

(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2004-529603

(P2004-529603A)

(43) 公表日 平成16年9月30日(2004.9.30)

(51) Int. Cl. ⁷	F I	テーマコード (参考)
C 1 2 N 15/09	C 1 2 N 15/00	Z N A A
A 6 1 K 38/51	A 6 1 K 45/00	2 G O 4 5
A 6 1 K 45/00	A 6 1 P 9/00	4 B O 2 4
A 6 1 P 9/00	A 6 1 P 15/00	4 B O 5 0
A 6 1 P 15/00	A 6 1 P 21/00	4 B O 6 3
		4 B O 6 4
	審査請求 未請求 予備審査請求 有	(全 207 頁) 最終頁に続く

(21) 出願番号	特願2002-507997 (P2002-507997)	(71) 出願人	301005050
(86) (22) 出願日	平成13年6月26日 (2001.6.26)		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12月24日 (2002.12.24)		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/US2001/020491	(74) 代理人	100089266
(87) 国際公開番号	W02002/002757		弁理士 大島 陽一
(87) 国際公開日	平成14年1月10日 (2002.1.10)	(72) 発明者	ガンディー、アミーナ・アール
(31) 優先権主張番号	60/215,476		アメリカ合衆国カリフォルニア州94025・メンロパーク・#1・ローブルアベニュー 837
(32) 優先日	平成12年6月29日 (2000.6.29)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/223,545		
(32) 優先日	平成12年8月4日 (2000.8.4)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/229,876		
(32) 優先日	平成12年8月31日 (2000.8.31)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 アデニルシクラーゼ及びグアニルシクラーゼ

(57) 【要約】

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

【特許請求の範囲】

【請求項 1】

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

(a) SEQ ID NO: 1 - 5 を有する群から選択したアミノ酸配列を含むポリペプチド

(b) SEQ ID NO: 1 - 5 からなる群から選択したアミノ酸配列と少なくとも 90% が同一であるような天然のアミノ酸配列を含むポリペプチド

(c) SEQ ID NO: 1 - 5 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) SEQ ID NO: 1 - 5 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】

SEQ ID NO: 1 - 5 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】

SEQ ID NO: 6 - 10 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。 20

【請求項 6】

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

【請求項 7】

7. 請求項 6 に記載の組換えポリヌクレオチドを用いて形質転換した細胞。

【請求項 8】

請求項 6 に記載の組換えポリヌクレオチドを含む遺伝形質転換体。

【請求項 9】

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

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

(b) そのように発現した前記ポリペプチドを回収する過程。

【請求項 10】

請求項 1 に記載のポリペプチドと特異的に結合する単離された抗体。

【請求項 11】

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

(a) SEQ ID NO: 6 - 10 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 40

(b) SEQ ID NO: 6 - 10 からなる群から選択したポリヌクレオチド配列と少なくとも 90% が同一であるような天然のポリヌクレオチドを含むポリヌクレオチド

(c) (a) のポリヌクレオチドに相補的なポリヌクレオチド

(d) (b) のポリヌクレオチドに相補的なポリヌクレオチド

(e) (a) ~ (d) の RNA 等価物

【請求項 12】

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

【請求項 13】

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

から検出する方法であって、

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

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

10

【請求項 14】

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

【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を含むことを特徴とする方法。

20

【請求項 16】

請求項 1 のポリペプチドと、薬剤として許容できる賦形剤とを含む組成物。

【請求項 17】

前記ポリペプチドが、SEQ ID NO: 1 - 5 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】

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

30

【請求項 19】

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

(a) 請求項 1 のポリペプチドを含むサンプルを化合物に曝露する過程と、

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

【請求項 20】

請求項 19 に記載の方法によって同定したアゴニスト化合物と薬剤として許容できる賦形剤とを含む組成物。

40

【請求項 21】

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

【請求項 22】

請求項 1 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 のポリペプチドを含むサンプルを化合物に曝露する過程と、

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

50

【請求項 2 3】

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

【請求項 2 4】

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

【請求項 2 5】

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

10

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

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

【請求項 2 6】

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

20

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

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

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

【請求項 2 7】

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

30

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

(b) 前記標的ポリヌクレオチドの変異発現を検出する過程と、

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

【請求項 2 8】

試験化合物の毒性を算定する方法であって、

(a) 核酸を含む生物学的サンプルを前記試験化合物で処理する過程と、

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

40

(c) ハイブリダイゼーション複合体の収量を定量するステップと、

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

【請求項 2 9】

生物学的サンプル中の A D G U C 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、

50

(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ADGUC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ADGUC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、SEQ ID NO: 1-5を有する群から選択したアミノ酸配列含むポリペプチドまたはその免疫抗原性断片を用いて動物を免疫化する過程と、

(b) 前記動物から抗体を単離する過程と、

(c) 前記単離された抗体をポリペプチドでスクリーニングし、それによって、SEQ ID NO: 1-5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。

30

【請求項36】

請求項35に記載の方法で産出した抗体。

【請求項37】

請求項36に記載の抗体及び適切なキャリアを含む化合物。

【請求項38】

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

(a) 抗体反応を誘発する条件下で、SEQ ID NO: 1-5を有する群から選択したアミノ酸配列含むポリペプチドまたはその免疫抗原性断片を用いて動物を免疫化する過程と、

(b) 前記動物から抗体産出細胞を単離する過程と、

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

(d) 前記ハイブリドーマ細胞を培養する過程と、

(e) SEQ ID NO: 1-5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】

50

請求項 39 に記載の抗体及び適切なキャリアを含む化合物。

【請求項 41】

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

【請求項 42】

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

【請求項 43】

S E Q I D N O : 1 - 5 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、

10

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

(b) 特異結合を検出する過程とを含み、該特異結合が、S E Q I D N O : 1 - 5 を有する群から選択したアミノ酸配列を有するポリペプチドがサンプル中に存在することを示唆することを特徴とする方法。

【請求項 44】

S E Q I D N O : 1 - 5 からなる群から選択したアミノ酸配列を有するポリペプチドを精製する方法であって、

20

(a) 前記抗体と前記ポリペプチドの特異結合を許容する条件下で、サンプルを用いて請求項 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、S E Q I D N O : 1 - 5 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とする方法。

【請求項 45】

S E Q I D N O : 1 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 46】

S E Q I D N O 2 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 47】

S E Q I D N O 3 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 48】

30

S E Q I D N O 4 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 49】

S E Q I D N O 5 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 50】

S E Q I D N O : 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。

【請求項 51】

S E Q I D N O : 7 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 52】

40

S E Q I D N O : 8 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 53】

S E Q I D N O : 9 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 54】

S E Q I D N O : 10 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】

50

(技術分野)

本発明は、アデニルシクラーゼ及びグアニルシクラーゼの核酸配列及びアミノ酸配列に関する。本発明はまた、これらの配列を利用した神経系疾患、心血管系疾患、視覚系疾患、生殖系疾患、平滑筋疾患、及び細菌感染の診断・治療・予防に関する。本発明はさらに、アデニルシクラーゼ及びグアニルシクラーゼの核酸配列及びアミノ酸配列の発現における外来性化合物の効果についての評価に関する。

【0002】

(発明の背景)

セカンドメッセンジャー分子は、細胞内シグナル伝達経路内での必須の役割を果たす。それは表面受容体にリガンドが結合すると活性化する媒介であり、下流エフェクター分子の活性剤として働く。サイクリックヌクレオチド、アデノシン3',5'-サイクリックリン酸(cAMP)及びグアノシン3',5'-サイクリックリン酸(cGMP)は、多様なシグナル伝達経路における重要なセカンドメッセンジャーである。cAMPおよびcGMPはATPとGTPの酵素であるアデニル(アデニル酸)シクラーゼ(AC)とグアニル(グアニル酸)シクラーゼ(GC)により生成される。従って細胞内のcAMPおよびcGMPレベルを調節する重要なステップは、ACおよびGC活性の調節である。

既知の哺乳動物ACのイソ型には9つある。すべては小さな細胞質内N末端ドメイン(N)、さらに6つの推定ヘリックス(M₁)を有する膜貫通ドメイン、大きな細胞質内ドメイン(C₁)、第2の膜貫通螺旋クラスタ(M₂)および第1の細胞質内ドメインに相
同的な第2細胞質ドメイン(C₂)から成る共通の構造を共有している(Simonds, W. F. (1999) Trends Pharmacol. Sci. 20:66-73)。C₁およびC₂ドメインは、約40%の同一性を共有し、また酵素の触媒中心を形成する~230アミノ酸領域(C_{1a}およびC_{2a})を含む。細胞質内ドメインの残りの部分は、C_{1b}またC_{2b}として知られる。C_{1a}およびC_{2a}ドメインの3次構造は、三層の / サンドイッチからなる。C_{1a}またC_{2a}ドメインは花輪のように列を作った配列で組まれる(Tang, W.-J及びJ. H. Hurley (1998) Mol. Pharmacol. 54:231-240)。すべての既知のGC触媒ドメインは哺乳動物ACのC₁またC₂領域に相通的であり、研究によると同じ構造を共有することを示唆する。膜貫通GCは1回の膜貫通交差および各タンパク質に1つ
の触媒ドメインタンパク質を含み、ホモ二量体として機能する。可溶性GCは、 および サブユニットの各々により寄与された1つの触媒ドメインを有し、 および サブユニットのヘテロ二量体として機能する。

【0003】

哺乳動物ACは、Gタンパク質、Ca²⁺シグナルおよびリン酸化による複雑な調節を受ける(前出のTang)。C_{1a}は抑制性Gタンパク質であるサブユニット(G_i)の主要な結合部位である。C_{2a}は刺激性Gタンパク質であるサブユニット(G_s)とG_qの主要な結合部位であり、プロテインキナーゼ(PK)Cとカルモジュリン(CaM)キナーゼIIのリン酸化部位を含む。C_{1b}はCa²⁺/CaM、Ca²⁺、PKAおよびCaMキナーゼIVによる調節のための部位である(Tesmer, J. G
. G及びS. R. Sprang (1998) Curr. Opin. Struct. Biol. 8:713-719)。9つのACイソ型の発現パターンと他の調節特性は多種多様である。例えば、AC4、AC7、AC9は様々な組織で発現するが、AC1とAC8は神経組織でのみ発現する。そしてAC5は主に心臓と脳で発現する(前出のSimonds)。すべてのACイソ型はG_sにより活性化するが、G_iまたG_qファミリーのサブユニットに示差反応を示す。またPKs、Ca²⁺、CaMに対して可変感度を示す。例えば、AC5とAC6はCaM非依存性的方法でCa²⁺のサブミクロモル濃度により抑制されるが、AC1、AC3、AC8はCa²⁺/CaMにより強く刺激を受ける。AC3、AC5およびAC6はG_{i1}による抑制に対する感受性を有するが、AC2はそうではない。この異質性により、Gタンパク質により伝送されたシグナルを細

10

20

30

40

50

胞内の Ca^{2+} レベルと PKC 活性に影響する他のソースからのシグナルとが統合されて、細胞外シグナルに対する組織特異的反応および細胞特異的反応を示すことを可能にする (前出の Simonds)。

【0004】

アデニリルシクラーゼは、ホルモン、神経伝達物質、臭気物質およびケモカインの細胞内シグナル伝達経路における重要な存在である (前出の Tang)。cAMP は、様々な細胞型における特異的な酵素の活性を変更する cAMP 依存性プロテインキナーゼを活性化させる。サイクリックヌクレオチド依存性イオンチャネルを活性化することにより、cAMP は細胞膜電位に影響を及ぼすことができる。cAMP は多様な組織特異的影響をも有する。cAMP のレベルの上昇によって、脂肪組織でのトリグリセリド加水分解の増加とアミノ酸吸収の減少、肝臓でのグリコーゲンからグルコースへの変換の増大、グリコーゲン合成の阻害、糖新生の増大、卵胞でのエストロゲンとプロゲステロンの合成の増大、副腎皮質でのアルドステロンとコルチゾールの合成の増大、心筋細胞での収縮速度の上昇、骨格筋でのグリコーゲンからグルコースへの変換、甲状腺でのサイロキシンの分泌、骨及び骨細胞からのカルシウムの再吸収、腸での流動分泌、腎臓での水分再吸収、血小板における凝集と分泌の阻害を引き起こす。(Lodish, H. 他 (1995) Molecular Cell Biology Scientific American Books New York NY, 871, 879886 ページを参照)。脳で発現する CaM 調節性 AC は、学習及び記憶の他にシナプス可塑性においても重要である。AC1 はメラトニン合成の調節に役割を果たすことも可能である (Xia, Z 及び D. R. Storm (1997) Curr. Opin. Neurobiol. 7:391-396)。

10

20

【0005】

膜 GC はアミノ末端で細胞外リガンド結合ドメインを有する。その既知のリガンドには、ナトリウム利尿ペプチドと細菌性エンテロトキシンが含まれる。ノックアウトしたマウスを使った実験では、GC-A は心臓機能および発達において役割を果たすことが明らかになった。ヒトの先天性失明の状態は GC-E の突然変異に関係がある。GC-G は骨格筋において高度に発現する。また GC-G は血流の調節に役割を果たす可能性がある (Wedel, B. J 及び D. L. Garbers (1998) Trends Endocrinol. Metab. 9:213-219)。可溶性グアニリルシクラーゼはヘムと結合しており、一酸化窒素により活性化される。一酸化窒素可溶性グアニリルシクラーゼ cGMP 経路は哺乳動物組織においてよく見られ、非血管系平滑筋また血管系平滑筋の弛緩、周辺及び中枢の神経伝達、血小板反応性、光伝達を含む数々の生理的プロセスを仲介する面で重要である。一酸化窒素-可溶性グアニリルシクラーゼ-cGMP 経路の過度な活動には、敗血ショックと片頭痛と関連がある可能性がある。また経路の低水準の活動にはインポテンス、高血圧、喘息と関連がある可能性がある (Hobbs, A. J. (1997) Trends Pharmacol. Sci. 18:484-491)。

30

【0006】

AC の既知のインヒビターには、降圧剤フォルスコリンと P 部位インヒビターがある。求められる特異性が達成可能な場合、心臓特有の AC インヒビターは鬱血性心不全の治療に有用であると提案されてきた。コレラその他の重症の下痢は細菌毒素による胃腸の AC と GC の活性に起因するので、酵素 AC と GC は有用な治療標的ともなる (Desseuer, C. W. ら (1999) Trends Pharmacol. Sci. 20:205-210)。

40

【0007】

新たなアデニリルシクラーゼ及びグアニリルシクラーゼ及びこれらをコードするポリヌクレオチドの発見は、神経系疾患、心血管系疾患、視覚系疾患、生殖系疾患、平滑筋疾患、及び細菌感染の診断、予防並びに治療において、またアデニリルシクラーゼ及びグアニリルシクラーゼの核酸及びアミノ酸配列の発現に対する外因性化合物の効果の評価において有用であるような新たな組成物を提供することにより、当分野における必要性を満たす。

50

【0008】

(発明の概要)

本発明は、総称して「ADGUC」、個別にはそれぞれ「ADGUC-1」、「ADGUC-2」、「ADGUC-3」、「ADGUC-4」、および「ADGUC-5」と呼ぶアデニルシクラーゼ及びグアニルシクラーゼである精製されたポリペプチドを提供する。或る実施態様において本発明は、(a)SEQ ID NO: 1-5を有する群から選択したアミノ酸配列からなるポリペプチド、(b)SEQ ID NO: 1-5を有する群から選択したアミノ酸配列と少なくとも90%が同一である天然のアミノ酸配列を有するポリペプチド、(c)SEQ ID NO: 1-5を有する群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または(d)SEQ ID NO: 1-5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、SEQ ID NO: 1-5の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。

10

【0009】

また、本発明は(a)SEQ ID NO: 1-5を有する群から選択したアミノ酸配列からなるポリペプチド、(b)SEQ ID NO: 1-5を有する群から選択したアミノ酸配列と少なくとも90%が同一である天然のアミノ酸配列からなるポリペプチド、(c)SEQ ID NO: 1-5を有する群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または(d)SEQ ID NO: 1-5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、を含む群から選択されたポリペプチドをコードする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SEQ ID NO: 1-5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SEQ ID NO: 6-10を有する群から選択される。

20

【0010】

本発明は更に、(a)SEQ ID NO: 1-5を有する群から選択したアミノ酸配列からなるポリペプチド、(b)SEQ ID NO: 1-5を有する群から選択したアミノ酸配列と少なくとも90%の相同性のある天然のアミノ酸配列からなるポリペプチド、(c)SEQ ID NO: 1-5を有する群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または(d)SEQ ID NO: 1-5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を有する群から選択したポリペプチドをコードするようなポリヌクレオチドと機能的に結合したプロモーター配列を有する組換えポリヌクレオチドを提供する。一実施態様では、本発明は組換えポリヌクレオチドを用いて形質転換した細胞を提供する。別の実施態様では、本発明は組換えポリヌクレオチドを含む遺伝形質転換体を提供する。

30

【0011】

更に本発明は、(a)SEQ ID NO: 1-5からなる一群から選択されたアミノ酸配列からなるポリペプチドと、(b)SEQ ID NO: 1-5からなる一群から選択されたアミノ酸配列と90%以上の配列同一性を有する天然のアミノ酸配列からなるポリペプチドと、(c)SEQ ID NO: 1-5からなる一群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d)SEQ ID NO: 1-5とからなる一群から選択されたアミノ酸配列を有するポリペプチドの免疫抗原性断片とで構成される一群から選択されたポリペプチドを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。

40

【0012】

本発明は更に、(a)SEQ ID NO: 1-5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、(b)SEQ ID NO: 1-5を有する群から選択したアミノ酸

50

配列と少なくとも90%同一である天然のアミノ酸配列を有するポリペプチド、(c)SEQ ID NO: 1-5を有する群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または(d)SEQ ID NO: 1-5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から構成される群から選択されたポリペプチドに特異結合するような実質上単離された抗体を提供する。

【0013】

本発明は更に、(a)SEQ ID NO: 6-10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と、(b)SEQ ID NO: 6-10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%が同一である天然のポリヌクレオチド配列を有するポリヌクレオチドと、(c)(a)のポリヌクレオチドに相補的なポリヌクレオチドと、(d)(b)のポリヌクレオチドに相補的なポリヌクレオチドと、(e)(a)~(d)のRNA等価物とで構成される群から選択した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。

10

【0014】

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し、該標的ポリヌクレオチドは(a)SEQ ID NO: 6-10を有する群から選択したポリヌクレオチド配列からなるポリヌクレオチド、(b)SEQ ID NO: 6-10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%同一の天然のポリヌクレオチド配列からなるポリヌクレオチド、(c)(a)のポリヌクレオチドに相補的なポリヌクレオチド、(d)(b)のポリヌクレオチドに相補的なポリヌクレオチド、および(e)(a)~(d)のRNA等価物を含む群から選択されたポリヌクレオチドの配列を有する。検出方法は、(a)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あるいはその断片との間でハイブリダイゼーション複合体が形成されるような条件下で、プローブは標的ポリヌクレオチドに特異的にハイブリダイズする。一実施態様では、プローブは少なくとも60の連続したヌクレオチドを含む。

20

【0015】

本発明はまた、サンプル中の標的ポリヌクレオチドを検出する方法を提供する。ここで、標的ポリヌクレオチドは、(a)SEQ ID NO: 6-10からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド、(b)SEQ ID NO: 6-10からなる群から選択したポリヌクレオチド配列と少なくとも90%の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、(c)(a)のポリヌクレオチドに相補的なポリヌクレオチド、(d)(b)のポリヌクレオチドに相補的なポリヌクレオチド、および(e)(a)~(d)のRNA等価物からなる群から選択された配列のポリヌクレオチドを有する。検出方法は、(a)ポリメラーゼ連鎖反応増幅を用いて標的ポリヌクレオチドまたはその断片を増幅する過程と、(b)増幅した標的ポリヌクレオチドまたはその断片の存在・不存在を検出し、該標的ポリヌクレオチドまたはその断片が存在する場合にはオプションでその量を検出する過程を含む。

30

40

【0016】

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し、有効量のポリペプチドは、(a)SEQ ID NO: 1-5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、(b)SEQ ID NO: 1-5を有する群から選択したアミノ酸配列と少なくとも90%同一である天然のアミノ酸配列を有するポリペプチド、(c)SEQ ID NO: 1-5を有する群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または(d)SEQ ID NO: 1-5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を含む。一実施例では、SEQ ID NO: 1-5からなる一群から選択されたアミノ酸配列を含む組成物を提供す

50

る。更に、本発明は、患者にこの組成物を投与することを含む、機能的ADGUCの発現の低下に関連した疾患やその症状の治療方法を提供する。

【0017】

本発明はまた、(a) SEQ ID NO: 1-5を有する群から選択したアミノ酸配列を含むポリペプチド、(b) SEQ ID NO: 1-5を有する群から選択したアミノ酸配列と少なくとも90%の相同性を有する天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1-5を有する群から選択したアミノ酸配列のポリペプチドの生物学的活性断片、または(d) SEQ ID NO: 1-5を有する群から選択したアミノ酸配列のポリペプチドの免疫抗原性断片を含む群から選択されたアゴニストとしてのポリペプチド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 10
ニング方法は、(a)ポリペプチドを有するサンプルを化合物に曝す過程と、(b)サンプル中のアゴニスト活性を検出する過程とを含む。別法では、本発明は、この方法によって同定されたアゴニスト化合物と好適な医薬用賦形剤とを含む組成物を提供する。更なる別法では、本発明は、この組成物の患者への投与を含む、機能的ADGUCの発現の低下 20
に関連した疾患やその症状の治療方法を提供する。

【0018】

本発明は更に、(a) SEQ ID NO: 1-5からなる群から選択したアミノ酸配列を含むポリペプチド、(b) SEQ ID NO: 1-5からなる群から選択したアミノ酸配列と少なくとも90%の同一性を有する天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1-5からなる群から選択したアミノ酸配列を含むポリペプチドの生物 20
学的活性断片、および(d) SEQ ID NO: 1-5からなる群から選択したアミノ酸配列を含むポリペプチドの免疫抗原性断片からなる群から選択されたポリペプチドのアンタゴニストとしての有効性を確認するために化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a)ポリペプチドを含むサンプルを化合物に曝す過程と、(b)サンプル中のアンタゴニスト活性を検出する過程とを含む。一実施態様で本発明は、この方法によって同定したアンタゴニスト化合物と薬剤として許容できる賦形剤とを含む成分を提供する。更なる別法では、本発明は、この組成物の患者への投与を含む、機能的ADGUCの過剰な発現に関連した疾患やその症状の治療方法を提供する。

【0019】

本発明は更に、(a) SEQ ID NO: 1-5を有する群から選択したアミノ酸配列を含むポリペプチド、(b) SEQ ID NO: 1-5を有する群から選択したアミノ酸配列と少なくとも90%の相同性を有する天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1-5を有する群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または(d) SEQ ID NO: 1-5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を含む群から選択したポリペプチドに特異結合する化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a) 30
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。

