリペプチドをコードする単離されたポリヌクレオチド。

【請求項 5】

配列番号 4 乃至 6 を有する群から選択した請求項 4 に記載の単離されたポリヌクレオチド。

【請求項 6】

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

20

【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 に記載のポリペプチドを製造する方法であって、

(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前記ポリペプチドの発現に適した条件下で培養する過程と、

30

(b) そのように発現した前記ポリペプチドを受容する過程とを有し、

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結合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】

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。

【請求項 11】

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

(a) 配列番号 4 乃至 6 を有す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) 配列番号 4 乃至 6 を有す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

40

(c) (a) のポリヌクレオチドに相補的なポリヌクレオチド

(d) (b) のポリヌクレオチドに相補的なポリヌクレオチド

(e) (a) ~ (d) の RNA 等価物

【請求項 12】

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

【請求項 13】

請求項 11 に記載のポリヌクレオチドの配列を有する標的ポリヌクレオチドをサンプル中から検出する方法であって、

50

(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とを有し、

前記プローブと前記標的ポリヌクレオチド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。

【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13に記載の方法。 10

【請求項15】

請求項11に記載のポリヌクレオチドの配列を有する標的ポリヌクレオチドをサンプル中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とを含むことを特徴とする方法。

【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。 20

【請求項17】

前記ポリペプチドが、配列番号1乃至3を有す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16に記載の成分。

【請求項18】

機能性HLYA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16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の有効性を確認するために化合物をスクリーニングする方法であって、

(a) 請求項1に記載のポリペプチドを含むサンプルを化合物に曝す過程と、

(b) 前記サンプル中のアゴニスト活性を検出する過程とを含むことを特徴とする方法。

【請求項20】

請求項19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HLYA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20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。 40

【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、

(a) 請求項1に記載のポリペプチドを含むサンプルを化合物に曝す過程と、

(b) 前記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ことを特徴とする方法。

【請求項23】

請求項22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 24】

機能性 H L Y A 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 23 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 25】

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

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

(b) 請求項 1 に記載のポリペプチドの試験化合物との結合を検出し、それによって請求項 1 に記載のポリペプチドに特異結合する化合物を同定する過程とを含むことを特徴とする方法。 10

【請求項 26】

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

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

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

(c) 試験化合物の存在下での請求項 1 に記載のポリペプチドの活性を、試験化合物の不存在下での請求項 1 に記載のポリペプチドの活性と比較する過程とを含み、 20

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 27】

請求項 5 に記載の配列を有する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であって、

(a) 前記標的ポリヌクレオチドの発現に適した条件下で、該標的ポリヌクレオチドを含むサンプルを化合物に曝す過程と、

(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。 30

【請求項 28】

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) 請求項 11 に記載の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と、請求項 11 に記載のポリヌクレオチドまたはその断片のポリヌクレオチド配列を有する前記生物学的サンプルの標的ポリヌクレオチドとの間に、特定のハイブリタイゼーション複合体が形成されるような条件下で、前記処理されたサンプルの核酸を前記プローブでハイブリタイズする過程と、

(c) 前記ハイブリタイゼーション複合体の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、 40

前記処理された生物学的サンプル中の前記ハイブリタイゼーション複合体の量の差が、前記試験化合物の毒性を標示することを特徴とする方法。

【請求項 29】

生物学的サンプル中の H L Y A 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 10 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 10 に記載の抗体。

10

【請求項 31】

請求項 10 に記載の抗体と、許容できる賦形剤とを含む化合物。

【請求項 32】

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

【請求項 33】

前記抗体が標識されることを特徴とする請求項 31 に記載の化合物。

【請求項 34】

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

20

【請求項 35】

請求項 10 に記載の抗体の特異性を有するポリクローナル抗体を調製する方法であって、
(a) 抗体反応を誘発する条件下で、配列番号 1 乃至 3 を有する群から選択したアミノ酸配列またはその免疫抗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、配列番号 1 乃至 3 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に特異結合するようなポリクローナル抗体を同定する過程とを含むことを特徴とする方法。

【請求項 36】

請求項 35 に記載の方法で産出した抗体。

30

【請求項 37】

請求項 36 に記載の抗体及び適切なキャリアを含む化合物。

【請求項 38】

請求項 10 に記載の抗体の特異性を有する抗体を用いてモノクローナル抗体を製造する方法であって、

(a) 抗体反応を誘発する条件下で、配列番号 1 乃至 3 を有する群から選択したアミノ酸配列またはその免疫抗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) 不滅の細胞を用いて前記抗体産出細胞を融合して、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) 配列番号 1 乃至 3 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

【請求項 39】

請求項 38 に記載の方法で産出したモノクローナル抗体。

【請求項 40】

請求項 39 に記載の抗体及び適切なキャリアを含む化合物。

【請求項 41】

50

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

【請求項 42】

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

【請求項 43】

配列番号 1 乃至 3 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 10 に記載の抗体をインキュベートする過程と、

10

(b) 特異結合を検出する過程とを含み、

該特異結合が、配列番号 1 乃至 3 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標示することを特徴とする方法。

【請求項 44】

配列番号 1 乃至 3 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を精製する方法であって、

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

(b) 前記サンプルから前記抗体を分離し、配列番号 1 乃至 3 を有す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

20

【請求項 45】

配列番号 1 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 46】

配列番号 2 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 47】

配列番号 3 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 48】

配列番号 4 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。

【請求項 49】

配列番号 5 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。

30

【請求項 50】

配列番号 6 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。

【発明の詳細な説明】

【0001】

(技術分野)

本発明は、ヒトリアーゼの核酸及びアミノ酸配列に関し、癌、自己免疫 / 炎症性疾患、細胞増殖異常及び眼病の診断、治療並びに予防において、また、ヒトリアーゼの核酸及びアミノ酸の発現に対する外因性化合物の効果の算定において、これらの配列を利用することに関する。

【0002】

40

(発明の背景)

リアーゼは、C - C、C - O、C - N、C - S、C - (ハロゲン化物)、P - Oまたはその他の結合の切断を非加水分解的にまたは非酸化的に触媒し、少なくとも一方は二重結合を有する2つの分子を形成する或るクラスの酵素である (S t r y e r , L . (1 9 9 5) B i o c h e m i s t r y , W . H . F r e e m a n a n d C o . , N e w Y o r k N Y , p . 6 2 0) 。 酵素の国際分類 (W e b b , E . C . (1 9 9 2) E n z y m e N o m e n c l a t u r e 1 9 9 2 : R e c o m m e n d a t i o n s o f t h e N o m e n c l a t u r e C o m m i t t e e o f t h e I n t e r n a t i o n a l U n i o n o f B i o c h e m i s t r y a n d M o l e c u l a r B i o l o g y o n t h e N o m e n c l a t u r e a n d C l a s s i f i c

50

ation of Enzymes, Academic Press, San Diego (CA) の下では、リアーゼは、最初の桁が 4 である酵素番号 (例えば EC 4 . x . x . x) により表されるような異なるクラスを形成する。

【0003】

リアーゼの更なる分類は、切断された基の特性と同様に結合切断のタイプを反映する。C - C リアーゼのグループには、カルボキシルリアーゼ (デカルボキシラーゼ)、アルデヒドリアーゼ (アルドラーゼ)、オキソ酸リアーゼ及びその他のリアーゼが含まれる。C - O リアーゼのグループには、多糖類に作用するリアーゼであるヒドロリアーゼ、及びその他のリアーゼが含まれる。C - N リアーゼのグループには、アンモニアリアーゼ、アミジンリアーゼ、アミンリアーゼ (デアミナーゼ) 及びその他のリアーゼが含まれる。リアーゼは、他の多様な酵素的プロセスと同様に脂肪酸代謝及びトリカルボン酸サイクルを含む代謝エネルギー生成において役割を有する細胞生化学の重要な成分である。

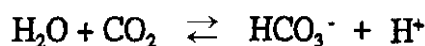
10

【0004】

リアーゼの重要なファミリーの 1 つには炭酸アンヒドラーゼ (CA) があり、これは炭酸デヒドラーゼとも呼ばれ、以下の反応において二酸化炭素の水和作用を触媒する。

【0005】

【数 1】



【0006】

CA は、タンパク質表面下約 15 Å の深さの切れ込みに位置し、3 つの His 残基のイミダゾール基に配位された亜鉛イオンにより 10^6 以上の因子によってこの反応を加速する。亜鉛イオンへの水結合は、急速に HCO_3^- に変換される。

20

【0007】

炭酸アンヒドラーゼの 8 つの酵素的な、進化的に関連する形状がヒト中に存在するものとして現在知られており、その 8 つとは、3 つのサイトゾルのアイソザイム (CA I、CA II、CA III)、2 つの膜結合形状 (CA IV 及び CA V)、ミトコンドリア形状 (CA V)、分泌唾液形状 (CA VI) 及び未だ特徴づけられていないアイソザイム (Prositate PDOCOO146 真核性タイプ炭酸アンヒドラーゼサイン) である。アイソエンザイム CA I、CA II 及びウシ CA III は類似の二次構造及びポリペプチド鎖の折り畳みを有しているが、CA I はトリプトファンを 6 つ、CA II は 7 つ、CA III は 8 つ有している (Boren, K. ら (1996) Protein Sci. 5: 2479 - 2484)。CA II は、哺乳動物の脳において優性な CA アイソエンザイムである。

30

【0008】

CA は、pH 調整、 CO_2 及び HCO_3^- 輸送、イオン輸送、水と電解質の均衡を含む様々な生理学的プロセスに参与する。例えば、CA II は、胃壁細胞による H^+ の分泌、腎尿細管細胞による H^+ の分泌、 H^+ を分泌して骨再吸収区画を酸性化する破骨細胞による H^+ の分泌に参与する。更に、CA II は膀胱細胞、毛様体 (ciliary body) 上皮、脈絡叢、唾液腺腺房細胞及び遠位結腸上皮による HCO_3^- の分泌を促進し、それによって涙液、眼房水、脳脊髄液及び唾液の産出において役割を果たし、水と電解質の均衡に参与する。CA II はまた、腎臓の近位尿細管、赤血球、肺における CO_2 交換を促進する。CA IV は、幾つかの組織において役割を有する。CA IV は、腎臓での HCO_3^- の再吸収を促進し、脳、骨格筋、心筋を含む組織における CO_2 流入を促進し、血液から肺胞への CO_2 交換を促進する。CA VI は、おそらく CA II を伴って唾液中の pH 調整において役割を果たし、食道及び胃において保護効果を有し得る。ミトコンドリアの CA V は、経路中の CA 抑制因子の効果に基づき、糖新生及び尿素形成 (ureagenesis) において重要な役割を果たすように見える (Sly, W. S. and flue, P. Y. (1995) Ann. Rev. Biochem. 64: 375 - 401)。

40

50

【0009】

多くの病状は、CA活性における変化によって特徴づけられる。CAIIにおける突然変異は、CAII欠失を導くものであり、腎尿細管アシドーシスを伴う大理石骨病(Online Medelian Inheritance in Man 259730 Osteopetrosis with Renal Tubular Acidosis)の原因である。脳脊髄液(CSF)中のCAIIの濃度は、脳損傷の患者において病原活性を特徴づけるようにみえる。脳梗塞の患者には高CA濃度が観察されてきた。一過性脳虚血発作、多発性硬化症または癲癇の患者のCAII濃度は、通常は正常な範囲にあるが、中枢神経系感染、痴呆または3叉神経痛の患者のCSF中ではより高いCAIIレベルが観察されてきた(Parkkilä, A.K. ら (1997) Eur. J. Clin. Invest. 27:392-397)。結腸の腺腫及び腺癌はCAを染色できないことが観察されてきたが、他方では、非腫瘍性の対照は結腸陰窩の上半分を裏打ちする円柱細胞の細胞質においてCAI及びCAIIを示した。新生物は、正常な陰窩の基盤を裏打ちするようなあまり成熟が進んでいない細胞に類似の染色パターンを示す(Gramlich T.L. ら (1990) Arch. Pathol. Lab. Med. 114:415-419)。

10

【0010】

多数の疾患への治療上の介入は、CA活性の変化に関与している。アセタゾールアミドなどのCA抑制因子は、緑内障(Stewart, W.C. (1999) Curr. Opin. Ophthalmol. 10:99-108)、本態性振戦及びパーキンソン病(Uitti, R.J. (1998) Geriatrics 53:46-48, 53-57)、間欠性運動失調(Singhvi, J.P. ら (2000) Neurology India 48:78-80)及び高所関連病(Kiocke, D.L. ら (1998) Mayo Clin. Proc. 73:988-992)の治療に用いられる。

20

【0011】

酵素が生理学的変化に比較的緩徐に反応するので、CA活性は、長期間の病状の指標として特に有用であり得る。CAT及び亜鉛濃度は、甲状腺機能亢進性グレース病において減少が観察されており(ヨシダ, K. (1996) Tohoku J. Exp. Med. 178:345-356)、グリコシル化CATは糖尿病において観察される(Kondo, T. ら (1987) Clin. Chim. Acta 166:227-236)。CAIとCAII間の反応性及び子宮内膜症に積極的な相関性が観察されてきた(Brinton, D.A. ら (1996) Ann. Clin. Lab. Sci. 26:409-420、D'Cruz, O.J. ら (1996) Fertil. Steril. 66:547-556)。

30

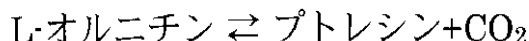
【0012】

リアーゼファミリーの別の重要なメンバーは、ポリアミン生合成における最初の律速酵素であるオルニチンデカルボキシラーゼ(ODC)である。ODCは、

【0013】

【数2】

40



【0014】

の反応においてオルニチンからプトレシンへの形質転換を触媒する。ポリアミンには、プトレシン及び後続の代謝経路産物であるスペルミジン及びスペルミンが含まれ、ポリアミンは、DNA合成、細胞分化及び増殖に不可欠な遍在性の細胞成分である。従って、ポリアミンは腫瘍増殖において役割を果たす(Medina, M.A. ら (1999) Biochem. Pharmacol. 57:1341-1344)。

【0015】

ODCは、ピリドキサル-5'-リン酸(PLP)依存酵素であり、これはホモ二量体

50

として活性である。保存された残基には、基質結合領域の一部と考えられるグリシン残基のひと配列及びPLP結合部位における残基がある(Prositate PDOC00685 Orn/DAP/Arg デカルボキシラーゼファミリー-2サイン)。哺乳動物のODCには、細胞内分解のためのシグナルとして作用するプロリングルタミン酸、セリン及びスレオニン残基中で濃縮された配列断片であるPEST領域も含まれる(前出のMedina)。

【0016】

多くの化学発癌物質及び腫瘍プロモーターは、ODCレベル及び活性を増加させる。幾つかの既知の腫瘍遺伝子はODC遺伝子の増強性転写によりODCレベルを増加させ得るものであり、ODC自体は非常に高レベルで発現される際に腫瘍遺伝子として作用し得る。多数の前癌性状態において高レベルのODCが見られ、ODCレベルの上昇は、腫瘍促進化合物に対するスクリーンの一部として用いられてきた(Pegg, A. E. ら (1995) J. Cell. Biochem. Suppl. 22: 132-138)。

10

【0017】

ODCの抑制因子は、動物モデル及びヒトの臨床試験において腫瘍を治療するために用いられてきたし、膀胱、脳、食道、消化管、肺、口腔、乳腺、胃、皮膚及び気管の腫瘍の発達を減少させることを示してきた(前出のPegg, McCann, P. P. and A. E. Pegg (1992) Pharmac. Ther. 54: 195-215)。ODCはまた、化学的予防のための標的としての見込みを示す(前出のPegg)。ODC抑制因子は、アフリカトリパノソーマ、マラリア、ニューモシスチスカリニによる感染を治療するのに用いられてきたし、リウマチ様関節炎などの自己免疫疾患の治療にも潜在的に有用である(前出のMcCann)。

20

【0018】

ピリドキサル依存デカルボキシラーゼの別のファミリーは、2群デカルボキシラーゼである。このファミリーには、グルタミン酸の脱炭酸を神経伝達物質GABAに触媒するグルタミン酸デカルボキシラーゼ(GAD)と、ヒスチジンの脱炭酸をヒスタミンに触媒するヒスチジンドカルボキシラーゼ(HDC)と、トリプトファンの脱炭酸をトリプタミンに触媒し、5-ヒドロキシ-トリプトファン及びジヒドロキシフェニルアラニン(レボドーパ)としても作用する、レボドーパデカルボキシラーゼまたはトリプトファンデカルボキシラーゼとしても知られている芳香族Lアミノ酸デカルボキシラーゼ(DDC)と、システインからタウリンを合成する際の速度制限酵素であるシステインスルフィン酸デカルボキシラーゼ(CSD)とがある(PROSITE PDOC00329 DDC/GAD/HDC/TyrDCピリドキサルリン酸付着部位)。タウリンは、脳で豊富であるスルホンアミノ酸であり、脳細胞で浸透圧調節器として働くと考えられている(Bitoun, M. and Tappaz, M. (2000) J. Neurochem. 75: 919-924)。

30

【0019】

新たなヒトリアーゼ及びそれらをコードするポリヌクレオチドの発見は、自己免疫/炎症性疾患、細胞増殖異常及び眼病の診断、治療並びに予防において、また、ヒトリアーゼの核酸及びアミノ酸の発現に対する外因性化合物の効果の算定において有用である新たな組成を提供することにより、当分野における要求を満たす。

40

【0020】

(発明の概要)

本発明は、集合的には「HLYA」、個別には「HLYA-1」、「HLYA-2」及び「HLYA-3」と呼ばれるような、実質上精製されたポリペプチドであるヒトリアーゼに特徴がある。或る実施態様において本発明は、(a)配列番号1乃至3を有する群から選択したポリペプチド、(b)配列番号1乃至3を有する群から選択したアミノ酸配列と少なくとも90%が同一であるアミノ酸配列を有する天然のポリペプチド、(c)配列番号1乃至3を有する群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または(d)配列番号1乃至3を有する群から選択したアミノ酸配列を有するポリペ

50

プチドの免疫抗原性断片を含む群から選択した実質上単離されたポリペプチドを提供する。一実施態様では、配列番号 1 乃至 3 を有する群から選択したアミノ酸配列を含む実質上単離されたポリペプチドを提供する。

【0021】

また、本発明は (a) 配列番号 1 乃至 3 を有する群から選択したアミノ酸配列、(b) 配列番号 1 乃至 3 を有する群から選択したアミノ酸配列と少なくとも 90 % の同一性を有する天然のアミノ酸配列、(c) 配列番号 1 乃至 3 を有する群から選択したアミノ酸配列の生物学的活性断片、または (d) 配列番号 1 乃至 3 を有する群から選択したアミノ酸配列の免疫抗原性断片を含むポリペプチドをコードするような実質上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配列番号 1 乃至 3 を有す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配列番号 4 乃至 6 を有する群から選択される。

10

【0022】

本発明は更に、(a) 配列番号 1 乃至 3 を有する群から選択したアミノ酸配列、(b) 配列番号 1 乃至 3 を有する群から選択したアミノ酸配列と少なくとも 90 % の同一性を有する天然のアミノ酸配列、(c) 配列番号 1 乃至 3 を有する群から選択したアミノ酸配列の生物学的活性断片、または (d) 配列番号 1 乃至 3 を有する群から選択したアミノ酸配列の免疫抗原性断片を含むポリペプチドをコードするような実質上単離されたポリヌクレオチドと機能的に結合したプロモーター配列を有する組換えポリヌクレオチドを提供する。一実施態様では、本発明は組換えポリヌクレオチドを用いて形質転換した細胞を提供する。別の実施態様では、本発明は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。

20

【0023】

また、本発明は (a) 配列番号 1 乃至 3 を有する群から選択したアミノ酸配列、(b) 配列番号 1 乃至 3 を有する群から選択したアミノ酸配列と少なくとも 90 % の同一性を有する天然のアミノ酸配列、(c) 配列番号 1 乃至 3 を有する群から選択したアミノ酸配列の生物学的活性断片、または (d) 配列番号 1 乃至 3 を有する群から選択したアミノ酸配列の免疫抗原性断片を含む実質上単離されたポリペプチドを製造する方法を提供する。製造方法は、(a) 組換えポリヌクレオチドを用いて形質転換した細胞をポリペプチドの発現に適した条件下で培養する過程と、(b) そのように発現したポリペプチドを受容する過程とを有し、組換えポリヌクレオチドはポリペプチドをコードするポリヌクレオチドに機能的に結合したプロモーター配列を有する。

30

【0024】

本発明は更に、(a) 配列番号 1 乃至 3 を有する群から選択したアミノ酸配列、(b) 配列番号 1 乃至 3 を有する群から選択したアミノ酸配列と少なくとも 90 % の同一性を有する天然のアミノ酸配列、(c) 配列番号 1 乃至 3 を有する群から選択したアミノ酸配列の生物学的活性断片、または (d) 配列番号 1 乃至 3 を有する群から選択したアミノ酸配列の免疫抗原性断片を含むポリペプチドに特異結合するような実質上単離された抗体を提供する。

【0025】

本発明は更に、(a) 配列番号 4 乃至 6 を有する群から選択したポリヌクレオチド配列、(b) 配列番号 4 乃至 6 を有す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、(c) (a) に相補的なポリヌクレオチド配列、(d) (b) に相補的なポリヌクレオチド配列、または (e) (a) ~ (d) の RNA 等価物を含む実質上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有する。

40

【0026】

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) 配列番号 4 乃至 6 を有する群から選択したポリヌクレオチド配列、(b) 配列番号 4 乃至 6 を有する群から選択したポリヌクレオチド配列と少なく

50

とも90%の相同性を有する天然のポリヌクレオチド配列、(c)(a)に相補的なポリヌクレオチド配列、(d)(b)に相補的なポリヌクレオチド配列、または(e)(a)~(d)のRNA等価物を含む実質上単離されたポリヌクレオチドを提供する。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からなる少なくとも20の連続したヌクレオチドを含むプローブを用いて該サンプルをハイブリダイズする過程と、(b)ハイブリダイゼーション複合体の存在・不存在を検出し、複合体が存在する場合にはオプションでその量を検出する過程からなり、プローブと標的ポリヌクレオチド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。

10

【0027】

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)配列番号4乃至6を有する群から選択したポリヌクレオチド配列、(b)配列番号4乃至6を有す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、(c)(a)に相補的なポリヌクレオチド配列、(d)(b)に相補的なポリヌクレオチド配列、または(e)(a)~(d)のRNA等価物を含む実質上単離されたポリヌクレオチドを提供する。検出方法は、(a)ポリメラーゼ連鎖反応増幅を用いて標的ポリヌクレオチドまたはその断片を増幅する過程と、(b)標的ポリヌクレオチドまたはその断片の存在・不存在を検出し、該標的ポリヌクレオチドまたはその断片が存在する場合にはオプションでその量を検出する過程を含む。

20

【0028】

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)配列番号1乃至3を有する群から選択したアミノ酸配列、(b)配列番号1乃至3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、(c)配列番号1乃至3を有する群から選択したアミノ酸配列の生物学的活性断片、または(d)配列番号1乃至3を有する群から選択したアミノ酸配列の免疫抗原性断片を含む。一実施態様では、成分は配列番号1乃至3を有する群から選択したアミノ酸配列を含む。更に本発明は、機能性HLYA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成分を投与する過程を有する方法を提供する。

30

【0029】

本発明はまた、(a)配列番号1乃至3を有する群から選択したアミノ酸配列、(b)配列番号1乃至3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、(c)配列番号1乃至3を有する群から選択したアミノ酸配列の生物学的活性断片、または(d)配列番号1乃至3を有する群から選択したアミノ酸配列の免疫抗原性断片を含む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。一実施態様では、本発明は機能性HLYA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成分を投与する過程を含む方法を提供する。

40

【0030】

本発明は更に、(a)配列番号1乃至3を有する群から選択したアミノ酸配列、(b)配列番号1乃至3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、(c)配列番号1乃至3を有する群から選択したアミノ酸配列の生物学的活性断片、または(d)配列番号1乃至3を有する群から選択したアミノ酸配列の免疫抗原性断片を含む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)サンプル中のアゴニスト活性を検出する

50

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。別の実施態様では、機能性 H L Y A の過剰発現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成分を投与する過程を含む方法を提供する。

【0031】

本発明は更に、(a) 配列番号 1 乃至 3 を有する群から選択したアミノ酸配列、(b) 配列番号 1 乃至 3 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、(c) 配列番号 1 乃至 3 を有する群から選択したアミノ酸配列の生物学的活性断片、または (d) 配列番号 1 乃至 3 を有する群から選択したアミノ酸配列の免疫抗原性断片を含むポリペプチドに特異結合する化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a) ポリペプチドを適切な条件下で少なくとも 1 つの試験化合物に結合させる過程と、(b) 試験化合物とのポリペプチドの結合を検出し、それによってポリペプチドに特異結合する化合物を同定する過程とを含む。

10

【0032】

本発明は更に、(a) 配列番号 1 乃至 3 を有する群から選択したアミノ酸配列、(b) 配列番号 1 乃至 3 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、(c) 配列番号 1 乃至 3 を有する群から選択したアミノ酸配列の生物学的活性断片、または (d) 配列番号 1 乃至 3 を有する群から選択したアミノ酸配列の免疫抗原性断片を含むポリペプチドの活性を調節する化合物をスクリーニングする方法を提供する。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、ポリペプチドを少なくとも 1 つの試験化合物に結合させる過程と、(b) ポリペプチドの活性を試験化合物の存在下で算定する過程と、(c) 試験化合物の存在下でのポリペプチドの活性を試験化合物の不存在下でのポリペプチドの活性と比較する過程とを含み、試験化合物の存在下でのポリペプチドの活性の変化は、ポリペプチドの活性を調節する化合物を標示する。

20

【0033】

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、配列番号 4 乃至 6 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。

30

【0034】

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) 配列番号 4 乃至 6 を有す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) 配列番号 4 乃至 6 を有す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% の相同性を有するポリヌクレオチド配列を含む天然のポリヌクレオチド、(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 等価物を含む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から構成されるプローブを用いて、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) 配列番号 4 乃至 6 を有す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) 配列番号 4 乃至 6 を有す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% の相同性を有するポリヌクレオチド配列を含む天然のポリヌクレオチド、(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、(v) (i) ~ (iv) の RNA 等価物を含む群から選択する。或いは、標的ポリヌクレオチドは、上記 (i) ~ (v) を含む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

40

50

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。

【0035】

(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、物質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、本発明の範囲を限定することを意図したものではないことも併せて理解されたい。本発明の範囲は、特許請求の範囲にのみ限定されることになる。

10

【0036】

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

【0037】

本明細書中で用いる全ての専門用語及び科学用語は、特に定義されている場合を除き、当業者に一般に理解されている意味と同じ意味を有する。本明細書で説明するものと類似あるいは同等の任意の装置、材料及び方法を用いて本発明の実施または試験を行うことができるが、ここでは好適な装置、材料、方法について説明する。本発明で言及する全ての刊行物は、刊行物中で報告されていて且つ本発明に関係があるであろう細胞、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。

20

【0038】

定義

「HLYA」は、実質上精製されたHLYAのアミノ酸配列であって、任意の種、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の種から得たもので、任意の天然物、合成物、半合成物或いは組換え物を起源とするものを指す。

30

【0039】

「アゴニスト」の語は、HLYAの生物学的活性を強化または擬態する分子を指す。このアゴニストは、HLYAと直接相互作用することによって、或いはHLYAが関与する生物学的経路の構成エレメントに作用することによって、HLYAの活性を調節するようなタンパク質、核酸、糖質、小分子その他の任意の化合物や成分を含み得る。

【0040】

「対立遺伝子変異体」は、HLYAをコードする遺伝子の別の形態である。対立遺伝子変異体は、核酸配列における少なくとも1の突然変異から作製し得る。また、変異RNAまたはポリペプチドからも作製し得る。ポリペプチド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若しくは数回生じ得る。

40

【0041】

HLYA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、HLYAと同一またはHLYA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。この定義に含まれるのは、HLYA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或いは検出困難な多型現象と、HLYAをコードするポリヌクレオチド配列のための正常

50

な染色体遺伝子座以外の遺伝子座に占めるような、対立遺伝子変異体に対する不適切或いは不測のハイブリダイゼーションである。コードされたタンパク質も「変異」し得るものであり、サイレント変化を生ぜしめて結果的に機能的に等価な H L Y A となるようなアミノ酸残基の欠失、挿入または置換を含み得る。計画的アミノ酸置換は、H L Y A の生物学的または免疫学的活性が保持される限りにおいて、残基の極性、電荷、溶解度、疎水性、親水性及び/または両親媒性特性の類似性に基づき行い得る。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。

10

【0042】

「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会合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。

【0043】

「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 (P C R) 技術を用いて行う。

20

【0044】

「アンタゴニスト」の語は、H L Y A の生物学的活性を阻害或いは弱める分子を指す。アンタゴニストとしては、抗体などのタンパク質、核酸、糖質、小分子またはその他の任意の化合物や成分を挙げることができるが、これらは H L Y A と直接相互作用することによって、或いは H L Y A が関与する生物学的経路の構成エレメントに作用することによって、H L Y A の活性を調節する。

【0045】

「抗体」の語は、無損傷免疫グロブリンやその断片、例えば F a 、 F (a b ') ₂ 及び F v 断片を指すが、これらはエピトープの決定基と結合することができる。H L Y A ポリペプチドを結合する抗体は、無損傷ポリペプチドを用いるか或いは免疫抗原として感心のある小ペプチドを含む断片を用いるかして調製することができる。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、翻訳または化学合成された R N A に由来し得るもので、好みに応じて担体タンパク質に接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キーホールリンペットヘモシアニン (K L H) 等がある。結合ペプチドは、動物を免疫化するために用いる。

30

【0046】

「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、抗体に結合するための無損傷抗原 (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。

40

【0047】

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

50

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。

【0048】

「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「免疫学的に活性」は、天然、組換えまたは合成のHLYA、或いはその任意のオリゴペプチドの能力であって、適切な動物または細胞において特定の免疫反応を誘導して特定の抗体と結合し得る能力を指す。

【0049】

「相補(的)」または「相補性」の語は、ポリヌクレオチドとポリヌクレオチドの、塩基対形成による自然結合を指す。例えば、配列「5' A - G - T 3'」は、その相補配列「3' T - C - A 5'」に結合する。

【0050】

「所定のポリヌクレオチド配列からなる成分」及び「所定のアミノ酸配列からなる成分」は、概して所定のポリヌクレオチド配列またはアミノ酸配列からなる任意の成分を指す。この成分には、乾燥製剤または水溶液が含まれ得る。HLYAまたはHLYA断片をコードするポリヌクレオチドからなる成分は、ハイブリダイゼーションプローブとして利用することができる。このプローブは凍結乾燥状態で保存し得るものであり、糖質等の安定化剤と会合し得る。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。

【0051】

「コンセンサス配列」は、不必要な塩基を分離するために再配列し、XL-PCRキット(Applied Biosystems, Foster City, CA)を用いて5'方向、3'方向のいずれか一方向或いは両方向に伸長させ、更に再配列した核酸配列を指す。或いは、断片アセンブルのコンピュータプログラムを用いて、1若しくは数個のIncycyteクローンの、場合によっては1若しくは数個のパブリックドメインESTの、オーバーラップした配列から組み立てた核酸配列を指す。コンピュータプログラムの例としては、GELVIEW断片アセンブリーシステム(GCG, Madison WI)やPhrap(University of Washington, Seattle WA)が挙げられる。伸長及びアセンブルを共に行ってコンセンサス配列を決定する配列もある。

【0052】

「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され得るアミノ酸と、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
Gln	Asn, Glu, His
Glu	Asp, Gln, His
Gly	Ala
His	Asn, Arg, Gln, Glu
Ile	Leu, Val

10

20

30

40

50

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

【0053】

10

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 螺旋構造、(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。

【0054】

「欠失」は、結果的に1若しくは数個のアミノ酸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。

【0055】

「誘導体」の語は、ポリペプチド配列またはポリヌクレオチド配列の化学修飾を指す。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(pegylation)、或いは任意の同様なプロセスであって誘導起源のポリペプチドから少なくとも1つの生物学的若しくは免疫学的機能を保持しているプロセスによって、修飾されたポリペプチドである。

20

【0056】

「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。

【0057】

「断片」は、H L Y AまたはH L Y Aをコードするポリヌクレオチドの固有部分であって、親配列と同一配列であるが親配列よりも長さが短い配列を指す。断片は、画定された配列の全長から1ヌクレオチド/アミノ酸残基を差し引いた長さよりも短い長さを有し得る。例えば或る断片は、5~1000の連続したヌクレオチドまたはアミノ酸残基を有し得る。プローブ、プライマー、抗原、治療用分子として、或いはその他の目的のために用いられる断片は、少なくとも5、10、15、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。

30

40

【0058】

配列番号4乃至6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、配列番号4乃至6を特異的に同定するものであり、例えば同一ゲノム中の配列番号4乃至6以外の配列とは異なるものである。配列番号4乃至6の断片は、例えば、ハイブリダイゼーション及び増幅技術において、或いは関連するポリヌクレオチド配列から配列番号4乃至6を区別する類似の方法において有用である。配列番号4乃至6の断片の正確な長さ及び断片に対応する配列番号4乃至6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。

【0059】

50

配列番号 1 乃至 3 の断片は、配列番号 4 乃至 6 の断片によってコードされる。配列番号 1 乃至 3 の断片には、配列番号 1 乃至 3 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、配列番号 1 乃至 3 の断片は、配列番号 1 乃至 3 を特異認識する抗体を産出するための免疫抗原性ペプチドとして有用である。配列番号 1 乃至 3 の断片及び断片に対応する配列番号 1 乃至 3 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。

【0060】

「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。

10

【0061】

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

【0062】

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。

【0063】

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の一部であり、一式の分子生物学分析プログラム (DNASTAR, Madison WI) である。CLUSTAL V については、Higgins, D. G. and P. M. Sharp (1989) CABIOS 5:151-153 及び Higgins, D. G. ら (1992) CABIOS 8:189-191 の文献に記載されている。ポリヌクレオチド配列の対をなすアラインメントの場合、デフォルトパラメータは、Ktuple = 2、gap penalty = 5、window = 4、「diagonals saved」= 4 と設定する。デフォルトとして「重みづけされた」残基の重みづけ表を選択する。CLUSTAL V は、アラインメントされたポリヌクレオチド配列対間の「類似率」として一致率を報告する。

20

30

【0064】

或いは、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が、国立バイオテクノロジー情報センター (NCBI) Basic Local Alignment Search Tool (BLAST) から提供されており (Altschul, S. F. ら (1990) J. Mol. Biol. 215:403-410)、これは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/>) を含む幾つかの情報源から入手可能である。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 (2

40

50

000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。

【0065】

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

10

【0066】

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。

【0067】

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。

20

【0068】

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、置換部位の酸性度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。

30

【0069】

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

40

【0070】

或いは、NCBI BLASTソフトウェア一式を用いてもよい。例えばポリペプチド配列を2つ1組で比較をする場合、デフォルトパラメータとして設定されたblastpと共に「BLAST 2 Sequences」ツールVersion 2.0.12(2000年4月21日)を使用してもよい。デフォルトパラメータの設定例を以下に示す。

【0071】

Matrix: BLOSUM62
 Open Gap: 11 及び Extension Gap: 1 penalties
 Gap x drop-off: 50

50

Expect : 10
Word Size : 3
Filter : on
【0072】

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、或る長さであってその長さに対し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。

【0073】

「ヒト人工染色体」（HAC）は直鎖状の小染色体であり、6 kb ~ 10 MbのサイズのDNA配列を含み、安定した有糸分裂染色体の分離及び維持に必要な全てのエレメントが含まれている。

【0074】

「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。

【0075】

「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。アニーリング許容条件は、例えば約6 × SSC、約1%（w/v）のSDS及び約100 µg/mlの変性サケ精子DNAの存在下で温度68において成立する。

【0076】

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

【0077】

本発明の、ポリヌクレオチドとポリヌクレオチドのハイブリダイゼーションに対する高ストリンジェンシー条件には、約0.2 × SSC及び約1%のSDS存在下で約68において1時間の洗浄条件が含まれる。或いは、65、60、55 または42 の温度で行ってもよい。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、遮断剤を用いて非特異ハイブリダイゼーションを阻止する。このような遮断剤には、例えば、約100 ~ 200 µg/mlの変性サケ精子DNAがあ

る。特定条件下で、例えばRNAとDNAのハイブリダイゼーションに有機溶剤、例えば約35～50% v/vの濃度のホルムアミドを用いることもできる。洗浄条件の有用なバリエーションは、当業者には自明であろう。ハイブリダイゼーションは、特に高ストリンジェント条件下では、ヌクレオチド間の進化的な類似性を示唆し得る。このような類似性は、ヌクレオチド及びヌクレオチドにコードされるポリペプチドに対する類似の役割を強く示唆している。

【0078】

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

10

【0079】

「挿入」及び「付加」の語は、1若しくは数個のアミノ酸残基またはヌクレオチド配列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。

【0080】

「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。

20

【0081】

「免疫抗原性断片」とは、哺乳動物等の生命体に導入されると免疫応答を誘発し得るようなHLYAのポリペプチドまたはオリゴペプチド断片である。「免疫抗原性断片」の語には、本明細書中で開示したような或いは当分野で既知であるような任意の抗体産出方法において有用なHLYAの任意のポリペプチドまたはオリゴペプチド断片も含まれる。

【0082】

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

【0083】

「エレメント」または「アレイエレメント」の語は、マイクロアレイの環境において、基質の表面上に配置されたハイブリダイズ可能なポリヌクレオチドを指す。

30

【0084】

「調節(する)」の語は、HLYAの活性の変化を指す。調節することによって例えば、HLYAのタンパク質活性、結合特性その他の生物学的、機能的または免疫学的特性が増大または低下し得る。

【0085】

「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。

40

【0086】

「機能的に結合した」は、第1核酸配列が第2核酸配列と機能的な関係があるように配置された状態を指す。例えば、プロモーターがコード配列の転写または発現に影響を及ぼす場合には、そのプロモーターはそのコード配列に機能的に結合している。同一のリーディングフレーム内で2つのタンパク質コード領域を結合する必要がある場合、一般に、機能的に結合したDNA配列は非常に近接するか、或いは連続し得る。

【0087】

「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

50

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。

【0088】

HLYA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、タンパク分解性開裂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、HLYAの酵素環境に依存し、細胞タイプによって変化し得る。

10

【0089】

「プローブ」は、HLYA、HLYAの相補配列またはこれらの断片をコードする核酸配列を指し、同一核酸配列、対立遺伝子核酸配列または関連する核酸配列の検出に用いられる。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。

【0090】

「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。

20

【0091】

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。

【0092】

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

30

40

【0093】

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までの大きめのポリヌクレオチドであって32キロボ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

50

医療センターのゲノムセンターから一般向けに入手可能)は、メガベース配列から特定のプライマーを選択することが可能であり、従ってゲノム全体の範囲でプライマーを設計するのに有用である。Primer3プライマー選択プログラム(マサチューセッツ州ケンブリッジのWhitehead Institute/MITゲノム研究センターから一般向けに入手可能)ではユーザーが「ミスプライミング・ライブラリ」をインプットすることができ、ここでプライマー結合部位として避けたい配列はユーザーが指定する。Primer3は特に、マイクロアレイのためのオリゴヌクレオチドの選択に有用である(後二者のプライマー選択プログラムのソースコードは、各自のソースから得てユーザー固有のニーズを満たすように変更してもよい)。PrimerGenプログラム(英国ケンブリッジ市の英国ヒトゲノムマッピングプロジェクト・リソースセンターから一般向けに入手可能)は、多数の配列アラインメントに基づいてプライマーを設計し、それによって、アラインメントされた核酸配列の最大保存領域または最小保存領域の何れかとハイブリダイズするようなプライマーの選択を可能にする。従って、このプログラムは、固有であって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。

10

20

30

40

50

【0094】

「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されるが、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、例えば前出のSambrookらの文献に記載されているような遺伝子工学的手法によって達成される。組換え核酸の語は、単に核酸の一部を付加、置換または欠失した変異核酸も含む。しばしば組換え核酸には、プロモーター配列に機能的に結合した核酸配列が含まれる。このような組換え核酸は、ベクターの不可欠なエレメントであって例えばある細胞を形質転換するために用いられるようなものであり得る。

【0095】

或いはこのような組換え核酸は、ウイルスベクターの不可欠なエレメント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。ワクシニアウイルスは組換え核酸が発現する哺乳動物のワクチン接種に用いるもので、哺乳動物の防御免疫応答を誘導する。

【0096】

「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またはウイルスタンパク質と相互作用する。

【0097】

「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。

【0098】

DNA配列に関する「RNA等価物」は、発生した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。

【0099】

「サンプル」の語は、その最も広い意味で用いられる。HLYAをコードする核酸若しく

はその断片、またはH L Y A自体を含む疑いのあるサンプルは、体液や、細胞から単離された細胞、染色体、細胞小器官または膜からの抽出物や、細胞や、溶解しているか基質に結合しているゲノムDNA、RNAまたはcDNAや、組織や、組織プリント等から構成され得る。

【0100】

「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ことを意味している。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）を含むポリヌクレオチドの存在が、抗体に結合している標識されたAの量を低減させる。

10

【0101】

「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。

【0102】

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

20

【0103】

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

【0104】

「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。

【0105】

「形質転換」は、外来性のDNAが宿主細胞に入り込み、宿主細胞を変化させるプロセスを表す。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、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」細胞の語には、限られた時間内に挿入されたDNAやRNAを発現するような一時的に形質転換された細胞のみならず、安定的に形質転換された細胞であってその中に挿入されたDNAが自律的に複製するプラスミドとして或いは宿主の染色体の一部として複製可能であるものも含まれる。

30

40

【0106】

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

50

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。

【0107】

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「多形性」変異体として説明し得る。スプライス変異体は参照分子とかなりの相同性を有し得るが、mRNAプロセッシング中のエキソンの交互スプライシングによって通常多数の或いは僅かな数のポリヌクレオチドを有することになる。対応するポリペプチドは、追加機能ドメインを有するか或いは参照分子に存在するドメインが欠落していることがある。種変異体は、種相互に異なるポリヌクレオチド配列である。結果的に生じるポリペプチドは通常、相互にかなりのアミノ酸相同性を有する。多形性変異体は、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、多形性変異体は、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「一塩基多型」(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。

10

20

【0108】

特定のポリペプチド配列の「変異体」は、ポリペプチド配列の1本の長さ全体で特定のポリペプチド配列に対して少なくとも40%の相同性を有するポリペプチド配列として画定される。ここで、デフォルトパラメータに設定した「BLAST 2 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%、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。

【0109】

発明

本発明は、新規なヒトリアーゼ(HLYA)と、HLYAをコードするポリヌクレオチドと、癌、自己免疫/炎症性疾患、細胞増殖異常及び眼病の診断、治療並びに予防にこれらの配列を利用する方法の発見に基づくものである。

【0110】

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)と相関している。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)の両方により表示した。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)の両方により表示した。

30

40

【0111】

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及び列2は本発明の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)及び対応するIncyteポリペプチド配列番号(IncyteポリペプチドID)を、列3は最も近いGenBank相同体のGenBank識別認識番号(Genbank ID NO)を、列4は各ポリペプチドとそのGenBank相

50

同体が一致する確率スコアを、列5はGenBank相同体の注釈と可能であればそれに関連する引用と共に示している。これらは全て特に引用することを以って本明細書の一部となす。

【0112】

表3は、本発明のポリペプチドの様々な構造的特徴を示している。列1及び列2は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(SEQ ID NO)及び対応するIncyteポリペプチド配列番号(IncyteポリペプチドID)を、列3は各ポリペプチドのアミノ酸残基の数を、列4はリン酸化部位を、列5はGCG配列分析ソフトウェアパッケージ(Genetics Computer Group, Madison WI)のMOTIFSプログラムにより決定したような潜在化グリコシル化部位を、列6はサイン配列、ドメイン及びモチーフからなるアミノ酸残基を、列7はタンパク質構造/機能分析及び場合によっては分析手法が適用される検索可能なデータベースを示している。

10

【0113】

合わせて、表2及び表3は、本発明のポリペプチドの特性を概略的に示している。これらの特性は、特許請求の範囲に記載されたポリペプチドがヒトリアーゼであることを証明する。例えば配列番号1は、Basic Local Alignment Search Tool (BLAST)により決定されるように、残基M1からM239まで59%がヒト炭酸アンヒドラーゼI (GenBank ID g179793)と同一である(表2を参照)。BLAST確率スコアは $6.9E-80$ であり、これは保存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。配列番号1はまた、保存されたタンパク質ファミリードメインの隠れマルコフモデル(HMM)ベースPFAMデータベースにおいて統計学的に有意な一致を検索することにより決定されるような真核型炭酸アンヒドラーゼドメインも含む(表3を参照)。BLIMPS、MOTIFS及びPROFILERESCAN分析から得たデータは、配列番号1がヒト炭酸アンヒドラーゼであることを裏付ける証拠を更に提供する。別の例では、配列番号2は、Basic Local Alignment Search Tool (BLAST)により決定されるように、残基M1からI459まで59%がアフリカツメガエルオルニチンデカルボキシラーゼ-2 (GenBank ID g6746571)と同一である(表2を参照)。BLAST確率スコアは $1.9E-148$ であり、これは保存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。配列番号2はまた、保存されたタンパク質ファミリードメインの隠れマルコフモデル(HMM)ベースPFAMデータベースにおいて統計学的に有意な一致を検索することにより決定されるようなオルニチンデカルボキシラーゼドメインも含む(表3を参照)。BLIMPS及びMOTIFS分析から得たデータは、配列番号2が炭酸アンヒドラーゼであることを裏付ける証拠を更に提供する。更に別の例では、配列番号3は、Basic Local Alignment Search Tool (BLAST)により決定されるように、残基E30からR314まで58%がヒトアデニル酸キナーゼ(GenBank ID g3283386)と同一である(表2を参照)。BLAST確率スコアは $5.7E-94$ であり、これは保存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。配列番号3はまた、保存されたタンパク質ファミリードメインの隠れマルコフモデル(HMM)ベースPFAMデータベースにおいて統計学的に有意な一致を検索することにより決定されるような真核型炭酸アンヒドラーゼドメインも含む(表3を参照)。BLAST-DOMO、BLAST-PRODOM、BLIMPS及びPROFILERESCAN分析から得たデータは、配列番号3が真核型炭酸アンヒドラーゼであることを裏付ける証拠を更に提供する。配列番号1乃至3の分析のためのアルゴリズム及びパラメータについては、表7に示した。

20

30

40

【0114】

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アセンブルした。列1及び列2は本発明の各ポリヌクレオチドに対するポリヌクレオチド配列識別番号(ポリヌクレオチドSEQ ID NO)及び対応するIncyte

50

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、列 3 は塩基対の各ポリヌクレオチド配列の長さを示している。列 4 は、配列番号 4 乃至 6 を同定するか或いは配列番号 4 乃至 6 と関連するポリヌクレオチド配列を区別するような例えばハイブリダイゼーションまたは増幅技術において有用であるようなポリヌクレオチド配列の断片を収載している。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アSEMBLするのに用いた。表 4 の列 6 及び列 7 は、各々の完全長配列に関連付けられた列 5 の cDNA 及び / またはゲノム配列のヌクレオチド開始位置 (5 ') 及び終止位置 (3 ') を示している。

10

【 0 1 1 5 】

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 ライブラリを示し得る。例えば、2698586F6 は *Incyte* cDNA 配列の識別番号であり、UTRSNOT12 は単離元の cDNA ライブラリである。cDNA ライブラリが示されていない *Incyte* cDNA は、プールされた cDNA ライブラリ (例えば 71672723V1 または 70996811V1) に由来する。或いは列 5 の識別番号は、完全長ポリヌクレオチド配列のアSEMBLに寄与する GenBank cDNA または EST (例えば g1191247) を示し得る。或いは列 5 の識別番号は、ゲノム DNA の Genscan 分析により予想されるコード領域を示し得る。Genscan 予想コード配列は、アSEMBL前に編集しておくことができる (実施例 4 を参照) 。また

列 5 の識別番号は、「エキソンスティッチング (*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示し得る (実施例 5 を参照) 。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 識別番号は示していない。

20

【 0 1 1 6 】

表 5 は、*Incyte* cDNA 配列を用いてアSEMBL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アSEMBL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。

30

【 0 1 1 7 】

本発明には、HLYA の変異体も含まれる。好適な HLYA の変異体のアミノ酸配列は、HLYA アミノ酸配列と少なくとも約 80 %、または少なくとも約 90 %、或いは少なくとも約 95 % もの一致率を有し、HLYA の機能的若しくは構造的特徴を少なくとも 1 つ有するような変異体である。

【 0 1 1 8 】

本発明には、HLYA をコードするポリヌクレオチドも含まれる。或る例では、HLYA をコードする配列番号 4 乃至 6 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。配列番号 4 乃至 6 のポリヌクレオチド配列は、配列表に示されているように等価 RNA 配列と同等の価値を有しているが、窒素塩基チミンの出現はウラシルに置換され、糖のバックボーンはデオキシリボースではなくシリボースから構成されている。

40

【 0 1 1 9 】

本発明には、HLYA をコードするポリヌクレオチド配列の変異配列も含まれる。具体的には、そのようなポリヌクレオチド配列の変異配列は、HLYA をコードするポリヌクレオチド配列と少なくとも約 70 %、或いは少なくとも約 85 %、または少なくとも約 95 % もの一致率を有する。本発明の或る実施態様では、配列番号 4 乃至 6 からなる群から選択されたアミノ酸配列と少なくとも約 70 %、或いは少なくとも約 85 %、または少なく

50

とも約95%もの一致率を有するような配列番号4乃至6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の変異配列を含む。上記の任意のポリヌクレオチドの変異体は、HLYAの機能的若しくは構造的特徴を少なくとも1つ有するアミノ酸配列をコードし得る。

【0120】

当業者であれば、遺伝暗号の縮重の結果、HLYAをコードする多数のポリヌクレオチド配列(既知の遺伝子または天然の遺伝子のポリヌクレオチド配列と最低限の類似性しか有しないポリヌクレオチド配列もある)を産出し得ることは理解できよう。したがって本発明には、可能コドン選択に基づく組合せの選択によって産出し得る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HLYA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。

10

【0121】

HLYA及びその変異体をコードするヌクレオチド配列は通常、好適に選択されたストリンジェントな条件下で天然のHLYAのヌクレオチド配列とハイブリダイズ可能であるが、HLYAまたはその誘導体であって実質上異なるコドンの使用法があるもの、例えば天然に存在しないコドンの封入があるものをコードするヌクレオチド配列を作り出すことは有益であろう。宿主が特定のコドンを利用する頻度に基づいて、特定の真核又は原核宿主に発生するペプチドの発現率を高めるようにコドンを選択することが可能である。コードされたアミノ酸配列を変えなくHLYA及びその誘導体をコードするヌクレオチド配列を実質上変更する別の理由には、天然の配列から作製される転写物より好ましい、例えば半減期が長いなどの特性を有するRNA転写物の作製がある。

20

【0122】

本発明には、HLYA、HLYA誘導体及びこれらの断片をコードするDNA配列又はこれらの断片を完全に合成化学によって作製することも含まれる。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HLYA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。

【0123】

更に本発明には、種々のストリンジェントな条件下で、請求項に記載のポリヌクレオチド配列、特に配列番号4乃至6で示される配列及びそれらの断片にハイブリダイズ可能なポリヌクレオチド配列も含まれる(Wahl, G. M. and 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条件については、「定義」に記載した。

30

【0124】

DNAシーケンシングの方法は当分野では公知であり、本発明のいずれの実施例もDNAシーケンシング方法を用いて実施可能である。DNAシーケンシング方法には酵素を用いることができ、例えばDNAポリメラーゼIのクレノウ断片、SEQUENASE (US Biochemical, Cleveland OH)、Taqポリメラーゼ (Applied Biosystems)、熱安定性T7ポリメラーゼ (Amersham, 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、MICROLAB2200液体転移システム (Hamilton, Reno, NV)、PTC200サーマルサイクラー (MJ Research, Watertown MA)及びABI CATALYT 800サーマルサイクラー (Applied Biosystems)等の装置を用いて配列の準備を自動化する。次に、ABI 373 或いは 377 DNAシーケンシ

40

50

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, pp. 856 - 853.等を参照)。

【0125】

HLYAをコードする核酸配列を、部分的ヌクレオチド配列を利用し且つ当分野でよく知られているPCR法をベースにした種々の方法を用いて伸長させ、プロモーター及び調節要素等の上流配列を検出することができる。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 (Sarkar, G. (1993) PCR Methods Applic 2: 318 - 322等を参照)。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からなる制限断片から得る (Triglia, T.ら (1988) Nucleic Acids Res 16: 8186等を参照)。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増幅する方法に關与している (Lagerstrom, M.ら (1991) PCR Methods Applic 1: 111 - 119等を参照)。この方法では、PCRを行う前に多重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。

【0126】

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、ランダムに初回抗原刺激を受けたライブラリは、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。

【0127】

市販のキャピラリー電気泳動システムを用いて、シーケンシングまたはPCR産物のサイズを分析し、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

10

20

30

40

50

ピラリ-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。

【0128】

本発明の別の実施例では、HLYAをコードするポリヌクレオチド配列またはその断片を、適切な宿主細胞内でHLYA、HLYAの断片またはその機能的等価物を発現させるような組換えDNA分子内でクローニングし得る。遺伝暗号固有の宿重に起因して、実質的同一或いは機能的等価のアミノ酸配列をコードするような別のDNA配列を作製してHLYAの発現に利用し得る。

【0129】

限定するものではないが遺伝子産物のクローニング、プロセッシング、発現の調節を含む種々の目的のために、当分野で通常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み換え、HLYAコード化配列を変えることができる。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定方向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。

10

【0130】

本発明のヌクレオチドは、Molecular Breeding (Maxygen Inc., Santa Clara CA. 米国特許第5,837,458号、Chang, C.-C.ら(1999) Nat. Biotechnol. 17:793-797、Christians, F.C.ら(1999) Nat. Biotechnol. 17:259-264、Cramer, A.ら(1996) Nat. Biotechnol. 14:315-319に記載)等のDNAシャッフリング技術の対象となり、HLYAの生物学的特性、例えば生物活性、酵素力、或いは他の分子や化合物との結合力等を変更または向上させ得る。DNAシャッフリングは、遺伝子断片のPCR仲介再組換えを用いて遺伝子変異体のライブラリを生成するプロセスである。ライブラリはその後、その遺伝子変異体を所望の特性に同定するような選択またはスクリーニングにかける。次にこれらの好適な変異体をプールし、更に反復してDNAシャッフリング及び選択/スクリーニングを行ってもよい。このように、遺伝の多様性は「人為的」品種改良及び急速な分子の進化を経て創生される。例えば、ランダムポイント突然変異を有する単一の遺伝子の断片を再結合し、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再結合し、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。

20

30

【0131】

別の実施例では、当分野でよく知られている化学的方法を用いて、HLYAをコードする配列の全部或いは一部を合成し得る(Caruthers, M.H.ら(1980) Nucleic Acids Symp. Ser. 7:215-223、Horn, T.ら(1980) Nucleic Acids Symp. Ser. 7:225-232等を参照)。或いは、化学的方法を用いてHLYAそれ自体またはその断片を合成し得る。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, pp. 55-60、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HLYAのアミノ酸配列またはその任意の一部は、直接合成する間、または他のタンパク質から得た配列またはその任意の一部と結合する間、或いはその両方を行っている間に変更し、変異型ポリペプチドを生成することが可能である。

40

50

【0132】

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. and 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, pp. 28-53等を参照)。

【0133】

生物学的に活性なHLYAを発現させるために、HLYA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とができる。好適な発現ベクターとは即ち好適な宿主内で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要素を含むベクターである。必要な要素には、ベクター及びHLYA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プロモーター、5'及び3'の非翻訳領域などの調節配列がある。このような要素は、長さ及び特異性が様々である。固有開始シグナルを用いて、HLYAをコードする配列をより効果的に翻訳することが可能もある。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。HLYA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。発現の効率は、用いられる特定の宿主細胞系に好適なエンハンサーを包含することによって高めることができる (Scharf, D. ら (1994) *Results Probl. Cell Differ.* 20:125-162.等を参照)。