【0020】

更に本発明は、(a) SEQ ID NO: 1-5からなる一群から選択されたアミノ酸配列からなるポリペプチドと、(b) SEQ ID NO: 1-5からなる一群から選択されたアミノ酸配列と90%以上の配列同一性を有する天然のアミノ酸配列からなるポリペプチドと、(c) SEQ ID NO: 1-5からなる一群から選択されたアミノ酸配列を有するポリペプチドの生物学的活性断片と、(d) SEQ ID NO: 1-5からなる一群から選択されたアミノ酸配列を有するポリペプチドの免疫抗原性断片とで構成される群から 40
選択されたポリペプチドの活性を調節する化合物をスクリーニングする方法を提供する。その方法は、(a)ポリペプチドの活性が許容された条件下で、ポリペプチドを少なくとも1つの試験化合物と混合させる過程と、(b)ポリペプチドの活性を試験化合物の存在下で算定する過程と、(c)試験化合物の存在下でのポリペプチドの活性を試験化合物の 50

不存在下でのポリペプチドの活性と比較する過程とを含み、試験化合物の存在下でのポリペプチドの活性の変化は、ポリペプチドの活性を調節する化合物であることを意味する。

【0021】

本発明は更に、標的ポリヌクレオチドの変異発現の有効性を確認するために化合物をスクリーニングする方法を提供する。標的ポリヌクレオチドは、SEQ ID NO: 6 - 10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。

【0022】

本発明は更に、(a)核酸を含む生物学的サンプルを試験化合物で処理する過程と、(b) (i) SEQ ID NO: 6 - 10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) SEQ ID NO: 6 - 10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を含むポリヌクレオチド、(iii) (i)に相補的な配列を有するポリヌクレオチド、(iv) (ii)のポリヌクレオチドに相補的なポリヌクレオチド、(v) (i) ~ (iv)のRNA等価物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) SEQ ID NO: 6 - 10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) SEQ ID NO: 6 - 10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を含むポリヌクレオチド、(iii) (i)のポリヌクレオチドに相補的な配列を有するポリヌクレオチド、(iv) (ii)のポリヌクレオチドに相補的なポリヌクレオチド、(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。

【0023】

(発明を実施するための形態)

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

【0024】

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

【0025】

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

、試薬及びベクターについて説明及び開示する目的で引用しているものである。本明細書のいかなる開示内容も、本発明が先行技術の効力によってこのような開示に対して先行する権利を与えられていないことを認めるものではない。

【0026】

定義

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

【0027】

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

10

【0028】

「対立遺伝子変異体」は、ADGUC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つの突然変異から作製し得る。また、変異mRNAまたはポリペプチドを作製し得る。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個若しくは数個の天然の対立遺伝子変異体を有し得る。一般に対立遺伝子変異体を生じさせる通常の突然変異性変化は、ヌクレオチドの自然欠失、付加または置換に帰するものである。これら各変化は、単独或いは他の変化と共に、所定の配列内で1回若しくは数回生じ得る。

20

【0029】

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

30

【0030】

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

40

【0031】

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

【0032】

用語「アンタゴニスト」は、ADGUCの生物学的活性を阻害或いは減弱する分子である

50

。アンタゴニストは、A D G U C に直接相互作用するか、或いは A D G U C が関与する生物学的経路の成分と作用して、A D G U C の活性を調節する抗体、核酸、糖質、小分子、任意の他の化合物や組成物などのタンパク質を含み得る。

【0033】

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

10

【0034】

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

【0035】

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

20

30

【0036】

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

【0037】

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

40

【0038】

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

50

させることができる。

【0039】

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

10

【0040】

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

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

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

【0041】

「欠失」は、結果的に1個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。

【0042】

「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

40

50

ある。

【0043】

「検出可能な標識」は、測定可能な信号を発生することができ、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。

【0044】

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

【0045】

「断片」は、A D G U CまたはA D G U C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、16、20、25、30、40、50、60、75、100、150、250若しくは500の連続したヌクレオチド或いはアミノ酸残基長さであり得る。断片は、分子の特定領域から優先的に選択し得る。例えば、ポリペプチド断片は、所定の配列に示すようなポリペプチドの最初の250または500アミノ酸（または最初の25%または50%）から選択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。

10

20

【0046】

SEQ ID NO: 6 - 10の断片には、固有のポリヌクレオチド配列領域が含まれる。

この領域は、SEQ ID NO: 6 - 10を特異的に同定するものであり、例えばこの断片を得たゲノム中のSEQ ID NO: 6 - 10以外の配列とは異なるものである。SEQ ID NO: 6 - 10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SEQ ID NO: 6 - 10を区別する類似の方法において有用である。SEQ ID NO: 6 - 10の断片の正確な長さ及び断片に対応するSEQ ID NO: 6 - 10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。

30

【0047】

SEQ ID NO: 1 - 5の断片は、SEQ ID NO: 6 - 10の断片によってコードされる。SEQ ID NO: 1 - 5の断片には、SEQ ID NO: 1 - 5を特異的に同定する固有のアミノ酸配列領域が含まれている。例えば、SEQ ID NO: 1 - 5の断片は、SEQ ID NO: 1 - 5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。SEQ ID NO: 1 - 5の断片及び断片に対応するSEQ ID NO: 1 - 5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。

【0048】

「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。

40

【0049】

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

【0050】

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

50

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。

【0051】

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

【0052】

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

【0053】

Matrix:BLOSUM62
 Reward for match: 1
 Penalty for mismatch: -2
 :5 及び:ペナルティ 2
 Gap x drop-off: 50
 Expect: 10
 Word Size: 11
 Filter:on

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。

【0054】

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。

【0055】

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

【0056】

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み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、Ktupple = 1、gap penalty = 3、window = 5、「diagonals saved」= 5と設定する。デフォルトの残基重み付け表としてPAM250マトリクスを選択する。ポリヌクレオチドアラインメントと同様に、CLUSTAL Vは、アラインメントされたポリペプチド配列対間の「類似率」として一致率を報告する。

【0057】

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

【0058】

Matrix: BLOSUM62
 Open Gap: 11 及び Extension Gap: ペナルティ 1
 Gap x drop-off: 50
 Expect: 10
 Word Size: 3
 Filter: on

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。

【0059】

「ヒト人工染色体(HAC)」は、約6kb(キロベース)~10MbのサイズのDNA

10

20

30

40

50

配列を含み得る、染色体の複製、分離及び維持に必要な全ての要素を含む直鎖状の小染色体である。

【0060】

「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。

【0061】

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

10

20

【0062】

一般に、ハイブリダイゼーションのストリンジェンシーは或る程度、洗浄ステップを実行する温度を基準にして表すことができる。このような洗浄温度は通常、所定のイオン強度及びpHにおける特定の配列の融点（ T_m ）より約5～20℃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。

30

【0063】

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

40

【0064】

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

50

【0065】

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

【0066】

「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等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。

【0067】

「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A D G U Cのポリペプチドまたはオリゴペプチド断片である。用語「免疫抗原性断片」はまた、本明細書で開示するまたは当分野で周知のあらゆる抗体生産方法に有用なA D G U Cのポリペプチド断片またはオリゴペプチド断片を含む。

10

【0068】

「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。

【0069】

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

20

【0070】

用語「調節」は、A D G U Cの活性の変化を指す。調節することによって例えば、A D G U C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。

【0071】

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

30

【0072】

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

【0073】

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

40

【0074】

A D G U C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、A D G U Cの酵素環境に依存し、細胞タイプによって変化し得る。

【0075】

「プローブ」とは、同一配列或いは対立遺伝子核酸配列、関連する核酸配列の検出に用い

50

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

【0076】

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。

【0077】

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。

【0078】

プライマーとして用いるオリゴヌクレオチドの選択は、そのような目的のために本技術分野でよく知られているソフトウェアを用いて行う。例えばOLIGO 4.06ソフトウェアは、各100ヌクレオチドまでのPCRプライマー対の選択に有用であり、オリゴヌクレオチド及び最大5,000ヌクレオチドまでの大きめのポリヌクレオチドであって32キロ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み込まれている。例えば、PrimOUプライマー選択プログラム(テキサス州ダラスにあるテキサス大学南西部医療センターのゲノムセンターから一般向けに入手可能)は、メガベース配列から特定のプライマーを選択することが可能であり、従ってゲノム全体の範囲でプライマーを設計するのに有用である。Primer3プライマー選択プログラム(マサチューセッツ州ケンブリッジのWhitehead Institute/MITゲノム研究センターから一般向けに入手可能)ではユーザーが「ミスプライミング・ライブラリ」をインプットすることができ、ここでプライマー結合部位として避けたい配列はユーザーが指定する。Primer3は特に、マイクロアレイのためのオリゴヌクレオチドの選択に有用である。(後二者のプライマー選択プログラムのソースコードは、各自のソースから得てユーザー固有のニーズを満たすように変更してもよい。)PrimerGenプログラム(英国ケンブリッジ市の英国ヒトゲノムマッピングプロジェクト・リソースセンターから一般向けに入手可能)は、多数の配列アラインメントに基づいてプライマーを設計し、それによって、アラインメントされた核酸配列の最大保存領域または最小保存領域の何れかとハイブリダイズするようなプライマーの選択を可能にする。従って、このプログラム

は、固有な、および保存されたオリゴヌクレオチド及びポリヌクレオチドの断片の同定に有用である。上記選択方法のいずれかによって同定したオリゴヌクレオチド及びポリヌクレオチドの断片は、ハイブリダイゼーション技術において、例えばPCRまたはシーケンシングプライマーとして、マイクロアレイエレメントとして、或いは核酸のサンプルにおいて完全または部分的相補的ポリヌクレオチドを同定する特異プローブとして有用である。オリゴヌクレオチドの選択方法は、上記の方法に限定されるものではない。

【0079】

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

10

【0080】

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

20

【0081】

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

【0082】

「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。

30

【0083】

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

【0084】

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

【0085】

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

40

【0086】

「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成分の少なくとも約60%

50

、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。

【0087】

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

【0088】

「基質」は、任意の好適な固体または半固体の支持体を指すものであって、膜、フィルタ、チップ、スライド、ウエハ、ファイバー、磁性または非磁性ビーズ、ゲル、管、プレート、ポリマー、微細粒子、毛管が含まれる。基質は、凹み、溝、ピン、チャンネル、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。

10

【0089】

「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。

【0090】

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

20

【0091】

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

30

【0092】

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

40

50

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「単一ヌクレオチド多形性」(SNP)を含み得る。SNPの存在は、例えば特定の個体群、病状または病状性向を示し得る。

【0093】

特定のポリペプチド配列の「変異体」は、ポリペプチド配列の1本の長さ全体で特定のポリペプチド配列に対して少なくとも40%の相同性を有するポリペプチド配列として画定される。ここで、デフォルトパラメータに設定した「BLAST 2 Sequence s」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、ポリペプチドの所定の長さに対して、例えば少なくとも50%、60%、70%、80%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%またはそれ以上の配列同一性を示し得る。

10

【0094】

発明

本発明は、新規のヒトアデニリルシクラーゼ及びグアニリルシクラーゼ(ADGUC)及びADGUCをコードするポリヌクレオチドの発見に基づき、これらの組成物を利用した神経系疾患、心血管系疾患、視覚系疾患、生殖系疾患、平滑筋疾患、及び細菌感染の診断、治療、及び予防に関する。

20

【0095】

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

【0096】

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

30

40

【0097】

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

50

した検索可能なデータベースを示す。

【0098】

表2及び表3は共に、本発明の各々のポリペプチドの特性を要約しており、それらの特性が請求の範囲に記載されたポリペプチドがアデニルシクラーゼ及びグアニルシクラーゼであることを確立している。例えば、SEQ ID NO: 1はラットアデニルシクラーゼ型IV (GenBank ID g202676)と89%の同一性を有することがBasic Local Alignment Search Tool (BLAST)によって示された。(表2参照)BLAST確率スコアは0.0であるが、これは観測されたポリペプチド配列アラインメントが偶然に得られる確率を示す。SEQ ID NO: 1はまた、アデニル酸/グアニル酸シクラーゼ触媒ドメインを有するが、これは、隠れマルコフモデル(HMM)を基にした保存されたタンパク質ファミリードメインのPFAMデータベースにおいて、統計的に有意な一致を検索して決定された。(表3参照)BLIMPS、MOTIFS及びPROFILESCAN解析よりのデータは、SEQ ID NO: 1がアデニル酸/グアニル酸シクラーゼである、さらに実証的な証拠を提供する。別の例において、SEQ ID NO: 2はラットのアデニルシクラーゼV型 (GenBank ID g1758332)のアミノ酸384から836に対して99%の同一性があることがBasic Local Alignment Search Tool (BLAST)によって確認された。(表2参照)BLAST確率スコアは0.0であるが、これは観測されたポリペプチド配列アラインメントが偶然に得られる確率を示す。SEQ ID NO: 2はまた、アデニル酸/グアニル酸シクラーゼ触媒ドメインを有するが、これは、隠れマルコフモデル(HMM)を基にした保存されたタンパク質ファミリードメインのPFAMデータベースにおいて、統計的に有意な一致を検索して決定された。(表3参照)BLIMPS、MOTIFS及びPROFILESCAN解析よりのデータは、SEQ ID NO: 2がアデニル酸/グアニル酸シクラーゼである、さらに実証的な証拠を提供する。別の例において、SEQ ID NO: 3はBasic Local Alignment Search Tool (BLAST)によって同定されるミミズのアデニルシクラーゼ (GenBank ID g868200)に46%の同一性を有する。(表2参照)BLAST確率スコアは 2.2×10^{-21} であり、これは観測されたポリペプチド配列アラインメントが偶然に得られる確率を示している。別の例において、SEQ ID NO: 4はラットの酵母のアデニルシクラーゼに類似のヒトタンパク質 (GenBank ID g1504042)のアミノ酸294から769に対して100%の同一性があることがBasic Local Alignment Search Tool (BLAST)によって確認された。(表2参照)BLAST確率スコアは 7.5×10^{-254} であり、これは観測されたポリペプチド配列アラインメントが偶然に得られる確率を示している。SEQ ID NO: 4はまた、多数のロイシンリッチリピートドメインを有するが、これは、隠れマルコフモデル(HMM)を基にした保存されたタンパク質ファミリードメインのPFAMデータベースにおいて、統計的に有意な一致を検索して決定された。(表3参照)別の例において、SEQ ID NO: 5はウシのアデニルシクラーゼI型 (GenBank ID g162613)に91%の同一性を有するが、これはBasic Local Alignment Search Tool (BLAST)によって決定される。(表2参照)BLAST確率スコアは0.0であるが、これは観測されたポリペプチド配列アラインメントが偶然に得られる確率を示す。SEQ ID NO: 5はまた、アデニル酸/グアニル酸シクラーゼ触媒ドメインを有するが、これは、隠れマルコフモデル(HMM)を基にした保存されたタンパク質ファミリードメインのPFAMデータベースにおいて、統計的に有意な一致を検索して決定された。(表3参照)BLIMPS、MOTIFS及びPROFILESCAN解析よりのデータは、SEQ ID NO: 5がアデニル酸/グアニル酸シクラーゼである、さらに実証的な証拠を提供する。SEQ ID NO: 1-5の解析のためのアルゴリズム及びパラメータが表7で記述されている。

【0099】

10

20

30

40

50

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て構築した。列1および列2はそれぞれ、本発明の各ポリヌクレオチドのポリヌクレオチド配列識別番号(Polynucleotide SEQ ID NO:)およびそれに対応するIncyteポリヌクレオチドコンセンサス配列番号(Incyte Polynucleotide ID)を示す。列3は、各ポリヌクレオチド配列の長さ(塩基対単位)を示す。列4は、例えば、SEQ ID NO: 6-10を同定するため、或いはSEQ ID NO: 6-10と関連するポリヌクレオチド配列とを区別するためのハイブリダイゼーションまたは増幅技術に有用なポリヌクレオチド配列の断片を示す。列5はcDNA配列、ゲノムDNAから予想されたコード配列(エキソン)及び/またはcDNA及びゲノムDNAを共に有する配列集合に対応する識別番号を示している。これらの配列は、本発明の完全長ポリヌクレオチド配列をアセンブル構築するのに用いた。表4の列6および列7はそれぞれ完全長配列に対して、列5の配列に対応するcDNA配列およびゲノム配列の開始ヌクレオチド(5')位置および終了ヌクレオチド(3')位置を示す。

10

20

30

40

50

【0100】

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ライブラリを特異的に指す場合もある。例えば、6766950J1はIncyte cDNA配列の識別番号であり、BRAUNOR01はそれが由来するcDNAライブラリの識別番号である。cDNAライブラリが示されていないIncyte cDNAは、プールされているcDNAライブラリ(例えば、71190306V1)に由来する。または、列5の識別番号は、完全長ポリヌクレオチド配列の組み立てに用いたGenBankのcDNAすなわちEST(例えば、g1379036)の識別番号の場合もある。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指す場合もある。例えば、GNN.g8567699__0000030__02は、Genscan推定コード配列の識別番号であって、g8567699がGenscan分析によって得られたGenBankの配列の識別番号である。このGenscan推定コード配列は、配列を構築する前に編集されている場合がある。(実施例4を参照)。または列5の識別番号は、「エキソンステッチング(exon-stitching)」アルゴリズムにより結び合わせたcDNA及びGenscan予想エキソンの両方からなる群を指す場合もある。または、列5の識別番号は、「エキソンストレッチング」アルゴリズムによってcDNAおよびGenscan推定エキソンの両方からなる群の場合もある(実施例5を参照)。例えば、FL3685652__g8346725__g202676は「ストレッチされた」配列の識別番号である。3685652はIncyteプロジェクト識別番号であり、g8346725は「エキソンストレッチング」アルゴリズムが適用されたヒトゲノム配列のGenBankの識別番号であり、またg202676は最も近いGenBankタンパク質相同体のGenBankの識別番号である。

【0101】

(実施例5を参照)。場合によっては、列5に示されている配列の範囲と重複するIncyte cDNAの範囲が得られ、最終的なコンセンサスポリヌクレオチド配列が決定されるが、それに相当するIncyte cDNAの識別番号は示されていない。

【0102】

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

【0103】

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

に対して少なくとも約80%のアミノ酸配列同一性、或いは少なくとも約90%のアミノ酸配列同一性、更には少なくとも約95%のアミノ酸配列同一性を有する。

【0104】

本発明には、ADGUCをコードするポリヌクレオチドも含まれる。特定の実施例において、本発明は、ADGUCをコードするSEQ ID NO: 6-10からなる一群から選択された配列を含むポリヌクレオチド配列を提供する。SEQ ID NO: 6-10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。

【0105】

本発明はまた、ADGUCをコードするポリヌクレオチド配列の変異配列を含む。詳細には、このようなポリヌクレオチド配列の変異配列は、ADGUCをコードするポリヌクレオチド配列と少なくとも70%のポリヌクレオチド配列同一性、或いは少なくとも85%のポリヌクレオチド配列同一性、更には少なくとも95%ものポリヌクレオチド配列同一性を有する。本発明の或る実施態様では、SEQ ID NO: 6-10からなる群から選択されたアミノ酸配列と少なくとも約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ようなSEQ ID NO: 6-10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、ADGUC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。

【0106】

当業者であれば、遺伝暗号の縮重の結果、ADGUC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ADGUC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。

【0107】

ADGUC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ADGUCのヌクレオチド配列とハイブリダイズ可能であるが、ADGUC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ることなくADGUC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。

【0108】

本発明には、ADGUC、ADGUC誘導体及びこれらの断片をコードするDNA配列又はそ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ADGUC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。

【0109】

更に本発明には、種々のストリンジェントな条件下で、請求項に記載されたポリヌクレオチド配列、特に、SEQ ID NO: 6-10及びそれらの断片とハイブリダイズ可能なポリヌクレオチド配列が含まれる（例えば、Wahl, G.M.及びS.L. Berger (1987) Methods Enzymol. 152: 399-407、Kimmel, A.R. (1987) Methods Enzymol. 152: 507

10

20

30

40

50

- 511等を参照)。アニーリング及び洗浄条件を含むハイブリダイゼーションの条件は、「定義」に記載されている。

【0110】

DNAシーケンシングの方法は当分野でよく知られており、本発明の何れの実施例もDNAシーケンシング方法を用いて実施可能である。DNAシーケンシング方法には酵素を用いることができ、例えばDNAポリメラーゼIのクレノウ断片、SEQUENASE (US Biochemical, Cleveland OH)、Taqポリメラーゼ (Applied Biosystems)、熱安定性T7ポリメラーゼ (Amersham, 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 CATALYST 800サーマルサイクラー (Applied Biosystems)等の装置を用いて配列の調製を自動化する。次に、ABI 373 或いは377 DNAシーケンシングシステム (Applied Biosystems)、MEGABACE 1000 DNAシーケンシングシステム (Molecular Dynamics, Sunnyvale CA)または当分野でよく知られている他の方法を用いてシーケンシングを行う。結果として得られた配列を当分野でよく知られている種々のアルゴリズムを用いて分析する。 (Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7、Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, 856-853ページ等を参照)。

【0111】

ADGUC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、使用し得る方法の一つである制限部位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. 2:318322を参照)。別の方法に逆PCR法があり、これは広範な方向に伸長させたプライマーを用いて環状化した鋳型から未知の配列を増幅する方法である。鋳型は、既知のゲノム遺伝子座及びその周辺の配列を含む制限酵素断片から得る (Triglia, T. 他 (1988) Nucleic Acids Res. 16:8186.を参照)。第3の方法としてキャプチャPCR法があり、これ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増幅する方法に関与している。 (Lagerstrom, M. 他 (1991) PCR Methods Applic. 1:111119を参照)。この方法では、PCRを行う前に複数の制限酵素の消化及び連結反応を用いて未知の配列領域内に組換え二本鎖配列を挿入することが可能である。また、未知の配列を検索するために用い得る別の方法については当分野で知られている。 (Parker, J.D.ら (1991) Nucleic Acids Res. 19:30553060を参照)。更に、PCR、ネステッドプライマー及びPromoter Finder (商標)ライブラリ (Clontech, Palo Alto CA)を用いてゲノムDNAをウォーキングすることができる。この手順は、ライブラリをスクリーニングする必要がなく、イントロン/エキソン接合部を見付けるのに有用である。全てのPCRベースの方法に対して、市販されているソフトウェア、例えばOLIGO 4.06プライマー分析ソフトウェア (National Biosciences, Plymouth MN) 或いは別の好適なプログラムを用いて、長さが約22~30ヌクレオチド、GC含有率が約50%以上、温度約