【0134】

当業者によく知られている方法を用いて、HLYAをコードする配列と、好適な転写及び翻訳調節要素とを含む発現ベクターを構築することが可能である。この方法には、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ある (Sambrook, J. ら (1989) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Press, Plainview NYの4, 8, 16-17章、Ausubel, F. M. ら (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NYの9, 13, 16章等を参照)。

【0135】

種々の発現ベクター/宿主系を利用して、HLYAをコードする配列を保持及び発現し得る。限定するものではないがこのような発現ベクター/宿主系には、組換えバクテリオファージ、プラスミドまたはコスミドDNA発現ベクターで形質転換させた細菌や、酵母菌発現ベクターで形質転換させた酵母菌や、ウイルス発現ベクター (例えばバキュロウイルス) に感染した昆虫細胞系や、ウイルス発現ベクター (例えばカリフラワーモザイクウイルスCaMVまたはタバコモザイクウイルスTMV) または細菌発現ベクター (例えばTiまたはpBR322プラスミド) で形質転換させた植物細胞系、動物細胞系などの微生物等がある (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、タカマツ, N. (1987) *EMBOJ.* 6:307-311、Coruzzi, G. ら (1984) *EMBOJ.* 3:1671-1680、Broglie, R. ら (1984) *Science* 224:838-843、Winter, J. ら (1991) *Results Probl. Cell Differ.* 17:85-105、『

10

20

30

40

50

『マグローヒル科学技術年鑑』(The McGraw Hill Yearbook of Science and Technology) (1992) McGraw Hill New York NY, pp. 191-196、Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81: 3655-3659、Harrington, J. J. 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、ポリヌクレオチド配列を標的器官、組織または細胞集団へ輸送することができる(Di Nicola, M. ら (1998) Cancer Gen. Ther. 5(6): 350-356、Yu, M. ら (1993) Proc. Natl. Acad. Sci. USA 90(13): 6340-6344、Buller, R. M. ら (1985) Nature 317(6040): 813-815; McGregor, D. P. ら (1994) Mol. Immunol. 31(3): 219-226、Verma, I. M. and N. Somia (1997) Nature 389: 239-242等を参照)。本発明は、使用する宿主細胞によって限定されるものではない。

10

20

30

40

50

【0136】

細菌系では、HLYAをコードするポリヌクレオチド配列の使用目的に応じて多数のクローニングベクター及び発現ベクターを選択し得る。例えば、HLYA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、サブクローニング及び増殖には、PBLUES CRIPT (Stratagene, La Jolla CA) またはPSPORT1プラスミド(Life Technologies)などの多機能大腸菌ベクターを用いることができる。HLYAをコードする配列の、ベクターの多数のクローニング部位への連結反応によって、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を同定する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(Van Heeke, G. and S. M. Schuster (1989) J. Biol. Chem. 264: 5503-5509, 等を参照)。多量のHLYAが必要な場合、例えば抗体を生成する場合などには、HLYA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T5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。

【0137】

酵母の発現系を使用してHLYAを生成し得る。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を分泌或いは細胞内への保持のいずれかに誘導し、安定した増殖のために外来配列の宿主ゲノムへの組込みを可能にする(前出のAusubel (1995)、Bitter, G. A. ら (1987) Methods Enzymol. 153: 516-544、Scorer, C. A. ら (1994) Bio/Technology 12: 181-184等を参照)。

【0138】

植物系を使用してHLYAを発現することも可能である。HLYA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(タカマツ, 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(前出のCoruzzi、前出のBroglie、前出のWinter等を参照)。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 McGra

w Hill Yearbook of Science and Technology) (1992) McGraw Hill New York NY, pp. 191 - 196 等を参照)。

【0139】

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。発現ベクターとしてアデノウイルスを用いる場合、HLYAをコードする配列は、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/翻訳複合体に連結反応され得る。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HLYAを発現する感染ウイルスを得ることが可能である(Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. 81: 3655 - 3659 等を参照)。

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。

【0140】

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。治療目的のために約6kb~10MbのHAC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(Harrington, J. J. ら(1997) Nat Genet. 15: 345 - 355. 等を参照)。

【0141】

長期にわたり哺乳動物系内で組換えタンパク質を生成するためには、細胞株内のHLYAの安定発現が望ましい。例えば、発現ベクターであって複製発現因子、内在性発現因子、或いはその両者のウイルス起源を含むものと、同一或いは別のベクター上の選択可能マーカー遺伝子とを用いて、HLYAをコードする配列を細胞株に形質転換することが可能である。ベクターの導入後、選択培地に移す前に強化培地で約1~2日間細胞を増殖させることができる。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。

【0142】

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(Wigler, M. ら(1977) Cell 11: 223 - 232、Lowy, I. ら(1980) Cell 22: 817 - 823 等を参照)。また、選択の基礎として代謝拮抗物質、抗生物質或いは除草剤への耐性を用いることができる。例えばdhfrはメトトレキセートに対する耐性を与え、neoはアミノグリコシッドネオマイシン及びG-418に対する耐性を与え、alsはクロルスルフロンに対する耐性を、patはホスフィノトリシンアセチルトランスフェラーゼに対する耐性を各々与える(Wigler, M. ら(1980) Proc. Natl. Acad. Sci. USA 77: 3567 - 3570、Colbere-Garapin, F. ら(1981) J. Mol. Biol. 150: 1 - 14 等を参照)。その他の選択可能遺伝子、例えば、代謝のための細胞要求を変えるtrpB及びhisDは、文献に記載されている(Hartman, S. C. and R. C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85: 8047 - 8051 等を参照)。可視マーカー、例えばアノトシアニン、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) Me

10

20

30

40

50

ethods Mol. Biol. 55:121-131等を参照)。

【0143】

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、HLYAをコードする配列がマーカー遺伝子配列内に挿入された場合、HLYAをコードする配列を含む形質転換細胞は、マーカー遺伝子機能の欠落により同定することが可能である。または単一プロモーター制御下で、HLYAをコードする配列とタンデムにマーカー遺伝子を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。

【0144】

通常、HLYAをコードする核酸配列を含み且つHLYA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。

【0145】

特異的ポリクローナルまたはモノクローナル抗体を用いてHLYA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。HLYA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部位モノクローナルベースのイムノアッセイ(two-site, monoclonal-based immunoassay)が好ましいが、競合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. 著(1990) Serological Methods, a Laboratory Manual, APS Press, St Paul, MN, Sect. IV, Coligan, J. E. 著(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)。

【0146】

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。HLYA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プローブを産出する方法には、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法がある。或いは、HLYAをコードする配列またはその任意の断片を、mRNAプローブを産出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。

【0147】

HLYAをコードするヌクレオチド配列を用いて形質転換した宿主細胞は、細胞培地からのタンパク質の回収及び発現に適した条件下で培養し得る。形質転換細胞から製造された

10

20

30

40

50

タンパク質が分泌されるか細胞内に留まるかは、使用される配列及び/またはベクターに依存する。当業者であれば理解し得るように、H L Y Aをコードするポリヌクレオチドを含む発現ベクターを、原核細胞膜または真核細胞膜を透過するH L Y Aの直接分泌を誘導するシグナル配列を含むように設計し得る。

【0148】

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(CHO、HeLa、MDCK、MEK293、WI38等)は、American Type Culture Collection(ATCC, Bethesda, 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。

【0149】

本発明の別の実施例では、H L Y Aをコードする天然の核酸配列、修飾核酸配列または組換え核酸配列を、上記任意の宿主系において融合タンパク質の翻訳をもたらす異種配列に連結反応させることができる。例えば、市販されている抗体を用いて認識可能な異種部分を含むキメラH L Y Aタンパク質は、H L Y A活性阻害剤に対するペプチドライブラリのスクリーニングを促進し得る。また、異種タンパク質部分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、融合タンパク質を遺伝子操作し、H L Y Aが精製後に異種部分から切断され得るように、H L Y Aコード配列と異種タンパク質配列の間にタンパク質分解性開裂部位を含めることもできる。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。

【0150】

本発明の更に別の実施例では、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、放射能標識したH L Y Aの合成が*in vitro*で可能である。これらの系は、T7、T3またはSP6プロモーターと機能的に結合したタンパク質コード配列の転写及び翻訳を結合する。翻訳は、例えば³⁵Sメチオニンのような放射能標識したアミノ酸前駆体の存在下で起こる。

【0151】

本発明のH L Y Aまたはその断片を用いて、H L Y Aに特異結合する化合物をスクリーニングし得る。少なくとも1個から複数個の試験化合物を用いて、H L Y Aへの特異結合をスクリーニングし得る。試験化合物の例としては、抗体、オリゴヌクレオチド、タンパク質(例えば受容体)または小分子が挙げられる。

【0152】

一実施例では、このように同定された化合物は、例えばリガンドまたはその断片などのH L Y Aの天然リガンド、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.ら(1991) Curren

10

20

30

40

50

t Protocols in Immunology 1 (2) の5章等を参照)。同様にして化合物は、HL Y A が結合する天然受容体に関連し得るか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断片に密接に関連し得る。いずれの場合にも、化合物は既知の技術を用いて合理的にデザインし得る。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は、分泌タンパク質としてまたは細胞膜上のいずれかでHL Y A を発現する好適な細胞の生成に関与している。好適な細胞には、哺乳動物、酵母、シヨウジョウバエまたは大腸菌からの細胞がある。HL Y A を発現する細胞またはHL Y A を含有する細胞膜断片を試験化合物と接触させ、HL Y A または化合物のいずれかの結合、刺激または阻害を分析する。

【0153】

10

アッセイは、ポリペプチドへの試験化合物の結合を単純に試験し得る。結合は、フルオロフォア、放射性同位体、酵素抱合体またはその他の検出可能な標識により検出する。例えば、アッセイは少なくとも1つの試験化合物を溶液中でHL Y A と結合するか固体支持体に固定するかのいずれかのステップ及びHL Y A の化合物への結合を検出するステップを有し得る。或いはアッセイは、標識された競争相手の存在下で試験化合物の結合を検出または測定し得る。更にアッセイは、細胞遊離製剤、化学ライブラリまたは天然の生成混合物を用いて実行することができ、試験化合物は、溶液中で遊離させるか固体支持体に固定し得る。

【0154】

20

本発明のHL Y A またはその断片を用いて、HL Y A の活性を調整する化合物をスクリーニングし得る。このような化合物には、アゴニスト、アンタゴニスト、或いは部分的または逆アゴニスト等がある。一実施例においては、HL Y A が少なくとも1つの試験化合物と結合しているような、HL Y A の活性を許容する条件下でアッセイが実行され、試験化合物存在下でのHL Y A の活性が試験化合物不存在下でのHL Y A の活性と比較される。試験化合物存在下でのHL Y A の活性の変化は、HL Y A の活性を調整する化合物を示す。或いは、試験化合物はHL Y A の活性に適した条件下で活性に適した条件下でHL Y A を含む *in vitro* または細胞遊離系と結合し、アッセイが実行される。これらアッセイのいずれかにおいて、HL Y A の活性を調整する試験化合物は間接的にそのようにすることができ、試験化合物と直接接触する必要がなくなる。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。

30

【0155】

別の実施例では、HL Y A またはその哺乳類同族体をコードするポリヌクレオチドは、胚幹 (ES) 細胞において相同的組換えを用いて動物モデル系内で「ロックアウト」される。このような技術は当技術分野において公知であり、ヒト疾病の動物モデルの生成に有用である (米国特許第5, 175, 383号及び同第5, 767, 337号等を参照)。例えば129/SvJ株化細胞等のマウスES細胞は、初期のマウス胎仔に由来し、培養液中で成長する。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等のマーカー遺伝子により分裂させた対象遺伝子 (gene of interest) を含むベクターを用いて形質転換される (neo: Capecchi, M. R. (1989) Science 244: 1288-1292)。ベクターは、相同的組換えにより宿主ゲノムの対応する領域に統合される。或いは、組織特異的または発達段階特異的な様式で対象遺伝子をロックアウトするCre-LoxP系を用いて相同的組換えが発生する (Marth, J. D. (1996) Clin. Invest. 97: 1999-2002; Wagner, K. U. (1997) Nucleic Acids Res. 25: 4323-4330)。形質転換されたES細胞を同定し、例えばC57BL/6マウス系統から採取したマウス細胞胚盤胞に微量注入する。胚盤胞を偽妊娠種雌に外科的に導入し、結果として得られるキメラ子孫の遺伝形質を決め、これを繁殖させてヘテロ接合性系統またはホモ接合性系統を生成する。このようにして産出した遺伝子導入動物は、潜在的治療薬または毒性薬剤を用いて試験し得る。

40

【0156】

50

H L Y A をコードするポリヌクレオチドは、ヒト胚盤胞由来の E S 細胞における i n v i t r o でも操作し得る。ヒト E S 細胞は、内胚葉、中胚葉及び外胚葉の細胞タイプを含む少なくとも 8 つの別々の細胞系統に分化する可能性を有する。この細胞系統は、例えば神経細胞、造血系統及び心筋細胞に分化する (T h o m s o n , J . A . ら (1 9 9 8) S c i e n c e 2 8 2 : 1 1 4 5 - 1 1 4 7) 。

【 0 1 5 7 】

H L Y A をコードするポリヌクレオチドは、モデルヒト疾病への「ノックイン」ヒト化動物 (ブタ) または遺伝子導入動物 (マウスまたはラット) も生成し得る。ノックイン技術を用いて、H L Y A をコードするポリヌクレオチドの或る領域を動物 E S 細胞に注入し、注入された配列は動物細胞ゲノムに統合する。形質転換された細胞を胞胚に注入し、胞胚を上記のように移植する。ヒトの疾病の治療に関する情報を得るために、遺伝子導入子孫または近交系について研究し、強力な医薬品を用いて遺伝子導入子孫または近交系を処理する。或いは、H L Y A を過剰発現させるべく例えば H L Y A を乳内に分泌するなどして同系交配させた哺乳動物は、タンパク質の簡便な源としても役立ち得る (J a n n e , J . ら (1 9 9 8) B i o t e c h n o l . A n n u . R e v . 4 : 5 5 - 7 4) 。

10

【 0 1 5 8 】

治療

化学的及び構造的類似性、例えば配列及びモチーフとの関連における類似性が、H L Y A の領域と細胞骨格関連タンパク質間に存在する。更に H L Y A の発現は、細胞増殖、癌及び炎症に密接に関連している。従って H L Y A は、癌、自己免疫 / 炎症性疾患、細胞増殖異常及び眼病において或る役割を果たすものと考えられる。H L Y A の発現または活性の増大に関連する疾患の治療においては、H L Y A の発現または活性を低下させることが望ましい。また、H L Y A 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、H L Y A の発現または活性を増大させることが望ましい。

20

【 0 1 5 9 】

従って、或る実施例において、H L Y A 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 H L Y A またはその断片や誘導体を投与することが可能である。限定するものではないがこのような疾患の例として癌が含まれ、その中には腺癌、白血病、リンパ腫、黒色腫、骨髄腫、肉腫、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、子宮の癌等が含まれ、また自己免疫 / 炎症疾患も含まれ、その中には後天性免疫不全症候群 (A I D S) 、アジソン病、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、自己免疫多発性内分泌腺障害症 (A P E C E D) 、気管支炎、胆嚢炎、接触皮膚炎、クローン病、アトピー性皮膚炎、皮膚筋炎、糖尿病、肺気腫、リンパ球毒素性一時性リンパ球減少症、赤芽球症、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、過好酸球増加症、過敏性大腸症候群、多発性硬化症、重症筋無力症、心筋または心膜の炎症、骨関節炎、骨粗しょう症、膵炎、多発性筋炎、乾癬、ライター症候群、リウマチ様関節炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性紅斑性狼瘡、全身性硬化症、血小板減少症、潰瘍性大腸炎、ブドウ膜炎、ウェルナー症候群、癌の合併症、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、蠕虫の感染症及び外傷が含まれ、また神経障害も含まれ、その中には癲癇、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、パーキンソン病その他の錐体外路障害、筋萎縮性側索硬化症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、多発性硬化症その他の脱髄疾患、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と、クールー、クロイツフェルト ヤコブ病及びガストマン ストラウスラー シャ

30

40

50

インカー症候群を含むプリオン病と、致死性家族性不眠症、神経系の栄養病及び代謝病、神経線維腫症、結節硬化症、小脳網膜性血管芽腫症 (cerebelloretinal hemangioblastomatosis)、脳3叉神経血管症候群、精神薄弱及びダウン症候群を含むその他の中枢神経系発達障害、脳性麻痺、神経骨格異常、自律神経系障害、脳神経障害、脊髄病、筋ジストロフィーその他の神経筋疾患、末梢神経疾患、皮膚筋炎及び多発性筋炎と、遺伝性、代謝性、内分泌性及び中毒性ミオパシーと、重症筋無力症、周期性四肢麻痺と、気分障害、不安障害及び精神分裂病を含む精神障害と、季節型感情障害 (SAD) と、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、トゥレット病、進行性核上麻痺、皮質基底核変性症 (corticobasal degeneration)、家族性前頭側頭骨痴呆が含まれ、また眼病も含まれ、その中には高眼圧症及び緑内障が含まれる。

10

【0160】

別の実施例では、HLYA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HLYA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。

【0161】

更に別の実施例では、実質的に精製されたHLYAを含む成分を好適な医薬用担体と共に患者に投与して、限定するものではないが上記した疾患を含むHLYA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。

20

【0162】

更に別の実施例では、HLYAの活性を調節するアゴニストを患者に投与して、限定するものではないが上記した疾患を含むHLYA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。

【0163】

更に別の実施例では、患者にHLYAのアンタゴニストを投与して、HLYAの発現または活性の増大に関連した疾患を治療または予防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、上記した癌、自己免疫/炎症性疾患、細胞増殖異常及び眼病がある。一実施態様においては、アンタゴニストとして直接的に、或いはHLYAを発現する細胞または組織に薬剤を輸送するターゲティングまたは輸送機構として間接的にHLYAと特異結合する抗体を用いることができる。

30

【0164】

別の実施例では、HLYAをコードするポリヌクレオチドの相補体を発現す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HLYAの発現または活性の増大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。

【0165】

別の実施例では、本発明の任意のタンパク質、アンタゴニスト、抗体、アゴニスト、相補配列、またはベクターを、別の好適な治療薬と組み合わせて投与することもできる。併用療法で用いる好適な治療薬は、当業者が従来 of 医薬原理に従ってを選択し得る。治療薬と組み合わせることにより、上記した種々の疾患の治療または予防に相乗効果をもたらし得る。このアプローチ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。

40

【0166】

HLYAのアンタゴニストは、本技術分野で一般的に知られている方法を用いて製造し得る。具体的には、精製されたHLYAを用いて抗体を作るか、治療薬のライブラリをスクリーニングして、HLYAと特異結合するものを同定することが可能である。HLYAの抗体も、本技術分野で一般的に知られている方法を用いて製造することが可能である。限定するものではないがこのような抗体には、ポリクローナル抗体、モノクローナル抗体、キメラ抗体、一本鎖抗体、Fab断片及びFab発現ライブラリによって作られた断片が含まれ得る。中和抗体 (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

50

。

【0167】

抗体を産生するために、HL Y A、またはHL Y Aの任意の断片またはオリゴペプチドであって免疫抗原性の特性を有するものを注入することによって、ヤギ、ウサギ、ラット、マウス、ヒト、その他を含む種々の宿主を免疫化し得る。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、プルロニックポリオール、ポリアニオン、ペプチド、油性乳剤、キーホールリンペットヘモシニアン、ジニトロフェノール等の界面活性剤とがある。ヒトに用いられるアジュバントの中では、BCG(カルメット ゲラン桿菌)及びコリネバクテリウム パルヴムが特に好ましい。

10

【0168】

HL Y Aに対する抗体を誘導するために用いるオリゴペプチド、ペプチドまたは断片は、少なくとも約5のアミノ酸から構成され、通常少なくとも約10のアミノ酸から構成されることになるアミノ酸配列を有することが好ましい。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。HL Y Aアミノ酸の短い伸長部は、別のタンパク質、例えばキーホールリンペットヘモシニアンの配列と融合し、キメラ分子に対する抗体が産生され得る。

【0169】

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

20

【0170】

更に、「キメラ抗体」を作製するために開発した技術、例えば好適な抗原特異性及び生物学的活性を有する分子を得るためのマウス抗体遺伝子のヒト抗体遺伝子へのスプライシングを用いることが可能である(Morrison, S. L. ら (1984) Proc. Natl. Acad. Sci. USA 81: 6851-6855、Neuberger, M. S. ら (1984) Nature 312: 604-608、タケダ, S. ら (1985) Nature 314: 452-454等を参照)。或いは、一本鎖抗体を産生するために説明された技術を適用し、本技術分野で知られている方法を用いて、HL Y A特異性一本鎖抗体を産生し得る。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

30

40

【0171】

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

【0172】

HL Y Aのための特異結合部位を有する抗体を産生することもできる。例えば、限定する

50

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 $F(a b')_2$ 断片と、 $F(a b')_2$ 断片のジスルフィド架橋を減らすことによって作製される $F a b$ 断片とがある。或いは、 $F a b$ 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 $F a b$ 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 (H u s e , W . D . ら (1 9 8 9) *Science* 256 : 1275 - 1281 等を参照)。

【0173】

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、H L Y A 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。2つの非干渉性H L Y A エピトープに反応するモノクローナル抗体を用いるような、2部位モノクローナルベースのイムノアッセイが一般に利用されるが、競合結合アッセイを利用してよい (前出の P o u n d の文献)。

【0174】

ラジオイムノアッセイ技術と共に様々な方法、例えばスキッチャード分析を用いて、H L Y A に対する抗体の親和性を評価する。親和性は結合定数 $K a$ で表す。 $K a$ は、平衡状態においてH L Y A 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た値であると定義する。ポリクローナル抗体は多様なH L Y A エピトープに対する親和性が不均一であり、ポリクローナル抗体試薬のために決定した $K a$ は、H L Y A 抗体の平均親和性または結合活性を表す。モノクローナル抗体は特定のH L Y A エピトープに対して単一特異的であり、モノクローナル抗体試薬のために決定した $K a$ は、親和性の真の測定値を表す。 $K a$ 値が約 $10^9 \sim 10^{12}$ L / m o l の範囲にあるような高親和性抗体試薬は、H L Y A 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。 $K a$ 値が約 $10^6 \sim 10^7$ L / m o l の範囲にあるような低親和性抗体試薬は、H L Y A 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及び類似の処理に用いるのが好ましい (C a t t y , D . (1 9 8 8) *Antibodies, Volume I: A Practical Approach*, I R L P r e s s , W a s h i n g t o n , D C ; L i d d e l l , J . E . a n d C r y e r , A . (1 9 9 1) *A Practical Guide to Monoclonal Antibodies*, J o h n W i l e y & S o n s , N e w Y o r k N Y) 。

【0175】

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、或る下流の適用例に対するこのような試薬の品質及び適性を決定することができる。例えば、少なくとも $1 \sim 2$ m g / m l 、好ましくは $5 \sim 10$ m g / m l の特異抗体を含むポリクローナル抗体試薬は、H L Y A 抗体複合体を沈殿させる必要がある処理において通常用いられる。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 C a t t y の文献、同 C o l i g a n らの文献等を参照)。

【0176】

本発明の別の実施例では、H L Y A をコードするポリヌクレオチド、H L Y A の任意の断片または相補配列を治療目的で使用することができる。ある実施形態では、H L Y A をコードする遺伝子のコード化または制御領域に相補的な配列またはアンチセンス分子 (D N A 、 R N A 、 P N A または修飾されたオリゴヌクレオチド) によって遺伝子発現の修飾を達成することができる。そのような技術は等分野では既知であり、アンチセンスオリゴヌクレオチドまたはより大きな断片を、H L Y A をコードする配列のコード領域または制御領域に延在する様々な位置からデザインすることが可能である (A g r a w a l , S . , e d . (1 9 9 6) *Antisense Therapeutics*, H u m a n a P r e s s I n c . , T o t a w a N J 等を参照)。

10

20

30

40

50

【0177】

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. ら (1998) *J. Allergy Clin. Immunol.* 102(3): 469-475、Scanlon, K. J. ら (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)。その他の遺伝輸送機構には、リポソーム系、人工的なウイルスエンベロプ及び当分野で公知のその他の系が含まれる (Rossi, J. J. (1995) *Br. Med. Bull.* 51(1): 217-225; Boado, R. J. ら (1998) *J. Pharm. Sci.* 87(11): 1308-1315、Morris, M. C. ら (1997) *Nucleic Acids Res.* 25(14): 2730-2736. 等を参照)。

【0178】

本発明の別の実施例では、HLYA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遺伝子治療に用いることが可能である。遺伝子治療を行うことにより、(i) 遺伝子欠損症 (例えばX染色体鎖遺伝 (Cavazzana-Calvo, M. ら (2000) *Science* 288: 669-672) により特徴付けられる重度の複合型免疫欠損 (SCID) - X1の場合)、先天性アデノシンデアミナーゼ (ADA) 欠損症に関連する重度の複合型免疫欠損 (Blaese, R. M. ら (1995) *Science* 270: 475-480、Bordignon, C. ら (1995) *Science* 270: 470-475)、嚢胞性繊維症 (Zabner, J. ら (1993) *Cell* 75: 207-216; Crystal, R. G. ら (1995) *Hum. Gene Therapy* 6: 643-666、Crystal, R. G. ら (1995) *Hum. Gene Therapy* 6: 667-703)、サラセミア (thalassaemia)、家族性高コレステロール血症、第VII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. and Somia, N. (1997) *Nature* 389: 239-242) を治療し、(ii) 条件的致死性遺伝子産物を発現させ (例えば制御不能な細胞増殖に起因する癌の場合)、(iii) 細胞内の寄生虫 (例えばヒト免疫不全ウイルス (HIV) (Baltimore, D. (1988) *Nature* 335: 395-396、Poeschl, E. ら (1996) *Proc. Natl. Acad. Sci. USA.* 93: 11395-11399)、B型若しくはC型肝炎ウイルス (HBV、HCV)、カンジダアルビカンス及びパラコクシジオイデス、ブラジリエンシス等の真菌寄生虫、並びに熱帯熱マラリア原虫及びトリパノソーマ、クルージ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。HLYAの発現若しくは調節に必要な遺伝子の欠損が疾患を発生させる場合、形質導入した細胞の好適な集団からHLYAを発現することにより、遺伝子欠損に起因する症状の発現を緩和し得る。

【0179】

本発明の更なる実施例では、HLYAをコードする哺乳動物発現ベクターを作製し、これらのベクターを機械的手段によってHLYA欠損細胞に導入することによって、HLYAの欠損による疾患や異常症を治療する。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) 弾道的金粒子の打ち込み (ballistic gold particle delivery)、(iii) リポソーム仲介形質移入、(iv) 受容体仲介遺伝子導入、及び (

v) DNAトランスポソンの使用 (Morgan, R. A. and W. F. Anderson (1993) *Annu. Rev. Biochem.* 62: 191 - 217、Ivics, Z. (1997) *Cell* 91: 501 - 510; Boulay, J - L. and H. Recipon (1998) *Curr. Opin. Biotechnol.* 9: 445 - 450) がある。

【0180】

HLYAの発現に影響を及ぼし得る発現ベクターには、限定するものではないが、PCDNA3.1、EPITAG、PRCCMV2、PREP、PVAXベクター (Invitrogen, Carlsbad CA)、PCMV-SCRIPT、PCMV-TAG、PEGSH/PERV (Stratagene, La Jolla CA) 及びPTE-T-OFF、PTET-ON、PTRE2、PTRE2-LUC、PTK-HYG (Clontech, Palo Alto CA) がある。HLYAは、(i) 恒常的に活性なプロモーター (例えばサイトメガロウイルス (CMV)、ラウス肉腫ウイルス (RSV)、SV40ウイルス、チミジンキナーゼ (TK) または アクチン遺伝子)、(ii) 誘導性プロモーター (例えばテトラサイクリン調節性プロモーター (Gossen, M. and H. Bujard (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89: 5547 - 5551、Gossen, M. ら (1995) *Science* 268: 1766 - 1769、Rossi, F. M. V. and H. M. Blau (1998) *Curr. Opin. Biotechnol.* 9: 451 - 456)、市販のInvitrogen社のT-REXプラスミドに含まれる)、エクジソン誘導性プロモーター (Invitrogen社のプラスミドPVGRXR及びPIN-Dから得られる)、FK506/ラパマイシン誘導性プロモーターまたはRU486/ミフェプリストーン誘導性プロモーター (前出のRossi, F. M. V. and H. M. Blau)、または(iii) 正常個体由来の、HLYA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て、発現させることができる。