68 ~ 72 で鋳型に対してアニーリングするようにプライマーを設計し得る。

【0112】

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

【0113】

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

【0114】

本発明の別の実施例では、ADGUCをコードするポリヌクレオチド配列またはその断片を組換えDNA分子にクローニングして、適切な宿主細胞内にADGUC、その断片または機能的等価物を発現させることが可能である。遺伝暗号固有の縮重により、実質的に同じ或いは機能的に等価のアミノ酸配列をコードする別のDNA配列が作られ得り、これらの配列をADGUCの発現に利用可能である。

【0115】

種々の目的でADGUCがコードする配列を変えるために、本発明のヌクレオチド配列を当分野で通常知られている方法を用いて組み換えることができる。ここで目的には、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、発現の調節がある。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特定部位突然変異誘導を利用して、新規な制限部位の生成、グリコシル化パターンの変更、コドン優先の変更、スプライス変異体の生成等を行う突然変異を導入し得る。

【0116】

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、ADGUC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られる。次にこれらの好適な変異体をプールし、更に反復してDNAシャッフリング及び選択/スクリーニングを行ってもよい。このように、遺伝の多様性は「人為的」品種改良及び急速な分子の進化を経て創生される。例えば、ランダムポイント突然変異を有する単一の遺伝子の断片を組み換えて、スクリーニングし、その後所望の特性が最適化されるまでシャッフリングすることができる。或いは、所定の遺伝子の断片を同種または異種のいずれかから得た同一遺伝子ファミリーの相同遺伝子の断片と組み換えて、それによって天然に

10

20

30

40

50

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。

【0117】

別の実施例によれば、ADGUCをコードする配列は、当分野で周知の化学的方法を用いて、全体或いは一部が合成可能である（例えば、Caruthers, M.H.ら(1980) *Nucl. Acids Res. Symp. Ser 7: 215-223*、Horn, T.他(1980) *Nucl. Acids Res. Symp. Ser 7: 225-232*等を参照)。或いは、化学的方法を用いてADGUC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, 55-60ページ、Roberge, J.Y.ら(1995) *Science* 269:202-204等を参照)。自動合成はABI 431Aペプチドシンセサイザ（Applied Biosystems）を用いて達成し得る。更にADGUCのアミノ酸配列または任意のその一部は、直接的な合成の際の変更、及び/または化学的方法を用いた他のタンパク質または任意のその一部からの配列との組み合わせにより、天然のポリペプチド配列を有するポリペプチドまたは変異体ポリペプチドを製作することが可能である。

10

【0118】

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, 28-53ページ等を参照）。

20

【0119】

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

30

40

【0120】

当業者によく知られている方法を用いて、ADGUC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。これらの方法には、in vitro組換えDNA技術、合成技術、及びin vivo遺伝子組換え技術が含まれる。（例えば、Sambrook, J.他.(1989) *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Press, Plainview NY, 4章及び8章, 及び16-17章、Ausubel, F.M.他.(1995) *Current Protocol*

50

s in Molecular Biology, John Wiley & Sons, New York NY, ch. 9章、13章及び16章等を参照)。

【0121】

種々の発現ベクター/宿主系を利用して、ADGUCをコードする配列の保持及び発現が可能である。限定するものではないがこのような発現ベクター/宿主系には、組換えバクテリオファージ、プラスミドまたはコスミドDNA発現ベクターで形質転換させた細菌や、酵母菌発現ベクターで形質転換させた酵母菌や、ウイルス発現ベクター(例えばバキュロウイルス)に感染した昆虫細胞系や、ウイルス発現ベクター(例えばカリフラワーモザイクウイルスCaMVまたはタバコモザイクウイルスTMV)または細菌発現ベクター(例えばTiまたはpBR322プラスミド)で形質転換させた植物細胞系、動物細胞系などの微生物等がある。(前出のSambrook、前出のAusubel、Van Heeke, G及びS.M. Schuster (1989) J. Biol. Chem. 264:5503-5509、Engelhard, E.K.ら(1994) Proc. Natl. Acad. Sci. USA 91:3224-3227、Sandig, V.ら(1996) Hum. Gene Ther. 7:1937-1945、Takamatsu, N. (1987) EMBOJ. 6:307-311、; 『マグローヒル科学技術年鑑』(The McGraw Hill Yearbook of Science and Technology) (1992) McGraw Hill New York NY, 191-196ページ、Logan, J及びT. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659、Harrington, J.J.ら(1997) Nat. Genet. 15:345-355等を参照)。レトロウイルス、アデノウイルス、ヘルペスウイルスまたはワクシニアウイルス由来の発現ベクター、または種々の細菌性プラスミド由来の発現ベクターを用いて、ヌクレオチド配列を標的器官、組織または細胞集団へ輸送することができる(Di Nicola, M. 他(1998) Cancer Gen. Ther. 5(6):350-356、Yu, M. 他(1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344、Buller, R.M. 他(1985) Nature 317(6040):813-815、McGregor, D.P. 他(1994) Mol. Immunol. 31(3):219-226、Verma, I.M及びN. Somia (1997) Nature 389:239-242等を参照)。本発明は使用される宿主細胞によって限定されるものではない。

【0122】

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

【0123】

ADGUCの発現に酵母の発現系の使用が可能である。因子、アルコールオキシダーゼ、PGHプロモーター等の構成型或いは誘導型のプロモーターを含む多数のベクターが、

酵母菌サッカロミセス - セレビジエまたは *Pichia pastoris* に使用可能である。更に、このようなベクターは、発現したタンパク質の分泌か細胞内への保持のどちらかを誘導し、安定した増殖のために宿主ゲノムの中に外来配列を組み込む。(例えば、Ausubel, 1995, 前出、Bitter, G. A. ら (1987) *Methods Enzymol.* 153: 516 - 544、Scorer, C. A. ら (1994) *Bio/Technology* 12: 181 - 184 等を参照)。

【0124】

植物系も ADGUC の発現に使用可能である。ADGUC をコードする配列の転写は、ウイルスプロモーター、例えば単独或いは TMV (*Takamatsu, N.* (1987) *EMBO J.* 6: 307 - 311) 由来のオメガリーダー配列と組み合わせて用いられるような CaMV 由来の 35S 及び 19S プロモーターによって促進される。或いは、RUBISCO の小サブユニット等の植物プロモーターまたは熱ショックプロモーターを用いてもよい (例えば、Coruzzi, G. ら (1984) *EMBO J.* 3: 1671 - 1680、Broglie, R. ら (1984) *Science* 224: 838 - 843、Winter, J. ら (1991) *Results Probl. Cell Differ.* 17: 85 - 105 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 McGraw Hill Yearbook of Science and Technology*) (1992) McGraw Hill New York NY. 191 - 196 ページ等を参照)。

10

20

【0125】

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

30

【0126】

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

【0127】

哺乳動物系の組換えタンパク質の長期にわたる産生のためには、細胞株における ADGUC の安定した発現が望ましい。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、ADGUC 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。

40

【0128】

任意の数の選択系を用いて、形質転換細胞株を回収できる。限定するものではないがこのような選択系には、tk⁻ 単純細胞のために用いられるヘルペスウイルスチミジンキナーゼ遺伝子と、apr⁻ 細胞のために用いられるアデニンホスホリボシルトランスフェラー

50

ゼ遺伝子がある(例えば、Wigler, M. 他(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)。この他の選択可能な遺伝子、例えば、代謝のための細胞の必要条件を変えるtrpB及びhisDは、文献に記載されている(Hartman, S.C.及びR.C. Mulligan(1988) Proc. Natl. Acad. Sci. USA 85:8047-8051等を参照)。可視マーカー、例えばアントシアニン、緑色蛍光タンパク質(GFP; Clontech)、グルクロニダーゼ及びその基質グルクロニド、またはルシフェラーゼ及びその基質ルシフェリン等を用いてもよい。これらのマーカーを用いて、トランスフォーマントを特定するだけでなく、特定のベクター系に起因する一過性或いは安定したタンパク質発現を定量することが可能である(Rhodes, C.A. (1995) Methods Mol. Biol. 55:121131等を参照)。

【0129】

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、ADGUC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、ADGUC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、1つのプロモーターの制御下でマーカー遺伝子がADGUCをコードする配列と一列に配置することも可能である。誘導または選択に応答したマーカー遺伝子の発現は通常、タンデム遺伝子の発現も示す。

【0130】

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

【0131】

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

【0132】

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。ADGUCをコードするポリヌクレオチドに関連する配列を検出するための、標識されたハイブリダイゼーションプローブ或いはPCRプローブを生成する方法には、オリゴ標識化、ニックトランスレーション、末端標識化、または標識されたヌクレオチドを用いるPCR増幅が含まれる。別法として、ADGUCをコードする配列、またはその任意の断片をmRNAプローブの生成の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。

10

20

30

40

50

【0133】

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

【0134】

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

【0135】

本発明の別の実施例では、ADGUC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ADGUCタンパク質は、ADGUC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、ADGUCが精製後に異種部分から切断され得るようになり、ADGUC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 (19

95) 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。

【0136】

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

【0137】

本発明のADGUCまたはその断片を用いて、ADGUCに特異結合する化合物をスクリーニングすることができる。少なくとも1つまたは複数の試験化合物を用いて、ADGUCへの特異的な結合をスクリーニングすることが可能である。試験化合物の例には、抗体、オリゴヌクレオチド、タンパク質(例えば受容体)または小分子が挙げられる。

【0138】

一実施例では、このように同定された化合物は、例えばリガンドやその断片などのADGUCの天然のリガンド、または天然の基質、構造的または機能的な擬態性または自然結合パートナーに密接に関連している(Coligan, J. E. 他(1991) Current Protocols in Immunology 1(2)の5章等を参照)。同様に、化合物は、ADGUCが結合する天然受容体、或いは例えばリガンド結合部位などの少なくとも受容体のある断片に密接に関連し得る。何れの場合も、既知の技術を用いてこの化合物を合理的に設計することができる。一実施例では、このような化合物に対するスクリーニングには、分泌タンパク質或いは細胞膜上のタンパク質の何れか一方としてADGUCを発現する好適な細胞の作製が含まれる。好適な細胞には、哺乳動物、酵母、ショウジョウバエ、大腸菌からの細胞が含まれる。ADGUCを発現する細胞またはADGUCを含有する細胞膜断片を試験化合物と接触させて、ADGUCまたは化合物の何れかの結合、刺激または阻害を分析する。

【0139】

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

【0140】

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

【0141】

別の実施例では、胚性幹細胞(ES細胞)における相同組換えを用いて動物モデル系内で、ADGUCまたはその哺乳動物相同体をコードするポリヌクレオチドを「ロックアウト

」する。このような技術は当技術分野において周知であり、ヒト疾患動物モデルの作製に有用である（米国特許第5,175,383号及び第5,767,337号等を参照）。例えば129/SvJ細胞株等のマウスES細胞は初期のマウス胚に由来し、培地で増殖させることができる。このES細胞は、ネオマイシンホスホトランスフェラーゼ遺伝子（neo:Capecchi, M. R. (1989) Science 244:1288-1292）等のマーカー遺伝子で破壊した目的の遺伝子を含むベクターで形質転換する。このベクターは、相同組換えにより宿主ゲノムの対応する領域に組み込まれる。別法では、Cre-loxP系を用いて相同組換えを行い、組織特異的または発生段階特異的に目的遺伝子をノックアウトする（Marth, J. D. (1996) Clin. Invest. 97:1999-2002、Wagner, K. U. 他 (1997) Nucleic Acids Res. 25:4323-4330）。形質転換したES細胞を同定し、例えばC57BL/6マウス系等から採取したマウス細胞胚盤胞に微量注入する。胚盤胞を偽妊娠メスに外科的に導入し、得られるキメラ子孫の遺伝形質を決め、これを交配させてヘテロ接合性系またはホモ接合性系を作製する。このようにして作製した遺伝子組換え動物は、可能性のある治療薬や毒性薬剤で検査することができる。

【0142】

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

【0143】

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

【0144】

治療

ADGUCのある領域とアデニリルシクラーゼ及びグアニリルシクラーゼのある領域との間に、例えば配列及びモチーフの内容における化学的及び構造的類似性が存在する。さらにADGUCの発現は前立腺癌、脳組織、回腸組織、子宮内膜組織と密接に関連する。よってADGUCは神経系疾患、心血管系疾患、視覚系疾患、生殖系疾患、平滑筋疾患及び細菌感染においてある役割を果たすと考えられる。ADGUC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、ADGUCの発現または活性を低下させることが望ましい。また、ADGUC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、ADGUCの発現または活性を亢進させることが望ましい。

【0145】

従って、一実施例において、ADGUC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ADGUC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、神経障害として、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 (cerebelloretinal hemangioblastomatosis)、脳3叉神経血管症候群、精神薄弱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、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、トゥーレット病、動静脈瘤などの心疾患、アテローム性動脈硬化症、高血圧、血管炎、レイノー病、動脈瘤、動脈乖離、拡張蛇行静脈、血栓、静脈血栓症、血管の腫瘍、血栓崩壊の合併症、バルーン血管形成術 (balloon angioplasty)、血管置換術、大動脈冠動脈バイパス術移植手術 (coronary artery bypass graft surgery)、うっ血性心不全、虚血性心疾患、狭心症、心筋梗塞、高血圧性心疾患、変性弁膜性心疾患、石灰化大動脈弁狭窄症、先天性2尖大動脈弁、僧帽弁輪状石灰化 (mitral annular calcification)、僧帽弁脱出、リウマチ熱、リウマチ性心疾患、感染性心内膜炎、非細菌性血栓性心内膜炎、全身性紅斑性狼瘡の心内膜炎、カルチノイド心疾患、心筋症、心筋炎、心膜炎、腫瘍性心疾患、先天性心臓疾患、心臓移植の合併症などの心疾患、先天性肺異常、肺拡張不全、肺うっ血及び肺水腫、肺動脈塞栓症、肺出血、肺梗塞、肺高血圧症、血管硬化症、閉塞性肺疾患、拘束型肺疾患、慢性閉塞性肺疾患、肺気腫、慢性気管支炎、気管支喘息、細気管支拡張症、細菌性肺炎、ウイルス性肺炎及びマイコプラズマ肺炎、肺膿瘍、肺結核、びまん性間質性疾患、塵肺症、サルコイド症、特発性肺線維症、剥離性間質性肺炎、過敏症肺炎、肺好酸球増加閉塞性細気管支炎 器質性肺炎 (pulmonary eosinophilia bronchiolitis obliterans - organizing pneumonia)、肺移植の合併症があり、視覚系疾患の中には、結膜炎、乾性角結膜炎、角膜炎、上強膜炎、虹彩炎、後部ブドウ膜炎、緑内障、一過性黒内障、虚血性視神経症、視神経炎、レーバー遺伝性視神経症、中毒性視覚神経炎、硝子体剥離、網膜剥離、白内障、黄斑変性症、中心性漿液性脈絡網膜症、色素性網膜炎、脈絡膜黒色腫、球後腫瘍、交叉腫瘍 (chiasmaltumor)、生殖系疾患としてプロラクチン産生障害があり、卵管疾患、

30 排卵欠損症、子宮内膜症、性周期障害、月経周期障害、多嚢胞性卵巣症候群、卵巣過刺激症候群、子宮内膜癌、卵巣癌、子宮筋腫、自己免疫疾患、子宮外妊娠、催奇形、乳がん、繊維嚢胞性乳房疾患、乳漏症、精子形成破壊、精子異常生理、精巣癌、前立腺癌、良性前立腺肥大、前立腺炎、パイロニー病、性交不能、男性乳癌、女性化乳房が含まれる。平滑筋疾患には、限定されるものではないが平滑筋の正常な活動の欠陥或いは変性すべてが含まれ、狭心症、アナフィラキシー性ショック、不整脈、喘息、心臓血管性ショック、クッシング症候群、高血圧、低血糖症、心筋梗塞、偏頭痛、およびクロム親和細胞腫 (褐色細胞腫)、および心筋症、脳疾患、癲癇、カーンズ・セイヤー症候群、乳酸アシドーシス、筋クロノス性疾患、および眼筋麻痺などのミオパシーが含まれる。そして細菌による感染は肺炎球菌感染症、ブドウ球菌、連鎖球菌、バシラス菌、コリネバクテリウム、クロストリジウム、髄膜炎菌、淋菌、リステリア、モラクセラ、キングラ、ヘモフィルス、レジオネラ、百日咳菌、シゲラ、サルモネラ、カンピロバクターを含むグラム陰性菌腸内細菌、シュドモナス、ピブリオ、ブルセラ、野兎病菌、エルシニア、バルトネラ、nocardium、放線菌、ミコバクテリウム、スピロヘータ、リケッチア、クラミジア、マイコプラズマに分類される。

40

【0146】

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

【0147】

更に別の実施例では、限定するものではないが上に列記した疾患を含むA D G U Cの発現または活性の低下に関連した疾患の治療または予防のために、実質的に精製されたA D G U Cを含む組成物を好適な医薬用担体と共に患者に投与することも可能である。

【0148】

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

【0149】

更なる実施例では、A D G U Cの発現または活性の増大に関連した疾患の治療または予防のために、患者にA D G U Cのアンタゴニストを投与することが可能である。限定するものではないが、このような疾患の例には、上記の神経系疾患、心血管系疾患、視覚系疾患、生殖系疾患、平滑筋疾患、及び細菌感染が含まれる。一実施態様では、A D G U Cと特異的に結合する抗体が直接アンタゴニストとして、或いはA D G U Cを発現する細胞または組織に薬剤を運ぶターゲティング或いは送達機構として間接的に用いられ得る。

【0150】

別の実施例では、限定するものではないが上に列記した疾患を含むA D G U Cの発現または活性の増大に関連した疾患の治療または予防のために、A D G U Cをコードするポリヌクレオチドの相補配列を発現するベクターを患者に投与することも可能である。

【0151】

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

【0152】

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

【0153】

抗体の産生のためには、ヤギ、ウサギ、ラット、マウス、ヒト及びその他のものを含む種々の宿主が、A D G U Cまたは任意の断片、または免疫原性の特性を備えるそのオリゴペプチドの注入によって免疫化され得る。宿主の種に応じて、種々のアジュバントを用いて免疫応答を高めることもできる。限定するものではないがこのようなアジュバントには、フロイントアジュバントと、水酸化アルミニウム等のミネラルゲルアジュバントと、リゾレシチン、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、B C G（カルメット ゲラン杆菌）及びコリネバクテリウム パルヴムが特に好ましい。

【0154】

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

10

20

30

40

50

列と融合し、キメラ分子に対する抗体が産生され得る。

【0155】

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

10

【0156】

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

20

【0157】

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)。

【0158】

A D G U C に対する特異的な結合部位を含む抗体も産生される。例えば、限定するものではないが、このような断片には、抗体分子のペプシン消化によって作製される $F(ab')_2$ 断片と、 $F(ab')_2$ 断片のジスルフィド架橋を還元することによって作製される Fab 断片とがある。或いは、Fab 発現ライブラリを作製することによって、モノクローナル Fab 断片を所望の特異性と迅速且つ容易に同定することが可能となる(Huse, W. D. ら (1989) Science 246: 1275-1281等を参照)。

30

【0159】

種々のイムノアッセイを用いてスクリーニングし、所望の特異性を有する抗体を同定することができる。隔離された特異性を有するポリクローナル抗体またはモノクローナル抗体の何れかを用いる競合的な結合、または免疫放射線活性のための数々のプロトコルが、当分野では周知である。通常このようなイムノアッセイには、A D G U C とその特異性抗体との間の複合体形成の計測が含まれる。2つの非干渉性A D G U C 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もよい(前出のPoundの文献)。

40

【0160】

ラジオイムノアッセイ技術と共にScatchard分析などの様々な方法を用いて、A D G U C に対する抗体の親和性を評価する。親和性を結合定数 K_a で表すが、この K_a は、平衡状態の下でA D G U C 抗体複合体のモル濃度を遊離抗体と遊離抗原のモル濃度で除して得られる値である。多数のA D G U C エピトープに対して親和性が不均一なポリク

50

ーナル抗体試薬の K_a は、ADGUC に対する抗体の平均親和性または結合活性を表す。特定の ADGUC エピトープに単一特異的なモノクローナル抗体試薬の K_a は、親和性の真の測定値を表す。 K_a 値が $10^9 \sim 10^{12}$ liter/mol の高親和性抗体試薬は、ADGUC 抗体複合体が激しい操作に耐えなければならないイムノアッセイに用いるのが好ましい。 K_a 値が $10^6 \sim 10^7$ liter/mol の低親和性抗体試薬は、ADGUC が抗体から最終的に活性化状態で解離する必要がある免疫精製 (immunopurification) 及び類似の処理に用いるのが好ましい。(Catty, D. (1988) Antibodies, Volume I: A Practical Approach. IRL Press, Washington, DC, Liddell, J. E 及び Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY)。

10

【0161】

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

20

【0162】

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

30

【0163】

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 及び W. Walther (1994) Pharmacol. Ther. 63 (3): 323 - 347 等を参照)。その他の遺伝子送達機構には、リポソーム系、人工的なウイルスエンベロープ及び当分野で公知のその他の系が含まれる (Rossi, J. J. (1995) Br. Med. Bull. 51 (1): 217 - 225、Boado, R. J. 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)。

40

【0164】

本発明の別の実施例では、ADGUC をコードするポリヌクレオチドを、体細胞若しくは生殖細胞の遺伝子治療に用いることが可能である。遺伝子治療を行うことにより、(i)

50

遺伝子欠損症（例えばX染色体鎖遺伝（Cavazzana-Calvo, M. ら（2000）Science 288:669-672）により特徴付けられる重度の複合型免疫欠損（SCID）-X1の場合）、先天性アデノシンデアミナーゼ（ADA）欠損症に関連する重度の複合型免疫欠損（Blaese, R.M. ら（1995）Science 270:475-480、Bordignon, C. ら（1995）Science 270:470-475）、嚢胞性繊維症（Zabner, J. ら（1993）Cell 75:207-216；Crystal, R.G. ら（1995）Hum. Gene Therapy 6:643-666、Crystal, R.G. ら（1995）Hum. Gene Therapy 6:667-703）、サラセミア（thalassaemia）、家族性高コレステロール血症、第VIII因子若しくは第IX因子欠損に起因する血友病（Crystal, R.G.（1995）Science 270:404-410、Verma, I.M. 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）、Candida albicans及びParacoccidioides brasiliensis等の真菌寄生虫、並びにPlasmodium falciparum及びTrypanosoma cruzi等の原虫寄生体に対する防御機能を有するタンパク質を発現させることができる。ADGUCの発現若しくは調節に必要な遺伝子の欠損が疾患を引き起こす場合、導入した細胞の好適な集団からADGUCを発現させて、遺伝子欠損によって起こる症状の発現を緩和することが可能である。

【0165】

本発明の更なる実施例では、ADGUC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ADGUC欠損細胞に導入することによって、ADGUC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) 遺伝子銃、(iii) リポソームを介した形質移入、(iv) 受容体を介した遺伝子導入、及び(v) DNAトランスポソンの使用（Morgan, R.A. およびW. F. Anderson（1993）Annu. Rev. Biochem. 62:191-217、Ivics, Z.（1997）Cell 91:501-510、Boulay, J-L. およびH. Recipon（1998）Curr. Opin. Biotechnol. 9:445-450）がある。