【0181】

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. and A. J. Eb (1973) *Virology* 52: 456 - 467) 若しくは電気穿孔法 (Neumann, B. ら (1982) *EMBO J.* 1: 841 - 845) を用いて形質転換を行う。初代細胞に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。

【0182】

本発明の別の実施例では、HLYA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) プロモーターまたは独立プロモーターの制御下でHLYA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ド配列を伴うRev応答性エレメント (RRE) とからなるレトロウイルスベクターを作製して治療することができる。レトロウイルスベクター (例えばPFB及びPFBNEO) はStratagene社から市販されており、刊行データ (Riviere, I. ら (1995) *Proc. Natl. Acad. Sci. U.S.A.* 92: 6733 - 6737) に基づいている。上記データを引用することをもって本明細書の一部とする。ベクターは、好適なベクター産生細胞系 (VPCCL) において増殖され、VPCCLは、標的細胞上の受容体に対する向性を有するエンペロープ遺伝子またはVSVg等の乱交雑エンペ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. and A. D. Miller (1988) *J. Virol.* 62: 3802 - 3806、Dull, T. ら (1998) *J. Virol.* 72: 8463 - 8471、Zufferey, R. ら (1998) *J. Virol.* 72: 9873 - 9880)。Riggに付与された米国特許第5,910,434号("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant")は、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について開示しており、これを引用することをもって本明細書の一部とする。レトロウイルスベクターの増殖、細胞集団(例えばCD4⁺T細胞)の形質導入、及び形質導入した細胞の患者への戻しは、遺伝子治療の分野では当業者に公知の方法であり、多数の文献に記載されている(Ranga, U. ら (1997) *J. Virol.* 71: 7020 - 7029、Bauer, G. ら (1997) *Blood* 89: 2259 - 2267、Bonyhadi, M. L. (1997) *J. Virol.* 71: 4707 - 4716、Ranga, U. ら (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95: 1201 - 1206、Su, L. (1997) *Blood* 89: 2283 - 2290)。

【0183】

別法では、アデノウイルス系遺伝子治療の輸送系を用いて、HLYAの発現に関連する1若しくは複数の遺伝子異常を有するような細胞にHLYAをコードするポリヌクレオチドを輸送する。アデノウイルス系ベクターの作製及びパッケージングについては、当業者に公知である。複製欠損型アデノウイルスベクターは、免疫調節タンパク質をコードする遺伝子を臍臓の無損傷の臍島内に導入するために可変性であることが証明された(Csete, M. E. ら (1995) *Transplantation* 27: 263 - 268)。使用できる可能性のあるアデノウイルスベクターは、Armentanoに付与された米国特許第5,707,618号("Adenovirus vectors for gene therapy")に記載されており、引用することをもって本明細書の一部とする。アデノウイルスベクターについては、Antinozzi, P. A. ら (1999) *Annu. Rev. Nutr.* 19: 511 - 544 及び Verma, I. M. and N. Somia (1997) *Nature* 18: 389: 239 - 242も参照されたい。両文献は、引用を以って本明細書の一部となす。

【0184】

別法では、ヘルペス系遺伝子治療の輸送系を用いて、HLYAの発現に関連する1若しくは複数の遺伝子異常を有する標的細胞にHLYAをコードするポリヌクレオチドを輸送する。HSVが向性を有するような中枢神経系の細胞にHLYAを導入する際には、単純ヘルペスウイルス(HSV)系のベクターの使用は特に役立つ。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輸送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つの内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継代

、ヘルペスウイルスの成長及び増殖、並びにヘルペスウイルスの細胞への感染は、当業者に公知の技術である。

【0185】

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HLYA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（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. and K. - J. Li (1998) *Cun. Opin. Biotech.* 9: 464 - 469）。ウイルスRNAの複製中に、通常はウイルスカプシドタンパク質をコードするサブゲノム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カプシドタンパク質が過剰産生される。同様に、HLYAに対するコード配列をカプシドコード領域のウイルスゲノムに導入することにより、ベクター導入細胞において多数のHLYAコードRNAが産生され、高レベルのHLYA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）。ウイルスの宿主の範囲が広いことにより、様々な細胞タイプへのHLYAの導入が可能になる。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。

10

20

30

40

【0186】

例えば開始部位から数えて約 - 10 と約 + 10 の間に位置する転写開始部位由来のオリゴヌクレオチドを用いて、遺伝子発現を阻害し得る。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. and B. I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, NY, pp. 163 - 177*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の翻訳を阻止するべくデザインすることができる。

【0187】

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、ヌクレオチド鎖切断に先立つ相補的標的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に参与している。例えば、組換え型のハンマーヘッド型リボザイム分子は、HLYAをコードする配列のヌクレオチド鎖切断を特異的且つ効果的に触媒する。

【0188】

任意の潜在的RNA標的内の特異的リボザイム切断部位は、GUA、GUU、GUC配列を含めたり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される。一度同定されると、オリゴヌクレオチドを機能不全にするような2次構造の特徴に対して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配列を、評価することが可能になる。候補標的の適合性の評価も、リボヌクレアーゼ保護アッセイを用いて相補的オリゴヌクレオチドとのハイブリダイゼーションの実施容易性をテストすることによって行うことができる。

【0189】

50

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、固相ホスホラ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、HRIP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。

【0190】

RNA分子を修飾することで、細胞内の安定性を高め、半減期を長くし得る。限定するものではないが可能な修飾には、分子の5'末端及び/または3'末端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リチオネートまたは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に加えて、非従来型塩基、例えばイノシン、クエオシン(queosine)、ワイブトシン(wybutosine)等を包含することによる。

【0191】

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

【0192】

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。HLYA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され、このように得られる。サンプルには例えば、無傷細胞、透過処理した細胞、細胞遊離系または再構成生化学系があり得る。HLYAをコードするポリヌクレオチドの発現における変化は、当分野で通常知られている任意の方法でアッセイする。通常、HLYA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、特異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若しくは複数の試験化合物に曝露される及び曝露されないポリヌクレオチドの発現の比較に対する基礎を形成し得る。試験化合物に曝露されるポリヌクレオチドの発現における変化の検出は、ポリヌクレオチドの発現を変異する際に試験化合物が有効であることを示している。特異ポリヌクレオチドの変異発現に有効な化合物に対して、例えば分裂酵母(*Schizosaccharomyces pombe*)遺伝子発現系(Atkins, D.ら(1999)米国特許第5,932,435号、Arn

10

20

30

40

50

dt, 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) の組合せライブラリをスクリーニングすることに関与している (Bruice, T. W. ら (1997) 米国特許第 5,686,242 号、Bruice, T. W. ら (2000) 米国特許第 6,022,691 号)。

【0193】

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肝細胞内に導入し、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輸送は、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)。

【0194】

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。

【0195】

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、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, Easton PA) の最新版に記載されている。このような成分は、HLYA、HLYA に対する抗体、擬態、アゴニスト、アンタゴニスト、または HLYA インヒビターから構成し得る。

【0196】

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。

【0197】

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重量有機薬) の場合には、速効製剤のエアロゾル輸送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胞領域を介しての肺輸送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. らの米国特許第 5,997,848 号等を参照)。肺輸送は、針注射なしに投与する点で優れており、潜在的に有毒な浸透エンハンサーの必要性をなくす。

【0198】

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。

【0199】

特殊形状の成分は、HLYA またはその断片を含む高分子を直接細胞内輸送するために調製される。例えば、細胞不透過性高分子を含むリボソーム製剤は、細胞融合及び高分子の細胞内輸送を促進し得る。或いは、HLYA またはその断片を HIV Tat-1 タンパク質から陽イオン 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

10

20

30

40

50

- 1572)。

【0200】

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。

【0201】

治療上の有効投与量は、症状や容態を回復させるような活性成分量を参考にする。そのような活性成分の例としては、HLYAまたはその断片、HLYAの抗体、HLYAのアゴニスト、アンタゴニストまたはインヒビターがある。薬用有効度及び毒性は、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%の医薬的有效量）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。毒性効果の治療効果に対する投与量の比は、治療指数であり、LD₅₀/ED₅₀比として表すことができる。高い治療指数を示すような成分が望ましい。細胞培養アッセイ及び動物実験から得られたデータは、ヒトに用いるための投与量の範囲を調剤するのに用いられる。このような組成物が含まれる投与量は、毒性を殆ど或いは全く含まず、ED₅₀を含むような血中濃度の範囲にあることが好ましい。用いられる投与形態、患者の感受性及び投与の経路によって、投与量はこの範囲内で様々に変わる。

10

【0202】

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

20

【0203】

通常、投与量は、投与の経路にもよるが約0.1～100,000μgであり、合計で約1gまでとする。特定の投与量及び輸送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輸送は、特定の細胞、状態、位置等に特異的なものとなる。

30

【0204】

診断

別の実施例では、HLYA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HLYAやHLYA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、HLYA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。HLYAの診断アッセイには、抗体及び標識を利用してヒトの体液において或いは細胞や組織のエキスにおいてHLYAを検出する方法が含まれる。抗体は、修飾して或いは修飾しないで使用し、レポーター分子の共有結合性或いは非共有結合性の接着によって標識化し得る。多様なレポーター分子が本技術分野で知られており、それらを用いることができる。幾つかのレポーター分子については上記した。

40

【0205】

HLYAを測定するための様々なプロトコル、例えばELISA、RIA、FACS等が本技術分野において知られており、HLYA発現の修正レベル或いは異常レベルを診断する基準を提供する。複合体の形成に適した条件下でヒト対象等の正常な哺乳動物対象から採取した体液または細胞とHLYAに対する抗体とを結合させることにより、HLYA発

50

現の正常値または標準値が決定される。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。対象内で発現したHL YAの量、制御、検体からの病変サンプルを標準値と比較する。標準値と対象との偏差が疾患を診断するパラメータとなる。

【0206】

別の実施例によれば、HL YA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。用いられ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。ポリヌクレオチドは、検体におけるHL YAの発現が疾患と相関し得るような該検体における遺伝子発現の検出及び定量に用いることができる。診断アッセイは、HL YAの不在、存在及び過剰発現を測定するために、そして治療インターベンション中にHL YAレベルの調製をモニターするために用いることができる。

10

【0207】

ある実施形態では、HL YAをコードする核酸配列を同定するために、HL YAまたは近縁の分子をコードする、ゲノム配列を含むポリヌクレオチド配列を検出可能なPCRプローブとのハイブリダイゼーションを用いることができる。プローブが、5'調節領域のような高特異領域を有するにせよ、保存されたモチーフのような低特異領域を有するにせよ、GC REC、突然変異体または関連配列をコードする天然の配列しか同定しないのかどうかは、プローブの特異性及びハイブリダイゼーション或いは増幅のストリンジェンシーが決定することになる。

【0208】

プローブはまた、関連する配列の検出にも用いることができ、その配列はHL YAをコードする任意の配列と少なくとも50%の相同性を有し得る。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、配列番号56乃至110の配列、或いはHL YA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。

20

【0209】

HL YAをコードするDNAに対する特異的ハイブリダイゼーションプローブを作製する方法には、HL YAまたはHL YA誘導体をコードするポリヌクレオチド配列を、mRNAプローブを作製するためのベクターにクローニングする方法が含まれる。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。

30

【0210】

HL YAをコードするポリヌクレオチド配列は、HL YAの発現に関係する疾患の診断の為に用い得る。限定するものではないがこのような疾患の例として癌が含まれ、その中には腺癌、白血病、リンパ腫、黒色腫、骨髄腫、肉腫、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、子宮の癌等が含まれ、また自己免疫/炎症疾患も含まれ、その中には後天性免疫不全症候群(AIDS)、アジソン病、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、自己免疫多発性内分泌腺障害症(APECED)、気管支炎、胆嚢炎、接触皮膚炎、クローン病、アトピー性皮膚炎、皮膚筋炎、糖尿病、肺気腫、リンパ球毒素性一時的リンパ球減少症、赤芽球症、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、過好酸球増加症、過敏性大腸症候群、多発性硬化症、重症筋無力症、心筋または心膜の炎症、骨関節炎、骨粗しょう症、膵炎、多発性筋炎、乾癬、ライター症候群、リウマチ様関節炎、強皮症、シェーグレン症候群、全身性アナ

40

50

フィラキシー、全身性紅斑性狼瘡、全身性硬化症、血小板減少症、潰瘍性大腸炎、ブドウ膜炎、ウェルナー症候群、癌の合併症、血液透析、体外循環、ウイルス感染症、細菌性感染症、真菌性感染症、寄生虫感染症、原虫感染症、蠕虫の感染症及び外傷が含まれ、また神経障害も含まれ、その中には癲癇、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、パーキンソン病その他の錐体外路障害、筋萎縮性側索硬化症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、多発性硬化症その他の脱髄疾患、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と、クールー、クロイツフェルト ヤコブ病及びガストマン ストラウスラー シャインカー症候群を含むプリオン病と、致死性家族性不眠症、神経系の栄養病及び代謝病、神経線維腫症、結節硬化症、小脳網膜性血管芽腫症 (cerebellor etinal hemangioblastomatosis)、脳3叉神経血管症候群、精神薄弱及びダウン症候群を含むその他の中枢神経系発達障害、脳性麻痺、神経骨格異常、自律神経系障害、脳神経障害、脊髄病、筋ジストロフィーその他の神経筋疾患、末梢神経疾患、皮膚筋炎及び多発性筋炎と、遺伝性、代謝性、内分泌性及び中毒性ミオパシーと、重症筋無力症、周期性四肢麻痺と、気分障害、不安障害及び精神分裂病を含む精神障害と、季節型感情障害 (SAD) と、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、トゥレット病、進行性核上麻痺、皮質基底核変性症 (corticobasal degeneration)、家族性前頭側頭骨痴呆が含まれ、また眼病も含まれ、その中には高眼圧症及び緑内障が含まれる。HLYAをコードするポリヌクレオチド配列は、サザン法、ノーザン法、ドットプロット法やその他の膜ベースの技術と、PCR法と、ディップスティック (dipstick) 法、ピン及びマルチフォーマットELISA様アッセイと、変異HLYAの発現を検出するために患者から採取した体液または組織を利用するマイクロアレイとにおいて使用し得る。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。

【0211】

或る形態では、関連する疾患、特に上記した疾患を検出するアッセイにおいて、HLYAをコードするヌクレオチド配列が有用であり得る。HLYAをコードするヌクレオチド配列は標準法で標識化され、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または組織のサンプルに添加することができる。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。患者サンプルのシグナル量が制御サンプルと比べて著しく変化している場合は、サンプル内のHLYAをコードするヌクレオチド配列の変異レベルは関連する疾患の存在を示している。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。

【0212】

HLYAの発現に関連する疾患の診断基準を提供するために、発現のための正常あるいは標準概要を確立する。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下で、動物或いはヒトの正常な対象から抽出した体液或いは細胞を、HLYAをコードする配列またはその断片と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証明する。

【0213】

疾患の存在が証明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。

【0214】

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、実際に臨床的症狀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。

【0215】

H L Y A をコードする配列から設計されたオリゴヌクレオチドを追加的に診断上利用することは、P C R の利用に關与し得る。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 L Y A をコードするポリヌクレオチドの断片、或いは H L Y A をコードするポリヌクレオチドと相補的ポリヌクレオチドの断片を含み、最適条件下で特定の遺伝子や条件を識別するべく利用される。また、オリゴマーは、やや緩いストリンジェント条件下で、密接に關連している D N A 或いは R N A 配列の検出または定量に用いることが可能である。

【0216】

或る態様において、H L Y A 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、制限酵素切断法 (S S C P) 及び蛍光 S S C P (f S S C P) がある。S S C P では、H L Y A 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、ポリメラーゼ連鎖反応法 (P C R) を用いた D N A の増幅を行う。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 シークエンシング機などのハイスループット機器でアンプリマー (a m p l i m e r) の検出が可能になる。更に、インシリコ S N P (i n s i l i c o 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、多形性を同定し得る。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クエ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、例えばハイスループット M A S S A R R A Y システム (S e q u e n o m , I n c . , S a n D i e g o C A) を用いた質量分析により S N P を検出し、特徴付ける。

【0217】

H L Y A 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 (c o a m p l i f i c a t i o n) 及び標準曲線から得た結果の補間もある (M e l b y , P . C . ら (1 9 9 3) J . I m m u n o l . M e t h o d s , 1 5 9 : 2 3 5 - 2 4 4 , D u p l a a , C . ら (1 9 9 3) A n a l . B i o c h e m . 2 1 2 : 2 2 9 - 2 3 6 等を参照)。目的のオリゴマーが種々の希釈液中に存在し、分光光度法または非色応答によって定量が迅速になるようなハイスループット型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。

【0218】

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標的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、突然変異及び多形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

10

20

30

40

50

機能としての疾病の進行/後退をモニターし、疾病治療における薬剤の活性を開発及びモニターすることができる。特に、患者にとって最もふさわしく、有効的な治療法を選択するために、この情報を用いて患者の薬理ゲノムプロファイルを開発することができる。例えば、患者の薬理ゲノムプロファイルに基づき、患者に対して高度に有効的で副作用を殆ど示さない治療薬を選択することができる。

【0219】

別の実施例では、HLYAに特異的な抗体、HLYAまたはその断片をマイクロアレイ上で要素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間相互作用、薬剤-標的相互作用及び遺伝子発現プロファイルをモニターまたは測定することが可能である。

10

【0220】

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、所与の条件下で所与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し得る (Seilliamerらの米国特許第5,840,484号 "Comparative Gene Transcript Analysis" を参照。該特許は特に引用することを以って本明細書の一部となす)。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要素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。

20

【0221】

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。

【0222】

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示す、しばしば分子指紋または毒性サイン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. and 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サインと同一のサインを有する場合には、毒性特性を共有している可能性がある。指紋またはサインは、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サインを提供する。自己の発現が任意の試験された化合物により変化しない遺伝子が同様に重要であっても、このような遺伝子の発現レベルを用いて残りの発現データを規準化する。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。毒性サインの要素への遺伝子機能を割り当てること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サイン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日に米国環境衛生研究所 (National Institute of Environmental Health Sciences) より発行された Press Release 00-02 を参照されたい。これについては <http://www.niehs.nih.gov/oc/news/toxchip.htm> で入手可能である)。従って、中毒学的スクリーニングの際に毒性サインを用いて、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

50

【0223】

一実施例では、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を、本発明のポリヌクレオチドに特異的な1若しくは複数のプローブでハイブリダイズし、それによって本発明のポリヌクレオチドに対応する転写レベルを定量し得る。処理した生物学的サンプル中の転写レベルと非処理生物学的サンプル中のレベルとを比較する。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。

【0224】

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 and 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同一性は、その部分配列を、好適には少なくとも5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。

【0225】

タンパク質の(proteomic)プロフィールは、HLYAに特異的な抗体を用いてHLYA発現レベルを定量することによっても作成し得る。或る実施例では、マイクロアレイ上で要素として抗体を用い、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、各アレイ要素における蛍光結合の量を検出し得る。

【0226】

プロテオームレベルでの毒性サイン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サイン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. and 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サインは有用たり得る。更に、mRNAの急速な分解により体液中での転写の分析は困難であるので、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。

【0227】

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

10

20

30

40

50

は、各タンパク質の量を定量し得るように分離する。各タンパク質の量を、非処理生物学的サンプル中の対応するタンパク質の量と比較する。両サンプルのタンパク質量の差は、処理されたサンプル中での試験化合物への毒性反応を示す。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。

【0228】

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

10

【0229】

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、He

20

【0230】

本発明の別の実施例では、天然のゲノム配列をマッピングする際に有効なハイブリダイゼーションプローブを産出するため、HLYA をコードする核酸配列を用いることが可能である。コード配列または非コード配列のいずれかを用いることができ、或る例では、コード配列全体で非コード配列が好ましい。例えば、多重遺伝子ファミリーのメンバー内での

30

【0231】

蛍光原位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, pp. 965-968. 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。物理的染色体地図上の HLYA 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は、その疾患に関係する DNA の領域を画定するのに役立つものであり、従って更に位置クローニングする試みとなり得る。

40

【0232】

50

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、特定のヒト染色体の数或いはアームが分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遺伝的疾患を調査する研究者にとって価値がある。疾患または症候群が、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、該領域に対する任意のマッピングにより更なる調査のための関連遺伝子或いは調節遺伝子を表すことができる(Gatti, R.A.ら(1988) Nature 336:577-580等を参照)。転座、反転等に起因する、健常者、保有者、感染者の三者間における染色体位置の相違を発見するために本発明のヌクレオチド配列を用いてもよい。 10

【0233】

本発明の別の実施例では、種々の薬剤スクリーニング技術を以って化合物のライブラリをスクリーニングするために、HLYA、HLYAの触媒作用断片、免疫原断片、またはそのオリゴペプチドを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。HLYAとテストされる薬剤との結合複合の形成は計測できる。

【0234】

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,らの(1984) PCT出願第WO84/03564号等を参照)。この方法においては、多数の異なる小さな試験用化合物を固体基質上で合成する。試験用化合物は、HLYA或いはその断片と反応させ、洗浄する。次に、本技術分野でよく知られている方法で、結合したHLYAを検出する。精製したHLYAはまた、上記した薬剤のスクリーニング技術において用いるプレート上で直接コーティングすることもできる。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

【0235】

別の実施例では、競合薬スクリーニングアッセイを用いることができる。このアッセイでは、HLYAを結合することができる中和抗体が、HLYAを結合するための試験化合物と特異的に競合する。この方法では、抗体が、1若しくは数個の抗原決定因子をHLYAと共有するペプチドの存在を検出する。 30

【0236】

別の実施例では、新規技術が現在知られているヌクレオチド配列の特性(限定するものではないがトリプレット遺伝暗号及び特異的塩基対の相互作用等を含む)に依存するのであれば、依然として発展すべきいかなる分子生物学技術においても、HLYAをコードするヌクレオチド配列を用いることができる。

【0237】

更に詳細説明をしなくとも、当業者であれば以上の説明を以って本発明を最大限に利用できるであろう。従って、以下に記載する実施例は単なる例示目的であり、本発明を限定するものではない。 40

【0238】

本明細書において開示した全ての特許、特許出願及び刊行物、特に米国特許第60/213,383号、同第60/215,544号、及び同第60/222,818号は、言及することをもって本明細書の一部となす。

【0239】

(実施例)

1 cDNAライブラリの作製

Incyte cDNAは、LIFESEQ GOLDデータベース(Incyte Genomics, Palo Alto CA)に記載されたcDNAライブラリに由来するものであり、表4の列5に列記した。ホモジナイズしてグアニジンイソチオシアネー 50

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、塩化セシウムにおいて遠心分離するかクロロホルムで抽出した。イソプロパノールか、酢酸ナトリウムとエタノールか、いずれか一方、或いは別の方法を用いて、溶解物からRNAを沈殿させた。

【0240】

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

10

【0241】

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、本技術分野で公知の推奨方法または類似の方法を用いて、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。cDNAは、好適なプラスミドのポリリンカーの適合性制限酵素部位に連結反応させた。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)またはpINCY(Incyte Pharmaceuticals, Palo Alto CA)、またはその派生物等である。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を含むコンピテント大腸菌細胞に形質転換した。

20

30

【0242】

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、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。

40

【0243】

別法では、ハイスルーブット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. B

50

iochem. 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, Eugene 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。

【0244】

3 シークエンシング及び分析

cDNAのシークエンス反応は、標準的方法或いはハイスルーブット装置、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。

【0245】

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)配列の除去し、あいまいな塩基対をマスクすることによって有効性を確認した。その際、BLAST、動的プログラミング及び隣接ジヌクレオチド頻度分析に基づくアルゴリズム及びプログラムを用いた。次に、BLAST、FASTA及びBLIMPSに基づくプログラムを用いて、プログラム中の注釈を得るべく、公共のデータベース、例えばGenBankの霊長類及びげっ歯類、哺乳動物、脊椎動物、真核生物のデータベースと、BLOCKS、PRINTS、DOMO、PRODOM及びPFAM等のHidden Markov Model(HMM)ベースのタンパク質ファミリーデータベース(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)の選択に対するIncyte cDNA配列またはその翻訳を問い合わせた。問合せは、BLAST、FASTA、BLIMPS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アSEMBLした。或いは、GenBank cDNA、GenBank EST、縫合配列、伸展配列または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アSEMBLを実行し、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Prositate等のデータベース、PFAM等のHidden Markov Model(HMM)ベースのタンパク質ファミリー

データベースに対する問合せによって完全長ポリペプチド配列を分析した。完全長ポリヌクレオチド配列はまた、MACDNASIS PROソフトウェア（日立ソフトウェアエンジニアリング、South San Francisco CA）及びLASERGENEソフトウェア（DNASTAR）を用いて分析した。ポリヌクレオチド及びポリペプチド配列アラインメントは、アラインメントした配列と配列の一致率も計算するMEGALIGNマルチシーケンスアラインメントプログラム（DNASTAR）に組み込まれているようなCLUSTALアルゴリズムによって特定されるデフォルトパラメータを用いて生成する。

【0246】

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5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2配列間の相同性が高くなる）。

10

【0247】

完全長ポリヌクレオチド及びポリペプチド配列の構築及び分析に用いる上記のプログラムは、配列番号4乃至6からのポリヌクレオチド配列の断片を同定するためにも使用できる。ハイブリダイゼーション及び増幅技術に有用な約20～約4000のヌクレオチドの断片は、表4の列4に記載した。

20

【0248】

4 ゲノムDNAからのコード配列の同定及び編集

Genscan遺伝子同定プログラムを公衆のゲノム配列データベース（例えばgbpri及びgbhtg）に対して実行することにより、推定上のヒトリーゼを先ず同定した。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. and 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、メチオニンから終止コドンに及びアセンブルされた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。いずれのGenscan予測cDNA配列がヒトリ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ヒトリーゼのためのPFAMモデムに対して問い合わせることにより分析した。Incyte cDNA配列の相同体をヒトリーゼとして注釈を付けてきたことにより、潜在的ヒトリーゼも同定した。こうして選択されたGenscan予測配列は、次にBLAST分析により公衆データベース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予測配列の公衆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に記載されたアセンブリプロセスを用いて、Incyte cDNA配列及び/または公衆のcDNA配列でGenscan予測コード配列をアセンブル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。

30

40

【0249】

5 ゲノム配列データのcDNA配列データへのアセンブリ

50

「縫合 (S t i t c h e d) 」配列

部分 c D N A 配列は、実施例 4 に記載の G e n s c a n 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アセンブルされた部分 c D N A は、ゲノム D N A にマッピングし、関連する c D N A 及び 1 若しくは複数のゲノム配列から予測された G e n s c a n エキソンを含むクラスタに分解した。c D N A 及びゲノム情報を統合するべくグラフ理論及び動的プログラミングに基づくアルゴリズムを用いて各クラスタを分析し、引き続いて確認、編集または伸長して完全長配列を産出するような潜在的スプライス変異体を生成した。間隔全体の長さがクラスタ中の 2 以上の配列に存在するような配列を同定し、そのように同定された間隔は推移により等しいと考えられた。例えば、1 つの c D N A 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 c D N A 配列により結び合わせて架橋し得る。このように同定された間隔は、親配列に沿って出現した順番にスティッチングアルゴリズムにより「縫合」され、配列変異体と同様に可能な限り長い配列が生成される。1 種類の親配列に沿って発生した間隔と間隔との連鎖 (c D N A - c D N A またはゲノム配列 - ゲノム配列) は、親の種類を変える連鎖 (c D N A - ゲノム配列) に優先した。結果として得られる縫合配列は、B L A S T 分析により公衆データベース g e n p e p t 及び g b p r i に翻訳されて比較された。G e n s c a n により予測された不正確なエキソンは、g e n p e p t からヒットしたトップの B L A S T と比較することにより修正した。必要な場合には、追加 c D N A 配列を用いるかゲノム D N A の検査により配列を更に伸長させた。

10

20

【 0 2 5 0 】

「伸展 (S t r e c h e d) 」配列

部分 D N A 配列は、B L A S T 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。まず、G e n B a n k 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衆データベースに B L A S T プログラムを用いて、実施例 3 に記載されたようにアセンブルされた部分 c D N A を問い合わせた。次に、最も近い G e n B a n k タンパク質相同体を B L A S T 分析により I n c y t e c D N A 配列または実施例 4 に記載の G e n S c a n 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 (H S P) を用いてキメラタンパク質を産出し、翻訳した配列を G e n B a n k タンパク質相同体上にマッピングした。元の G e n B a n k 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。G e n B a n k 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、公衆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、相同ゲノム配列を追加することにより部分 D N A 配列を「伸展」または伸長させた。結果として得られる伸展配列を検査し、完全遺伝子を含んでいるか否かを決定した。

30

【 0 2 5 1 】

6 ポリヌクレオチドをコードする H L Y A の染色体マッピング

配列番号 4 乃至 6 を配列するために用いた c D N A 配列は、B L A S T 及びその他のスミス ウォーターマンアルゴリズムのインプリメンテーションを用いて、I n c y t e L I F E S E Q のデータベース及びパブリックドメインのデータベースから得た配列と比較した。配列番号 4 乃至 6 に適合するデータベースから得た配列は、P h r a p (表 7) 等のアセンブリアルゴリズムを用いて隣接する配列及びオーバーラップする配列のクラスタに配列した。スタンフォード・ヒトゲノムセンター (S H G C)、ホワイトヘッド・ゲノム研究所 (W I G R)、G e n e t h o n 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測定した。マッピングされた配列がクラスタに含まれている結果、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。

40

【 0 2 5 2 】

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 (c M)

50

は、染色体マーカー間の組換え頻度に基づく計測単位である。平均して、1 cMは、ヒト中の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、配列が各クラスター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。ヒトゲノム地図及びその他の公衆に利用可能なリソース、例えばNCBIの「GeneMap'99」ウェブサイト (<http://www.ncbi.nlm.nih.gov/genemap/>) を用いて、以前に同定した病変遺伝地図が上記の間隔内またはその近傍にあるかどうかを決定することができる。

【0253】

10

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、標識されたヌクレオチド配列の、特定の細胞種または組織からのRNAが結合される膜へのハイブリダイゼーションに関与している。(前出のSambrook, 7章、同Ausubel, F.M.ら, 4章及び16章等を参照。)

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 Pharmaceuticals) 等のヌクレオチド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、多数の膜系ハイブリダイゼーションよりも断然速い。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。検索の基準はプロダクトスコアであり、次式で定義される。

20

【0254】

【数3】

(BLASTスコア×配列一致率)

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

【0255】

プロダクトスコアは、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。プロダクトスコアは、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(ギャップにより離隔され得る)。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プロダクトスコアを計算する。プロダクトスコアは、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。例えばプロダクトスコア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。プロダクトスコア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。プロダクトスコア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。

30

40

【0256】

或いは、HLAをコードするポリヌクレオチド配列を、ポリヌクレオチド配列が派生した組織源に関連して分析した。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アセンブルされる。各cDNA配列は、ヒト組織から構築された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

50

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、細胞系、発達、炎症、神経性、外傷、心血管、貯留 (pooled)、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。演算結果の割合は、HLYAをコードするcDNAの組織特異発現または疾患特異発現を反映する。cDNA配列及びcDNAライブラリ/組織情報については、LIFESEQ GOLDデータベース (Incyte Genomics, Palo Alto CA) で見られる。

【0257】

8 ポリヌクレオチドをコードするHLYA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。開始プライマーの設計は、長さが約22~30ヌクレオチド、GC含有率が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。

【0258】

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であるか或い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。

【0259】

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。反応混合液には、DNA鋳型、各プライマー200nmolと、 Mg^{2+} 、 $(NH_4)_2SO_4$ 及びβ-メルカプトエタノールを含む反応緩衝液と、Taq DNAポリメラーゼ (Amersham Pharmacia Biotech) と、ELONGASE酵素 (Life Technologies) と、Pfu DNAポリメラーゼ (Stratagene) が含まれていた。プライマー対PCI A、PCI Bに対して用いたパラメータは次の通りである。

ステップ1: 94 で3分間

ステップ2: 94 で15秒

ステップ3: 60 で1分間

ステップ4: 68 で2分間

ステップ5: ステップ2、3、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

【0260】

1X TEに溶解したPICOGREEN定量試薬 (0.25% (v/v) PICOGREEN、Molecular Probes, Eugene OR) 100μlと、希釈していないPCR産物0.5μlとを不透明な蛍光光度計プレート (Coming Co

10

20

30

40

50

star, Acton MA)の各穴に分配し、DNAを試薬と結合可能なようにさせることによって、各穴のDNA濃度を測定した。サンプルの蛍光を計測してDNAの濃度を定量するべくプレートをFluoroskan II (Labsystems Oy, Helsinki, Finland)でスキャンした。反応混合物のアリコート5~10 µlを1%アガロースミニゲル上で電気泳動法によって解析し、どの反応が配列の伸長に成功したかを決定した。

【0261】

伸長させたヌクレオチドは、脱塩及び濃縮して384穴プレートに移し、CviJIコラウイルスエンドヌクレアーゼ(Molecular Biology Research, Madison WI)を用いて消化し、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(Stratagene)で処理して制限部位のオーバーハングを満たし、コンピテント大腸菌細胞に形質移入した。形質移入した細胞を抗生物質含有培地上で選択し、個々のコロニーを選択してLB/2x carb液体培地の384穴プレート内において37 °Cで一晩培養した。

10

20

【0262】

細胞を溶解し、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PCRによってDNAを増幅した。その際用いたパラメータは次の通りである。

ステップ1: 94 °C で3分間

ステップ2: 94 °C で15秒

ステップ3: 60 °C で1分間

ステップ4: 72 °C で2分間

ステップ5: ステップ2、3、4を29回繰り返す

ステップ6: 72 °C で5分間

ステップ7: 4 °C で保存

30

【0263】

DNAは、上記のPICOGREEN試薬(Molecular Probes)によって定量した。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 energy transfer sequencing primer及びDYENAMIC DIRECT kit(Amersham Pharmacia Biotech)またはABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit(Applied Biosystems)を用いてシーケンシングした。

40

【0264】

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。

【0265】

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

配列番号4乃至6から得たハイブリダイゼーションプローブを利用して、cDNA、ゲノムDNAまたはmRNA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(Nat

50

ional Biosciences)等の最新ソフトウェアを用いて設計し、各オリゴマー50 pmolと、 $[\gamma - ^{32} P]$ 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)を結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)を用いて実質的に精製する。Ase I、Bgl II、Eco RI、Pst I、Xba IまたはPvu II (DuPont NEN)のいずれか1つのエンドヌクレアーゼで消化されたヒトゲノムDNAの典型的な膜ベースのハイブリダイゼーション解析において、毎分 10^7 カウントの標識されたプローブを含むアリコートを用いる。

10

【0266】

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°C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。

【0267】

10 マイクロアレイ

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

20

30

【0268】

完全長cDNA、発現遺伝子配列断片 (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を構成し得る。ハイブリダイゼーションに好適な断片またはオリゴマーを、LASERGENEソフトウェア (DNASTAR)等の本技術分野で公知のソフトウェアを用いて選択することが可能である。アレイエレメントは、生物学的サンプル中でポリヌクレオチドを用いてハイブリダイズされる。生物学的サンプル中のポリヌクレオチドは、検出を容易にするために蛍光標識またはその他の分子タグに接合される。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、レーザー脱着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。

40

【0269】

組織または細胞サンプルの準備

グアニジウムチオシアネート法を用いて組織サンプルから全RNAを単離し、オリゴ(dT)セルロース法を用いてポリ(A)⁺RNAを精製する。各ポリ(A)⁺RNAサン

50

ルは、MMLV逆転写酵素、0.05 pg/μlのオリゴ(dT)プライマー(21mer)、1×第1鎖緩衝液、0.03 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。特異制御ポリ(A)⁺RNAは、370 で2時間インキュベートした後、in vitro転写により非コード酵母ゲノムDNAから合成する。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、850 で20分間インキュベートし、反応を停止させてRNAを減成する。サンプルは、2つの連続するCHROMA SPIN 30ゲル濾過スピカラム(CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA)を用いて精製する。結合後、2つの反応サンプルは、1 mlのグリコーゲン(1 mg/ml)を用いて析出させたエタノール、60 mlの酢酸ナトリウム及び300 mlの100%エタノールである。サンプルは次に、Speed VAC(Savant Instruments Inc., Holbrook NY)を用いて乾燥して仕上げ、14 μlの5×SSC/0.2% SDS中で再懸濁する。

【0270】

マイクロアレイの準備

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

【0271】

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大量の蒸留水洗液を用いて0.1%のSDS及びアセトン中で超音波により洗浄する。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、蒸留水中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 の天火で硬化させる。

【0272】

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を高速ロボット装置により開口キャピラリープリントエレメントに充填する。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デポジットする。

【0273】

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。

【0274】

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 Cy 3 及び Cy 5 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キャピティを有する防水チェンバーに移行させる。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。

【0275】

検出

レポーター標識ハイブリダイゼーション複合体は、Cy 3 の励起のためには 488 nm、Cy 3 の励起のためには 632 nm でスペクトル線を生成し得る Innova 70 混合ガス 10 W レーザー (Coherent, Inc., Santa Clara CA) を備えた顕微鏡で検出する。20 × 顕微鏡対物レンズ (Nikon, Inc., Melville NY) を用いて、アレイ上に励起レーザー光を集中させる。アレイを含むスライドを顕微鏡のコンピュータ制御 X-Y ステージに置き、対物レンズを通過してラスタスキャンする。本実施例で用いた 1.8 cm × 1.8 cm のアレイは、20 µm の解像度でスキャンした。

【0276】

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ーは 2 つの蛍光体を連続的に励起する。放射された光は、2 つの蛍光体に応じて波長に基づき 2 つの光電子増倍管検出器 (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) に分割される。アレイと光電子増倍管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リングする。用いる蛍光体の最大発光は、Cy 3 では 565 nm、Cy 5 では 650 nm である。装置は両方の蛍光体からのスペクトルを同時に記録し得るが、レーザー源において好適なフィルターを用いて各アレイを通常 2 度スキャンし、蛍光体 1 つにつき 1 度スキャンする。

【0277】

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 制御種により生成されるシグナル強度を用いて較正する。アレイ上の特定の位置には相補的 DNA 配列が含まれ、その位置におけるシグナルの強度をハイブリダイジング種の重量比 1:100,000 に相関させる。異なる源 (例えば試験及び制御細胞を表す) からの 2 つのサンプルを、各々異なる蛍光体で標識し、他と異なって発現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、2 つの蛍光体を有する較正 cDNA の標識サンプルにより較正し、ハイブリダイゼーション混合体に各々等量を加える。

【0278】

光電子増倍管の出力は、IBM 互換 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 つの異なる蛍光体を同時に励起及び測定する場合には、各蛍光体の発光スペクトルを用いて、データは先ず蛍光体と蛍光体の間の光学磁気プリンティング (発光スペクトルの重畳に起因する) に集められる。

【0279】

グリッドは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。

【0280】

10

20

30

40

50

1.1 相補的ポリヌクレオチド

H L Y A 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 H L Y A 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。O l i g o 4 . 0 6 ソフトウェア (N a t i o n a l B i o s c i e n c e s) と H L Y A をコードする配列とを用いて、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、H L Y A をコードする転写物にリボソームが結合しないように相補的オリゴヌクレオチドをデザインする。

10

【0281】

1.2 H L Y A の発現

H L Y A 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 H L Y A を発現するために、抗生物質耐性及び c D N A 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 c D N A をサブクローニングする。このようなプロモーターには、l a c オペレーター調節エレメントに関連する T 5 または T 7 バクテリオファージプロモーター及び t r p - l a c (t a c) 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、B L 2 1 (D E 3) 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 (I P T G) で誘導されると H L Y A を発現する。真核細胞での H L Y A の発現は、一般にバキュロウイルスとして知られている A u t o g r a p h i c a c a l i f o r n i c a 核多面性ウイルス (A c M N P V) を昆虫細胞株または哺乳動物細胞株に感染させて行う。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中間体を用いる細菌仲介遺伝子転移のどちらかによって、H L Y A をコードする c D N A と置換する。ウイルスの感染力は維持され、強いポリヘドリンプロモーターによって高いレベルの c D N A 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 S p o d o p t e r a f r u g i p e r d a (S f 9) 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。(E n g e l h a r d , E . K . ら (1 9 9 4) P r o c . N a t l . A c a d . S c i . U S A 9 1 : 3 2 2 4 - 3 2 2 7 , S a n d i g , V . ら (1 9 9 6) H u m . G e n e T h e r . 7 : 1 9 3 7 - 1 9 4 5 . 等を参照) 。

20

30

【0282】

殆どの発現系では、融合タンパク質として H L Y A を合成するのに例えばグルタチオン S トランスフェラーゼ (G S T) またはペプチドエピトープ標識、例えば F L A G や 6 - H i s を用いる。これらを用いることにより、未精製細胞溶解物から組換え融合タンパク質の親和性ベースの精製を迅速に 1 ステップで行うことができる。G S T は日本住血吸虫からの 26 k D a 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 (A m e r s h a m P h a r m a c i a B i o t e c h) 。精製後、G S T の部分を特定の開発部位において H L Y A からタンパク質分解的に切断することが可能である。F L A G 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 F L A G 抗体 (E a s t m a n K o d a k) を用いて免疫親和性精製を可能にする。6 つの連続したヒスチジン残基の伸長である 6 - H i s は、金属キレート樹脂 (Q I A G E N) 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 A u s u b e l (1 9 9 5) 10 章、16 章に記載されている。これらの方法で精製した H L Y A を直接用いて、適用可能な場合には実施例 16、17 及び 18 のアッセイを行うことができる。

40

【0283】

1.3 機能的アッセイ

H L Y A 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 H L Y A

50

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

【0284】

遺伝子発現におけるHLYAの影響は、HLYAをコードする配列とCD64またはCD64-GFPのいずれ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。形質転換細胞と非形質転換細胞は、ヒトIgGまたはCD64に対する抗体(DYNAL, Lake Success, NY)で覆われた磁気ビーズを用いて有効に分離することができる。mRNAは、本技術分野で公知の方法で細胞から精製することができる。HLYAその他の目的の遺伝子をコードするmRNAの発現は、ノーザン分析或いはマイクロアレイ技術で分析することができる。

【0285】

1.4 HLYA特異抗体の産生

ポリアクリルアミドゲル電気泳動法(PAGE; Harrington, M. G. (1990) Methods Enzymol. 182: 488-495等を参照)または他の精製技術を用いて実質上精製されたHLYAを用いて、標準プロトコルでウサギを免疫化して抗体を産出する。

【0286】

或いは、LASERGENEソフトウェア(DNA STAR)を用いてHLYAアミノ酸配列を解析し、免疫抗原性の高い領域を決定する。そして対応するオリゴペプチドを合成し、このオリゴペプチドを用いて当業者によく知られている方法で抗体を生成する。適切なエピトープ、例えばC末端付近或いは隣接する親水性領域にあるエピトープの選択については、当分野で公知である(前出のAusubel, 1995, 11章等を参照)。

【0287】

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、免疫抗原性を高める(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 - K L M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H L Y A活性を検査するには、ペプチドまたはH L Y Aを基質に結合し、1% B S Aを用いてブロックし、ウサギ抗血清と反応させて洗浄し、さらに放射性ヨウ素で標識したヤギ抗ウサギI g Gと反応させる。

【0288】

1.5 特異抗体を用いた天然のH L Y Aの精製

天然または組換えH L Y Aを、H L Y A特異抗体を用いた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、抗H L Y A抗体を活性化クロマトグラフィ用樹脂、例えばC N B r活性化セファロース (A m e r s h a m P h a r m a c i a B i o t e c h) と共有結合させることにより構築する。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。

10

【0289】

H L Y Aを含む培養液をイムノアフィニティークラムに通し、H L Y Aを優先的に吸着する条件下 (例えば洗浄剤が存在する高イオン強度緩衝液) でカラムを洗浄する。抗体とH L Y Aの結合を破壊する条件 (例えばp H 2 ~ 3の緩衝液、或いは尿素またはチオシアン酸塩イオン等の高濃度のカオトロップ剤) でカラムを溶出させ、H L Y Aを回収する。

【0290】

1.6 H L Y Aと相互作用する分子の同定

H L Y Aまたは生物学的に活性であるH L Y A断片を、¹²⁵Iボルトンハンター試薬で標識する (B o l t o n A . E . a n d W . M . H u n t e r (1 9 7 3) B i o c h e m . J . 1 3 3 : 5 2 9 - 5 3 9 等を参照)。マルチウェルプレートの穴の中に予め配列しておいた候補分子を、標識したH L Y Aと共にインキュベートして洗浄し、標識したH L Y A複合体を有する任意の穴をアッセイする。H L Y A濃度を変えて得たデータを用いて、候補分子とのH L Y Aの数、親和性及び会合の値を計算する。

20

【0291】

或いは、H L Y Aと相互作用する分子を、F i e l d s , S . 及び O . S o n g (1 9 8 9) N a t u r e 3 4 0 : 2 4 5 - 2 4 6 に記載されているような酵母2ハイブリッドシステムまたはM A T C H M A K E Rシステム (C l o n t e c h) 等の2ハイブリッドシステムに基づき、市販のキットを用いて分析する。

【0292】

H L Y Aはまた、ハイスループットな方法で酵母2ハイブリッドシステムを使用するP A T H C A L L I N Gプロセス (C u r a G e n C o r p . , N e w H a v e n C T) に用いて、遺伝子の2大ライブラリにコードされる遺伝子間の全ての相互作用を決定することができる (N a n d a b a l a n , K . ら (2 0 0 0) 米国特許第6, 057, 101号)。

30

【0293】

1.7 H L Y A活性の実証

H L Y Aの炭酸アンヒドラーゼ活性に対するアッセイは、ストップフロー蛍光定量法と組み合わせて蛍光p H指示薬8 - ヒドロキシピレン - 1 , 3 , 6 - トリスルホン酸 (ピラニン) を用い、炭酸アンヒドラーゼ活性を測定する (S h i n g l e s , ら 1 9 9 7 , A n a l . B i o c h e m . 2 5 2 : 1 9 0 - 1 9 7)。p H 6 . 0の溶液をp H 8 . 0の溶液と混合し、重炭酸塩の初期脱水速度を測定する。p H 6 . 0の溶液に炭酸アンヒドラーゼを添加すると、2 m sの溶解時間で生理的温度における活性の初期速度の測定が可能になる。S h i n g l e s らは、このアッセイを用いて、哺乳動物の炭酸アンヒドラーゼアイソフォームを比較することにより、活性とスルホンアミドに対する感度の差を決定した。蛍光技術の感度は、6 5 n g / m l程度に低いタンパク質濃度で初期速度を決定することができる。

40

【0294】

L - [1 - ¹⁴C] - オルニチンからのC O ₂ の放出を測定することにより、H L Y Aのオルニチンデカルボキシラーゼ活性をアッセイする (R e d d y , S . G e t a l .

50

(1996) J. Biol. Chem. 271:24945-24953)。活性は、0.5 mM L-オルニチン及び0.5 μ Ci L-[1-¹⁴C]-オルニチンを含む200 μ lのアッセイ緩衝液(50 mM Tris/HCl、pH 7.5、0.1 mM EDTA、2 mM ジチオスレイトール、5 mM NaF、0.1% ブリッジ35、1 mM PMSF、60 μ M ピリドキサル-5-リン酸)中で測定する。15~30分後に1 M クエン酸の添加により反応が停止し、進化した¹⁴CO₂を、20 μ lの2 N NaOHで飽和したペーパーディスクフィルターに捕捉する。ディスクの放射活性は、液体シンチレーションスペクトログラフィーにより決定する。放出された¹⁴CO₂の量は、HLYAのオルニチンデカルボキシラーゼ活性に比例する。

【0295】

10

1.8 HLYA活性の同定

HLYA活性または阻害のアゴニストまたはアンタゴニストは、17章に記載のアッセイを用いて試験し得る。アゴニストはHLYA活性を増加させ、アンタゴニストはHLYA活性を減少させる。

【0296】

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に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。

20

【0297】

(表の簡単な説明)

表1は、本発明の完全長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。

【0298】

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。

【0299】

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ヌクレオチド配列の構造的な特徴を、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。

30

【0300】

表4は、本発明のポリヌクレオチド配列をアセンブルするために用いたcDNA及びゲノム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。

【0301】

表5は、本発明のポリヌクレオチドのための代表的なcDNAライブラリを示す。

【0302】

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。

40

【0303】

表7は、HLYAの分析に用いたツール、プログラム、アルゴリズムを、適用可能な説明、引用文献及び閾値パラメータと共に示す。

【表1】

表 1

Incyte プロジェクト ID	ポリペプチド SEQ ID NO:	Incyte ポリペプチド ID	ポリヌクレオチド SEQ ID NO:	Incyte ポリヌクレオチド ID
6338333	1	6338333CD1	4	6338333CB1
1415322	2	1415322CD1	5	1415322CB1
3267634	3	3267634CD1	6	3267634CB1

10

20

30

40

【表 2】

冊 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	6338333CD1	g13249093	1.00E-130	[f1] [ハツカネズミ] 炭酸アンヒドラーゼXIII Hewett-Emmett, D. (2000) (in) Chegwidden, W.R., Carter, N.D. and Edwards, Y.H. (Eds.); THE CARBONIC ANHYDRASES: NEW HORIZONS: 29-76; Birkhauser Verlag, Basel [ヒト] 炭酸アンヒドラーゼ (EC 4.2.1.1) Lowe, N. et al. (1990) Structure and methylation patterns of the gene encoding human carbonic anhydrase I. Gene 93 (2), 277-283 [アフリカツメガエル] オルニチンデカルボキシラーゼ2 Bernstein, H.G. and Muller, M. (1999) Prog. Neurobiol. 57:485-505 Medina, M.A. et al. (1999) Biochem Pharmacol 57:1341-1344 [f1] [ヒト] 炭酸アンヒドラーゼ関連タンパク質X Hewett-Emmett, D. and Tashian, R.E. (1996) Mol. Phylogenet. Evol. 5: 50-77 [ヒト] 炭酸アンヒドラーゼ関連タンパク質2
2	1415322CD1	g6746571	1.90E-148	
3	3267634CD1	g13094699	0	
		g3283386	5.70E-94	

【表 3】

10

20

30

40

表 3 - 1

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸 残基	潜在的リン酸化部位	潜在的グリコ シル化部位	サイン(signature)配列、 ドメイン、モチーフ	分析方法及び データベース
1	6338333CD1	242	S100 S126 S131 S167 S30 S44 S49 S56 S88 T178 Y41	N218 N76	真核型炭酸アンヒドラーゼドメイン: W6-A239 真核型炭酸アンヒドラーゼサイン: BL00162: W17-E47; Y52-T74; Y89- N125; K128-G152; D191-Q223 真核型炭酸アンヒドラーゼサイン: G82-A143 炭酸アンヒドラーゼドラーゼ 炭酸リアーゼ亜前駆体 シグナルタンパク質糖タンパク質 PD000865: E10-A238 炭酸アンヒドラーゼ DM00356 JN0836 25-261: D25-A239 真核炭酸アンヒドラーゼモチーフ: S106-V122	HMMER_PFAM BLIMPS_BLOCKS PROFILES SCAN BLAST_PRODUM BLAST_DOMO MOTIFS MOTIFS
2	1415322CD1	460	S18 S149 S168 S174 S177 T311 S346 S6 S18 T34 S212 S258 Y352	N308	Y67-G85 Orn/DAP/Arg デカルボキシラーゼ 2_1 M230-G241 Orn/DAP/Arg デカルボキシラーゼ 2_2 V80-C90, V193-C203, M394-C404 Prokar_Lipoprotein 付着部位 Orn_DAP_Arg_deC: F41-T401 ピリドキサーール依存デカルボキシラーゼ BL00878A: D45-P61 BL00878B: H189-C203 BL00878C: E225-G244 BL00878D: V270-A288 BL00878E: L356-D365 BL00878F: G379-T401 Orn/DAP/Arg デカルボキシラーゼ	HMMER_PFAM BLIMPS_BLOCKS

10

20

30

40

【表 4】

表 4

ポリヌクレオチド SEQ ID NO:	Incyte ポリヌクレオチド ID	配列長さ	選択断片	配列断片	5' 位置	3' 位置
4	6338333CB1	911	1-28, 741- 766, 870- 911	71672723V1 70513545V1	429 1	911 572
5	1415322CB1	2064	1058-1146, 430-1032	2698586F6 (UTRSNOT12) 91191247 1401396H1 (BRAITUT08) 6854836H1 (BRAIFEN08) 5273763H1 (OVARIN02) 6312295H1 (NERDYDN03) 5109549F6 (PROSTUS19) 3245804F6 (BRAINOT19) 70996811V1 6447115H1 (BRAINOC01) 6328316H1 (BRANDIN01) 3246722F7 (BRAINOT19) 3267634H1 (BRAINOT20) 6926620H1 (SINITMR01) 1295926T6 (PGANNOT03) 70995406V1	1758 1550 1887 1075 714 877 312 1 1060 770 1695 404 1 216 2372 1550	2064 2064 2064 1781 987 1540 900 416 1635 1489 2395 1060 236 834 2938 2217
6	3267634CB1	2938	1862-2093, 1-786			

10

20

30

40

【表 6】

表 5

ポリヌクレオチド SEQ_ID NO:	Incyte プロジェクト ID	代表的なライブラリ
4	6338333CB1	BRSTNOT07
5	1415322CB1	BRAXNOT03
6	3267634CB1	BRAINOT09

10

20

30

40

【表 7】

表 5

ポリヌクレオチド SEQ_ID NO:	Incyte プロジェクト ID	代表的なライブラリ
4	6338333CB1	BRSTNOT07
5	1415322CB1	BRAXNOT03
6	3267634CB1	BRAINOT09

10

20

30

40

【表 8】

表 9

ライブラリ	ベクター	ライブラリの説明
BRAINOT09	pINCY	BRAINOT09 ライブラリは、妊娠 23 週目に死亡した白人の男胎児から採取した脳組織から単離した RNA を用いて作製した。
BRAXNOT03	pINCY	BRAXNOT03 ライブラリは、心不全で死亡した 35 歳の白人男性の脳から採取した感覚運動皮質組織から単離した RNA を用いて作製した。病理学的には、中程度の軟膜線維症及び大脳新皮質の複数の微小梗塞を示していた。患者の病歴には、拡張型心筋症、うっ血性心不全、心臓肥大、脾腫及び肝肥大が含まれる。
BRSTNOT07	pINCY	BRSTNOT07 ライブラリは、43 歳の白人女性から拡大単純片側乳房切除の際に採取した病変乳房組織から単離した RNA を用いて作製した。病理学的には、上皮過形成、乳頭腫症及び管拡張を伴う軽度の増殖性線維細胞の変化を示していた。関連する腫瘍組織は、病理学的には面皰の壊死を伴った浸潤性グレード 4、核グレード 3 の乳腺腫を示していた。家族歴には、癲癇、心血管疾患及び II 型糖尿病が含まれる。

【表 9】

10

20

30

40

表 7-1

プログラム名	説明	引用文献	パラメーター國値
ABI FACTURA	核酸配列においてベクター配列を除去し、不明瞭な塩基をマスクするプログラム	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	アミノ酸又は核酸配列の比較や注釈付けに有用な Fast Data Finder	Applied Biosystems, Foster City, CA.	ミスマッチ<50%
ABI AutoAssembler	核酸配列を構築するプログラム	Applied Biosystems, Foster City, CA.; Paracel inc., Pasadena, CA.	
BLAST	アミノ酸及び核酸配列の配列類似性検索に有用な Basic Local Alignment Search Tool. BLAST には、blastp 及び blastn, blastx, tblastn, tblastx の 5 つの機能がある。	Altschul, S.F.ら(1990) J. Mol. Biol. 215:403-410; Altschul, S.F.ら(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	照会配列と同種の配列のグループとの類似性を検索する Pearson 及び Lipman アルゴリズム。FASTA は、少なくとも 5 つの機能 fasta, tfasta, fastx, tfastx, ssearch からなる。	Pearson, W.R.及び D.J. Lipman (1988) Proc.Natl. Acad Sci. 85:2444-2448; Pearson, W.R.(1990) Methods Enzymol. 183: 63-98; Smith, T.F.及び M. S. Waterman (1981) Adv.Appl. Math. 2:482-489.	ESTs : fasta E 値=1.06E-6 Assembled ESTs : fasta 一致率=95%以上 一致長さ=200 塩基以上 fastx E 値=1.0E-8 以下 完全長配列 : fastx スコア=100 以上
BLIMPS	或る配列を、BLOCKS 及び PRINTS データベースにおける配列と一致させて、遺伝子ファミリー及び配列相同性、構造的フィンガープリント領域を検索する BLocks IMProved Searcher	Itemikoff, S 及び J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff 及び S.Henikoff (1996) Methods Enzymol. 266:88-105; Artwood, T.K.ら(1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	確率値=1.0E-3 以下
HMMER	PFAM 等の、タンパク質ファミリーコンセンサス配列の Hidden Markov Model に基づいたデータベースに対する問合せ配列を探すアルゴリズム	Krogh, A.ら(1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L.ら(1988) Nucleic Acids Res. 26:320-322.	PFAM ヒット : 確率値=1.0E-3 以下 シグナルペプチドヒット : スコア=0 以上