【0166】

限定するものではないがADGUCの発現に影響を及ぼし得る発現ベクターには、PCDNA 3.1、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。ADGUCを発現させるために、(i) 恒常的に活性なプロモーター（例えば、サイトメガロウイルス（CMV）、ラウス肉腫ウイルス（RSV）、SV40ウイルス、チミジンキナーゼ（TK）、若しくは - アクチン遺伝子等）、(ii) 誘導性プロモーター（例えば、市販されているT-REXプラスミド（Invitrogen）に含まれている、テトラサイクリン調節性プロモーター（Gossen, M及びH. Bujard（1992）Proc. Natl. Acad. Sci. U.S.A. 89:5547-5551；Gossen, M. 他（1995）Science 268:1766-1769、Rossi, F.M.V. 及びH.M. Blau（1998）Curr. Opin. Biotechnol. 40

9 : 4 5 1 - 4 5 6))、エクジソン誘導性プロモーター (市販されているプラスミド P V G R X R 及び P I N D に含まれている : I n v i t r o g e n)、F K 5 0 6 / ラパマイシン誘導性プロモーター、または R U 4 8 6 / ミフェプリストーン誘導性プロモーター (R o s s i , F . M . V . 及び H . M . B l a u , 前出)、または (i i i) 正常な個体に由来する A D G U C をコードする内在性遺伝子の天然のプロモーター若しくは組織特異的プロモーターを用いることが可能である。

【 0 1 6 7 】

市販のリポソーム形質転換キット (例えば I n v i t r o g e n 社の P e r F e c t L i p i d T r a n s f e c t i o n K i t) 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 (G r a h a m , F . L . 及び A . J . E b (1 9 7 3) V i r o l o g y 5 2 : 4 5 6 - 4 6 7) 若しくは電気穿孔法 (N e u m a n n , B . ら (1 9 8 2) E M B O J . 1 : 8 4 1 - 8 4 5) を用いて形質転換を行う。初代培養細胞に D N A 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。

10

【 0 1 6 8 】

本発明の別の実施例では、A D G U C 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 (L T R) プロモーターまたは独立プロモーターの制御下で A D G U C をコードするポリヌクレオチドと、(i i) 好適な R N A パッケージングシグナルと、(i i i) 追加レトロウイルス・シス作用性 R N A 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 R e v 応答性エレメント (R R E) 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 P F B 及び P F B N E O) は S t r a t a g e n e 社から市販されており、刊行データ (R i v i e r e , I . ら (1 9 9 5) P r o c . N a t l . A c a d . S c i . U . S . A . 9 2 : 6 7 3 3 - 6 7 3 7) 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 (V P C L) において増殖され、V P C L は、標的細胞上の受容体に対する向性を有するエンベロープ遺伝子または V S V g 等の乱交雑エンベロープタンパク質を発現する (A r m e n t a n o , D . ら (1 9 8 7) J . V i r o l . 6 1 : 1 6 4 7 - 1 6 5 0 、 B e n d e r , M . A . ら (1 9 8 7) J . V i r o l . 6 1 : 1 6 3 9 - 1 6 4 6 、 A d a m , M . A . 及び A . D . M i l l e r (1 9 8 8) J . V i r o l . 6 2 : 3 8 0 2 - 3 8 0 6 、 D u l l , T . ら (1 9 9 8) J . V i r o l . 7 2 : 8 4 6 3 - 8 4 7 1 、 Z u f f e r e y , R . ら (1 9 9 8) J . V i r o l . 7 2 : 9 8 7 3 - 9 8 8 0) 。 R I G G に付与された米国特許第 5 , 9 1 0 , 4 3 4 号 (「 M e t h o d f o r o b t a i n i n g r e t r o v i r u s p a c k a g i n g c e l l l i n e s p r o d u c i n g h i g h t r a n s d u c i n g e f f i c i e n c y r e t r o v i r a l s u p e r n a t a n t 」) において、レトロウイルスパッケージング細胞系を得るための方法が開示されており、引用することをもって本明細書の一部とする。レトロウイルスベクターの増殖、細胞集団 (例えば C D 4 ⁺ 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 (R a n g a , U . ら (1 9 9 7) J . V i r o l . 7 1 : 7 0 2 0 - 7 0 2 9 、 B a u e r , G . ら (1 9 9 7) B l o o d 8 9 : 2 2 5 9 - 2 2 6 7 、 B o n y h a d i , M . L . (1 9 9 7) J . V i r o l . 7 1 : 4 7 0 7 - 4 7 1 6 、 R a n g a , U . ら (1 9 9 8) P r o c . N a t l . A c a d . S c i . U . S . A . 9 5 : 1 2 0 1 - 1 2 0 6 、 S u , L . (1 9 9 7) B l o o d 8 9 : 2 2 8 3 - 2 2 9 0) 。

20

30

40

【 0 1 6 9 】

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

50

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

10

【0170】

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

20

30

【0171】

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ADGUCをコードするポリヌクレオチドを標的細胞に送達する。プロトタイプのウイルスであるセムリキ森林熱ウイルス (Semliki Forest Virus, SFV) の生物学的研究が広範に行われており、遺伝子導入ベクターがSFVゲノムに基づいていることが分かった (Garoff, H. および K. - J. Li (1998) *Cun. Opin. Biotech.* 9: 464 - 469)。ウイルスRNAの複製中に、通常はウイルスのキャプシッドタンパク質をコードするサブゲノムRNAが作り出される。このサブゲノムRNAは、完全長のゲノムRNAより高いレベルに複製されるため、酵素活性 (例えばプロテアーゼ及びポリメラーゼ) を有するウイルスタンパク質に比べてキャプシッドタンパク質が過剰産生される。同様に、ADGUCに対するコード配列をキャプシッ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ADGUCコードRNAが産生され、高レベルのADGUCが合成される。通常はウイルスの感染が数日以内での細胞溶解に関係する一方で、シンドビスウイルス (SIN) の変異体を有するハムスター正常腎臓細胞 (BHK-21) の持続的な感染を確立する能力は、ウイルスの溶解複製を遺伝子治療に適用できるように好適に変更可能であることを示唆している (Dryga, S. A. ら. (1997) *Virology* 228: 74 - 83)。様々な宿主にウイルスを導入できることから、様々なタイプの細胞にADGUCを

40

50

導入することができる。或る集団におけるサブセットの細胞の特定形質導入は、形質導入前に細胞の選別を必要とし得る。ウイルスの感染性cDNAクローンの処置方法、ウイルスのcDNA及びRNAの形質移入方法及びウイルスの感染方法は、当業者に公知である。

【0172】

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

10

【0173】

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

20

【0174】

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

30

【0175】

本発明の相補リボ核酸分子及びリボザイムは、核酸分子合成のために当分野でよく知られている任意の方法を用いて作製し得る。任意の方法には、固相フォスフォアミダイト化合物合成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、ADGUC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、細胞株、細胞または組織内に導入することができる。

40

【0176】

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

50

【0177】

本発明の追加実施例は、A D G U C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発現変化を起こすの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プロモータ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、A D G U Cの発現または活性の増加に関連する疾患の治療においては、A D G U Cをコードするポリヌクレオチドの発現を特異的に阻害する化合物が治療上有用であり、A D G U C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、A D G U Cをコードするポリヌクレオチドの発現を特異的に促進する化合物が治療上有用であり得る。

10

【0178】

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。A D G U Cをコードするポリヌクレオチドを含むサンプルは、少なくとも1つの試験化合物に曝露して得る。サンプルには例えば、無傷細胞、透過化処理した細胞、無細胞再構成系または再構成生化学系があり得る。A D G U C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、A D G U C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つ以上の試験化合物に曝露される及び曝露されないポリヌクレオチドの発現の比較に対する基礎を形成し得る。試験化合物に曝露されるポリヌクレオチドの発現における変化の検出は、ポリヌクレオチドの発現を変異する際に試験化合物が有効であることを示している。特異ポリヌクレオチドの変異発現に有効な化合物に対して、例えば Schizosaccharomyces pombe 遺伝子発現系 (Atkins, D. ら (1999) 米国特許第5,932,435号、Arndt, G.M. ら (2000) *Nucleic Acids Res.* 28:E15) または HeLa 細胞等のヒト細胞系 (Clarke, M.L. ら (2000) *Biochem. Biophys. Res. Commun.* 268:8-13) を用いてスクリーニングを実行する。本発明の特定の実施例は、特異的ポリヌクレオチド配列に対するアンチセンス活性のためのオリゴヌクレオチド (デオキシリボヌクレオチド、リボヌクレオチド、ペプチド核酸、修飾オリゴヌクレオチド) の組み合わせライブラリをスクリーニングすることに関する (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)。

20

30

40

【0179】

ベクターを細胞または組織に導入する多数の方法が利用可能であり、in vivo、in vitro 及び ex vivo の使用に対して同程度に適している。ex vivo 治療の場合、ベクターを患者から採取した幹細胞内に導入し、クローニング増殖して同一患者に自家移植で戻すことができる。トランスフェクション、リボソーム注入またはポリカチオンアミノポリマーによる輸送は、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を参照)。

【0180】

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

50

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。

【0181】

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

【0182】

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。

【0183】

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

【0184】

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。

【0185】

ADGUCまたはその断片を含む高分子を直接細胞内に送達するべく、特殊な形態に組成物が調製されるのが好ましい。例えば、細胞不透過性高分子を含むリポソーム製剤は、細胞融合及び高分子の細胞内送達を促進し得る。或いは、ADGUC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: 1569-1572)。

【0186】

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、サル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。

【0187】

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、ADGUCまたはその断片、ADGUCの抗体、ADGUC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。治療有効度及び毒性は、細胞培養または動物実験における標準的な薬剤手法によって、例えばED₅₀（集団の50%の医薬的有効量）またはLD₅₀（集団の50%の致死量）を測定するなどして決定することができる。毒性効果の薬用効果に対する投与量の比は、治療指数であり、LD₅₀/ED₅₀比として表すことができる。高い治療指数を示すような成分が望ましい。細胞培養アッセイ及び動物実験から得られたデータは、ヒトに用いるための投与量の範囲を調剤するのに用いられる。このような組成物が含まれる投与量は、毒性を殆ど或いは全く含まず、ED₅₀を含むような血中濃度の範囲にあることが好ましい。用いられる投与形態

10

20

30

40

50

、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。

【0188】

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。

【0189】

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

【0190】

(診断)

別の実施例では、ADGUC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ADGUCやADGUC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、ADGUC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。ADGUC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ADGUCを検出する方法が含まれる。この抗体は修飾されたものもされていないものも可能であり、レポーター分子との共有結合または非共有結合で標識化できる。レポーター分子としては広くさまざまな種類が本分野で知られており、また使用可能であるが、そのうちのいくつかは上記で説明されている。

【0191】

ADGUC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、ADGUC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。正常或いは標準的なADGUCの発現の値は、複合体の形成に適した条件下で、正常な哺乳動物、例えばヒトなどの被験者から採取した体液または細胞とADGUCに対する抗体とを結合させることによって決定する。標準複合体形成量は、種々の方法、例えば測光法で定量できる。被験者、対照及び疾患の生検組織からのサンプルのADGUCの発現の量が基準値と比較される。基準値と被験者との偏差が疾患を診断するパラメータとなる。

【0192】

別の実施例によれば、ADGUCをコードするポリヌクレオチドを診断目的に用いることもできる。用いることができるポリヌクレオチドには、オリゴヌクレオチド配列、相補的RNA及びDNA分子、そしてPNAが含まれる。このポリヌクレオチドを用いて、疾患と相関し得るADGUCを発現する生検組織における遺伝子の発現を検出し定量する。診断アッセイは、ADGUC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ADGUC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。

【0193】

ある実施形態では、ADGUCをコードする核酸配列を同定するために、ADGUC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。プローブが同定するのはADGUC、突然変異体または関連配列をコードするような、天然に存在する配列のみであるか否かは、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によって決定されることになる。ここで、プローブの特異性とは、プローブが高特

10

20

30

40

50

異領域（例えば5'調節領域）からなるのか、低特異領域（例えば保存されたモチーフ）からなるのかということである。

【0194】

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A D G U C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。目的の本発明のハイブリダイゼーションプローブには、DNAあるいはRNAが可能であり、SEQ ID NO: 6-10の配列、或いはA D G U C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。

【0195】

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

【0196】

A D G U Cをコードするポリヌクレオチド配列を用いて、A D G U Cの発現に関連する疾患を診断することが可能である。限定するものではないがこのような疾患の例として、神経障害として、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(cerebelloretinal hemangioblastomatosis)、脳3叉神経血管症候群、精神薄弱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、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、動静脈瘤などの心疾患、アテローム性動脈硬化症、高血圧、血管炎、レイノー病、動脈瘤、動脈乖離、拡張蛇行静脈、血栓、静脈血栓症、血管の腫瘍、血栓崩壊の合併症、バルーン血管形成術(balloon angioplasty)、血管置換術、大動脈冠動脈バイパス術移植手術(coronary artery bypass graft surgery)、うっ血性心不全、虚血性心疾患、狭心症、心筋梗塞、高血圧性心疾患、変性弁膜性心疾患、石灰化大動脈弁狭窄症、先天性2尖大動脈弁、僧帽弁輪状石灰化(mitral annular calcification)、僧帽弁脱出、リウマチ熱、リウマチ性心疾患、感染性心内膜炎、非細菌性血栓性心内膜炎、全身性紅斑性狼瘡の心内膜炎、カルチノイド心疾患、心筋症、心筋炎、心膜炎、腫瘍性心疾患、先天性心臓疾患、心臓移植の合併症などの心疾患、先天性肺異常、肺拡張不全、肺うっ血及び肺水腫、肺動脈塞栓症、肺出血、肺梗塞、肺高血圧症、血管硬化症、閉塞性肺疾患、拘束型肺疾患、慢性閉塞性肺疾患、肺気腫、慢性気管支炎、気管支喘息、細気管支拡張症、細菌性肺炎、ウイルス性肺炎及びマイコプラズマ肺炎、肺膿瘍、肺結核、びまん性間質性疾患、塵肺症、サルコイド症、特発性肺線維症、剥離性間質性肺炎、過敏症肺炎、肺好酸球増加閉塞性細気管支炎 器質性肺炎(pulmonary eos

10

20

30

40

50

inophilia bronchiolitis obliterans - organizing pneumonia)、肺移植の合併症があり、視覚系疾患の中には、結膜炎、乾性角結膜炎、角膜炎、上強膜炎、虹彩炎、後部ブドウ膜炎、緑内障、一過性黒内障、虚血性視神経症、視神経炎、レーバー遺伝性視神経症、中毒性視覚神経炎、硝子体剥離、網膜剥離、白内障、黄斑変性症、中心性漿液性脈絡網膜症、色素性網膜炎、脈絡膜黒色腫、球後腫瘍、交叉腫瘍(chiasmaltumor)、生殖系疾患としてプロラクチン産生障害があり、卵管疾患、排卵欠損症、子宮内膜症、性周期障害、月経周期障害、多嚢胞性卵巣症候群、卵巣過刺激症候群、子宮内膜癌、卵巣癌、子宮筋腫、自己免疫疾患、子宮外妊娠、催奇形、乳がん、繊維嚢胞性乳房疾患、乳漏症、精子形成破壊、精子異常生理、精巣癌、前立腺癌、良性前立腺肥大、前立腺炎、パイロニー病、性交不能、男性乳癌、女性化乳房が含まれる。平滑筋疾患には、限定されるものではないが平滑筋の正常な活動の欠陥或いは変性すべてが含まれ、狭心症、アナフィラキシー性ショック、不整脈、喘息、心臓血管性ショック、クッシング症候群、高血圧、低血糖症、心筋梗塞、偏頭痛、およびクロム親和細胞腫(褐色細胞腫)、および心筋症、脳疾患、癲癇、カーンズ・セイヤー症候群、乳酸アシドーシス、筋クローヌス性疾患、および眼筋麻痺などのミオパシーが含まれる。そして細菌による感染は肺炎球菌感染症、ブドウ球菌、連鎖球菌、バシラス菌、コリネバクテリウム、クロストリジウム、髄膜炎菌、淋菌、リステリア、モラクセラ、キングセラ、ヘモフィルス、レジオネラ、百日咳菌、シゲラ、サルモネラ、カンピロバクターを含むグラム陰性菌腸内細菌、シュードモナス、ビブリオ、ブルセラ、野兔病菌、エルシニア、バルトネラ、nocardium、放線菌、ミコバクテリウム、spirochaetale、リケッチア、クラミジア、マイコプラズマに分類される。ADGUC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ADGUC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。

【0197】

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

【0198】

ADGUCの発現に関連する疾患の診断の基準となるものを提供するために、発現の正常すなわち標準的なプロファイルが確立される。これは、ハイブリダイゼーション或いは増幅に好適な条件下で、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、ADGUC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確定する。

【0199】

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

ら数ヶ月の期間にわたる治療の効果を示し得る。

【0200】

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。

【0201】

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

10

【0202】

或る実施態様において、A D G U C 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、一塩基多型性 (S N P) を検出し得る。S N P は、多くの場合にヒトの先天性または後天性遺伝病の原因となるような置換、挿入及び欠失である。限定するものではないが S N P の検出方法には、S S C P (s i n g l e - s t r a n d e d c o n f o r m a t i o n p o l y m o r p h i s m) 及び蛍光 S S C P (f S S C P) 法がある。S S C P では、A D G U C をコードするポリヌクレオチド配列由来のオリゴヌクレオチドプライマーを用いたポリメラーゼ連鎖反応 (P C R) で D N A を増幅する。D N A は例えば、病変組織または正常組織、生検サンプル、体液その他に由来し得る。D N A 内の S N P は、一本鎖形状の P C R 生成物の 2 次及び 3 次構造に差異を生じさせる。差異は非変性ゲル中でのゲル電気泳動法を用いて検出可能である。f S C 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高処理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、例えば高処理 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。

20

30

【0203】

A D G U C 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 (例えば、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 を参照)。目的のオリゴマーが種々の希釈液中に存在し、分光光度法または比色反応によって定量が迅速になるような高処理フォーマットのアッセイを行うことによって、複数のサンプルの定量速度を加速することができる。

40

【0204】

更に別の実施例では、本明細書に記載した任意のポリヌクレオチド配列由来のオリゴヌクレオチドまたはより長い断片を、マイクロアレイにおけるエレメントとして用いることができる。多数の遺伝子の関連発現レベルを同時にモニターする転写イメージング技術にマ

50

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。

【0205】

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

10

【0206】

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

20

【0207】

転写イメージは、組織、細胞株、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。

30

【0208】

本発明のポリヌクレオチドの発現のプロファイルを作製する転写イメージはまた、工業的または天然の環境化合物の毒性試験のみならず、in vitro モデル系及び薬剤の前臨床評価に関連して使用し得る。全ての化合物は、作用及び毒性のメカニズムを暗示する、しばしば分子フィンガープリントまたは毒性シグネチャと称されるような特徴的な遺伝子発現パターンを惹起する (Nuwaysir, E. F. ら、(1999) Mol. Carcinog. 24:153-159、Steiner, S. 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シグネチャと同一のシグネチャを有する場合には、毒性特性を共有している可能性がある。フィンガープリントまたはシグネチャは、多数の遺伝子及び遺伝子ファミリーからの発現情報を含んでいる場合には、最も有用且つ正確である。理想的には、発現のゲノム全域にわたる測定が最高品質のシグネチャを提供する。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。毒性シグネチャの要素に遺伝子機能を割り当てることが毒性メカニズムの解釈に役立つが、毒性の予測につながるシグネチャの統計的に一致させるのに遺伝子機能の知識は必要とされない (例えば2000年2月29日にNational Institute of Environmental Health Sciencesより発行されたPress Release 00-02を参照

40

50

されたい。これについては <http://www.niehs.nih.gov/oc/news/toxchip.htm> で入手可能である)。従って、中毒学的スクリーニングの際に毒性シグネチャを用いて、全ての発現した遺伝子配列を含めることは重要且つ望ましいことである。

【0209】

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

10

【0210】

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

20

30

【0211】

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

40

【0212】

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. および J. Seilhamer (1997) Electrophoresis 18:533-537)、転写イメージにはそれ程影響しないがタンパク質のプロフィールを変化させるような化合物の分析においてプロテオーム毒性シグネチャは有用たり得る。更に、体液中での転写の分析はmRNA急速な分解により困難であるので、タンパク質のプロ

50

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。

【0213】

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

【0214】

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

【0215】

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. U S A 93:10614-10619、Baldeschweiler らの (1995) PCT 出願第 WO95/251116 号、Shalon, D. らの (1995) PCT 出願第 WO95/35505 号、Heller, R. A. ら (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155、Heller, M. J. らの (1997) 米国特許第 5,605,662 号等を参照)。様々なタイプのマイクロアレイが公知であり、詳細については、DNA Microarrays: A Practical Approach, M. Schena, 編集。(1999) Oxford University Press, London に記載されている。該文献は、特に引用することを以って本明細書の一部となす。

【0216】

本発明の別の実施例ではまた、ADGUC をコードする核酸配列を用いて、天然のゲノム配列をマッピングするのに有用なハイブリダイゼーションプローブを作製することが可能である。コード配列または非コード配列のいずれかを用いることができ、或る例では、コード配列全体で非コード配列が好ましい。例えば、多重遺伝子ファミリーのメンバー内でのコード配列の保存により、染色体マッピング中に望ましくないクロスハイブリダイゼーションが生じる可能性がある。核酸配列は、特定の染色体、染色体の特定領域または人工形成の染色体、例えば、ヒト人工染色体 (HAC)、酵母人工染色体 (YAC)、細菌人工染色体 (BAC)、細菌 P1 産物、或いは単一染色体 cDNA ライブラリに対してマッピングされる (Harrington, J. J. ら (1997) Nat Genet. 15:345-355、Price, C. M. (1993) Blood Rev. 7:127-134、Trask, B. J. (1991) Trends Genet. 7:149-154 等を参照)。一度マッピングすると、本発明の核酸配列を用いて例えば病状の遺伝を特定の染色体領域の遺伝または制限酵素断片長多型 (RFLP) と関連させるような遺伝子連鎖地図を発生させ得る。(例えば、Lander, E. S 及び D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357 を参照)。

【0217】

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, 965-968 ページ、等を参照)。遺伝地図データの例は、種々の科学雑誌あるいは Onl

10

20

30

40

50

ine Mendelian Inheritance in Man (OMIM)のウェブサイトに見ることができる。物理的な染色体地図上のADGUCをコードする遺伝子の位置と特定の疾患との相関性、或いは特定の疾患に対する素因が、このような疾患と関連するDNA領域の決定に役立つため、更なる位置を決定するクローニングが行われる。

【0218】

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

【0219】

本発明の別の実施例では、任意の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、ADGUC、ADGUC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。ADGUCと検査する薬剤との結合による複合体の形成を測定可能である。

20

【0220】

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

30

【0221】

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

【0222】

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

40

【0223】

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

【0224】

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60/215,476、60/223,545、60/229,876、60/234,838、60/236,483は、言及することをもって本明細書の一部となす。

【0225】

50

(実施例)

1 cDNAライブラリの作製

Incyte cDNAは、LIFESEQ GOLDデータベース (Incyte Genomics, Palo Alto CA) に記載されたcDNAライブラリに由来するものであり、表4の列5に列記した。ホモジナイズしてグアニジニウムイソチオシアネート溶液に溶解した組織もあり、また、ホモジナイズしてフェノールまたは好適な変性剤の混合液に溶解した組織もある。変性剤の混合液は、例えばフェノールとグアニジニウムイソチオシアネートの単相溶液であるTRIZOL (Life Technologies) 等である。結果として得られた溶解物は、塩化セシウムクッション上で遠心分離するかクロロホルムで抽出した。イソプロパノールか、酢酸ナトリウムとエタノールか、いずれか一方、或いは別の方法を用いて、溶解物からRNAを沈殿させた。 10

【0226】

RNAの純度を高めるため、RNAのフェノールによる抽出及び沈殿を必要な回数繰り返した。場合によっては、DNアーゼでRNAを処理した。殆どのライブラリでは、オリゴd(T)連結常磁性粒子 (Promega)、OLIGOTEXラテックス粒子 (QIAGEN, Chatsworth CA) またはOLIGOTEX mRNA精製キット (QIAGEN) を用いて、ポリ(A+) RNAを単離した。別法では、別のRNA単離キット、例えばPOLY(A) PURE mRNA精製キット (Ambion, Austin TX) を用いて組織溶解物からRNAを直接単離した。 20

【0227】

場合によってはStratagene社へのRNA提供を行い、対応するcDNAライブラリをStratagene社が作製することもあった。そうでない場合は、UNIZAPベクターシステム (Stratagene) またはSUPERSCRIPTプラスミドシステム (Life Technologies) を用いて本技術分野で公知の推奨方法または類似の方法でcDNAを合成し、cDNAライブラリを作製した (前出のAusubel, 1997, unit 5.1 - 6.6等を参照)。逆転写は、オリゴd(T) またはランダムプライマーを用いて開始した。合成オリゴヌクレオチドアダプターを二本鎖cDNAに連結反応させ、好適な制限酵素でcDNAを消化した。殆どのライブラリに対して、cDNAのサイズ (300 ~ 1000 bp) 選択は、SEPHACRYL S1000、SEPHAROSE CL2BまたはSEPHAROSE CL4Bカラムクロマトグラフィー (Amersham Pharmacia Biotech)、或いは調製用アガロースゲル電気泳動法を用いて行った。合成オリゴヌクレオチドアダプターを二本鎖cDNAに連結反応させ、好適な制限酵素または酵素でcDNAを消化した。好適なプラスミドは、例えばPBLUESCRIPTプラスミド (Stratagene)、pSPORT1プラスミド (Life Technologies)、PCDNA2.1プラスミド (Invitrogen, Carlsbad CA)、PBK-CMVプラスミド (Stratagene) あるいはpINCY (Incyte Genomics, Palo Alto CA) またはその誘導体である。組換えプラスミドは、Stratagene社のXL1-Blue、XL1-Blue MRFまたはSOLR、或いはLife Technologies社のDH5、DH10BまたはELECTROMAX DH10Bを含むコンピテント大腸菌細胞に形質転換した。 30 40

【0228】

2 cDNAクローンの単離

UNIZAPベクターシステム (Stratagene) を用いた*in vivo*切除によって、或いは細胞溶解によって、実施例1のようにして得たプラスミドを宿主細胞から回収した。MagicまたはWIZARD Minipreps DNA精製システム (Promega)、AGTC Miniprep精製キット (Edge Biosystems, Gaithersburg MD)、QIAGEN社のQIAWELL 8 Plasmid、QIAWELL 8 Plus Plasmid及びQIAWELL 8 Ultra Plasmid 精製システム、R.E.A.L. Prep 96プラスミドキットの 50

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。

【0229】

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

10

【0230】

3 シークエンシング及び分析

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

20

30

【0231】

Incyte cDNA配列に由来するポリヌクレオチド配列は、ベクター、リンカー及びポリ(A)配列を除去し、あいまいな塩基対をマスクすることによって有効性を確認した。その際、BLAST、動的プログラミング及び隣接ジヌクレオチド頻度分析に基づくアルゴリズム及びプログラムを用いた。次に、BLAST、FASTA及びBLIMPSに基づくプログラムを用いて、プログラム中の注釈を得るべく、公共のデータベース、例えばGenBankの霊長類及びげっ歯類、哺乳動物、脊椎動物、真核生物のデータベースと、BLOCKS、PRINTS、DOMO、PRODOM及びPFAM等の隠れMarkovモデル(HMM)ベースのタンパク質ファミリーデータベースの選択に対するIncyte cDNA配列またはその翻訳を問い合わせた(HMMは、遺伝子ファミリーのコンセンサス1次構造を分析する確率的アプローチである。例えば、Eddy, S. R. (1996) Curr. Opin. Struct. Biol. 6:361-365を参照のこと)。問合せは、BLAST、FASTA、BLIMPS及びHMMEERに基づくプログラムを用いて行った。Incyte cDNA配列は、完全長のポリヌクレオチド配列を産出するように構築した。或いは、GenBank cDNA、GenBank EST、ステッチされた配列、ストレッチされた配列またはGenscan予測コード配列(実施例4及び5を参照)を用いてIncyte cDNAの集団を完全長まで伸長させた。Phred、Phrap及びConsedに基づくプログラムを用いて構築し、GenMark、BLAST及びFASTAに基づくプログラムを用いてcDNAの集団をオープンリーディングフレームに対してスクリーニングした。対応する完全長

40

50

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等の隠れマルコフモデル (HMM) ベースのタンパク質ファミリーデータベースに対する問合せによって完全長ポリペプチド配列を分析した。完全長ポリヌクレオチド配列はまた、MACDNASIS PROソフトウェア (日立ソフトウェアエンジニアリング, South San Francisco CA) 及びLASERGENEソフトウェア (DNASTAR) を用いて分析した。ポリヌクレオチド及びポリペプチド配列アラインメントは、アラインメントした配列と配列の一致率も計算するMEGALIGNマルチシークエ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。

【0232】

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

【0233】

完全長ポリヌクレオチド配列およびポリペプチド配列の構築及び分析に用いる上記のプログラムは、SEQ ID NO: 6 - 10のポリヌクレオチド配列断片の同定にも利用できる。ハイブリダイゼーション及び増幅技術に有用である約20~約4000ヌクレオチドの断片を表4の列4に示した。

【0234】

4 ゲノムDNAからのコード配列の同定及び編集

推定上のアデニルシクラーゼ及びグアニルシクラーゼは、公共のゲノム配列データベース (例えば、gbpriやgbhtg) においてGenscan遺伝子同定プログラムを実行して初めに同定した。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. およびS. Karlin (1997) J. Mol. Biol. 268: 78 - 94、Burge, C. 及びS. Karlin (1998) Cuff. Opin. Struct. Biol. 8: 346 - 354等を参照)。プログラムは予測エキソンを連結し、メチオニンから停止コドンに及ぶ構築されたcDNA配列を形成する。Genscanの出力は、ポリヌクレオチド及びポリペプチド配列のFASTAデータベースである。Genscanが一度に分析する配列の最大範囲は、30kbに設定した。これらのGenscan推定cDNA配列の内、どの配列がアデニルシクラーゼ及びグアニルシクラーゼをコードするかを決定するために、コードされたポリペプチドをPFAMモデルにおいてアデニルシクラーゼ及びグアニルシクラーゼについて問合せて分析した。Incyte cDNA配列の相同体をアデニルシクラーゼ及びグアニルシクラーゼとして注釈を付けてきたことにより、潜在的アデニルシクラーゼ及びグアニルシクラーゼも同定した。こうして選択されたGenscan予測配列は、次にBLAST分析により公共データベースgbpri及びgbhtgと比較した。必要であれば、genpeptからヒットしたトップのBLASTと比較することによりGenscan予測配列を編集し、余分なまたは取り除かれたエキソンなどのGenscanにより予測された配列のエラーを修正する。BLAST分析はまた、任意のIncyte cDNAまたはGenscan予測配列の公共cDNA適用範囲の発見に用いられるので、転写の証拠を提供する。Incyte cDNA適用範囲が利用できる場合には、この情報を用いてGenscan予測配列を修正また

は確認した。完全長ポリヌクレオチド配列は、実施例 3に記載された構築プロセスを用いて、Incyte cDNA 配列及び/または公共の cDNA 配列で Genscan 予測コード配列を構築することにより得た。或いは、完全長ポリヌクレオチド配列は編集または非編集の Genscan 予測コード配列に完全に由来する。

【0235】

5 cDNA 配列データを使ったゲノム配列データの構築
ステッチ配列 (Stitched Sequence)

部分 cDNA 配列は、実施例 4に記載の Genscan 遺伝子同定プログラムにより予測されたエキソンを用いて伸長させた。実施例 3に記載されたように構築された部分 cDNA は、ゲノム DNA にマッピングし、関連する cDNA 及び 1 つ若しくは複数のゲノム配列から予測された Genscan エキソンを含むクラスタに分解した。cDNA 及びゲノム情報を統合するべくグラフ理論及び動的プログラミングに基づくアルゴリズムを用いて各クラスタを分析し、引き続いて確認、編集または伸長して完全長配列を産出するような潜在的スプライス変異体を生成した。間隔全体の長さがクラスタ中の 2 以上の配列に存在するような配列を同定し、そのように同定された間隔は推移により等しいと考えられた。例えば、1 つの cDNA 及び 2 つのゲノム配列に間隔が存在する場合、3 つの間隔は全て等しいと考えられる。このプロセスは、無関係であるが連続したゲノム配列を cDNA 配列により結び合わせて架橋し得る。このようにして同定された区間を、親配列 (parent sequence) に沿って現われるようにステッチアルゴリズムで縫い合わせ、可能な最も長い配列および変異配列を作製する。1 種類の親配列に沿って発生した間隔と間隔との連鎖 (cDNA - cDNA またはゲノム配列 - ゲノム配列) は、親の種類を変える連鎖 (cDNA - ゲノム配列) に優先した。結果として得られるステッチ配列は、BLAST 分析により公共データベース genpept 及び gbpr1 に翻訳されて比較された。Genscan により予測された不正確なエキソンは、genpept からヒットしたトップの BLAST と比較することにより修正した。必要な場合には、追加 cDNA 配列を用いるかゲノム DNA の検査により配列を更に伸長させた。

【0236】

ストレッチ配列 (Stretched Sequence)

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

【0237】

6 ADGUC をコードするポリヌクレオチドの染色体マッピング

SEQ ID NO: 6 - 10 を構築するために用いた配列を、BLAST 及び Smith-Waterman アルゴリズムを用いて、Incyte LIFESEQ データベース及び公共のドメインデータベースの配列と比較した。SEQ ID NO: 6 - 10 に適合するデータベースから得た配列は、Phrap (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 (SHGC)、ホワイトヘッド・ゲノム研究所 (WIGR)、Genethon 等の公的な情報源から入手可能な放射線ハイブリッド及び遺伝地図データを用

いて、クラスタ化された配列が前もってマッピングされたかを測定した。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。

【0238】

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

10

【0239】

7 ポリヌクレオチド発現の分析

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

20

【0240】

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

【数1】

(BLASTスコア×配列一致率)

5 × (長さ(配列1), 長さ(配列2))の最小値

30

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

40

【0241】

或いは、ADGUCをコードするポリヌクレオチド配列は、由来する組織に対して分析する。例えば或る完全長配列は、Incycyte cDNA配列(実施例3を参照)と少なくとも一部は重畳するように構築される。各cDNA配列は、ヒト組織から作製されたcDNAライブラリに由来する。各ヒト組織は、心血管系、結合組織、消化系、胚構造、内分

50

泌系、外分泌腺、女性生殖器、男性生殖器、生殖細胞、血液および免疫系、肝臓、筋骨格系、神経系、膵臓、呼吸器系、感覚器官、皮膚、顎口腔系、分類不能/混合、または尿管などの生物/組織のカテゴリのーつに分類される。各カテゴリのライブラリ数を数えて、全カテゴリ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/病状カテゴリ即ち癌、細胞株、発達、炎症、神経性、外傷、心血管、鬱血、その他の1つに分類される。各カテゴリのライブラリ数を数えて、全カテゴリの総ライブラリ数で除する。得られるパーセンテージは、ADGUCをコードするcDNAの疾患特異的な発現と組織を反映する。cDNA配列およびcDNAライブラリ/組織の情報は、LIFESEQ GOLD データベース (Incyte Genomics, Palo Alto, CA) から得ることができる。

10

【0242】

8 ポリヌクレオチドをコードするADGUC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区間は全て回避した。

20

【0243】

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

【0244】

高忠実度の増幅が、当業者によく知られている方法を利用したPCR法によって得られた。PCRは、PTC-200 サーマルサイクラー (MJ Research, Inc.) を用いて96穴プレート内で行った。反応混合液は、鋳型DNA及び200 nmolの各プライマー、 Mg^{2+} と $(NH_4)_2SO_4$ と β -メルカプトエタノールを含むバッファー、Taq DNAポリメラーゼ (Amersham Pharmacia Biotech)、ELONGASE 酵素 (Life Technologies)、Pfu DNAポリメラーゼ (Stratagene) を含む。プライマーの組、PCI AとPCI Bに対して以下のパラメータで増幅を行った。ステップ 1: 94 で3分間、ステップ 2: 94 で15秒間、ステップ 3: 60 で1分間、ステップ 4: 68 で2分間、ステップ 5: ステップ 2、3、及び4を20回繰り返す。ステップ 6: 68 で5分間、ステップ 7: 4 で保存。別法では、プライマー対、T7とSK+に対して以下のパラメータで増幅を行った。ステップ 1: 94 で3分間、ステップ 2: 94 で15秒間、ステップ 3: 57 で1分間、ステップ 4: 68 で2分間、ステップ 5: ステップ 2、3、及び4を20回繰り返す。ステップ 6: 68 で5分間、ステップ 7: 4 で保存。

30

40

【0245】

各ウェルのDNA濃度は、1X TE及び0.5 μ lの希釈していないPCR産物に溶解した100 μ lのPICOGREEN定量試薬 (0.25 (v/v) PICOGREEN; Molecular Probes, Eugene OR) を不透明な蛍光光度計プレート (Coming Costar, Acton MA) の各ウェルに分配してDNAが試薬と結合できるようにして測定する。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 μ l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。

50

【0246】

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コレ
ラウイルスエンドヌクラーゼ (Molecular Biology Research
, Madison WI) を用いて消化し、pUC 18ベクター (Amersham P
harmacia 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・シークエンシングのために、消化したヌクレオチドを低濃度 (0.6~0.
8%) のアガロースゲル上で分離し、断片を切除し、寒天をAgar ACE (Prom
ega) で消化した。伸長させたクローンをT4リガーゼ (New England Bi
olabs, Beverly MA) を用いてpUC 18ベクター (Amersham
Pharmacia Biotech) に再連結し、Pfu DNAポリメラーゼ (Str
atagene) で処理して制限部位の張出部 (overhang) を満たし、大腸菌細
胞に形質移入した。形質移入した細胞を選択して抗生物質を含む培地に移し、それぞれの
コロニーを切りとってLB/2Xカルベニシリン培養液の384ウェルプレートに37
で一晩培養した。

【0247】

細胞を溶解して、Taq DNAポリメラーゼ (Amersham Pharmacia
Biotech) 及びPfu DNAポリメラーゼ (Stratagene) を用いて以
下の手順でDNAをPCR増幅した。ステップ 1: 94 で3分間、ステップ 2: 94
で15秒間、ステップ 3: 60 で1分間、ステップ 4: 72 で2分間、ステップ
5: ステップ2、3及び4を29回繰り返す。ステップ 6: 72 で5分間、ステップ
7: 4 で保存。上記したようにPICOGREEN試薬 (Molecular Pro
bes) でDNAを定量化した。DNAの回収率が低いサンプルは、上記と同一の条件を
用いて再増幅した。サンプルは20%ジメチルスルホキシド (1:2, v/v) で希釈
し、DYENAMIC エネルギートランスファー シークエンシングプライマー、及びD
YENAMIC DIRECT kit (Amersham Pharmacia Biot
ech) またはABI PRISM BIGDYE ターミネーターサイクル シークエンシ
ング反応キット (Terminator cycle sequencing ready
reaction kit) (Applied Biosystems) を用いてシークエ
ンシングした。

【0248】

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。

【0249】

9 個々のハイブリダイゼーションプローブの標識及び使用
SEQ ID NO: 6-10 から得たハイブリダイゼーションプローブを利用して、cD
NA、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を、OLIGO4.06ソフトウェア
(National Bioscience) のような最新式のソフトウェアを用いてデ
ザインし、50 pmolの各オリゴマーと、250 µCiの[³²P] アデノシン三
リン酸 (Amersham Pharmacia Biotech),) 及びT4ポリヌ
クレオチドキナーゼ (DuPont NEN, Boston MA) とを組み合わせて用い
ることにより標識する。標識したオリゴヌクレオチドは、SEPHADEX G-25超
細繊維分子サイズ排除デキストラン ビードカラム (Amersham Pharmacia
Biotech) を用いて十分に精製する。Ase I、Bgl II、Eco RI、P
st I、Xba IまたはPvu II (DuPont NEN) のいずれか1つのエンド
ヌクラーゼで消化されたヒトゲノムDNAの典型的な膜ベースのハイブリダイゼーショ
ン解析において、毎分10⁷ カウントの標識されたプローブを含むアリコットを用いる。

【0250】

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

【0251】

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、10 圧電印刷(インクジェット印刷、前出のB a l d e s c h w e i l e r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(S c h e n a (1 9 9 9) 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(S c h e n a , M . ら (1 9 9 5) S c i e n c e 2 7 0 : 4 6 7 - 4 7 0 , S h a l l o n . D . ら (1 9 9 6) G e n o m e R e s . 6 : 6 3 9 - 6 4 5 , M a r s h a l l , A . 及び J . H o d g s o n (1 9 9 8) N a t . B i o t e c h n o l . 1 6 : 2 7 - 3 1等を参照)。

【0252】

完全長cDNA、発現配列タグ(EST)、またはその断片またはオリゴマーは、マイクロアレイのエレメントと成り得る。ハイブリダイゼーションに好適な断片またはオリゴマーを、レーザGENEソフトウェア(DNA STAR)等の本技術分野で公知のソフトウェアを用いて選択することが可能である。アレイエレメントは、生物学的サンプル中でポリヌクレオチドを用いてハイブリダイズされる。生物学的サンプル中のポリヌクレオチドは、検出を容易にするために蛍光標識またはその他の分子タグに抱合される。ハイブリダイゼーション後、生物学的サンプルからハイブリダイズされていないヌクレオチドを除去し、30 蛍光スキャナを用いて各アレイエレメントにおいてハイブリダイゼーションを検出する。或いは、レーザ脱離及び質量スペクトロメトリを用いてもハイブリダイゼーションを検出し得る。マイクロアレイ上のエレメントにハイブリダイズする各ポリヌクレオチドの相補性の度合及び相対存在度は、算定し得る。一実施例におけるマイクロアレイの調整及び使用について、以下に詳述する。

【0253】

組織または細胞サンプルの調製

グアニジウムチオシアネート法を用いて組織サンプルから全RNAを単離し、オリゴ(dT)セルロース法を用いてポリ(A)⁺RNAを精製する。各ポリ(A)⁺RNAサンプルを、40 MMLV逆転写酵素、0.05pg/μlのオリゴ(dT)プライマー(21mer)、1×第一鎖合成バッファー、0.03unit/μlのRNアーゼ阻害因子、500μMのdATP、500μMのdGTP、500μMのdTTP、40μMのdCTP、40μMのdCTP-Cy3(BDS)またはdCTP-Cy5(Amer sham Pharmacia Biotech)を用いて逆転写する。逆転写反応は、GEMBR IGH Tキット(Incyte)を用いて200ngのポリ(A)⁺RNA含有の25体積ml内で行う。特異的対照ポリ(A)⁺RNAは、非コード酵母ゲノムDNAから i n v i t r o 転写により合成する。各反応サンプル(1つはCy3、もう1つはCy5標識)は、2.5mlの0.5M水酸化ナトリウムで処理し、37 で2時間インキュベートし、85 で20分間インキュベートして反応を停止させてRNAを分解させる。サンプルは、2つの連続するCHROMA SPIN 30ゲル濾過スピンカラム(CLON 50

TECH 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 μ l の5xSSC/0.2% SDS 中で再懸濁する。

【0254】

マイクロアレイの調製

本発明の配列を用いて、アレイエレメントを生成する。各アレイエレメントは、クローン化 cDNA インサートによりベクター含有細菌性細胞から増幅する。PCR 増幅は、cDNA インサートの側面に位置するベクター配列に相補的なプライマーを用いる。30 サイクルの PCR で1~2 ng の初期量から5 μ g 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。

10

【0255】

精製したアレイエレメントは、ポリマーコートされたスライドガラス上に固定する。顕微鏡スライドガラス (Corning) は、処理中及び処理後に0.1%のSDS及びアセトン中で超音波をかけ、蒸留水で非常に良く洗って洗浄する。スライドガラスは、4%フッ化水素酸 (VWR Scientific Products Corporation (VWR), West Chester PA) 中でエッチングし、蒸留水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 のオープンで硬化させる。

20

【0256】

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/ μ lのアレイエレメントDNA 1 μ l を高速機械装置により開放型キャピラリープリンティングエレメント (open capillary printing element) に充填する。装置はここで、スライド毎に約5 nlのアレイエレメントサンプルを加える。

30

【0257】

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。

【0258】

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、5xSSC, 0.2%SDSハイブリダイゼーション緩衝液中のCy3及びCy5標識したcDNA合成生成物を各0.2 μ g含む9 μ lのサンプル混合体を有する。サンプル混合体は、65 まで5分間加熱し、マイクロアレイ表面上で等分して1.8 cm²のカバーガラスで覆う。アレイは、顕微鏡スライドより僅かに大きい空洞を有する防水チェンバーに移す。チェンバーのコーナーに140 μ lの5xSSC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 (1xSSC, 0.1%SDS) において45 で10分間洗浄し、第2洗浄緩衝液中 (0.1xSSC) において45 で10分間各々3度洗浄して乾燥させる。