10

20

30

40

表7-2

プログラム名	説明	引用文献	パラメーター関値
ProfileScan	Prosite で定義された配列パターンと一致するタンパク質配列における構造的及び配列モチーフを検索するアルゴリズム	Gribkov, M.ら(1988) CABIOS 4:61-66; Gribkov ら(1989) Methods Enzymol. 183:146-159; Bairoch, A.ら(1997) Nucleic Acids Res. 25:217-221.	標準化された質のスコア \geq 特定の Prosite モチーフのための GCG 指定された「高い」値 通常、スコア=1.4-2.1
Phred	高い感度及び確率で自動の配列のトレースを検査する塩基呼び出しアルゴリズム	Ewing, B.ら(1998) Genome Res. 8:175-185; Ewing, B.及び P. Green (1998) Genome Res. 8:186-194.	
Phrap	配列の相同性の検索及び DNA 配列の構築に有用であるような、Smith-Waterman アルゴリズムを効率的に利用したプログラムである SWAT 及び CrossMatch を含む Phis Revised Assembly プログラム	Smith, T.F.及び M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F.及び M.S. Waterman (1981) J. Mol. Biol. 147:195-197; 並びに Green, P., University of Washington, Seattle, WA.	スコア=120 以上 一致長さ=56 以上
Consed	Phrap assemblies の編集及び観察用のグラフィックツール	Gordon, D.ら(1998) Genome Res. 8:195-202.	
SPScan	分泌シグナルペプチドの存在するかどうかタンパク質配列をスクリーンするウエイトマトリックス分析プログラム	Nielson, H.ら(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
Motifs	Prosite で定義されたものと一致するパターンに対してアミノ酸配列を検索するプログラム	Bairoch ら、前出; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

10

20

30

40

【国際公開パンフレット】

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
3 January 2002 (03.01.2002)

PCT

(10) International Publication Number
WO 02/00840 A2

- (51) International Patent Classification: C12N [US/US]; 837 Roble Avenue, #1, Menlo Park, CA 94025 (US). BURFORD, Neil [GB/US]; 105 Wildwood Circle, Durham, CT 06422 (US). THANGAVELU, Kavitha [IN/US]; 1950 Montecito Avenue, 23, Mountain View, CA 94043 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US).
- (21) International Application Number: PCT/US01/19166
- (22) International Filing Date: 13 June 2001 (13.06.2001)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/213,383 23 June 2000 (23.06.2000) US
60/215,544 30 June 2000 (30.06.2000) US
60/222,818 4 August 2000 (04.08.2000) US
- (71) Applicant (for all designated States except US): INCYTE GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).
- (72) Inventors; and
(75) Inventors/Applicants (for US only): THORNTON, Michael [US/US]; 9 Madway Road, Woodside, CA 94062 (US). RAMKUMAR, Jayalaxmi [IN/US]; 34359 Maybird Circle, Fremont, CA 94555 (US). TRIBOLEY, Catherine, M. [FR/US]; 1121 Tennessee Street, #5, San Francisco, CA 94107 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). NGUYEN, Dannie, B. [US/US]; 1403 Ridgewood Drive, San Jose, CA 95118 (US). YAO, Monique, G. [US/US]; 111 Frederick Court, Mountain View, CA 94043 (US). PATTERSON, Chandra [US/US]; 490 Sherwood Way #1, Menlo Park, CA 94025 (US). GANDHI, Ameena, R.
- (74) Agents: HAMLET-COX, Diana et al., Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TI, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- Published:
— without international search report and to be republished upon receipt of that report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 02/00840 A2

(54) Title: HUMAN LYASES

(57) Abstract: The invention provides human lyases (HLXA) and polynucleotides which identify and encode HLXA. 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 HLXA.

WO 02/00840

PCT/US01/19166

HUMAN LYASES**TECHNICAL FIELD**

This invention relates to nucleic acid and amino acid sequences of human lyases and to the use of these sequences in the diagnosis, treatment, and prevention of cancers and autoimmune/inflammatory, neurological, and eye disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of human lyases.

BACKGROUND OF THE INVENTION

Lyases are a class of enzymes that catalyze the cleavage of C-C, C-O, C-N, C-S, C-(halide), P-O, or other bonds without hydrolysis or oxidation to form two molecules, at least one of which contains a double bond (Stryer, L. (1995) Biochemistry, W.H. Freeman and Co., New York NY, p.620). Under the International Classification of Enzymes (Webb, E.C. (1992) Enzyme Nomenclature 1992: Recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the Nomenclature and Classification of Enzymes, Academic Press, San Diego CA), lyases form a distinct class designated by the numeral 4 in the first digit of the enzyme number (i.e., EC 4.x.x.x).

Further classification of lyases reflects the type of bond cleaved as well as the nature of the cleaved group. The group of C-C lyases includes carboxyl-lyases (decarboxylases), aldehyde-lyases (aldolases), oxo-acid-lyases, and other lyases. The C-O lyase group includes hydro-lyases, lyases acting on polysaccharides, and other lyases. The C-N lyase group includes ammonia-lyases, amidine-lyases, amine-lyases (deaminases), and other lyases. Lyases are critical components of cellular biochemistry, with roles in metabolic energy production, including fatty acid metabolism and the tricarboxylic acid cycle, as well as other diverse enzymatic processes.

One important family of lyases are the carbonic anhydrases (CA), also called carbonate dehydratases, which catalyze the hydration of carbon dioxide in the reaction $H_2O + CO_2 \rightleftharpoons HCO_3^- + H^+$. CA accelerates this reaction by a factor of over 10^6 by virtue of a zinc ion located in a deep cleft about 15Å below the protein's surface and co-ordinated to the imidazole groups of three His residues. Water bound to the zinc ion is rapidly converted to HCO_3^- .

Eight enzymatic and evolutionarily related forms of carbonic anhydrase are currently known to exist in humans: three cytosolic isozymes (CAI, CAII, and CAIII), two membrane-bound forms (CAIV and CAVII), a mitochondrial form (CAV), a secreted salivary form (CAVI) and a yet uncharacterized isozyme (Prosite PDOC00146 Eukaryotic-type carbonic anhydrases signature).

WO 02/00840

PCT/US01/19166

Though the isoenzymes CAI, CAII, and bovine CAIII have similar secondary structure and polypeptide-chain fold, CAI has 6 tryptophans, CAII has 7 and CAIII has 8 (Boren, K. et al. (1996) *Protein Sci.* 5:2479-2484). CAII is the predominant CA isoenzyme in the brain of mammals.

CAs participate in a variety of physiological processes that involve pH regulation, CO₂ and HCO₃⁻ transport, ion transport, and water and electrolyte balance. For example, CAII contributes to H⁺ secretion by gastric parietal cells, by renal tubular cells, and by osteoclasts that secrete H⁺ to acidify the bone-resorbing compartment. In addition, CAII promotes HCO₃⁻ secretion by pancreatic duct cells, ciliary body epithelium, choroid plexus, salivary gland acinar cells, and distal colonic epithelium, thus playing a role in the production of pancreatic juice, aqueous humor, cerebrospinal fluid, and saliva, and contributing to electrolyte and water balance. CAII also promotes CO₂ exchange in proximal tubules in the kidney, in erythrocytes, and in lung. CAIV has roles in several tissues: it facilitates HCO₃⁻ reabsorption in the kidney; promotes CO₂ flux in tissues including brain, skeletal muscle, and heart muscle; and promotes CO₂ exchange from the blood to the alveoli in the lung. CAVI probably plays a role in pH regulation in saliva, along with CAII, and may have a protective effect in the esophagus and stomach. Mitochondrial CAV appears to play important roles in gluconeogenesis and ureagenesis, based on the effects of CA inhibitors on these pathways. (Sly, W.S. and Hu, P.Y. (1995) *Ann. Rev. Biochem.* 64:375-401.)

A number of disease states are marked by variations in CA activity. Mutations in CAII which lead to CAII deficiency are the cause of osteopetrosis with renal tubular acidosis (Online Medelian Inheritance in Man 259730 Osteopetrosis with Renal Tubular Acidosis). The concentration of CAII in the cerebrospinal fluid (CSF) appears to mark disease activity in patients with brain damage. High CA concentrations have been observed in patients with brain infarction. Patients with transient ischemic attack, multiple sclerosis, or epilepsy usually have CAII concentrations in the normal range, but higher CAII levels have been observed in the CSF of those with central nervous system infection, dementia, or trigeminal neuralgia (Parkkila, A.K. et al. (1997) *Eur. J. Clin. Invest.* 27:392-397). Colonic adenomas and adenocarcinomas have been observed to fail to stain for CA, whereas non-neoplastic controls showed CAI and CAII in the cytoplasm of the columnar cells lining the upper half of colonic crypts. The neoplasms show staining patterns similar to less mature cells lining the base of normal crypts. (Gramlich T.L. et al. (1990) *Arch. Pathol. Lab. Med.* 114:415-419).

Therapeutic interventions in a number of diseases involve altering CA activity. CA inhibitors such as acetazolamide are used in the treatment of glaucoma (Stewart, W.C. (1999) *Curr. Opin. Ophthalmol.* 10:99-108), essential tremor and Parkinson's disease (Uitti, R.J. (1998) *Geriatrics* 53:46-48, 53-57), intermittent ataxia (Singhvi, J.P. et al. (2000) *Neurology India* 48:78-80), and altitude related

WO 02/00840

PCT/US01/19166

illnesses (Klocke, D.L. et al. (1998) Mayo Clin. Proc. 73:988-992).

CA activity can be particularly useful as an indicator of long-term disease condition, since the enzyme reacts relatively slowly to physiological changes. CAI and zinc concentrations have been observed to decrease in hyperthyroid Graves' disease (Yoshida, K. (1996) Tohoku J. Exp. Med. 178:345-356) and glycosylated CAI is observed in diabetes mellitus (Kondo, T. et al. (1987) Clin. Chim. Acta 166:227-236). A positive correlation has been observed between CAI and CAII reactivity and endometriosis (Brinton, D.A. et al. (1996) Ann. Clin. Lab. Sci. 26:409-420; D'Cruz, O.J. et al. (1996) Fertil. Steril. 66:547-556).

Another important member of the lyase family is ornithine decarboxylase (ODC), the initial rate-limiting enzyme in polyamine biosynthesis. ODC catalyses the transformation of ornithine into putrescine in the reaction $L\text{-ornithine} \rightarrow \text{putrescine} + \text{CO}_2$. Polyamines, which include putrescine and the subsequent metabolic pathway products spermidine and spermine, are ubiquitous cell components essential for DNA synthesis, cell differentiation, and proliferation. Thus the polyamines play a key role in tumor proliferation (Medina, M.A. et al. (1999) Biochem. Pharmacol. 57:1341-1344).

ODC is a pyridoxal-5'-phosphate (PLP)-dependent enzyme which is active as a homodimer. Conserved residues include those at the PLP binding site and a stretch of glycine residues thought to be part of a substrate binding region (Prosite PDOC00685 Orn/DAP/Arg decarboxylase family 2 signatures). Mammalian ODCs also contain PEST regions, sequence fragments enriched in proline, glutamic acid, serine, and threonine residues that act as signals for intracellular degradation (Medina, *supra*).

Many chemical carcinogens and tumor promoters increase ODC levels and activity. Several known oncogenes may increase ODC levels by enhancing transcription of the ODC gene, and ODC itself may act as an oncogene when expressed at very high levels. A high level of ODC is found in a number of precancerous conditions, and elevation of ODC levels has been used as part of a screen for tumor-promoting compounds (Pegg, A.E. et al. (1995) J. Cell. Biochem. Suppl. 22:132-138).

Inhibitors of ODC have been used to treat tumors in animal models and human clinical trials, and have been shown to reduce development of tumors of the bladder, brain, esophagus, gastrointestinal tract, lung, oral cavity, mammary gland, stomach, skin and trachea (Pegg, *supra*; McCann, P.P. and A.E. Pegg (1992) Pharmac. Ther. 54:195-215). ODC also shows promise as a target for chemoprevention (Pegg, *supra*). ODC inhibitors have also been used to treat infections by African trypanosomes, malaria, and *Pneumocystis carinii*, and are potentially useful for treatment of autoimmune diseases such as lupus and rheumatoid arthritis (McCann, *supra*).

Another family of pyridoxal-dependent decarboxylases are the group II decarboxylases. This

WO 02/00840

PCT/US01/19166

family includes glutamate decarboxylase (GAD) which catalyzes the decarboxylation of glutamate into the neurotransmitter GABA; histidine decarboxylase (HDC), which catalyzes the decarboxylation of histidine to histamine; aromatic-L-amino-acid decarboxylase (DDC), also known as L-dopa decarboxylase or tryptophan decarboxylase, which catalyzes the decarboxylation of tryptophan to tryptamine and also acts on 5-hydroxy-tryptophan and dihydroxyphenylalanine (L-dopa); and cysteine sulfonic acid decarboxylase (CSD), the rate-limiting enzyme in the synthesis of taurine from cysteine (PROSITE PDOC00329 DDC/GAD/HDC/TyrDC pyridoxal-phosphate attachment site). Taurine is an abundant sulfonic amino acid in brain and is thought to act as an osmoregulator in brain cells (Bitoun, M. and Tappaz, M. (2000) J. Neurochem. 75:919-924).

The discovery of new human lyases 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 cancers and autoimmune/inflammatory, neurological, and eye disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of human lyases.

15

SUMMARY OF THE INVENTION

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

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:1-3, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. In one alternative, the

WO 02/00840

PCT/US01/19166

polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-3. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:4-6.

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:1-3, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. 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:1-3, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3.

The invention further provides an isolated polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, c) a

WO 02/00840

PCT/US01/19166

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

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, 5 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:4-6, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the 10 polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of 15 said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID 20 NO:4-6, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, 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) 25 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:1-3, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group 30 consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3,

WO 02/00840

PCT/US01/19166

and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional HLYA, comprising administering to a patient in need of such treatment the composition.

5 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:1-3, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a
10 polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. 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
15 pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional HLYA, comprising administering to a patient in need of such treatment the composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an
20 amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and
25 d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of
30 treating a disease or condition associated with overexpression of functional HLYA, comprising administering to a patient in need of such treatment the composition.

The invention further provides a method of screening for a compound that specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a naturally occurring polypeptide

WO 02/00840

PCT/US01/19166

comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

The invention further provides a method of screening for a compound that modulates the activity of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence selected from the group consisting of SEQ ID NO:4-6, 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:4-6, ii) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, iii) a polynucleotide having a sequence complementary to i), iv) a polynucleotide complementary to the polynucleotide of

WO 02/00840

PCT/US01/19166

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:4-6, ii) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, 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.

15 BRIEF DESCRIPTION OF THE TABLES

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

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

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

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

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

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

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

DESCRIPTION OF THE INVENTION

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

WO 02/00840

PCT/US01/19166

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

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

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

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

An "allelic variant" is an alternative form of the gene encoding HLYA. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to
30 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 HLYA include those sequences with deletions,

WO 02/00840

PCT/US01/19166

insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as HLYA or a polypeptide with at least one functional characteristic of HLYA. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding HLYA, and improper or unexpected hybridization to allelic variants, with a
5 locus other than the normal chromosomal locus for the polynucleotide sequence encoding HLYA. 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 HLYA. 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
10 or immunological activity of HLYA is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine;
15 and phenylalanine and tyrosine.

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

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity
25 of HLYA. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of HLYA either by directly interacting with HLYA or by acting on components of the biological pathway in which HLYA participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments
30 thereof, such as Fab, F(ab')₂, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind HLYA 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,

WO 02/00840

PCT/US01/19166

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.

5 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.

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

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

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

A "composition comprising a given polynucleotide sequence" and a "composition comprising a
30 given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding HLYA or fragments of HLYA may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be

WO 02/00840

PCT/US01/19166

associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

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

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

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

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.

WO 02/00840

PCT/US01/19166

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

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

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

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

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

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

A fragment of SEQ ID NO:4-6 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:4-6, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:4-6 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:4-6 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:4-6 and the region of SEQ ID NO:4-6 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

WO 02/00840

PCT/US01/19166

A fragment of SEQ ID NO:1-3 is encoded by a fragment of SEQ ID NO:4-6. A fragment of SEQ ID NO:1-3 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-3. For example, a fragment of SEQ ID NO:1-3 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-3. The precise length of a
5 fragment of SEQ ID NO:1-3 and the region of SEQ ID NO:1-3 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
10 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
15 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
20 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
25 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
30 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

WO 02/00840

PCT/US01/19166

Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

10 *Open Gap: 5 and Extension Gap: 2 penalties*

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

15 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.

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

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e

WO 02/00840

PCT/US01/19166

sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by

5 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:

10 *Matrix: BLOSUM62*

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

15 *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

20 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

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

25 chromosome replication, segregation and maintenance.

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

"Hybridization" refers to the process by which a polynucleotide strand anneals with a

30 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

WO 02/00840

PCT/US01/19166

stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml sheared, denatured salmon sperm DNA.

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

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

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

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

WO 02/00840

PCT/US01/19166

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
5 cellular and systemic defense systems.

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

10 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.

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

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

"Post-translational modification" of an HLYA 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

WO 02/00840

PCT/US01/19166

by cell type depending on the enzymatic milieu of HLYA.

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

5 Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

10 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

15 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

20 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

25 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

30 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

WO 02/00840

PCT/US01/19166

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 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

WO 02/00840

PCT/US01/19166

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 HLYA, nucleic acids encoding HLYA, or fragments thereof may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

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

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

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

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

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

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

WO 02/00840

PCT/US01/19166

transformed cells which express the inserted DNA or RNA for limited periods of time.

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

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternative splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides will generally have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of

WO 02/00840

PCT/US01/19166

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

THE INVENTION

The invention is based on the discovery of new human lyases (HLYA), the polynucleotides encoding HLYA, and the use of these compositions for the diagnosis, treatment, or prevention of cancers and autoimmune/inflammatory, neurological, and eye disorders.

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

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

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

WO 02/00840

PCT/US01/19166

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 human lyases. For example, SEQ ID NO:1 is 59% identical, from residue M1 to residue A239, to human carbonic anhydrase I (GenBank ID g179793) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 6.9e-80, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:1 also contains a eukaryotic-type carbonic anhydrase domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:1 is a carbonic anhydrase. In an alternative example, SEQ ID NO:2 is 59% identical, from residue M1 to residue I459, to *Xenopus laevis* ornithine decarboxylase-2 (GenBank ID g6746571) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 1.9e-148, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:2 also contains an ornithine decarboxylase 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 and MOTIFS analyses provide further corroborative evidence that SEQ ID NO:2 is a carbonic anhydrase. In yet another example, SEQ ID NO:3 is 58% identical, from residue E30 to residue R314, to human adenylate kinase (GenBank ID g3283386) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 5.7e-94, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:3 also contains a eukaryotic-type carbonic anhydrase 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 BLAST-DOMO, BLAST-PRODOM, BLIMPS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:3 is a eukaryotic-type carbonic anhydrase. The algorithms and parameters for the analysis of SEQ ID NO:1-3 are described in Table 7.

As shown in Table 4, the full length polynucleotide sequences of the present invention were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Columns 1 and 2 list the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and the corresponding Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) for each polynucleotide of the invention.

WO 02/00840

PCT/US01/19166

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

10 The identification numbers in Column 5 of Table 4 may refer specifically, for example, to Incyte cDNAs along with their corresponding cDNA libraries. For example, 2698586F6 is the identification number of an Incyte cDNA sequence, and UTRSNOT12 is the cDNA library from which it is derived. In an alternative example, 644715H1 is the identification number of an Incyte cDNA sequence, and BRAINOC01 is the cDNA library from which it is derived. Incyte cDNAs for
15 which cDNA libraries are not indicated were derived from pooled cDNA libraries (e.g., 71672723V1 or 70996811V1). Alternatively, the identification numbers in column 5 may refer to GenBank cDNAs or ESTs (e.g., g1191247) which contributed to the assembly of the full length polynucleotide sequences. Alternatively, the identification numbers in column 5 may refer to coding regions predicted by Genscan analysis of genomic DNA. The Genscan-predicted coding sequences may have been
20 edited prior to assembly. (See Example IV.) Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. (See Example V.) Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon-stretching" algorithm. (See Example V.) In some cases, Incyte cDNA coverage redundant with the
25 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
30 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 HLYA variants. A preferred HLYA variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid

WO 02/00840

PCT/US01/19166

sequence identity to the HLYA amino acid sequence, and which contains at least one functional or structural characteristic of HLYA.

The invention also encompasses polynucleotides which encode HLYA. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:4-6, which encodes HLYA. The polynucleotide sequences of SEQ ID NO:4-6, 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 HLYA. 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 HLYA. 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:4-6 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:4-6. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of HLYA.

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 HLYA, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring HLYA, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode HLYA and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring HLYA under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding HLYA or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding HLYA and its derivatives without altering the encoded amino acid sequences

WO 02/00840

PCT/US01/19166

include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode HLYA and HLYA derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding HLYA 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:4-6 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Applied Biosystems), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Applied Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Applied Biosystems), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) *Short Protocols in Molecular Biology*, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) *Molecular Biology and Biotechnology*, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding HLYA may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) *PCR Methods Applic.* 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising

WO 02/00840

PCT/US01/19166

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

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

20 Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate
25 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
30 encode HLYA may be cloned in recombinant DNA molecules that direct expression of HLYA, 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 HLYA.

WO 02/00840

PCT/US01/19166

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

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

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number
10 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 HLYA, 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
15 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,
20 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 HLYA may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids
25 Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, HLYA itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; and Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved
30 using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the amino acid sequence of HLYA, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

WO 02/00840

PCT/US01/19166

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

5 In order to express a biologically active HLYA, the nucleotide sequences encoding HLYA 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
10 encoding HLYA. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding HLYA. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding HLYA and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may
15 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.*
20 *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 HLYA 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*
25 *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 HLYA. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with
30 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

WO 02/00840

PCT/US01/19166

(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
 5 Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) *Cancer Gen. Ther.* 5(6):350-356; Yu, M. et al. (1993) *Proc. Natl. Acad. Sci. USA* 90(13):6340-6344; Buller, R.M. et al. (1985) *Nature* 317(6040):813-815; McGregor, D.P. et al. (1994)
 10 *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 HLYA. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding HLYA can be achieved using a
 15 multifunctional *E. coli* vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding HLYA 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
 20 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 HLYA are needed, e.g. for the production of antibodies, vectors which direct high level expression of HLYA 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 HLYA. A number of vectors
 25 containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast *Saccharomyces cerevisiae* or *Pichia pastoris*. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra, Bitter, G.A. et al. (1987) *Methods Enzymol.* 153:516-544; and Scorer, C.A. et al. (1994)
 30 *Bio/Technology* 12:181-184.)

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

WO 02/00840

PCT/US01/19166

6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. (1984) *Science* 224:838-843; and Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated
5 transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding HLYA may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader
10 sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses HLYA 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.

15 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.)

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

Any number of selection systems may be used to recover transformed cell lines. These
30 include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* and *apr* cells, respectively. (See, e.g., Wigler, M. et al. (1977) *Cell* 11:223-232; Lowy, I. et al. (1980) *Cell* 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to

WO 02/00840

PCT/US01/19166

methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorosulfuron 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.

10 (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 HLYA is inserted within a marker gene sequence, transformed cells containing sequences encoding HLYA can be identified by the absence of marker gene function. Alternatively,

15 a marker gene can be placed in tandem with a sequence encoding HLYA 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 HLYA and that express HLYA may be identified by a variety of procedures known to those of skill in the art. These

20 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 HLYA using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques

25 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 HLYA is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and

30 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

WO 02/00840

PCT/US01/19166

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 HLYA include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide.

Alternatively, the sequences encoding HLYA, 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 HLYA 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 HLYA may be designed to contain signal sequences which direct secretion of HLYA 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 HLYA may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric HLYA protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of HLYA 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

WO 02/00840

PCT/US01/19166

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 HLYA encoding sequence and the heterologous protein sequence, so that HLYA 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 HLYA 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.

HLYA of the present invention or fragments thereof may be used to screen for compounds that specifically bind to HLYA. At least one and up to a plurality of test compounds may be screened for specific binding to HLYA. 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 HLYA, 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 HLYA 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 HLYA, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or E. coli. Cells expressing HLYA or cell membrane fractions which contain HLYA are then contacted with a test compound and binding, stimulation, or inhibition of activity of either HLYA or the compound is analyzed.

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

WO 02/00840

PCT/US01/19166

or affixed to a solid support, and detecting the binding of HLYA 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.

5 HLYA of the present invention or fragments thereof may be used to screen for compounds that modulate the activity of HLYA. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for HLYA activity, wherein HLYA is combined with at least one test compound, and the activity of HLYA in the presence of a test compound is compared with the activity of HLYA in the absence of the test
10 compound. A change in the activity of HLYA in the presence of the test compound is indicative of a compound that modulates the activity of HLYA. Alternatively, a test compound is combined with an *in vitro* or cell-free system comprising HLYA under conditions suitable for HLYA activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of HLYA may do so indirectly and need not come in direct contact with the test compound. At least one
15 and up to a plurality of test compounds may be screened.

In another embodiment, polynucleotides encoding HLYA or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent Number 5,175,383 and U.S. Patent Number 5,767,337.) For
20 example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (*neo*; Capecchi, M.R. (1989) *Science* 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the
25 Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) *Clin. Invest.* 97:1999-2002; Wagner, K.U. et al. (1997) *Nucleic Acids Res.* 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce
30 heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding HLYA 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

WO 02/00840

PCT/US01/19166

lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) Science 282:1145-1147).

Polynucleotides encoding HLYA 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 HLYA is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress HLYA, e.g., by secreting HLYA in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) Biotechnol. Annu. Rev. 4:55-74).

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of HLYA and human lyases. In addition, the expression of HLYA is closely associated with diseased breast tissue as well as adult and fetal brain tissue. Therefore, HLYA appears to play a role in cancers and autoimmune/inflammatory, neurological, and eye disorders. In the treatment of disorders associated with increased HLYA expression or activity, it is desirable to decrease the expression or activity of HLYA. In the treatment of disorders associated with decreased HLYA expression or activity, it is desirable to increase the expression or activity of HLYA.

Therefore, in one embodiment, HLYA 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 HLYA. Examples of such disorders include, but are not limited to, a cancer such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable

WO 02/00840

PCT/US01/19166

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,

5 complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; 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

10 pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous

15 sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia

20 gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; and an eye disorder such as ocular hypertension and glaucoma.

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

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

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

In still another embodiment, an agonist which modulates the activity of HLYA may be administered to a subject to treat or prevent a disorder associated with decreased expression or

WO 02/00840

PCT/US01/19166

activity of HLYA including, but not limited to, those listed above.

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

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

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

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

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with HLYA or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and *Corynebacterium parvum* are especially preferable.