40

【0259】

検出

50

レポーター標識ハイブリダイゼーション複合体は、Cy3の励起のためには488nm、Cy5の励起のためには632nmでスペクトル線を発生し得るInnova70混合ガス10Wレーザ(Coherent, Inc., Santa Clara CA)を備えた顕微鏡で検出する。20×顕微鏡対物レンズ(Nikon, Inc., Melville NY)を用いて、アレイ上に励起レーザ光の焦点を当てる。アレイを含むスライドを顕微鏡のコンピュータ制御のX-Yステージに置き、対物レンズを通過してラスタースキャンする。本実施例で用いた1.8cm×1.8cmのアレイは、20μmの解像度でスキャンした。

【0260】

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

【0261】

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

【0262】

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

【0263】

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

【0264】

11 相補的ポリヌクレオチド

ADGUCをコードする配列或いはその任意の一部に対して相補的な配列は、天然のADGUCの発現を低下させるため即ち阻害するために用いられる。約15~30塩基対を含むオリゴヌクレオチドの使用について記すが、これより小さな或いは大きな配列の断片の場合でも本質的に同じ方法を用いることができる。Oligo4.06ソフトウェア(National Biosciences)とADGUC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、相補的なオリゴヌクレオチドを設計

して、リボソームが A D G U C をコードする転写物に結合するのを阻害する。

【 0 2 6 5 】

1 2 A D G U C の発現

A D G U C の発現及び精製は、細菌若しくはウイルスを基にした発現系を用いて行うことができる。細菌で A D G U C 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) で誘発されると A D G U C を発現する。真核細胞での A D G U C の発現は、昆虫細胞株または哺乳動物細胞株に一般にバキュロウイルスとして知られている A u t o g r a p h i c a c a l i f o r n i c a 核多角体病ウイルス (A c M N P V) を感染させて行う。バキュロウイルスの非必須の多角体遺伝子を、相同組換え或いは転移プラスミドの媒介を伴う細菌の媒介による遺伝子転移のどちらかによって、A D G U C 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、強い多角体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)。

【 0 2 6 6 】

殆どの発現系では、A D G U C が、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 F L A G や 6 - H i s などのペプチドエピトープ標識で合成された融合タンパク質となるため、未精製の細胞溶解物からの組換え融合タンパク質の親和性ベースの精製が素早く 1 回で行うことができる。G S T は日本住血吸虫からの 2 6 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 A D G U C からタンパク質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 ヒスチジン残基が連続して伸長した 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) 1 0 章、1 6 章に記載されている。これらの方法で精製した A D G U C を直接用いて以下の 実施例 1 6、1 7、及び 1 8 のアッセイを行うことができる。

【 0 2 6 7 】

1 3 機能的アッセイ

A D G U C 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 A D G U C をコードする配列の発現によって評価する。c D N A 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 c D N A をサブクローニングする。選り抜きのベクターには、p C M V S P O R T プラスミド (L i f e T e c h n o l o g i e s) 及び p C R 3 . 1 プラスミド (I n v i t r o g e n , C a r l s b a d C A) 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。リボソーム製剤或いは電気穿孔法を用いて、5 ~ 1 0 μ g の組換えベクターをヒト細胞株、例えば内皮由来または造血由来の細胞株に一時的に形質移入する。更に、標識タンパク質をコードする配列を含む 1 ~ 2 μ g のプラスミドを同時に形質移入する。標識タンパク質の発現により、形質移入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

10

20

30

40

50

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

10

【0268】

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

20

【0269】

14 ADGUCに特異的な抗体の作製

ポリアクリルアミドゲル電気泳動法(PAGE; Harrington, M. G. (1990) Methods Enzymol. 182: 488-495を参照)または他の精製技術を用いて実質上精製されたADGUC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。

30

【0270】

或いは、レーザGENEソフトウェア(DNA STAR)を用いてADGUC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)。

【0271】

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-KLM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ADGUC活性を検査するには、ペプチドまたはADGUCを基板に結合し、1%BSAを用いてブロッキング処理し、ウサギ抗血清と反応させて洗浄し、さらに放射性ヨウ素標識されたヤギ抗ウサギIgGと反応させる。

40

【0272】

15 特異的な抗体を用いる天然ADGUCの精製

天然または組換えADGUCを、ADGUC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ADGUC抗

50

体を活性化クロマトグラフィー用樹脂、例えばCNBr活性化セファロース (Amersham Pharmacia Biotech) 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。

【0273】

ADGUCを含む培養液をイムノアフィニティーカラムに通し、ADGUC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。そのカラムを、抗体とADGUCとの結合を切るような条件で (例えば、pH 2~3のバッファー、或いは高濃度の尿素またはチオシアン酸塩イオンのようなカオトロピック塩で) 溶出させ、ADGUCを回収する。

【0274】

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

【0275】

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

【0276】

ADGUC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 PATHCALLING プロセス (CuraGen Corp., New Haven CT) に用いて、遺伝子の2大ライブラリにコードされるタンパク質間の全ての相互作用を決定することができる (Nandabalan, K. ら (2000) 米国特許第6,057,101号)。

【0277】

17 ADGUC活性の実証
ADGUCのアデニルシクラーゼ活性はATPをcAMPに変換する能力により実証される (Mittal, C. K. (1986) Meth. Enzymol. 132: 422-428)。このアッセイでは、ADGUCは基質 [^{32}P] ATPと共にインキュベートされ、次に過剰基質が生成物であるサイクリック [^{32}P] AMPから分離される。ADGUC活性はpH 7.5の0.6 M Tris-HClを5 μl 、0.2 M MgCl₂を5 μl 、クレアチン・ホスホキナーゼを3 units含む150 mM クレアチンリン酸を5 μl 、4.0 mM 1-メチル-3-イソブチルキサンチンを5 μl 、20 mM cAMPを5 μl 、20 mM ジチオスレートールを5 μl 、10 mM ATPを5 μl 、[^{32}P] ATP ($2-4 \times 10^6$ cpm)を10 μl 、水を含む合計量100 μl を使い捨ての培養チューブ (12 x 75 mm) に入れて測定する。反応混合液は前もって30 $^\circ\text{C}$ に温められる。反応はADGUCを前もって温められた反応混合液に添加することにより開始する。30 $^\circ\text{C}$ で10-15分インキュベートした後、反応は30%のよく冷やしたトリクロロ酢酸 (TCA) を25 μl 加えることにより終了する。インキュベーション時間「0」の反応溶液とADGUCを入れないでインキュベートした反応溶液は、ネガティブコントロールとして使用される。産物はイオン交換クロマトグラフィーにより分離され、 ^{32}P -ラジオアイソトープカウンタを使用してサイクリック [^{32}P] AMPを定量する。ADGUC活性は、反応中に形成されるサイクリック [^{32}P] AMPの量に比例する。

10

20

30

40

50

【0278】

ADGUCのグアニリルシクラーゼ活性はGTPをcGMPに変換する能力により実証される(前出のMittal)。このアッセイでは、ADGUCは基質[γ - 32 P]GTPと共にインキュベートされ、次に過剰基質が、産物であるサイクリック[32 P]GMPから分離される。反応混合液はpH 7.5の1 M Tris-HClを5 μ l、80 mM MnCl₂あるいはMgCl₂を5 μ l、40 mM テオフィリンあるいは2.0 mM 1-メチル-3-イソブチルキサンチンを25 μ l、20 μ gのクレアチン・ホスホキナーゼを含む150 mM クレアチンリン酸を5 μ l(120-135 units/mgタンパク質)、20 mM cGMPを5 μ l、10 mM GTPを10 μ l、[γ - 32 P]GTP(2-4 \times 10⁶ cpm含有)を10 μ l、および水を含む、合計量100 μ lの溶液である。反応はADGUCの添加によって開始される。37°Cで10-15分インキュベートした後、反応は40%のよく冷やしたトリクロロ酢酸を20 μ l加えることにより終了する。Zero-time インキュベーションとADGUC不在でインキュベートした反応はネガティブコントロールとして使用される。産物はイオン交換クロマトグラフィにより分離され、 β -ラジオアイソトープカウンタを使用してサイクリック[32 P]GMPを定量する。ADGUC活性は、反応中に形成されるサイクリック[32 P]GMPの量に比例する。

【0279】

18 ADGUCアゴニスト及びアンタゴニストの同定
ADGUC活性化または抑制のアゴニストまたはアンタゴニストは、実施例17で記述したアッセイを使用してテストしてもよい。またはサイリックAMP Flash Plate AssayとサイリックGMP Flash Plate Assay(NEN Life Sciences Products)などのマルチウェルプレートフォーマットで高処理読出しを可能にするアッセイ技術を使用するアゴニストは、ADGUC活性の増加を引き起こし、アンタゴニストはADGUC活性の減少を引き起こす。

【0280】

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

【0281】

(表の簡単な説明)

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

【0282】

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。

【0283】

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

【0284】

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

【0285】

表5は、本発明のポリヌクレオチドの代表的なcDNAライブラリを示す。

【0286】

10

20

30

40

50

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

【 0 2 8 7 】

表 7 は、本発明のポリヌクレオチドとポリペプチド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 エクト ID	ポリバブチド No:	ポリバブチド SEQ ID	Incyte ID	ポリバブチド SEQ ID NO:	Incyte ポリバブチド ID
3685652	1	3685652CD1	3685652CD1	6	3685652CB1
7478216	2	7478216CD1	7478216CD1	7	7478216CB1
1505146	3	1505146CD1	1505146CD1	8	1505146CB1
1577526	4	1577526CD1	1577526CD1	9	1577526CB1
7478215	5	7478215CD1	7478215CD1	10	7478215CB1

10

20

30

40

【表 2】

表2

ポリペプチドSEQ ID NO:	Incyte ポリペプチド ID	GenBank ID NO:	確率スコア	GenBank 相同体
1	3685652CD1	g202676	0	[トブナスミ] アデニルシクラーゼ タイプIV Cao, B 及び Gilman, A.G. (1991) Cloning and expression of a widely distributed (Type IV) adenylyl cyclase. Proc. Natl. Acad. Sci. U.S.A. 88:10178-10182.
2	7478216CD1	g1758332	0	[トブナスミ] アデニルシクラーゼ タイプV Scholich, K.5(1989) Science 283:1328-1331 アデニルシクラーゼに類似の[緑虫]
3	1505146CD1	g868200	2.20E-21	Wilson, R.5(1994) Nature 368: 32-38 酵母のアデニルシクラーゼに類似[ヒト] (S56776)
4	1577526CD1	g1594042	7.50E-254	
5	7478215CD1	g162613	0	[ウシ] アデニルシクラーゼ I 型

10

20

30

40

表3-1

SEQ ID NO.	Incyte ポリペプチドID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグナチヤ配列、ドメインおよびモチーフ	分析方法及びデータベース
1	3685652CD1	1086	S6, S51, S115, T202, S212, T218, S253, S290, Y334, S342, T368, Y453, S474, S504, T535, S559, Y585, T615, T696, Y874, S887, S928, S950, T967, S977, S1024	N706, N713, N845, N947	アダニル酸シクラーゼとグアニル酸シクラーゼ触媒ドメイン:U264-D457, L873-T1073 グアニル酸シクラーゼタンパク質: BL00452A:E219-K240, BL00452B:L264-D306, BL00452C:V930-L946, BL00452D:A1003-L1045, BL00452E:R1055-F1070 グアニル酸シクラーゼシグナチヤ (guanylate_cyclase.prf): A367-G429, H985-S1047 シクラーゼリアーゼ アデニル アデニル酸型 ATP PYROPHOSPHATELYASE cAMP 合成 阻害 通PD002814:L631-H875 グアニル酸シクラーゼ:DM00173 E26770 759-1046:H771-C1067 シグナル切断M1-A50 膜貫通ドメイン:L29-W49, W97-L124, I590-N616, I627-M645, L675-F694, M800-N819 グアニル酸シクラーゼ:G386-E409, G1004-E1027 膜貫通ドメイン: S315-M331, スコア 14.27 Y355-L371, スコア 11.82 グアニル酸シクラーゼとグアニル酸シクラーゼ触媒ドメイン: guanylate_cyc:I512-R696 スコア 286.9; E=2.6 e-82 グアニル酸シクラーゼシグナチヤ guanylate_cyclases.prf:E605-G668 スコア 30.97	HMMER-PFAM BLIMPS-BLOCKS PROFILESKAN BLAST-PRODOM BLAST-DBOM SPSCAN HMMER MOTIFS HMMER HMMER-PFAM PROFILESKAN
2	7478216CD1	837	S124 S128 S150 S159 S171 S190 S202 S264 S293 S298 S376 S455 S473 S619 S806 T129 T538 T545 T698 T795 Y573	N313 N519		

10

20

30

40

【表4】

表3-2

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグナルペプチド、トメインおよびモチーフ	分析方法及びデータベース
3	1505146CD1	104	S10 S14 S34 T99 Y89		グアニル酸シグナルペプチド BL00452:C624-L666, E476-K497, I512-D554, C562-L578, P<4e-10 グアニル酸シグナルペプチド DM00173:A45195I276-552:G406-Y683, P_value 2.2e-148 アミノ酸シグナルペプチド PD030262:M76-E180 P_value 9.9e-50 PD003877:S376-Y513 P_value 4.7e-43 PD000360:Q515-A653 P_value 1.3e-24 PD002814:K754-Q837 P_value 1.1e-21 ライジンジッパー:L542-L563 Guanylate_Cyclases:G625-E648 signal_cleavage:M1-S84 ライジンリチドシグナルペプチド PR0019A:V20-I33 PR0019B:C40-L53 シグナルペプチド: M1-S16 膜貫通ドメイン: Y224-Y244	BLIMPS BLOCKS BLAST_DOMO BLAST_PRODOM MOTIFS MOTIFS SPSCAN BLIMPS_PRINTS SPSCAN HMMER HMMER PFAM HMMER HMMER-PFAM
4	1577526CD1	769	S112 S113 S143 S167 S190 S273 S307 S347 S351 S622 S69 S738 T131 T138 T406 T588 T717 T759 T764	N321 N44 N600 N715	ライジンリチドシグナルペプチド N623-T645, K646-S668, N669-K691, K692-S714, N715-Q737, N430-E452, N454-K476, N477-K504, N505-F527, N552-N574, N575-Q599, N600-S622 膜貫通ドメイン: T610-F631, L638-T658	SPSCAN HMMER HMMER PFAM HMMER
5	7478215CD1	1119	S1007 S1042 T27 S1096 S238 S248 S384 S479 S543 S599 S783 S660	N1061 N301 N548 N704	膜貫通ドメイン: アミノ酸シグナルペプチド P478, L859-G1056	HMMER HMMER-PFAM

10

20

30

40

【表 5】

表3-3

SEQ ID NO.	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
			S700 S781 S969 T235 T320 T389 T401 T529 T53 T541 T578 T589 T61 T950 Y1052 Y355 Y861 S28		グアニル酸シグラーゼモチーフ: G407-E430, G987-D1010 グアニル酸シグラーゼシグネチャ: Q968-R1030, E387-G450 グアニル酸シグラーゼタンパク質 BL00452: E259-K280, I294-D336, I916-I932, A986-L1028, R1038-F1053 グアニル酸シグラーゼ DM00173P19754 191-465: L189-S464, R810-E1022 グアニル酸シグラーゼ DM00173P19754 759-1052:L759- L1050, R240-Y452 グアニル酸シグラーゼ DM00173P32870 158-436: G188-F465, E822-T1021 グアニル酸シグラーゼ DM00173JA45195 276-552: G188-F465, I791-E1022 アデニル酸シグラーゼ PD000387DQ862-G1056, Q297- G433, V325-Y458 アデニル酸シグラーゼ PD088215Y458-L539 アデニル酸シグラーゼ PD000387T158-Y295 アデニル酸シグラーゼ I 型, EC 4.6.1.1: PD097355:P84-F140	MOTIFS ProfileScan BLIMPS-BLOCKS BLAST-DOMO BLAST-DOMO BLAST-DOMO BLAST-DOMO BLAST-PRODOM BLAST-PRODOM BLAST-PRODOM BLAST-PRODOM

10

20

30

40

表4-1

ポリヌクレオチド SEQ ID NO.	Incyte ポリヌクレオチド ID	配列長	選択された断片	配列断片	5'位置	3'位置
6	3685652CB1	3769	1-1652, 3727-3769, 2180-3019	7345578HI (SYNODIN02)	3170	3741
				724821T6 (SYNCOAT01)	3003	3699
				7709325HI (PANCNOE02)	22	557
				FL3685652 g8346725 g202676	410	3670
				q1379036	3282	3769
				6534126HI (CONFTDT02)	1	286
7	7478216CB1	3137	1-669, 1813-2019, 2427-3137, 753-820	1639357F6 (UTRSNOT06)	293	732
				71190306V1	1253	1874
				6766950J1 (BRAUNOR01)	1026	1670
				7357757HI (HEARNOR03)	679	1042
				71153228V1	2559	3137
				6754326HI (SINTER02)	1894	2496
				GNN.g8567699 000003 002	1	1290
				71151710V1	2350	2967
				71152591V1	1815	2332
				71874115V1	1	538
8	1505146CB1	1958	107-210, 1897-1958, 1-68, 352-927	71874576V1	476	1072
				71873077V1	1261	1958
				71873608V1	1172	1853
				71875066V1	525	1211
				52092666 (MMLR2DT01)	1235	1727
9	1577526CB1	2660	1-875	6783247F9 (SINLTMC01)	2105	2660
				6568376T8 (NCLDTXN05)	469	897
				71434255V1	1709	2404
				5302686HI (DRGNON04)	413	667
				6980889HI (BRAHTR04)	966	1552
				71441001V1	1925	2530
				5847513HI (BRAFN0T04)	757	1029
				GNN.g8176661 000002 002	1	2039
				7738563HI (BRAITVE01)	607	1357
				7738563HI (BRAITVE01)	607	1357

10

20

30

40

【表7】

表4-2

ポリヌクレオチド SEQ ID NO:	Incyte ポリヌクレオチド ID	配列長	選択された断片	配列断片	5'位置	3'位置
			90, 830- 1142, 2811- 3811, 1723- 2125	725266872 (BRAIFEE04)	1	616
				698456283 (BRAIFER05)	2035	2786
				723058081 (BRAXYDR15)	1320	1855
				729282481 (BRAIFER06)	2681	3268
				698122188 (BRAIFER05)	2042	2788
				729245989 (BRAIFER06)	3255	3811
				689402681 (BRAITDR03)	558	1143
				725266882 (BRAIFEE04)	1496	2055

10

20

30

40

表5

ポリシクレンオキド SEQ ID NO.	Inoxyte プロジェクト ID	代表的タイプリ
6	3685652CB1	PROSTUS23
7	7478216CB1	BRAUNOR01
8	1505146CB1	UTREMTU7
9	1577526CB1	SINITMC01
10	7478215CB1	BRAIFER05

10

20

30

40

表6-1

ライブラリ	バウター	ライブラリの説明
BRAALPER05	P-INCY	ライブラリは妊娠 23 週目で死産の左心低形成症児(白人男子)から採取した脳組織から単離した RNA を用いて作製した。
BRAUNOR01	P-INCY	このランダムライブラリは、アテローム性動脈硬化症により胸部大動脈破裂による出血で死亡の 81 才白人女性から得た線条体、淡蒼球および後部被蓋組織から単離した RNA を用いて作製した。病理学検査では、双方の内前運動脈に中程度の粥状アテローム性動脈硬化、前頭皮質および海馬に軽微な神経学的な障害、および散在した広汎性のアミロイド斑、および神経原線維縮小が認められた。これは年齢と矛盾しない。肉眼で見て、軟膜は上矢状静脈洞に沿った程度の硝子質化と厚化が見られるのみであった。残りの軟膜は薄く、充血した血管がいくらか見られた。主に前頭葉(双方)に軽度の萎縮が見られた。概型的には、アルツハイマータイプ II 星状細胞のペアが新皮質の深層にあった。中央前頭皮質の灰白質の奥深部のニューロンの周りでグリオトーシス(satellitosis)が増大していた。扁桃はまれな広汎性斑と神経原線維縮小を含んでいた。後部海馬は反応性グリオトーシスに囲まれたヒモジリン含有マクロファージを伴う微視的な炎症性の空洞化領域を含んでいた。患者の病歴は、敗血症、胆管炎、術後無気肺、肺炎冠動脈疾患(CAD)、左心室肥大に起因する心肥大、巨脾症、細動脈性腎硬化、結節性コロイド甲状腺腫、気腫、鬱血性心不全(CHF)、甲状腺機能低下症および味覚障害を含んでいた。
OUTREDMT07	P-INCY	ライブラリは 32 才の白人女性から採取した子宮内腫瘍組織から単離した RNA を用いて作製した。病理は腫瘍性で扁平円柱上皮境界 1、6、7 時の位置で重症の子宮頸部形成異常(CIN III)を示した。軽度のコロイドサイト異形成(heillocytotic dyslasia)が子宮頸管内の他の場所でも確認された。子宮頸部の線にはなかった。子宮内腫瘍は分泌腺にあり、子宮内腫瘍は診断上異常はなかった。
PROSTUS23	P-INCY	このサブストラクショナル前立腺腫瘍ライブラリは、プールされた前立腺腫瘍ライブラリの 1000 万のクローンと 2 回のサブストラクショナルハイブリダイゼーションにかけられた、プールされた前立腺腫瘍ライブラリの 1000 万のクローンをプールして作製した。サブストラクショナルハイブリダイゼーション(A)、61 才 (B)、66 才 (C)、および 68 才 (D)の白人男子からランダム前立腺切除術時に採取した前立腺腫瘍から単離した mRNA を用いた。4 個の前立腺腫瘍ライブラリから等しい数のクローンをプールすることによって作製した。病理はすべてのドナーにおいて腫瘍を示した。病歴は、ドナー A では、PSA の上昇、硬結、およびタバコの使用、ドナー B は、PSA の上昇、硬結、前立腺肥厚、腎不全、骨関節炎、腎動脈狭窄、良性の HTN、血小板減少症、高脂血症、タバコアルコール使用、および C 型肝炎(保有者)、ドナー C は、PSA の上昇、硬結、およびタバコの使用、また、ドナー D は、PSA の上昇、硬結、高コレステロール血症、および腎結石である。サブストラクショナルハイブリダイゼーショングループは、前立腺腫瘍、前立腺上皮細胞および三人のドナーによる前立腺ストローマからの線維芽細胞から由来する 3 個の前立腺腫瘍ライブラリから得られた同じ数の cDNA クローンをプールすることによって作製した。サブストラクショナルハイブリダイゼーションの条件は Swaroop ら, NAR 19 (1991):1854 および Bonaldo らの Genome Research 6 (1996):791 の方法に基づいたものである。
SINITWC01	P-INCY	この大型分画ライブラリは二人の供与者からのプールされた cDNA を用いて作製した。cDNA は 30 才白人女性(供与者 A)の部分結腸切除術、切開肝生検、および永久的結腸造設術時切除した回腸組織および 70 才白人女性(供与者 B)の右半結腸切除、切開肝生検、S 状結腸鏡検査、結腸鏡検査および永久的結腸造設術時に摘出された回腸組織から単離された mRNA を用いて作製した。一致する腫瘍組織の病理(供与者 A)は回腸末端部に生じたカルチノイド腫瘍(グレード 1)の神経内分泌腺癌を示した。この腫瘍は回腸壁

10

20

30

40

【表 10】

表6-2

ライブラリ	ベクター	<p>ライブラリの説明</p> <p>から腸間膜脂肪を貫通して、棲着した盲腸にまで広がっており、そこで腫瘍は腸壁を貫通して粘膜表面まで広がっている。複数のリンパ節が腫瘍に陽性であった。さらに2箇所のリンパ節にもまた直接の腫瘍拡大が見られた。供与者 B の病理は回腸の瘻は非腫瘍性であった。一致する腫瘍の病理(供与者 B)は侵襲度2の癌腫、回盲弁遠位部位に浸透性塊の形成を示した。腫瘍の侵襲は固有筋を通過して漿膜脂肪組織にまで及んでいた。一つの腸局性リンパ節は転移性癌腫のマイクロアークスに陽性であった。供与者 A は腸面紅潮を示し、非特定の上腹部/骨盤症状を示した。患者の病歴は子宮内腫瘍とカバロ中薬、アルコール中薬であった。供与者 B の病歴には、悪性乳房腫瘍、II 型糖尿病、高脂血症、ウイルス性肝炎、非特定甲状腺疾患、骨関節炎、悪性皮膚腫瘍が含まれている。供与者 B の服用薬剤はタモキシフェンである。</p>
-------	------	--

10

20

30

40

表7-1

プログラム	説明	引用文献	パラメータ間値
ABI FACTURA	核酸配列においてベクター配列を除去して、不定の塩基をマスクするプログラム	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	Fast Data Finderは、アミノ酸または核酸配列の比較、注釈付けに有用である。	Applied Biosystems, Foster City, CA. Paracel Inc., Pasadena, CA.	不一致<50%
ABIオートアセンブラ	核酸配列を構築するプログラム。	Applied Biosystems, Foster City, CA.	
BLAST	Basic Local Alignment Search Tool はアミノ酸および核酸配列の配列類似性検索に有用であり、BLASTには5つの機能がある: blastp, blastn, blastx, tblastnおよびtblastx。	Altschul, S.F. 他 (1990) J. Mol. Biol. 215:403-410, Altschul, S.F. 他 (1997) Nucleic Acids Res. 25:3389-3402.	ESTs 確率値=1.0E-8 以下 完全長配列 確率値= 1.0E-10以下
FASTA	問合せ配列と同種の配列群配との類似性を検索する Pearson および Lipman アルゴリズム。FASTAは最低5つの機能からなる: fasta, tfasta, fastx, tfastx および ssearch。	Pearson, W.R および Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448, Pearson, W.R. (1990) Methods Enzymol. 183:63-98, Smith, J.F. 及び M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs fasta E 値=1.06E-6 構築された ESTs: fasta 同一性=95%以上 一致した長さ=200塩基以上 fastx E 値=1.0E-8 以下 完全長配列: fastx スコア=100以上
BLIMPS	配列をBLOCKS, PRINTS, DOMO, PRODOMおよびPFAM データベースの配列と対応させて遺伝子ファミリー、配列相同性および構造的フィンガープリント領域を検索するBLocks IMPROved サーチャー。	Henikoff, S. 及び J.G. Henikoff (1991) Nucleic Acids Res. 19:6566-6572; Henikoff, J.G. 及び S. Henikoff (1996) Methods Enzymol. 266:88-105, Attwood, T.K. 他 (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	確率値= 1.0E-3以下
HMMER	PFAMのようなタンパク質ファミリーデータベースの隠れマルコフモデル(HMM)に基づいたデータベースに対して問合せ配列を検索するアルゴリズム。	Krogh, A. 他 (1994) J. Mol. Biol. 235:1501-1531, Sonnhammer, E.L.L. 他 (1988) Nucleic Acids Res. 26:320-322, Durbin, R. 他 (1998) Our World View. In a Nutshell, Cambridges Univ. Press 1-350ページ。	PFAMヒット 確率値=1.0E-3 以下 シングルペプチドヒット: スコア=0以上

10

20

30

40

表7-2

プログラム	説明	引用文献	パラメータ閾値
ProfileScan	Prositesで定義された配列パターンと一致するタンパク質配列の構造的モチーフおよび配列モチーフを検索するアルゴリズム。	Gribisov, M. 他 (1988) CABIOS 4:61-66; Gribisov, M. 他 (1989) Methods Enzymol.183:146-159; Bairoch, A. 他 (1997) Nucleic Acids Res. 25:217-221.	標準化された質のスコアとその特定の Prosite モチーフに対するGGG指定" HUGH"値 通常、スコア=1.4-2.1.
Phred	高い感度と精度で自動配列決定機のトレースを調べる塩基読み出しアルゴリズム。	Ewing, B. 他 (1998) Genome Res. 8:175-185; Ewing, B. 及びP. Green (1998) Genome Res. 8:186-194.	
Phrap	Smith-Waterman アルゴリズムの効率的なインプリメンテーションに基づくSWATやCrossMatchを含むPhils Revised Assembly プログラムで、配列相同性の検索やDNA配列の構築に有用である。	Smith, T.F. および M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. および M.S. Waterman (1981) J. Mol. Biol. 147:195-197. Green, P., University of Washington, Seattle, WA.	スコア=120以上 一致した長さ=56以上
Consed	Phrap で構築したものの表示、編集用グラフィカルツール。	Gordon, D. 他 (1998) Genome Res. 8:195-202.	
SPScan	タンパク質配列をスキャンして、分泌シグナルペプチドの存在を調べる重み付けマトリックス解析プログラム。	Nielson, H. 他 (1997) Protein Engineering 10:1-6; Claverie, J.M.及びS. Audic (1997) CABIOS 12:431-439.	スコア=3.5以上
TMAP	加重マトリックスを用いて蛋白質配列での置換通セグメントの描写および配向を決定するプログラム。	Persson, B. およびP. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. およびP. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	隠れMarkov モデル(HMM) を使って蛋白質配列の膜貫通セグメントを描写し、配向を決定するプログラム。	Sonnhammer, E.L. 他 (1998) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol. Glasgow 他編集. The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, 175-182ページ.	
Motifs	Prositesで定義された配列と一致したモチーフのアミノ酸配列を検索するプログラム。	Bairoch, A. 他 (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, 第9版, M51-59ページ 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
10 January 2002 (10.01.2002)

PCT

(10) International Publication Number
WO 02/02757 A2

(51) International Patent Classification: C12N 9/00

(21) International Application Number: PCT/US01/20491

(22) International Filing Date: 26 June 2001 (26.06.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/215,476 29 June 2000 (29.06.2000) US
60/223,545 4 August 2000 (04.08.2000) US
60/229,876 31 August 2000 (31.08.2000) US
60/234,838 22 September 2000 (22.09.2000) US
60/236,483 29 September 2000 (29.09.2000) US(71) Applicant (for all designated States except US): INCYTE
GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo
Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): GANDHI, Amcena,
R. [US/US]; 837 Roble Avenue, #1, Menlo Park, CA
94025 (US). TRIBOULEY, Catherine [FR/US]; 1121
Tennessee Street #5, San Francisco, CA 94107 (US).
DING, Li [CH/US]; 3353 Alma Street, #146, Palo Alto,
CA 94306 (US). LI, Dyrung, Aina, M. [US/US]; 233
Coy Drive, San Jose, CA 95123 (US). LEE, Ernestinae,
A. [US/US]; 624 Kains Street, Albany, CA 94706 (US).
YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale,
CA 94087 (US). YANG, Junming [CH/US]; 7125 Bank
Lane, San Jose, CA 95129 (US). BAUGHN, Mariah,
R. [US/US]; 14244 Santiago Road, San Leandro, CA
94577 (US). THORNTON, Michael [US/US]; 9 Medway
Road, Woodside, CA 94062 (US). YAO, Monique, G.[US/US]; 111 Frederick Court, Mountain View, CA 94043
(US). WALLA, Narinder, K. [US/US]; 890 Davis Street
#205, San Leandro, CA 94577 (US). TANG, Y. Tom
[US/US]; 4230 Ramwick Court, San Jose, CA 95118 (US).
ELLIOTT, Vicki, S. [US/US]; 3770 Polton Place Way,
San Jose, CA 95121 (US). LI, Yan [—US]; 3885 Corrina
Way, Palo Alto, CA 94303 (US).(74) Agents: HAMLET-COX, Diana et al.; Incyte Genomics,
Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, 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,
TL, 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, BI, 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
- with sequence listing part of description published sepa-
rately in electronic form and available upon request from
the International Bureau

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

WO 02/02757 A2

(54) Title: ADENYLYL AND GUANYLYL CYCLASES

(57) Abstract: The invention provides human adenylyl and guanylyl cyclases (ADGUC) and polynucleotides which identify and encode ADGUC. 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 ADGUC.

WO 02/02757

PCT/US01/20491

ADENYLYL AND GUANYLYL CYCLASES**TECHNICAL FIELD**

This invention relates to nucleic acid and amino acid sequences of adenylyl and guanylyl cyclases and to the use of these sequences in the diagnosis, treatment, and prevention of neurological, cardiovascular, vision, reproductive, and smooth muscle disorders, and bacterial infections, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of adenylyl and guanylyl cyclases.

BACKGROUND OF THE INVENTION

An essential role in intracellular signaling pathways is filled by second messenger molecules, intermediaries that are activated upon binding of ligands to surface receptors and serve as activators of downstream effector molecules. The cyclic nucleotides, adenosine 3',5'-cyclic monophosphate (cAMP) and guanosine 3',5'-cyclic monophosphate (cGMP) are critical second messengers in a wide variety of signaling pathways. cAMP and cGMP are generated by the enzymes adenylyl (adenylate) cyclase (AC) and guanylyl (guanylate) cyclase (GC) from ATP and GTP. Thus a key step in regulating intracellular cAMP and cGMP levels is modulation of AC and GC activity.

There are nine known isoforms of mammalian AC. All share a common structure comprising a small cytoplasmic N-terminal domain (N) followed by a membrane spanning domain having six predicted α -helices (M_1), a large cytoplasmic domain (C_1), a second transmembrane helical cluster (M_2), and a second cytoplasmic domain homologous to the first (C_2) (Simonds, W.F. (1999) Trends Pharmacol. Sci. 20:66-73). The C_1 and C_2 domains contain ~230 amino acid regions (C_{1a} and C_{2a}) that share approximately 40% identity and form the enzyme's catalytic core. The remaining portions of the cytoplasmic domains are known as C_{1b} and C_{2b} . The tertiary structures of the C_{1a} and C_{2a} domains consist of a three layer α/β sandwich, with the C_{1a} and C_{2a} domains arranged in a head to tail fashion as a wreath (Tang, W.-J. and J.H. Hurley (1998) Mol. Pharmacol. 54:231-240). All known GC catalytic domains are homologous to the mammalian AC C_1 and C_2 regions and studies suggest that they share the same structure. The transmembrane GCs contain a single transmembrane crossing and a single catalytic domain per protein, and function as homodimers. Soluble GCs function as heterodimers of α and β subunits, with one catalytic domain contributed by each of the two subunits.

Mammalian ACs are subjected to complex regulation by G-proteins, Ca^{2+} signals, and

WO 02/02757

PCT/US01/20491

phosphorylation (Tang, *supra*). C_{1s} is the primary binding site for the inhibitory G protein α subunit ($G_{i\alpha}$). C_{2s} is the primary binding site for the stimulatory G protein α subunit ($G_{s\alpha}$) and $G_{12\alpha}$ and contains phosphorylation sites for protein kinase (PK) C and calmodulin (CaM) kinase II. C_{1b} is the site of regulation by Ca^{2+} /CaM, Ca^{2+} , PKA, and CaM kinase IV (Tesmer, J.G.G. and S.R. Sprang (1998) *Curr. Opin. Struct. Biol.* 8:713-719). The expression patterns and other regulatory properties of the nine AC isoforms vary widely. For example, while AC4, AC7, and AC9 are expressed in a wide range of tissues, AC1 and AC8 are expressed only in neural tissues, while AC5 is expressed predominantly in heart and brain (Simonds, *supra*). All AC isoforms are activated by $G_{\alpha s}$, but display differential responses to subunits of the $G_{i\alpha}$ and $G_{12\alpha}$ families, as well as variable sensitivity to PKs, Ca^{2+} , and CaM. For example, AC1, AC3, and AC8 are strongly stimulated by Ca^{2+} /CaM, while AC5 and AC6 are inhibited by submicromolar concentrations of Ca^{2+} in a CaM-independent manner. AC3, AC5, and AC6 are sensitive to inhibition by G_{i1} , while AC2 is not. This heterogeneity allows for tissue- and cell-specific responses to extracellular signals, with integration of signals transmitted by G-proteins with signals from other sources that affect intracellular Ca^{2+} levels and PKC activity (Simonds, *supra*).

Adenylyl cyclases are key players in intracellular signaling pathways of hormones, neurotransmitters, odorants, and chemokines (Tang, *supra*). cAMP activates cAMP-dependent protein kinases which modify the activities of specific enzymes in various cell types. By activating cyclic nucleotide-gated ion channels, cAMP can affect the cell membrane potential. cAMP also has various tissue-specific effects. Increased levels of cAMP lead to an increase in triglyceride hydrolysis and a decrease in amino acid uptake in adipose tissue; an increase in conversion of glycogen to glucose, an inhibition of glycogen synthesis, and an increase in gluconeogenesis in liver; an increase in the synthesis of estrogen and progesterone in ovarian follicle; an increase in the synthesis of aldosterone and cortisol in adrenal cortex; an increase in the contraction rate in cardiac muscle cells; conversion of glycogen to glucose in skeletal muscle; secretion of thyroxine in thyroid; an increase in resorption of calcium from bone in bone cells; fluid secretion in intestine; resorption of water in kidney; and an inhibition of aggregation and secretion in blood platelets. (See Lodish, H. et al. (1995) *Molecular Cell Biology* Scientific American Books, New York NY, pp. 871, 879-886.) The CaM-regulated ACs expressed in brain are important for synaptic plasticity as well as learning and memory. AC1 may also play a role in the regulation of melatonin synthesis (Xia, Z. and D.R. Storm (1997) *Curr. Opin. Neurobiol.* 7:391-396).

The membrane GCs have extracellular ligand-binding domains at the amino terminus, for which some known ligands are the natriuretic peptides and bacterial enterotoxins. Experiments with

WO 02/02757

PCT/US01/20491

knock-out mice reveal that GC-A has a role in heart function and/or development. A form of congenital blindness in humans has been linked to mutations in GC-E. GC-G has high expression in skeletal muscle and may have a role in regulation of blood flow (Wedel, B.J. and D.L. Garbers (1998) Trends Endocrinol. Metab. 9:213-219). Soluble guanylyl cyclase is associated with heme and
5 activated by nitric oxide. The nitric oxide-soluble guanylyl cyclase-cGMP pathway is widespread in mammalian tissues and important in mediating numerous physiological processes including vascular and non-vascular smooth muscle relaxation, peripheral and central neurotransmission, platelet reactivity, and phototransduction. Overactivity of the nitric oxide-soluble guanylyl cyclase-cGMP pathway may be associated with septic shock and migraine, while underactivity of the pathway may
10 be associated with impotence, hypertension, and asthma (Hobbs, A.J. (1997) Trends Pharmacol. Sci. 18:484-491).

Known inhibitors of ACs include the hypotensive drug forskolin and the P-site inhibitors. If the required specificity could be achieved, cardiac-specific AC inhibitors have been proposed as useful for the treatment of congestive heart failure. As cholera and other serious diarrhoeal diseases result
15 from activation of gastrointestinal ACs and GCs by bacterial toxins, these enzymes would also be useful therapeutic targets (Dessauer, C.W. et al. (1999) Trends Pharmacol. Sci. 20:205-210).

The discovery of new adenylyl and guanylyl cyclases 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 neurological, cardiovascular, vision, reproductive, and smooth muscle disorders, and
20 bacterial infections, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of adenylyl and guanylyl cyclases.

SUMMARY OF THE INVENTION

25 The invention features purified polypeptides, adenylyl and guanylyl cyclases, referred to collectively as "ADGUC" and individually as "ADGUC-1," "ADGUC-2," "ADGUC-3," "ADGUC-4," and "ADGUC-5." 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:1-5, b) a polypeptide comprising a naturally occurring amino acid sequence
30 at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-5. In one alternative, the

WO 02/02757

PCT/US01/20491

invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-5.

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:1-5, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-5. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-5. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:6-10.

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

WO 02/02757

PCT/US01/20491

occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-5.

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:6-10, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

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

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b)

WO 02/02757

PCT/US01/20491

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

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-5. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional ADGUC, comprising administering to a patient in need of such treatment the composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group

WO 02/02757

PCT/US01/20491

consisting of SEQ ID NO:1-5. 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
5 treating a disease or condition associated with overexpression of functional ADGUC, 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:1-5, b) a polypeptide comprising a
10 naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-5. The method comprises a) combining the polypeptide with at least one
15 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:1-5, b) a polypeptide comprising a
20 naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-5. The method comprises a) combining the polypeptide with at least one
25 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:6-10, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target
30

WO 02/02757

PCT/US01/20491

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:6-10, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10, iii) a polynucleotide having a sequence complementary to i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of a polynucleotide sequence selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

BRIEF DESCRIPTION OF THE TABLES

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

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

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

Table 4 lists the cDNA and/or genomic DNA fragments which were used to assemble

WO 02/02757

PCT/US01/20491

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

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

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

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

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of

WO 02/02757

PCT/US01/20491

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

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

"Altered" nucleic acid sequences encoding ADGUC include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as ADGUC or a polypeptide with at least one functional characteristic of ADGUC. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding ADGUC, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding ADGUC. 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 ADGUC. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of ADGUC 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; 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 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

WO 02/02757

PCT/US01/20491

in the art.

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

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

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

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

The term "biologically active" refers to a protein having structural, regulatory, or biochemical

WO 02/02757

PCT/US01/20491

functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic" refers to the capability of the natural, recombinant, or synthetic ADGUC, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

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

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding ADGUC or fragments of ADGUC may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; 10 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 20 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 25 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

WO 02/02757

PCT/US01/20491

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

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

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

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

WO 02/02757

PCT/US01/20491

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.

5 A fragment of SEQ ID NO:6-10 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:6-10, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:6-10 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:6-10 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:6-10 and the region of SEQ ID NO:6-10 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

10 A fragment of SEQ ID NO:1-5 is encoded by a fragment of SEQ ID NO:6-10. A fragment of SEQ ID NO:1-5 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-5. For example, a fragment of SEQ ID NO:1-5 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-5. The precise length of a fragment of SEQ ID NO:1-5 and the region of SEQ ID NO:1-5 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

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

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

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

30 Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of

WO 02/02757

PCT/US01/20491

molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequences.

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

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

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

WO 02/02757

PCT/US01/20491

length over which percentage identity may be measured.

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

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

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

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

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment

WO 02/02757

PCT/US01/20491

length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

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

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

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

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

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,

WO 02/02757

PCT/US01/20491

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

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

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

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

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

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

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

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

WO 02/02757

PCT/US01/20491

antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

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

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

"Post-translational modification" of an ADGUC may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of ADGUC.

"Probe" refers to nucleic acid sequences encoding ADGUC, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

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

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

WO 02/02757

PCT/US01/20491

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).

5 Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU
10 primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to
15 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
20 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
25 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
30 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.

WO 02/02757

PCT/US01/20491

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.

5 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.

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

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

25 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.

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

WO 02/02757

PCT/US01/20491

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

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

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

10 15 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*.

20 25 30 A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater

WO 02/02757

PCT/US01/20491

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

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

THE INVENTION

The invention is based on the discovery of new human adenylyl and guanylyl cyclases (ADGUC), the polynucleotides encoding ADGUC, and the use of these compositions for the diagnosis, treatment, or prevention of neurological, cardiovascular, vision, reproductive, and smooth muscle disorders, and bacterial infections.

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

WO 02/02757

PCT/US01/20491

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.

5 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 15 shows analytical methods for protein structure/function analysis and in some cases, searchable databases to which the analytical methods were applied.

Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these properties establish that the claimed polypeptides are adenylyl and guanylyl cyclases. For example, SEQ ID NO:1 is 89% identical to rat adenylyl cyclase type IV (GenBank ID g202676) as determined 20 by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 0.0, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:1 also contains an adenylate/guanylate cyclase catalytic 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 25 PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:1 is an adenylate/guanylate cyclase. In an alternative example, SEQ ID NO:2 is 99% identical over amino acids 384 to 836 to rat adenylyl cyclase type V (GenBank ID g1758332) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 0.0, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ 30 ID NO:2 also contains an adenylate/guanylate cyclase catalytic 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:2 is an

WO 02/02757

PCT/US01/20491

adenylate/guanylate cyclase. In another alternative example, SEQ ID NO:3 is 46% identical to worm adenylate cyclase (GenBank ID g868200) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is $2.2e-21$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. In another alternative example,

5 SEQ ID NO:4 is 100% identical over amino acids 294 to 769 to human protein similar to yeast adenylate cyclase (GenBank ID g1504042) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is $7.5e-254$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:4 also contains multiple leucine rich repeat domains as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See

10 Table 3.) In another alternative example SEQ ID NO:5 is 91% identical to bovine adenylyl cyclase, type I (GenBank ID g162613) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 0.0, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:5 also contains an

15 adenylate/guanylate cyclase catalytic 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:5 is an adenylate/guanylate cyclase. The algorithms and parameters for the analysis of SEQ ID NO:1-5 are described in Table 7.

20 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.

25 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:6-10 or that distinguish between SEQ ID NO:6-10 and related polynucleotide sequences. Column 5 shows identification numbers corresponding to cDNA sequences, coding sequences (exons) predicted from genomic DNA, and/or sequence assemblages

30 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.

WO 02/02757

PCT/US01/20491

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, 6766950J1 is the identification number of an Incyte cDNA sequence, and BRAUNOR01 is the cDNA library from which it is derived. Incyte cDNAs for which cDNA libraries are not indicated were derived from pooled cDNA libraries (e.g., 71190306V1). Alternatively, the identification numbers in column 5 may refer to GenBank cDNAs or ESTs (e.g., g1379036) which contributed to the assembly of the full length polynucleotide sequences. Alternatively, the identification numbers in column 5 may refer to coding regions predicted by Genscan analysis of genomic DNA. For example, GNN.g8567699_0000030_02 is the identification number of a Genscan-predicted coding sequence, with g8567699 being the GenBank identification number of the sequence to which Genscan was applied. The Genscan-predicted coding sequences may have been edited prior to assembly. (See Example IV.) Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. (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. For example, FL3685652_g8346725_g202676 is the identification number of a "stretched" sequence, with 3685652 being the Incyte project identification number, g8346725 being the GenBank identification number of the human genomic sequence to which the "exon-stretching" algorithm was applied, and g202676 being the GenBank identification number of the nearest GenBank protein homolog. (See Example V.) In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in column 5 was obtained to confirm the final consensus polynucleotide sequence, but the relevant Incyte cDNA identification numbers are not shown.