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

WO 02/00840

PCT/US01/19166

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

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

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

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

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

30 Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between HLYA and its

WO 02/00840

PCT/US01/19166

specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering HLYA 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 HLYA. Affinity is expressed as an association constant, K_a , which is defined as the molar concentration of HLYA-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 HLYA epitopes, represents the average affinity, or avidity, of the antibodies for HLYA. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular HLYA epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^8 to 10^{12} L/mole are preferred for use in immunoassays in which the HLYA-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 HLYA, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of HLYA-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding HLYA, 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 HLYA. 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 HLYA. (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

WO 02/00840

PCT/US01/19166

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

In another embodiment of the invention, polynucleotides encoding HLYA may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) *Science* 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) *Science* 270:475-480; Bordignon, C. et al. (1995) *Science* 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) *Cell* 75:207-216; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:643-666; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:667-703), thalassamias, 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 HLYA expression or regulation causes disease, the expression of HLYA 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 HLYA are treated by constructing mammalian expression vectors encoding HLYA and introducing these vectors by mechanical means into HLYA-deficient cells. Mechanical transfer technologies for

WO 02/00840

PCT/US01/19166

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).

5 Expression vectors that may be effective for the expression of HLYA include, but are not limited to, the pCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). HLYA may be expressed
10 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
15 ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and Blau, H.M. *supra*), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding HLYA from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID
20 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
25 standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to HLYA expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding HLYA 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
30 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

WO 02/00840

PCT/US01/19166

an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) *J. Virol.* 61:1647-1650; Bender, M.A. et al. (1987) *J. Virol.* 61:1639-1646; Adam, M.A. and A.D. Miller (1988) *J. Virol.* 62:3802-3806; Dull, T. et al. (1998) *J. Virol.* 72:8463-8471; Zufferey, R. et al. (1998) *J. Virol.* 72:9873-9880). U.S. Patent Number 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., CD4⁺ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) *J. Virol.* 71:7020-7029; Bauer, G. et al. (1997) *Blood* 89:2259-2267; Bonyhadi, M.L. (1997) *J. Virol.* 71:4707-4716; Ranga, U. et al. (1998) *Proc. Natl. Acad. Sci. USA* 95:1201-1206; Su, L. (1997) *Blood* 89:2283-2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding HLYA to cells which have one or more genetic abnormalities with respect to the expression of HLYA. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) *Transplantation* 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent Number 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999) *Annu. Rev. Nutr.* 19:511-544 and Verma, I.M. and N. Somia (1997) *Nature* 18:389:239-242, both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding HLYA to target cells which have one or more genetic abnormalities with respect to the expression of HLYA. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing HLYA to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) *Exp. Eye Res.* 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent Number 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent Number 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a

WO 02/00840

PCT/US01/19166

cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) *J. Virol.* 73:519-532 and Xu, H. et al. (1994) *Dev. Biol.* 163:152-161, hereby incorporated by reference. The manipulation of
5 cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to
10 deliver polynucleotides encoding HLYA to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) *Curr. Opin. Biotechnol.* 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full length genomic RNA,
15 resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for HLYA into the alphavirus genome in place of the capsid-coding region results in the production of a large number of HLYA-coding RNAs and the synthesis of high levels of HLYA in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a
20 persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) *Virology* 228:74-83). The wide host range of alphaviruses will allow the introduction of HLYA into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of
25 manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can
30 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,

WO 02/00840

PCT/US01/19166

Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

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

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

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

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding HLYA. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not

WO 02/00840

PCT/US01/19166

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 HLYA expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding HLYA may be therapeutically useful, and in the treatment of disorders associated with decreased HLYA expression or activity, a compound which specifically promotes expression of the polynucleotide encoding HLYA 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 HLYA 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 HLYA are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding HLYA. 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).

WO 02/00840

PCT/US01/19166

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 HLYA, antibodies to HLYA, and mimetics, agonists, antagonists, or inhibitors of HLYA.

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

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

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

Specialized forms of compositions may be prepared for direct intracellular delivery of macromolecules comprising HLYA or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the

WO 02/00840

PCT/US01/19166

macromolecule. Alternatively, HLYA or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

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

10 A therapeutically effective dose refers to that amount of active ingredient, for example HLYA or fragments thereof, antibodies of HLYA, and agonists, antagonists or inhibitors of HLYA, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED₅₀ (the dose therapeutically effective in 50% of the population) or LD₅₀ (the dose
15 lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD₅₀/ED₅₀ ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED₅₀ with little or no toxicity.
20 The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the
25 severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

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

WO 02/00840

PCT/US01/19166

conditions, locations, etc.

DIAGNOSTICS

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

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

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

WO 02/00840

PCT/US01/19166

invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:4-6 or from genomic sequences including promoters, enhancers, and introns of the HLYA gene.

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

Polynucleotide sequences encoding HLYA may be used for the diagnosis of disorders associated with expression of HLYA. Examples of such disorders include, but are not limited to, a cancer such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hyper eosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial

WO 02/00840

PCT/US01/19166

thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; and an eye disorder such as ocular hypertension and glaucoma. The polynucleotide sequences encoding HLYA 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 HLYA expression. Such qualitative or quantitative methods are well known in the art.

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

WO 02/00840

PCT/US01/19166

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,

- 5 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.

- 10 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.

- 15 Additional diagnostic uses for oligonucleotides designed from the sequences encoding HLYA 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 HLYA, or a fragment of a polynucleotide complementary to the polynucleotide encoding HLYA, and will be employed under optimized conditions for identification of a specific gene or
20 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 HLYA may be used to detect single nucleotide polymorphisms (SNPs). SNPs are
25 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 HLYA 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
30 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

WO 02/00840

PCT/US01/19166

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

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

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

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

WO 02/00840

PCT/US01/19166

5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

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

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

In one embodiment, the toxicity of a test compound is assessed by treating a biological 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

WO 02/00840

PCT/US01/19166

quantified. The transcript levels in the treated biological sample are compared with levels in an untreated biological sample. Differences in the transcript levels between the two samples are indicative of a toxic response caused by the test compound in the treated sample.

Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, *supra*). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for 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 HLYA to quantify the levels of HLYA expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) *Anal. Biochem.* 270:103-111; 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

WO 02/00840

PCT/US01/19166

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

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

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

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

In another embodiment of the invention, nucleic acid sequences encoding HLYA 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

WO 02/00840

PCT/US01/19166

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 HLYA on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

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

WO 02/00840

PCT/US01/19166

synthesized on a solid substrate. The test compounds are reacted with HLYA, or fragments thereof, and washed. Bound HLYA is then detected by methods well known in the art. Purified HLYA 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
5 solid support.

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

10 In additional embodiments, the nucleotide sequences which encode HLYA 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
15 description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications and publications, mentioned above and below, in particular U.S. Ser. No. 60/213,383, U.S. Ser. No. 60/215,544, and U.S. Ser. No. 60/222,818, are
20 expressly incorporated by reference herein.

EXAMPLES

I. Construction of cDNA Libraries

Incyte cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD
25 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
30 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

WO 02/00840

PCT/US01/19166

(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 CL4E column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), or pINCY (Incyte Genomics, Palo Alto CA), or derivatives thereof. Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 α , DH10B, or ElectroMAX DH10B from Life Technologies.

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

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

III. Sequencing and Analysis

WO 02/00840

PCT/US01/19166

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

5 MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the

10 ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, *supra*, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

15 The polynucleotide sequences derived from Incyte cDNAs were validated by removing 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

20 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.

25 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

30 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

WO 02/00840

PCT/US01/19166

Markov model (HMM)-based protein family databases such as PFAM. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are generated using default parameters specified by the CLUSTAL algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

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

The programs described above for the assembly and analysis of full length polynucleotide and polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:4-6. 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 human lyases were initially identified by running the Genscan gene identification program against public genomic sequence databases (e.g., gbpri and gbhtg). Genscan is a general-purpose gene identification program which analyzes genomic DNA sequences from a variety of organisms (See Burge, C. and S. Karlin (1997) *J. Mol. Biol.* 268:78-94, and Burge, C. and S. Karlin (1998) *Curr. Opin. Struct. Biol.* 8:346-354). The program concatenates predicted exons to form an assembled cDNA sequence extending from a methionine to a stop codon. The output of Genscan is a FASTA database of polynucleotide and polypeptide sequences. The maximum range of sequence for Genscan to analyze at once was set to 30 kb. To determine which of these Genscan predicted cDNA sequences encode human lyases, the encoded polypeptides were analyzed by querying against PFAM models for human lyases. Potential human lyases were also identified by homology to Incyte cDNA sequences that had been annotated as human lyases. 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.

WO 02/00840

PCT/US01/19166

BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide sequences were obtained by assembling Genscan-predicted coding sequences with Incyte cDNA sequences and/or public cDNA sequences using the assembly process described in Example III. Alternatively, full length polynucleotide sequences were derived entirely from edited or unedited Genscan-predicted coding sequences.

V. Assembly of Genomic Sequence Data with cDNA Sequence Data

"Stitched" Sequences

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

WO 02/00840

PCT/US01/19166

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 HLYA Encoding Polynucleotides

10 The sequences which were used to assemble SEQ ID NO:4-6 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:4-6 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available 15 from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO., to that map location.

Map locations are represented by ranges, or intervals, of human chromosomes. The map 20 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 25 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

30 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.)

WO 02/00840

PCT/US01/19166

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

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

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

10 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
 15 (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
 20 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 HLYA are analyzed with respect to the tissue sources from which they were derived. For example, some full length sequences are
 25 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;
 30 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,

WO 02/00840

PCT/US01/19166

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 HLYA. cDNA sequences and cDNA library/tissue information are found in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA).

5 **VIII. Extension of HLYA 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
10 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
15 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
20 enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4
25 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

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

WO 02/00840

PCT/US01/19166

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 designed for such extension, and an appropriate genomic library.

IX. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:4-6 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 µCi of [γ -³²P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10⁷ counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or

WO 02/00840

PCT/US01/19166

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
5 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.

X. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing
10 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 (Sчена (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
15 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.)

20 Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection.

25 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
30 in detail below.

Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)⁺ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)⁺ RNA sample is

WO 02/00840

PCT/US01/19166

reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/ μ l oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/ μ l RNase inhibitor, 500 μ M dATP, 500 μ M dGTP, 500 μ M dTTP, 40 μ M dCTP, 40 μ M dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)⁺ RNA with

5 GEMBRIGHT kits (Incyte). Specific control poly(A)⁺ RNAs are synthesized by *in vitro* transcription from non-coding yeast genomic DNA. After incubation at 37°C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85°C to the stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc.

10 (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.

WO 02/00840

PCT/US01/19166

Microarray Preparation

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

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

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

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

WO 02/00840

PCT/US01/19166

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
5 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
10 emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on
15 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
20 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
25 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
30 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 HLYA-encoding sequences, or any parts thereof, are used

WO 02/00840

PCT/US01/19166

to detect, decrease, or inhibit expression of naturally occurring HLYA. 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 HLYA. 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 HLYA-encoding transcript.

XII. Expression of HLYA

Expression and purification of HLYA is achieved using bacterial or virus-based expression systems. For expression of HLYA in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac (tac)* hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express HLYA upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of HLYA 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 HLYA 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, HLYA 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 HLYA 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-

WO 02/00840

PCT/US01/19166

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 HLYA obtained by these methods can be used directly in the assays shown in Examples XVI, XVII, and XVIII, where applicable.

5 XIII. Functional Assays

HLYA function is assessed by expressing the sequences encoding HLYA at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include PCMV SPORT (Life Technologies) and PCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μ g of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μ g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of HLYA on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding HLYA 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 HLYA and other genes of interest can be analyzed by northern

WO 02/00840

PCT/US01/19166

analysis or microarray techniques.

XIV. Production of HLYA Specific Antibodies

HLYA 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 HLYA amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, *supra*, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Applied Biosystems) using Fmoc chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, *supra*.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for anti-peptide and anti-HLYA activity by, for example, binding the peptide or HLYA to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XV. Purification of Naturally Occurring HLYA Using Specific Antibodies

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

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

XVI. Identification of Molecules Which Interact with HLYA

HLYA, 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 HLYA, washed,

WO 02/00840

PCT/US01/19166

and any wells with labeled HLYA complex are assayed. Data obtained using different concentrations of HLYA are used to calculate values for the number, affinity, and association of HLYA with the candidate molecules.

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

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

XVII. Demonstration of HLYA activity

An assay for carbonic anhydrase activity of HLYA uses the fluorescent pH indicator 8-hydroxypyrene-1,3,6-trisulfonate (pyranine) in combination with stopped-flow fluorometry to measure carbonic anhydrase activity (Shingles, et al. 1997, *Anal. Biochem.* 252: 190-197). A pH 6.0 solution is mixed with a pH 8.0 solution and the initial rate of bicarbonate dehydration is measured. Addition of carbonic anhydrase to the pH 6.0 solution enables the measurement of the initial rate of activity at physiological temperatures with resolution times of 2 ms. Shingles et al. used this assay to resolve differences in activity and sensitivity to sulfonamides by comparing mammalian carbonic anhydrase isoforms. The fluorescent technique's sensitivity allows the determination of initial rates with a protein concentration as little as 65 ng/ml.

Ornithine decarboxylase activity of HLYA is assayed by measuring the release of CO₂ from L-[1-¹⁴C]-ornithine (Reddy, S.G et al. (1996) *J. Biol. Chem.* 271:24945-24953). Activity is measured in 200 µl assay buffer (50 mM Tris/HCl, pH 7.5, 0.1 mM EDTA, 2 mM dithiothreitol, 5 mM NaF, 0.1% Brij35, 1 mM PMSF, 60 µM pyridoxal-5-phosphate) containing 0.5 mM L-ornithine plus 0.5 µCi L-[1-¹⁴C]ornithine. The reactions are stopped after 15-30 minutes by addition of 1 M citric acid, and the ¹⁴CO₂ evolved is trapped on a paper disk filter saturated with 20 µl of 2 N NaOH. The radioactivity on the disks is determined by liquid scintillation spectrophotography. The amount of ¹⁴CO₂ released is proportional to ornithine decarboxylase activity of HLYA.

XVIII. Identification of HLYA Agonists and Antagonists

Agonists or antagonists of HLYA activation or inhibition may be tested using the assay described in section XVII. Agonists cause an increase in HLYA activity and antagonists cause a decrease in HLYA activity.

Various modifications and variations of the described methods and systems of the invention

WO 02/00840

PCT/US01/19166

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
5 to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

WO 02/00840

PCT/US01/19166

Table 1

Inocyte Project ID	Polyresinide SER ID NO.	Inocyte Polyresinide ID	Polyamidoamine SER ID NO.	Inocyte Polyamidoamine ID
6338333	1	6338333C01	4	6338333C01
1415322	2	1415322C01	5	1415322C01
3267634	3	3267634C01	6	3267634C01

WO 02/00840

PCT/US01/19166

Table 2

Polypeptide SEQ. ID. NO.	Incyte Peptide ID	GenBank ID NO.	Probability score	GenBank Homolog
1	636933CD1	91249093	1.00E-130	[Fl] [Mus musculus] carbonic anhydrase XIII Hewitt-Emmett, D. (2000) (in) Chegwidden, W.R., Carter, N.D. and Edwards, Y.H. (Eds.); THE CARBONIC ANHYDRASES. NEW HORIZONS: 29-76; Birkhauser, Verlag, Basel. [Homo sapiens] carbonic anhydrase I (CC 4.2.1.1)
2	1415322CD1	96746571	1.90E-148	Lowe, N. et al. (1990) Structure and methylation patterns of the gene encoding human carbonic anhydrase I. Gene 93 (2), 277-283 [Xenopus laevis] ornithine decarboxylase-2 Bernstein, H.G. and Muller, M. (1995) Prog. Neurobiol. 57:485-505 Medina, M.A. et al. (1999) Biochem Pharmacol 57:1341-1344 [Fl] [Homo sapiens] carbonic anhydrase-related protein X Hewitt-Emmett, D. and Ashlan, R.F. (1997) Mol. Phylogenet. Evol. 5: 50-77
3	3267634CD1	913094699	0	[Homo sapiens] carbonic anhydrase-related protein 2

WO 02/00840

PCT/US01/19166

Table 3

SEQ ID NO.	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1	6336333CD1	242	S100 S126 S131 S167 S20 S44 S49 S56 S88 T178 Y41	N218 N76	Eukaryotic-type carbonic anhydrase domain: W6-A239 Eukaryotic-type carbonic anhydrases signature BL00162: W17-S47; Y52-Y74; Y89-S125; K128-G152; D191-Q223 Eukaryotic-type carbonic anhydrases signature: G82-A143 LYASE CARBONATE ZINC PRECURSOR SIGNAL PROTEIN GLYCOPROTEIN PD000865: E10-A238 CARBONIC ANHYDRASE EM00356 NM0836125-261: D25-A239 Eukaryotic carbonic anhydrases motif: S106-V122	HMMER_FFAM BLIMPS_BLOCKS
2	1415322CD1	460	S18 S140 S169 S174 S177 S181 S216 S258 S314 S316 S317 S318 S319 S320 S321 S322 S323 S324 S325 S326 S327 S328 S329 S330 S331 S332 S333 S334 S335 S336 S337 S338 S339 S340 S341 S342 S343 S344 S345 S346 S347 S348 S349 S350 S351 S352 S353 S354 S355 S356 S357 S358 S359 S360 S361 S362 S363 S364 S365 S366 S367 S368 S369 S370 S371 S372 S373 S374 S375 S376 S377 S378 S379 S380 S381 S382 S383 S384 S385 S386 S387 S388 S389 S390 S391 S392 S393 S394 S395 S396 S397 S398 S399 S400 S401 S402 S403 S404 S405 S406 S407 S408 S409 S410 S411 S412 S413 S414 S415 S416 S417 S418 S419 S420 S421 S422 S423 S424 S425 S426 S427 S428 S429 S430 S431 S432 S433 S434 S435 S436 S437 S438 S439 S440 S441 S442 S443 S444 S445 S446 S447 S448 S449 S450 S451 S452 S453 S454 S455 S456 S457 S458 S459 S460	N308	767-G85 Orn/DAP/Arg Decarboxylase 2_1 R210-G241 Orn/DAP/Arg Decarboxylase 2_2 V80-C90, V193-C203, M394-C404 Prokar_Lipoprotein attachment site S11B Orn_DAP_Arg_Rec: F41-T401 Pyridoxal-dependent decarboxylase BL00178: F41-T401 BL00179: F41-T401 BL00180: H180-C203 BL00181: R235-C244 BL00182: V270-S288 BL00183: I356-D365 BL00184: G379-T401 Orn/DAP/Arg decarboxylase	BLAST_DOMO MOTIFS MOTIFS HMMER_FFAM BLIMPS_BLOCKS

WO 02/00840

PCT/US01/19166

Table 4

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment (s)	Sequence Fragments	5' Position	3' Position
4	633833CBI	911	1-28, 741- 876, 870- 911	71672723VI 70513545VI	429	911
5	141532CBI	2064	1058-1146, 430-1032	2694586F6 (PTESMOT12) 91191247 1401356HI (BRAL7UT08) 6854836HI (BRAIFEN08) 5273763HI (OVARIND02) 6312955HI (NEHD0DN03) 5109549F6 (BROSUS19) 3245804F6 (BRAINOT19)	1758 1550 1887 1075 714 877 312	2064 2064 1781 987 1540 900 416
6	3267634CBI	2938	1863-2033, 1-786	7099681IV1 6447115HI (BRAINOC01) 6328316HI (BRANDIND01) 3246722F7 (BRAINOT19) 3267634HI (BRAINOT20) 6926620HI (SINITMR01) 12952606 (FGANN0703) 70995406VI	1060 770 1695 404 1 216 2372 1550	1635 1489 2395 1060 236 834 2938 2217

WO 02/00840

PCT/US01/19166

Table 5

Polynucleotide SEQ. ID NO.	Project ID	Representative Library
4	638833CELL	BRAINOT07
5	1415222CELL	BRAINOT03
6	3267634CELL	BRAINOT09

WO 02/00840

PCT/US01/19166

Table 5

Polynucleotide SEQ. ID. NO.	Incyte Project ID	Representative Library
5	31523CB1	BRANNO7
6	326734CB1	BRANNO9

WO 02/00840

PCT/US01/19166

Table 6

Library	Vector	Library Description
BRALN0T09	pINCY	Library was constructed using RNA isolated from brain tissue removed from a Caucasian male fetus, who died at 23 weeks' gestation.
BRAXN0T03	pINCY	Library was constructed using RNA isolated from sensory-motor cortex tissue removed from the brain of a 39-year-old Caucasian male who died from cardiac failure. Pathology indicated severe atherosclerosis, atherosclerotic plaques with multiple microinfarctions of the cerebral neocortex. Pathology also included dilated cardiomyopathy, congestive heart failure, cardiomegaly and an enlarged spleen and liver.
BRSTN0T07	pINCY	Library was constructed using RNA isolated from diseased breast tissue removed from a 43-year-old Caucasian female during a unilateral simple mastectomy. Pathology indicated mildly proliferative fibrocystic changes with epithelial hyperplasia, papillomatosis, and duct ectasia. Pathology also indicated moderate ductal carcinoma in situ, nuclear grade 3 mammary adenocarcinoma with extensive comedo necrosis. Family history included epilepsy, cardiovascular disease, and type II diabetes.

WO 02/00840

PCT/US01/19166

Table 7

Program	Description	Reference	Parameter Threshold
ABLFACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABIPARACEL FDF	A Fast Data Finder useful in computing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	ESTs: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, tfasta, fastx, tfastx, and asearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 185:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E value=1.0E-6 Assembled ESTs: fasta.Identity= 95% or greater and Match length=200 bases or greater; fastx E values=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 Alwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Naishell, Cambridge Univ. Press, pp. 1-350.	PFAM hits: Probability value= 1.0E-3 or less Signal peptide hits: Score= 0 or greater

WO 02/00840

PCT/US01/19166

Table 7 (cont.)

Program	Description	Reference	Parameter Threshold
ProfitScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1990) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality scores, CCG-specified "HHH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Plus Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Scores= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Goodson, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielsen, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Scores=3.5 or greater
Modis	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

87

WO 02/00840

PCT/US01/19166

What is claimed is:

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

WO 02/00840

PCT/US01/19166

10. An isolated antibody which specifically binds to a polypeptide of claim 1.
11. An isolated polynucleotide selected from the group consisting of:
- 5 a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6,
b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6,
c) a polynucleotide complementary to a polynucleotide of a),
d) a polynucleotide complementary to a polynucleotide of b), and
10 e) an RNA equivalent of a)-d).
12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 11.
- 15 13. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:
- a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex
20 is formed between said probe and said target polynucleotide or fragments thereof, and
b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.
14. A method of claim 13, wherein the probe comprises at least 60 contiguous nucleotides.
25
15. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:
- a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and
30 b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.
16. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable

WO 02/00840

PCT/US01/19166

excipient.

17. A composition of claim 16, wherein the polypeptide has an amino acid sequence selected from the group consisting of SEQ ID NO:1-3.

5

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

19. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- b) detecting agonist activity in the sample.

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

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

22. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- b) detecting antagonist activity in the sample.

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

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

25. A method of screening for a compound that specifically binds to the polypeptide of claim

WO 02/00840

PCT/US01/19166

I, said method comprising the steps of:

- 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.

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

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

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

- a) exposing a sample comprising the target polynucleotide to a compound, 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.

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

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

WO 02/00840

PCT/US01/19166

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

5

29. A diagnostic test for a condition or disease associated with the expression of HLYA in a biological sample comprising the steps of:

a) combining the biological sample with an antibody of claim 10, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex; and
10 b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.

30. The antibody of claim 10, wherein the antibody is:

a) a chimeric antibody,
15 b) a single chain antibody,
c) a Fab fragment,
d) a F(ab')₂ fragment, or
e) a humanized antibody.

20 31. A composition comprising an antibody of claim 10 and an acceptable excipient.

32. A method of diagnosing a condition or disease associated with the expression of HLYA in a subject, comprising administering to said subject an effective amount of the composition of claim 31.

25

33. A composition of claim 31, wherein the antibody is labeled.

34. A method of diagnosing a condition or disease associated with the expression of HLYA in a subject, comprising administering to said subject an effective amount of the composition of claim 33.

30

35. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 10 comprising:

WO 02/00840

PCT/US01/19166

- a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, or an immunogenic fragment thereof, under conditions to elicit an antibody response;
- b) isolating antibodies from said animal; and
- 5 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:1-3.
36. An antibody produced by a method of claim 35.
- 10 37. A composition comprising the antibody of claim 36 and a suitable carrier.
38. A method of making a monoclonal antibody with the specificity of the antibody of claim 10 comprising:
- 15 a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, 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-
- 20 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:1-3.
- 25 39. A monoclonal antibody produced by a method of claim 38.
40. A composition comprising the antibody of claim 39 and a suitable carrier.
41. The antibody of claim 10, wherein the antibody is produced by screening a Fab expression
- 30 library.
42. The antibody of claim 10, wherein the antibody is produced by screening a recombinant immunoglobulin library.

WO 02/00840

PCT/US01/19166

43. A method for detecting a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3 in a sample, comprising the steps of:
- a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and
 - 5 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:1-3 in the sample.
44. A method of purifying a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3 from a sample, the method comprising:
- 10 a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and
 - b) separating the antibody from the sample and obtaining the purified polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3.
- 15 45. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:1.
46. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:2.
47. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:3.
- 20 48. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:4.
49. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:5.
- 25 50. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:6.

WO 02/00840

PCT/US01/19166

<110> INCYTE GENOMICS, INC.

THORNTON, Michael
 RAMKUMAR, Jayalaxmi
 TRIBOULEY, Catherine M.
 YUE, Henry
 NGUYEN, Dannie B.
 YAO, Monique G.
 PATTERSON, Chandra
 GANDHI, Ameena R.
 BURFORD, Neil
 THANGAVELU, Kavitha
 BAUGHN Mariah R.

<120> HUMAN LYASES

<130> PI-0137 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/213,383; 60/215,544; 60/222,818

<151> 2000-06-23; 2000-06-30; 2000-08-04

<160> 6

<170> PERL Program

<210> 1

<211> 242

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 6338333CD1

<400> 1

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

WO 02/00840

PCT/US01/19166

Ile Gly Glu Pro Asn Ser Gln Leu Gln Lys Ile Thr Asp Thr Leu	
155	160
Asp Ser Ile Lys Glu Lys Gly Lys Gln Thr Arg Phe Thr Asn Phe	165
170	175
Asp Leu Leu Ser Leu Leu Pro Pro Ser Trp Asp Tyr Trp Thr Tyr	180
185	190
Pro Gly Ser Leu Thr Val Pro Pro Leu Leu Glu Ser Val Thr Trp	195
200	205
Ile Val Leu Lys Gln Pro Ile Asn Ile Ser Ser Gln Gln Leu Ala	210
215	220
Lys Phe Arg Ser Leu Leu Cys Thr Ala Glu Gly Glu Ala Ala Ala	225
230	235
Phe Leu	240

<210> 2

<211> 460

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1415322CD1

<400> 2

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

WO 02/00840

PCT/US01/19166

```

230                235                240
Gly Thr Glu Gly Ala Lys Val Arg Phe Glu Glu Ile Ala Ser Val
245                250                255
Ile Asn Ser Ala Leu Asp Leu Tyr Phe Pro Glu Gly Cys Gly Val
260                265                270
Asp Ile Phe Ala Glu Leu Gly Arg Tyr Tyr Val Thr Ser Ala Phe
275                280                285
Thr Val Ala Val Ser Ile Ile Ala Lys Lys Glu Val Leu Leu Asp
290                295                300
Gln Pro Gly Arg Glu Glu Glu Asn Gly Ser Thr Ser Lys Thr Ile
305                310                315
Val Tyr His Leu Asp Glu Gly Val Tyr Gly Ile Phe Asn Ser Val
320                325                330
Leu Phe Asp Asn Ile Cys Pro Thr Pro Ile Leu Gln Lys Lys Pro
335                340                345
Ser Thr Glu Gln Pro Leu Tyr Ser Ser Ser Leu Trp Gly Pro Ala
350                355                360
Val Asp Gly Cys Asp Cys Val Ala Glu Gly Leu Trp Leu Pro Gln
365                370                375
Leu His Val Gly Asp Trp Leu Val Phe Asp Asn Met Gly Ala Tyr
380                385                390
Thr Val Gly Met Gly Ser Pro Phe Trp Gly Thr Gln Ala Cys His
395                400                405
Ile Thr Tyr Ala Met Ser Arg Val Ala Trp Glu Ala Leu Arg Arg
410                415                420
Gln Leu Met Ala Ala Glu Gln Glu Asp Asp Val Glu Gly Val Cys
425                430                435
Lys Pro Leu Ser Cys Gly Trp Glu Ile Thr Asp Thr Leu Cys Val
440                445                450
Gly Pro Val Phe Thr Pro Ala Ser Ile Met
455                460

```

<210> 3

<211> 328

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3267634CD1

<400> 3

```

Met Glu Ile Val Trp Glu Val Leu Phe Leu Leu Gln Ala Asn Phe
1          5          10          15
Ile Val Cys Ile Ser Ala Gln Gln Asn Ser Pro Lys Ile His Glu
20         25         30
Gly Trp Trp Ala Tyr Lys Glu Val Val Gln Gly Ser Phe Val Pro
35         40         45
Val Pro Ser Phe Trp Gly Leu Val Asn Ser Ala Trp Asn Leu Cys
50         55         60
Ser Val Gly Lys Arg Gln Ser Pro Val Asn Ile Glu Thr Ser His
65         70         75
Met Ile Phe Asp Pro Phe Leu Thr Pro Leu Arg Ile Asn Thr Gly
80         85         90
Gly Arg Lys Val Ser Gly Thr Met Tyr Asn Thr Gly Arg His Val
95         100        105

```

3/6

WO 02/00840

PCT/US01/19166

Ser Leu Arg Leu Asp Lys Glu His Leu Val Asn Ile Ser Gly Gly
 110 115 120
 Pro Met Thr Tyr Ser His Arg Leu Glu Glu Ile Arg Leu His Phe
 125 130 135
 Gly Ser Glu Asp Ser Gln Gly Ser Glu His Leu Leu Asn Gly Gln
 140 145 150
 Ala Phe Ser Gly Glu Val Gln Leu Ile His Tyr Asn His Glu Leu
 155 160 165
 Tyr Thr Asn Val Thr Glu Ala Ala Lys Ser Pro Asn Gly Leu Val
 170 175 180
 Val Val Ser Ile Phe Ile Lys Val Ser Asp Ser Ser Asn Pro Phe
 185 190 195
 Leu Asn Arg Met Leu Asn Arg Asp Thr Ile Thr Arg Ile Thr Tyr
 200 205 210
 Lys Asn Asp Ala Tyr Leu Leu Gln Gly Leu Asn Ile Glu Glu Leu
 215 220 225
 Tyr Pro Glu Thr Ser Ser Phe Ile Thr Tyr Asp Gly Ser Met Thr
 230 235 240
 Ile Pro Pro Cys Tyr Glu Thr Ala Ser Trp Ile Ile Met Asn Lys
 245 250 255
 Pro Val Tyr Ile Thr Arg Met Gln Met His Ser Leu Arg Leu Leu
 260 265 270
 Ser Gln Asn Gln Pro Ser Gln Ile Phe Leu Ser Met Ser Asp Asn
 275 280 285
 Phe Arg Pro Val Gln Pro Leu Asn Asn Arg Cys Ile Arg Thr Asn
 290 295 300
 Ile Asn Phe Ser Leu Gln Gly Lys Asp Cys Pro Asn Asn Arg Ala
 305 310 315
 Gln Lys Leu Gln Tyr Arg Val Asn Glu Trp Leu Leu Lys
 320 325

<210> 4
 <211> 911
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 6338333CEI

<400> 4
 cggaattcgg ctcgagttcc accccgaggg accatgtcga ggctcagctg gggataccgc 60
 gagcacaacg gtcctattca ctggaaggaa tttttcccta ttgctgatgg tgatcagcaa 120
 tctccaattg agattaaaac caaagaagtg aaatatgact cttccctccg accacttagt 180
 atcaagatg acccaagctc agctaaaatc atcagcaaca gggccattc ctccaatggt 240
 gactttgatg acacagagaa caaatcagtt ctggtgggtg gtcctctcac tggaaagtac 300
 agttacggc aggttcacct tcaactgggg tccgctgatg accaaggctc cgagcaacata 360
 gtagatggag tgagctatgc tgcagagctc catgttcttc actggaattc agacaatac 420
 cccagctttg ttgaggcagc tcatgaacca gatggactgg ctgtcttggg agtgttttta 480
 cagattgggtg aaccataatc ocaactgcaa aagattactg acactttgga ttccattaaa 540
 gaaaagggta aacaaactcg attcacaaat ttgacatcat tgtctctgct tccaccatcc 600
 tggactactt ggacatatcc tggttctctt acagttccac ctcttcttga gagtgtca 660
 tggattgttt taagcaacc tataaacatc agctctcaac agctggccaa atttgcagb 720
 ctctgtgca cagcggaggg tgaagcagca gctttctgt gatagagct cactctgtca 780
 cccagctgg agggcagtg tacaatctg gctaattgca gcctccaact cotggactca 840
 agtgatctc ccactcage ctccagagtc ctgaccactg gcatacttt tcccaaatgc 900

WO 02/00840

PCT/US01/19166

cagggggaaa a

911

<210> 5
<211> 2064
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1415322CB1

```

<400> 5
gtggtgctg ggctggggg gcycagggc cgggaccga gcccgggaa gcgagagac 60
ggaggcgcc aggatccgat tcaactccct gggagacct tggccgaag cctgttaat 120
gcgttttaag cagaggcctc ggctccgcaa ctgcactcc tctcggggg gtggcacaag 180
ttcggaggtc accggggacc cccctagca ggcgctggc ctctggccc cgggaagag 240
gacggagttt gtgtgttgca tactttctaa ggcggggct gcagcagcg ctccatccag 300
cccctcagct cctcctgcaa ggcattggct gctacctgag tgaatcgagc ttgtgatgg 360
tggagggagg ctccagtacc ctagacctgc tgaaggaact cactctgggg gcctcagag 420
ccaccacgga cggagtagct gccttctctg tggctgacct ggggtccata gtgaggagc 480
acttttctt lctgaagtgc ctgcccagag tccggccctt ttatgctgtc aagtgcaca 540
gcagcccagg tgtgctgaa gttctggccc agctggggc gggctttagc tgtccaaca 600
aggcagagat gtagttggtc cagcatattg gaatccctgc cagttagatc atctggcca 660
accctgttaa gcaaatgca cagatcaaat atgctgccaa gcattggatc cagctgtga 720
gctttgacaa tgagatggag ctggcaagg tggtaagag ccaccocagt gccaaatgg 780
ttctgtgcat tgcctaccgat gaactccact cctgagctg cctgagccta aagtttgag 840
tgtcactgaa atcctgcaga caectccttg aaaatgcgaa gaagcaccat gtggaggtg 900
tgggtgtgag ttttcacatt ggcagtgctt gtcctgacct tcaggcctat gctcagcca 960
tcgcagacgc ccgctcctg tttgaaatgg gcaccgagct ggtcacaag atgcacgttc 1020
tggacctggg tgggtgcttc cctggcacag aaggggcca agtgagattt gaagagattg 1080
cttccgtgat caactcagcc ttggacctgt acttcccaga gggctgtggc gtggacatct 1140
ttgtgagct ggggagctac taactgacct cggccttcac tgtggcagtc agcatcatg 1200
ccaagaagga ggttctgcta gaccagcctg gcaggggaga ggaanaatgt tccactcca 1260
agacctcgt gtaccacctt gatgagggcg tgtatgggat cttaactca gtectgttg 1320
acaacatctg cctaccctcc atcctgcaga agaaccatc cacggagcag ccctgtaca 1380
gcagcagcct gtggggcccg gcggttgatg gctgtgattg cgtggctgag ggcctgtggc 1440
tgccgcaact acactgagg gactgtctgg tctttgacaa catgggccc tacactgtg 1500
gcatgggttc cccctttgg gggaccaggg cctgccacat caactatgcc atgtcccgg 1560
tggcctggga agcctgcga aggcagctga tggctgcaga acagggagat gactggagg 1620
gtgtgtgcaa gectctgccc tgcggctggg agatcacaga caccctgtgc gtgggccc 1680
tcttcccacc agcagcagtc atgtgagttg gctccttccc ccccgagaa tcccagcgg 1740
gcctcagaga tgcactcggg agaggtgggg aagatggcag gcaagggtac ccttgcccag 1800
gactctggg cccaccctgc caccocctgc ctccactcgc agtgttctg cctgtcaat 1860
aggaccagtc ttacactcgc tgtagttcaa gtatgcaaca taaatcctg tecttccagc 1920
tgtctctgcc tctctgcag tgcagggggc ctggtcagcc aggtgtggg gtgttctgg 1980
ggtctccttt ggtctccttc ccactttgt aaatataat caataaata aatatttag 2040
tttttaaaa ctgaaaaaaa aaaa 2064

```

<210> 6
<211> 2938
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3267634CB1

WO 02/00840

PCT/US01/19166

```

<400> 6
gcattgagg tggcggagg ctgccgggg ctggcgggtt cgggcaataa tccctgectc 60
tctttctctg tgtgtctgct gtgtctgct cttcccggcc ccccgggaagc aggagaagaa 120
ctgccccgga ggcgagcagc caccctccga ccatgccccg ttgagggggg cggacttcca 180
gggcaacttg cgcgggactg cctgggctta gccagggagc taagcgetcc cgggagcccg 240
gaattgcaag ggcgagcccg gcggggggct atcgtctatg tcttcttggg ggcgcagacc 300
aatcgggggtc tegtthttgc gtggaaagac ccagtgttgg tggcttcagg tggctgtctc 360
cgcgcgcccgc gccgcccgcg ctgctagtgc ggtttccgcc gctggtgca agagaagaga 420
cagcggagcg gggagacctc caaggcagcg aggcacgga catgtgtcag cacatctggg 480
gocacatccg gtcgagcccg aggggagatt tgcgggaaca attcaactg cgatattgat 540
cttgggggtg actgtccctg gccggctgtc ggggtggagt gcgagtgtgc actcgtctgg 600
aagtgtgtgc gagtgtgtat gtgtgtgtgc cgtgtcgggc tcccccttc ccccgtttt 660
cccgtcagtg gatgcacttg gaatgagaat cagaggatgg aatagtctg ggaggtgctt 720
ttctctcttc aagccaattt catcgtctgc atatcagctc aacagaattc accaaaaatc 780
catgaaggct ggtgggcata caaggagggt gtccagggaa gctttgttcc agttccttct 840
ttctggggat tggtaacctc agcttgaat ctttctctg tgggaaacg gcagtgcca 900
gtcaacatag agaccagtca catgatcttc gacccttct tgacacctct tgcatacaac 960
acggggggca ggaaggtcag tggaccactg tacaacactg gaagacactg atccctctgc 1020
ctggacaagg agcacttggt caacatactc ggaggcccca tgacatacag ccacggctg 1080
gagggagacc gactacactt tgggagtgag gacagccaa ggtcggagca cctcctcaat 1140
ggcaggctct tctctgggga ggtgcagctc atccactata accatgagct atatacgaat 1200
gtcacagaag ctgcaaaagc tccaaatgga ttggtggtag ttctatatat tataaaagt 1260
tctgatccat caaacccatt ttttaactga atgctcaaca gagatactat cacaagaata 1320
acataaaaa atgatgcata tttactacag gggcttaata tagaggaact atatccagag 1380
acctctagtt tcatcactta cgtgggtctg atgactatcc caccctgcta tgagacagca 1440
agtgtgatca taatgaacaa acctgtctat ataaccagga tgagatgca ttctctgccc 1500
ctgctcagcc agaacccagc atctcagatc ttctgagca tgagtgaaca cttcaggcct 1560
gtccagccac tcaaaccccg ctgctccgcc accaatatca acttcaagtt acaggggag 1620
gactgtccaa acaaccggag ccagagctt cagtatagag taatgaatg gctctcaag 1680
tagggaacaa agccaagagc aatcccactc cagtgaatg ctcaactgt gaattgact 1740
aacctagaat gtccccttc tttctctct cctcttctt ccccgaagcc tcaattcatc 1800
ttgggattgg ccttttctc atgaaaagtg tctgcaaac catggcagag gaatacattc 1860
ctcacacata ctcaacaaca cacacacaag cacttgcaca tacatacaaa cacatgcaaa 1920
catacctaca cacacacaca ctcttacaac ctccatcatg ggaagtcaag ttccagaaac 1980
aaaagtctca ttcataagag gtcttagaag aaaataacca gttaacctga ttcaatttt 2040
gataccgttt tctgaacta ataatctac caatgagac ttttcagcct ttgtacatac 2100
aaaattcttc caaaagagag aggggaaat acagctctga tggcatcaaa cggactttgc 2160
atcaagtaat ttcagatagt gtcttaggat cctttgaggg tgctggtagc aggtgagcag 2220
gcaaaagtgg accaaggaca cttatttcta gattatgatt cttctgttta ctcaacaatt 2280
tacaagaaa aaaaggacag acattgaaga gctacacatt gtatataat caccacagac 2340
tataaggaaa tggaaattat tccctctttg tccatctctc gtatagatg ttgccaagat 2400
cagaatgat ccatltgctg tttctgtttt tccaaaggtc atacattgtg ttgtgtatt 2460
gttacccagc caataaatgt gtttaacagc ttaatttcat tttctggct ttggtctgtt 2520
ctctctctt acaggtaag ccctggctcc atgcaactgc attctttgat tteactgtt 2580
ccttcatcta catgttttgt tcaattgcag ccagtttlla ctgagttgt ggcaatcag 2640
aatgcatttg ctaagcaagt atgactttaa tccaactcca tggctcaatc atccacatga 2700
ggtgagcttc agcctgagat agcagccgac agacttcttg cgtttcaaaa ctgccatgcc 2760
cccctgtgat gctcccgta aggaatgcac tttgcttctg aagttcctgg gaaaggggta 2820
tgttttctct ccaggtgcag ccagatctca caaagtacaa aacgaatgcc tttctttct 2880
tgttataat ggtcaactcac tgtgtttggt tactgtcaag aatcaataa atgtgta 2938

```

【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
3 January 2002 (03.01.2002)

PCT

(10) International Publication Number
WO 02/000840 A3(51) International Patent Classification: C12N 15/60,
9/88, 5/10, C07K 16/40, C12Q 1/68, G01N 33/68, A01K
67/027[US/US]; 837 Roble Avenue, #1, Menlo Park, CA 94025
(US). **BURFORD, Neil** [GB/US]; 105 Wildwood Circle,
Durham, CT 06422 (US). **THANGAVELU, Kavitha**
[IN/US]; 1950 Montecito Avenue, 23, Mountain View,
CA 94043 (US). **BAUGHN, Mariah, R.** [US/US]; 14244
Santiago Road, San Leandro, CA 94577 (US).

(21) International Application Number: PCT/US01/19166

(22) International Filing Date: 13 June 2001 (13.06.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/213,383 23 June 2000 (23.06.2000) US
60/215,544 30 June 2000 (30.06.2000) US
60/222,818 4 August 2000 (04.08.2000) US(74) Agents: **HAMLET-COX, Diana** et al., Incyte Genomics,
Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).(81) Designated States (national): AF, AG, AI, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CI, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GI, GM,
HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK,
LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,
MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL,
TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.(71) Applicant (for all designated States except US): **INCYTE
GENOMICS, INC.** [US/US]; 3160 Porter Drive, Palo
Alto, CA 94304 (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, LI, MC, NL, PT, SI, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **THORNTON,
Michael** [US/US]; 9 Medway Road, Woodside, CA 94062
(US). **RAMKUMAR, Jayalaxmi** [IN/US]; 34359 May-
bird Circle, Fremont, CA 94555 (US). **TRIBOULEY,
Catherine, M.** [FR/US]; 1121 Tennessee Street, #5, San
Francisco, CA 94107 (US). **YUE, Henry** [US/US]; 826
Leis Avenue, Sunnyvale, CA 94087 (US). **NGUYEN,
Daniel, B.** [US/US]; 1403 Ridgewood Drive, San Jose,
CA 95118 (US). **YAO, Monique, G.** [US/US]; 111
Frederick Court, Mountain View, CA 94043 (US). **PAT-
TERSON, Chandra** [US/US]; 490 Sherwood Way #1,
Menlo Park, CA 94025 (US). **GANDHI, Ameena, R.**Published:
— with international search report(88) Date of publication of the international search report:
8 August 2002For two-letter codes and other abbreviations, refer to the "Guidance
Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

WO 02/000840 A3

(54) Title: HUMAN LYASES

(57) Abstract: The invention provides human lyases (HLYA) and polynucleotides which identify and encode HLYA. 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 HLYA.

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		Int. Application No. PCT/US 01/19166
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/68 C12N9/88 C12N5/10 C07K16/40 C12Q1/68 G01N33/68 A01K67/027		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 C12N C07K C12Q G01N A01K		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) WPI Data, PAJ, CAB Data, SEQUENCE SEARCH, BIOSIS, EPO-Internal		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO 00 61748 A (HUMAN GENOME SCIENCES INC ; ROSEN CRAIG A (US); RUBEN STEVEN M (US)) 19 October 2000 (2000-10-19) SEQ ID Nos. 37 and 120; claims 1-23	1-19,22, 25-44
E	WO 01 62927 A (INCYTE GENOMICS, INC.) 30 August 2001 (2001-08-30) SEQ ID Nos. 4 and 215; claims 1-27	1-19,22, 25-44

	-/--	
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another claim or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "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 "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "*" document member of the same patent family		
Date of the actual completion of the international search	Date of mailing of the international search report	
21 February 2002	22. 05. 2002	
Name and mailing address of the ISA European Patent Office, P.B. 5618 Patentstrasse 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-9016	Authorized officer HORNIG H.	

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/19166

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	LOWE N ET AL: "STRUCTURE AND METHYLATION PATTERNS OF THE GENE ENCODING HUMAN CARBONIC ANHYDRASE I" GENE (AMSTERDAM), vol. 93, no. 2, 1990, pages 277-283, XP002190991 ISSN: 0378-1119 the whole document ---	
A	US 5 972 684 A (CORLEY NEIL C ET AL) 26 October 1999 (1999-10-26) the whole document ---	
A	WO 88 04688 A (FLOREY HOWARD INST) 30 June 1988 (1988-06-30) the whole document ---	
A	HEWETT-EMMETT DAVID ET AL: "Functional diversity, conservation, and convergence in the evolution of the alpha-, beta-, and gamma-carbonic anhydrase gene families." MOLECULAR PHYLOGENETICS AND EVOLUTION, vol. 5, no. 1, 1996, pages 50-77, XP002190992 ISSN: 1055-7903 cited in the application the whole document -----	

INTERNATIONAL SEARCH REPORT	International application No. PCT/US 01/19166
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:	
<p>1. <input checked="" type="checkbox"/> Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely: see FURTHER INFORMATION sheet PCT/ISA/210</p> <p>2. <input checked="" type="checkbox"/> Claims Nos.: 20, 21, 23 and 24 because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically: see FURTHER INFORMATION sheet PCT/ISA/210</p> <p>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).</p>	
Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)	
This International Searching Authority found multiple inventions in this International application, as follows:	
see additional sheet	
<p>1. <input type="checkbox"/> As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.</p> <p>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.</p> <p>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.:</p> <p>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.: 1-44 (partially); 45, 48 (completely)</p>	
<p>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.</p>	

INTERNATIONAL SEARCH REPORT

International Application No. PCT/US 01/19166

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

Although claims 18, 35 and 38 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Although claim(s) 32, 33 and 34 are directed to a diagnostic method practised on the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Continuation of Box I.2

Claims Nos.: 20,21,23 and 24

Present claim 20, 21, 23 and 34 relate to a composition comprising an agonist and /or antagonist and a method for treating a disease, comprising administering to a patient in need of such treatment said agonist and/or antagonist without giving a true technical characterization. Moreover no such compounds are defined in the application. In consequence, the scope of said claims is ambiguous and vague, and their subject-matter is not sufficiently disclosed and supported (Art. 5 and 6 PCT). No search can be carried out for such purely speculative claims whose wording is, in fact, a mere recitation of the results to be achieved. Consequently, the search has been carried out for those parts of the claims which appear to be clear, supported and disclosed, namely those parts relating to antibodies, antisense molecules, ribozymes, polypeptides and nucleic acids, the structure which can be directly derived from SEQ ID Nos. 1 and 4.

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 11 procedure.

INTERNATIONAL SEARCH REPORT

International Application No. PCT/US 01/19166

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

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

1. Claims: (45,48)-complete; (1-44)-partially

An isolated polypeptide sequence of SEQ ID Nos. 1; an isolated polynucleotide sequence of SEQ ID No. 4 encoding said polypeptide; a cell transformed with said polynucleotide sequence; a transgenic organism comprising said polynucleotide sequence; a method for producing said polypeptide encoded by said polynucleotide sequence; an isolated polyclonal and/or monoclonal antibody which binds to said polypeptide sequence; a method for detecting a target polynucleotide sequence comprising SEQ ID No. 4; a composition comprising said polypeptide and a pharmaceutical acceptable excipient; a method for screening a compound for effectiveness as an agonist and/or antagonist of said polypeptide; a method for assessing toxicity of a test compound using said polynucleotide sequence; a diagnostic test for a condition or disease using said antibody; a method for detecting said polypeptide in a sample;

2. Claims: (46,49)-complete, (1-44)-partially

Idem as invention 1 but limited to SEQ ID Nos. 2 and 5;

3. Claims: (47,50)-complete, (1-44)-partially

Idem as invention one but limited to SEQ ID Nos. 3 and 6;

INTERNATIONAL SEARCH REPORT
Information on patent family members

International Application No
PCT/US 01/19166

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0061748 A	19-10-2000	AU 4072200 A	14-11-2000
		EP 1171591 A1	16-01-2002
		WO 0061748 A1	19-10-2000
WO 0162927 A	30-08-2001	AU 4175401 A	03-09-2001
		WO 0162927 A2	30-08-2001
		AU 4175301 A	03-09-2001
		AU 4170901 A	03-09-2001
		WO 0162922 A2	30-08-2001
		AU 3663101 A	03-09-2001
		WO 0162918 A3	18-04-2002
US 5972684 A	26-10-1999	NONE	
WO 8804688 A	30-06-1988	AU 1101888 A	15-07-1988
		WO 8804688 A1	30-06-1988

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 35/00	A 6 1 P 37/06	4 B 0 6 5
A 6 1 P 37/06	A 6 1 P 43/00	4 C 0 8 4
A 6 1 P 43/00	C 0 7 K 16/40	4 H 0 4 5
C 0 7 K 16/40	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 9/88	
C 1 2 N 5/10	C 1 2 P 21/02	C
C 1 2 N 9/88	C 1 2 P 21/08	
C 1 2 P 21/02	C 1 2 Q 1/68	A
C 1 2 P 21/08	G 0 1 N 33/15	Z
C 1 2 Q 1/68	G 0 1 N 33/50	Z
G 0 1 N 33/15	G 0 1 N 33/53	D
G 0 1 N 33/50	G 0 1 N 33/53	M
G 0 1 N 33/53	G 0 1 N 33/566	
G 0 1 N 33/566	C 1 2 N 5/00	A
	A 6 1 K 37/48	

(81) 指定国 AP(GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW

- (72) 発明者 ランクマール、ジャヤラクシミ
アメリカ合衆国カリフォルニア州 9 4 5 5 5 ・ フレモント ・ メイバードサークル 3 4 3 5 9
- (72) 発明者 トリボーリー、キャサリン・エム
アメリカ合衆国カリフォルニア州 9 4 1 0 7 ・ サンフランシスコ ・ # 5 ・ テネシーストリート 1 1 2 1
- (72) 発明者 ユエ、ヘンリー
アメリカ合衆国カリフォルニア州 9 4 0 8 7 ・ サニーベイル ・ ルイスアベニュー 8 2 6
- (72) 発明者 ニュエン、ダニエル・ビー
アメリカ合衆国カリフォルニア州 9 5 1 1 8 ・ サンノゼ ・ リッジウッドドライブ 1 4 0 3
- (72) 発明者 ヤオ、モニク・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 パターソン、チャンドラ
アメリカ合衆国カリフォルニア州 9 4 0 2 5 ・ メンロパーク ・ # 1 ・ シャーウッドウェイ 4 9 0
- (72) 発明者 ガンディー、アミーナ・アール
アメリカ合衆国カリフォルニア州 9 4 0 2 5 ・ メンロパーク ・ # 1 ・ ローブルアベニュー 8 3 7
- (72) 発明者 バーフォード、ニール
アメリカ合衆国コネチカット州 0 6 4 2 2 ・ ダーハム ・ ワイルドウッドサークル 1 0 5
- (72) 発明者 サンガベル、カピサ
アメリカ合衆国カリフォルニア州 9 4 0 4 3 ・ マウンテンビュー ・ 2 3 ・ モンテシトアベニュー 1 9 5 0
- (72) 発明者 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
- F ターム(参考) 2G045 AA34 AA35 BB14 BB20 BB50 CB01 DA13 DA36 FA12 FB02

专利名称(译)	人类崛起		
公开(公告)号	JP2004501632A	公开(公告)日	2004-01-22
申请号	JP2002505956	申请日	2001-06-13
[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	ソーントンマイケル ランクマールジャヤラクシミ トリポーリーキャサリンエム ユエヘンリー ニュエンダニエルビー ヤオモニクジー パターソンチャンドラ ガンディーアミーナアール バーフォードニール サンガベルカピサ ボーグンマライアアール		
发明人	ソーントン、マイケル ランクマール、ジャヤラクシミ トリポーリー、キャサリン・エム ユエ、ヘンリー ニュエン、ダニエル・ビー ヤオ、モニク・ジー パターソン、チャンドラ ガンディー、アミーナ・アール バーフォード、ニール サンガベル、カピサ ボーグン、マライア・アール		
IPC分类号	G01N33/50 A61K38/43 A61K45/00 A61P17/02 A61P29/00 A61P35/00 A61P37/06 A61P43/00 C07K16 /40 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N9/88 C12N15/09 C12N15/60 C12P21/02 C12P21 /08 C12Q1/68 G01N33/15 G01N33/53 G01N33/566		
CPC分类号	A61P17/02 A61P29/00 C12N9/88		
FI分类号	C12N15/00.ZNA.A A61K45/00 A61P17/02 A61P29/00 A61P35/00 A61P37/06 A61P43/00.111 C07K16 /40 C12N1/15 C12N1/19 C12N1/21 C12N9/88 C12P21/02.C C12P21/08 C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.D G01N33/53.M G01N33/566 C12N5/00.A A61K37/48		
F-TERM分类号	2G045/AA34 2G045/AA35 2G045/BB14 2G045/BB20 2G045/BB50 2G045/CB01 2G045/DA13 2G045 /DA36 2G045/FA12 2G045/FB02 2G045/FB03 2G045/FB05 2G045/FB07 4B024/AA01 4B024/AA11 4B024/BA07 4B024/BA44 4B024/CA04 4B024/DA02 4B024/EA02 4B024/GA11 4B024/HA01 4B024 /HA14 4B050/CC03 4B050/DD11 4B050/LL01 4B050/LL03 4B063/QA01 4B063/QA13 4B063/QQ42 4B063/QR18 4B063/QR32 4B063/QR56 4B063/QS25 4B063/QS34 4B063/QS36 4B063/QX02 4B063 /QX05 4B064/AG27 4B064/CA10 4B064/CE12 4B064/DA01 4B064/DA13 4B065/AA90X 4B065/AA90Y 4B065/AB01 4B065/BA02 4B065/CA24 4B065/CA25 4B065/CA44 4B065/CA46 4C084/AA02 4C084 /AA07 4C084/AA17 4C084/BA01 4C084/BA08 4C084/BA22 4C084/CA18 4C084/DC26 4C084/NA14 4C084/ZA33 4C084/ZB05 4C084/ZB11 4C084/ZB26 4C084/ZC02 4C084/ZC20 4H045/AA10 4H045 /AA11 4H045/AA20 4H045/AA30 4H045/BA10 4H045/CA40 4H045/DA76 4H045/DA89 4H045/EA20 4H045/EA50 4H045/FA74		

优先权

60/213383 2000-06-23 US

60/215544 2000-06-30 US

60/222818 2000-08-04 US

外部链接

[Espacenet](#)

摘要(译)

本发明提供了鉴定和编码HLYA的人裂解酶 (HLYA) 和多核苷酸。 本发明还提供表达载体， 宿主细胞， 抗体， 激动剂和拮抗剂。 本发明还提供了用于诊断， 治疗或预防与HLYA异常表达有关的疾病的方法。

The image shows a table that is severely rotated and distorted, making the content completely illegible. It appears to be a technical table, possibly containing sequence data or experimental results, but the specific information cannot be discerned.