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

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

The invention also encompasses polynucleotides which encode ADGUC. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected

WO 02/02757

PCT/US01/20491

from the group consisting of SEQ ID NO:6-10, which encodes ADGUC. The polynucleotide sequences of SEQ ID NO:6-10, 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.

5 The invention also encompasses a variant of a polynucleotide sequence encoding ADGUC. 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 ADGUC. 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:6-10 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:6-10. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of ADGUC.

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

25 Although nucleotide sequences which encode ADGUC and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring ADGUC under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding ADGUC 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 ADGUC and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

30 The invention also encompasses production of DNA sequences which encode ADGUC and ADGUC derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the

WO 02/02757

PCT/US01/20491

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 ADGUC or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of
5 hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:6-10 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

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

The nucleic acid sequences encoding ADGUC may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences,
25 such as promoters and regulatory elements. For example, one method which may be employed, 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
30 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

WO 02/02757

PCT/US01/20491

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).

5 Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about
10 68°C to 72°C.

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

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

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

30 The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter ADGUC-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic

WO 02/02757

PCT/US01/20491

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

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number 5,837,458; Chang, C.-C. et al. (1999) *Nat. Biotechnol.* 17:793-797; Christians, F.C. et al. (1999) *Nat. Biotechnol.* 17:259-264; and Cramer, A. et al. (1996) *Nat. Biotechnol.* 14:315-319) to alter or improve the biological properties of ADGUC, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

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

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

WO 02/02757

PCT/US01/20491

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

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

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

WO 02/02757

PCT/US01/20491

Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659; and Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) Cancer Gen. Ther. 5(6):350-356; Yu, M. et al. (1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344; Buller, R.M. et al. (1985) Nature 317(6040):813-815; McGregor, D.P. et al. (1994) Mol. Immunol. 31(3):219-226; and Verma, I.M. and N. Somia (1997) Nature 389:239-242.) The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding ADGUC. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding ADGUC can be achieved using a multifunctional *E. coli* vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding ADGUC into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for *in vitro* transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of ADGUC are needed, e.g. for the production of antibodies, vectors which direct high level expression of ADGUC 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 ADGUC. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast *Saccharomyces cerevisiae* or *Pichia pastoris*. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, *supra*; Bitter, G.A. et al. (1987) *Methods Enzymol.* 153:516-544; and Scorer, C.A. et al. (1994) *Bio/Technology* 12:181-184.)

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

WO 02/02757

PCT/US01/20491

constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding ADGUC may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses ADGUC in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

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

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

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* and *apr* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981)

WO 02/02757

PCT/US01/20491

J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), β glucuronidase and its substrate β -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding ADGUC is inserted within a marker gene sequence, transformed cells containing sequences encoding ADGUC can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding ADGUC 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 ADGUC and that express ADGUC may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of ADGUC using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on ADGUC 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 Wiley-Interscience, New York NY; and Pound, J.D. (1998) *Immunochemical Protocols*, Humana Press, Totowa NJ.)

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

WO 02/02757

PCT/US01/20491

Alternatively, the sequences encoding ADGUC, 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 ADGUC 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 ADGUC may be designed to contain signal sequences which direct secretion of ADGUC 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 ADGUC 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 ADGUC protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of ADGUC activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins,

WO 02/02757

PCT/US01/20491

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 ADGUC encoding sequence and the heterologous protein sequence, so that

5 ADGUC 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 ADGUC may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These
10 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.

ADGUC of the present invention or fragments thereof may be used to screen for compounds that specifically bind to ADGUC. At least one and up to a plurality of test compounds may be
15 screened for specific binding to ADGUC. 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 ADGUC, 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):
20 Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which ADGUC 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 ADGUC, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or E. coli.
25 Cells expressing ADGUC or cell membrane fractions which contain ADGUC are then contacted with a test compound and binding, stimulation, or inhibition of activity of either ADGUC 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
30 assay may comprise the steps of combining at least one test compound with ADGUC, either in solution or affixed to a solid support, and detecting the binding of ADGUC 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

WO 02/02757

PCT/US01/20491

libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

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

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

Polynucleotides encoding ADGUC may also be manipulated *in vitro* in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate

WO 02/02757

PCT/US01/20491

into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) Science 282:1145-1147).

Polynucleotides encoding ADGUC 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 ADGUC 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 ADGUC, e.g., by secreting ADGUC 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 ADGUC and adenylyl and guanylyl cyclases. In addition, the expression of ADGUC is closely associated with prostate cancer, brain tissue, ileum tissue and endometrial tissue. Therefore, ADGUC appears to play a role in neurological, cardiovascular, vision, reproductive, and smooth muscle disorders, and bacterial infections. In the treatment of disorders associated with increased ADGUC expression or activity, it is desirable to decrease the expression or activity of ADGUC. In the treatment of disorders associated with decreased ADGUC expression or activity, it is desirable to increase the expression or activity of ADGUC.

Therefore, in one embodiment, ADGUC 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 ADGUC. Examples of such disorders include, but are not limited to, a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease; prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome; fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord

WO 02/02757

PCT/US01/20491

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; akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, 5 postherpetic neuralgia, and Tourette's disorder; a cardiovascular disorder such as arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebotrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery, congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, 10 degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, and complications of cardiac transplantation, congenital lung 15 anomalies, atelectasis, pulmonary congestion and edema, pulmonary embolism, pulmonary hemorrhage, pulmonary infarction, pulmonary hypertension, vascular sclerosis, obstructive pulmonary disease, restrictive pulmonary disease, chronic obstructive pulmonary disease, emphysema, chronic bronchitis, bronchial asthma, bronchiectasis, bacterial pneumonia, viral and mycoplasma pneumoniae pneumonia, lung abscess, pulmonary tuberculosis, diffuse interstitial diseases, pneumoconioses, sarcoidosis, idiopathic 20 pulmonary fibrosis, desquamative interstitial pneumonitis, hypersensitivity pneumonitis, pulmonary eosinophilia bronchiolitis obliterans-organizing pneumonia, diffuse pulmonary hemorrhage syndromes, Goodpasture's syndromes, idiopathic pulmonary hemosiderosis, pulmonary involvement in collagen-vascular disorders, pulmonary alveolar proteinosis, lung tumors, inflammatory and noninflammatory pleural effusions, pneumothorax, pleural tumors, drug-induced lung disease, radiation- 25 induced lung disease, and complications of lung transplantation; a vision disorder such as conjunctivitis, keratoconjunctivitis sicca, keratitis, episcleritis, iritis, posterior uveitis, glaucoma, amaurosis fugax, ischemic optic neuropathy, optic neuritis, Leber's hereditary optic neuropathy, toxic optic neuropathy, vitreous detachment, retinal detachment, cataract, macular degeneration, central serous chorioretinopathy, retinitis pigmentosa, melanoma of the choroid, retrolubar tumor, and chiasmal 30 tumor; a reproductive disorder such as disorders of prolactin production; infertility, including tubal disease, ovulatory defects, and endometriosis; disruptions of the estrous cycle, disruptions of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, endometrial and ovarian tumors, uterine fibroids, autoimmune disorders, ectopic pregnancies, and teratogenesis; cancer

WO 02/02757

PCT/US01/20491

of the breast, fibrocystic breast disease, and galactorrhea; disruptions of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, and gynecomastia; a smooth muscle disorder such as any impairment or alteration in the normal action of smooth muscle including, but not limited to, angina, anaphylactic shock, arrhythmias, asthma, cardiovascular shock, Cushing's syndrome, hypertension, hypoglycemia, myocardial infarction, migraine, and pheochromocytoma, and myopathies including cardiomyopathy, encephalopathy, epilepsy, Kearns-Sayre syndrome, lactic acidosis, myoclonic disorder, and ophthalmoplegia; and an infection by a bacterial agent classified as pneumococcus, staphylococcus, streptococcus, bacillus, corynebacterium, clostridium, meningococcus, gonococcus, listeria, moraxella, kingella, haemophilus, legionella, bordetella, gram-negative enterobacterium including shigella, salmonella, and campylobacter, pseudomonas, vibrio, brucella, francisella, yersinia, bartonella, norcardium, actinomyces, mycobacterium, spirochaetae, rickettsia, chlamydia, and mycoplasma.

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

In a further embodiment, a composition comprising a substantially purified ADGUC in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of ADGUC including, but not limited to, those provided above.

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

In a further embodiment, an antagonist of ADGUC may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of ADGUC. Examples of such disorders include, but are not limited to, those neurological, cardiovascular, vision, reproductive, and smooth muscle disorders, and bacterial infections described above. In one aspect, an antibody which specifically binds ADGUC 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 ADGUC.

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

WO 02/02757

PCT/US01/20491

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 ADGUC may be produced using methods which are generally known in the art. In particular, purified ADGUC may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind ADGUC. Antibodies to ADGUC 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 ADGUC 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 ADGUC have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein. Short stretches of ADGUC amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

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

WO 02/02757

PCT/US01/20491

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

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

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

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

WO 02/02757

PCT/US01/20491

ADGUC epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the ADGUC-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

5 procedures which ultimately require dissociation of ADGUC, 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
10 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 ADGUC-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
15 Coligan et al. *supra*.)

In another embodiment of the invention, the polynucleotides encoding ADGUC, 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
20 encoding ADGUC. 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 ADGUC. (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
25 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
30 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

WO 02/02757

PCT/US01/20491

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 ADGUC 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), thalassemias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) Science 270:404-410; Verma, I.M. and N. Somia (1997) Nature 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) Nature 335:395-396; Poeschla, E. et al. (1996) Proc. Natl. Acad. Sci. USA. 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as Candida albicans and Paracoccidioides brasiliensis; and protozoan parasites such as Plasmodium falciparum and Trypanosoma cruzi). In the case where a genetic deficiency in ADGUC expression or regulation causes disease, the expression of ADGUC 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 ADGUC are treated by constructing mammalian expression vectors encoding ADGUC and introducing these vectors by mechanical means into ADGUC-deficient cells. Mechanical transfer technologies for use with cells in vivo or ex vitro include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) Annu. Rev. Biochem. 62:191-217; Ivics, Z. (1997) Cell 91:501-510; Boulay, J-L. and H. Récipon (1998) Curr. Opin. Biotechnol. 9:445-450).

Expression vectors that may be effective for the expression of ADGUC 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). ADGUC may be

WO 02/02757

PCT/US01/20491

expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or β -actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. USA 89:5547-5551; Gossen, M. et al. (1995) Science 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and Blau, H.M. *supra*), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding ADGUC from a normal individual.

10 Commercially available liposome transformation kits (e.g., the PERFECT LIPID TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) Virology 52:456-467), or by electroporation (Neumann, E. et al. 15 (1982) EMBO J. 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to ADGUC expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding ADGUC under the control of an independent promoter or the retrovirus long 20 terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) Proc. Natl. Acad. Sci. USA 92:6733-6737), incorporated by reference herein. The vector is propagated in 25 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 30 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

WO 02/02757

PCT/US01/20491

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
5 polynucleotides encoding ADGUC to cells which have one or more genetic abnormalities with respect to the expression of ADGUC. 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
10 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
15 polynucleotides encoding ADGUC to target cells which have one or more genetic abnormalities with respect to the expression of ADGUC. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing ADGUC 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
20 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
25 cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) *J. Virol.* 73:519-532 and Xu, H. et al. (1994) *Dev. Biol.* 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of
30 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.

WO 02/02757

PCT/US01/20491

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding ADGUC 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
5 alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for ADGUC into the
10 alphavirus genome in place of the capsid-coding region results in the production of a large number of ADGUC-coding RNAs and the synthesis of high levels of ADGUC in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in 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
15 allow the introduction of ADGUC into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

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

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

WO 02/02757

PCT/US01/20491

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

WO 02/02757

PCT/US01/20491

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

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

WO 02/02757

PCT/US01/20491

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.

5 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 ADGUC, antibodies to ADGUC, and mimetics, agonists, antagonists, or inhibitors of ADGUC.

10 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. 15 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. 20 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.

25 Specialized forms of compositions may be prepared for direct intracellular delivery of macromolecules comprising ADGUC or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, ADGUC 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 30 transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys,

WO 02/02757

PCT/US01/20491

or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example ADGUC or fragments thereof, antibodies of ADGUC, and agonists, antagonists or inhibitors of ADGUC, 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 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. 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 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 about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind ADGUC may be used for the diagnosis of disorders characterized by expression of ADGUC, or in assays to monitor patients being treated with ADGUC or agonists, antagonists, or inhibitors of ADGUC. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics.

WO 02/02757

PCT/US01/20491

Diagnostic assays for ADGUC include methods which utilize the antibody and a label to detect ADGUC 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.

5 A variety of protocols for measuring ADGUC, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of ADGUC expression. Normal or standard values for ADGUC expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibodies to ADGUC under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of ADGUC 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.

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

20 In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding ADGUC or closely related molecules may be used to identify nucleic acid sequences which encode ADGUC. 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 ADGUC, 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 ADGUC encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:6-10 or from genomic sequences including promoters, enhancers, and introns of the ADGUC gene.

30 Means for producing specific hybridization probes for DNAs encoding ADGUC include the cloning of polynucleotide sequences encoding ADGUC or ADGUC derivatives into vectors for the

WO 02/02757

PCT/US01/20491

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 ADGUC may be used for the diagnosis of disorders associated with expression of ADGUC. Examples of such disorders include, but are not limited to, 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 neuronal muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease; prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Strausler-Scheinker syndrome; fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, 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; akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, and Tourette's disorder; a cardiovascular disorder such as arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery, congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, and complications of cardiac transplantation, congenital lung anomalies, atelectasis, pulmonary congestion and edema,

WO 02/02757

PCT/US01/20491

pulmonary embolism, pulmonary hemorrhage, pulmonary infarction, pulmonary hypertension, vascular sclerosis, obstructive pulmonary disease, restrictive pulmonary disease, chronic obstructive pulmonary disease, emphysema, chronic bronchitis, bronchial asthma, bronchiectasis, bacterial pneumonia, viral and mycoplasmal pneumonia, lung abscess, pulmonary tuberculosis, diffuse interstitial diseases,

5 pneumoconioses, sarcoidosis, idiopathic pulmonary fibrosis, desquamative interstitial pneumonitis, hypersensitivity pneumonitis, pulmonary eosinophilia bronchiolitis obliterans-organizing pneumonia, diffuse pulmonary hemorrhage syndromes, Goodpasture's syndromes, idiopathic pulmonary

10 hemosiderosis, pulmonary involvement in collagen-vascular disorders, pulmonary alveolar proteinosis, lung tumors, inflammatory and noninflammatory pleural effusions, pneumothorax, pleural tumors, drug-induced lung disease, radiation-induced lung disease, and complications of lung transplantation; a vision

15 disorder such as conjunctivitis, keratoconjunctivitis sicca, keratitis, episcleritis, iritis, posterior uveitis, glaucoma, amaurosis fugax, ischemic optic neuropathy, optic neuritis, Leber's hereditary optic neuropathy, toxic optic neuropathy, vitreous detachment, retinal detachment, cataract, macular degeneration, central serous chorioretinopathy, retinitis pigmentosa, melanoma of the choroid,

20 retinobulbar tumor, and chiasmal tumor; a reproductive disorder such as disorders of prolactin production; infertility, including tubal disease, ovulatory defects, and endometriosis; disruptions of the estrous cycle, disruptions of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation

25 syndrome, endometrial and ovarian tumors, uterine fibroids, autoimmune disorders, ectopic pregnancies, and teratogenesis; cancer of the breast, fibrocystic breast disease, and galactorrhea;

30 disruptions of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, and gynecomastia; a smooth muscle disorder such as any impairment or alteration in the normal action of smooth muscle including, but not limited to, angina, anaphylactic shock, arrhythmias, asthma, cardiovascular shock, Cushing's syndrome, hypertension, hypoglycemia, myocardial infarction,

35 migraine, and pheochromocytoma, and myopathies including cardiomyopathy, encephalopathy, epilepsy, Kearns-Sayre syndrome, lactic acidosis, myoclonic disorder, and ophthalmoplegia; and an infection by a bacterial agent classified as pneumococcus, staphylococcus, streptococcus, bacillus, corynebacterium, clostridium, meningococcus, gonococcus, listeria, moraxella, kingella, haemophilus, legionella, bordetella, gram-negative enterobacterium including shigella, salmonella, and campylobacter,

40 pseudomonas, vibrio, brucella, francisella, yersinia, bartonella, norcardium, actinomycetes, mycobacterium, spirochaetae, rickettsia, chlamydia, and mycoplasma. The polynucleotide sequences encoding ADGUC 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

WO 02/02757

PCT/US01/20491

microarrays utilizing fluids or tissues from patients to detect altered ADGUC expression. Such qualitative or quantitative methods are well known in the art.

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

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

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

WO 02/02757

PCT/US01/20491

Additional diagnostic uses for oligonucleotides designed from the sequences encoding ADGUC 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 ADGUC, or a fragment of a polynucleotide complementary to the polynucleotide encoding ADGUC, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding ADGUC may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from the polynucleotide sequences encoding ADGUC are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSSCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed *in silico* SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

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

WO 02/02757

PCT/US01/20491

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

A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time. (See Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent Number 5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

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

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with *in vitro* model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed

WO 02/02757

PCT/US01/20491

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.

5 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
10 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
15 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
20 invention, so that transcript levels corresponding to the polynucleotides of the present invention may be quantified. The transcript levels in the treated biological sample are compared with levels in an untreated biological sample. Differences in the transcript levels between the two samples are indicative of a toxic response caused by the test compound in the treated sample.

Another particular embodiment relates to the use of the polypeptide sequences of the present
25 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
30 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

WO 02/02757

PCT/US01/20491

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 ADGUC to quantify the levels of ADGUC 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 to each array element.

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

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

WO 02/02757

PCT/US01/20491

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

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

WO 02/02757

PCT/US01/20491

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

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

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

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

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

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

In additional embodiments, the nucleotide sequences which encode ADGUC may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on

WO 02/02757

PCT/US01/20491

properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

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

The disclosures of all patents, applications and publications, mentioned above and below, including U.S. Ser. No. 60/215,476, 60/223,545, 60/229,876, 60/234,838 and 60/236,483, are hereby expressly incorporated by reference.

10

EXAMPLES

I. Construction of cDNA Libraries

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

20 Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A)+ RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the 25 POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPRT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, 30 units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column

WO 02/02757

PCT/US01/20491

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 XLI-Blue, XLI-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

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

WO 02/02757

PCT/US01/20491

frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, *supra*, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

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

Table 7 summarizes the tools, programs, and algorithms used for the analysis and assembly of
30 Incyte cDNA and full length sequences and provides applicable descriptions, references, and threshold parameters. The first column of Table 7 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where

WO 02/02757

PCT/US01/20491

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:6-10. 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 adenyllyl and guanylyl cyclases were initially identified by running the Genscan gene identification program against public genomic sequence databases (e.g., gbpr1 and gbhtg). Genscan is a general-purpose gene identification program which analyzes genomic DNA sequences from a variety of organisms (See Burge, C. and S. Karlin (1997) *J. Mol. Biol.* 268:78-94, and Burge, C. and S. Karlin (1998) *Curr. Opin. Struct. Biol.* 8:346-354). The program concatenates predicted exons to form an assembled cDNA sequence extending from a methionine to a stop codon. The output of Genscan is a FASTA database of polynucleotide and polypeptide sequences. The maximum range of sequence for Genscan to analyze at once was set to 30 kb. To determine which of these Genscan predicted cDNA sequences encode adenyllyl and guanylyl cyclases, the encoded polypeptides were analyzed by querying against PFAM models for adenyllyl and guanylyl cyclases. Potential adenyllyl and guanylyl cyclases were also identified by homology to Incyte cDNA sequences that had been annotated as adenyllyl and guanylyl cyclases. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbpr1 public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from genpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide sequences were obtained by assembling Genscan-predicted coding sequences with Incyte cDNA sequences and/or public cDNA sequences using the assembly process described in Example III. Alternatively, full length polynucleotide sequences were derived entirely from edited or unedited Genscan-predicted coding sequences.

V. Assembly of Genomic Sequence Data with cDNA Sequence Data

"Stitched" Sequences

WO 02/02757

PCT/US01/20491

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 gbpi public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from genpept. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

20 **"Stretched" Sequences**

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

VI. Chromosomal Mapping of ADGUC Encoding Polynucleotides

WO 02/02757

PCT/US01/20491

The sequences which were used to assemble SEQ ID NO:6-10 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:6-10 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and 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 position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (<http://www.ncbi.nlm.nih.gov/genemap/>), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

VII. Analysis of Polynucleotide Expression

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

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

WO 02/02757

PCT/US01/20491

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

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

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

VIII. Extension of ADGUC Encoding Polynucleotides

WO 02/02757

PCT/US01/20491

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
5 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
10 extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+} , $(NH_4)_2SO_4$, and 2-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE
15 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
20 repeated 20 times; Step 6: 68 °C, 5 min; Step 7: storage at 4 °C.

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

The extended nucleotides were desalted and concentrated, transferred to 384-well plates,
30 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

WO 02/02757

PCT/US01/20491

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:6-10 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 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 & Schnell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate.

WO 02/02757

PCT/US01/20491

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

X. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, supra), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Schna (1999), supra). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Schna, M. et al. (1995) *Science* 270:467-470; Shalon, D. et al. (1996) *Genome Res.* 6:639-645; Marshall, A. and J. Hodgson (1998) *Nat. Biotechnol.* 16:27-31.)

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorption and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)⁺ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)⁺ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/ μ l oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/ μ l RNase inhibitor, 500 μ M dATP, 500 μ M dGTP, 500 μ M dTTP, 40 μ M dCTP, 40 μ M dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)⁺ RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)⁺ RNAs are synthesized by in vitro transcription

WO 02/02757

PCT/US01/20491

from non-coding yeast genomic DNA. After incubation at 37° C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85° C to stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 µl 5X SSC/0.2% SDS.

Microarray Preparation

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

15 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 µl of the array element DNA, at an average concentration of 100 ng/µl, 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.

25 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

30 Hybridization reactions contain 9 µl of sample mixture consisting of 0.2 µg 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

WO 02/02757

PCT/US01/20491

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 µl of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60°C. The arrays are washed for 10 min at 45°C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45°C in a second wash buffer (0.1X SSC), and dried.

Detection

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

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

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

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

WO 02/02757

PCT/US01/20491

measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

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

XI. Complementary Polynucleotides

Sequences complementary to the ADGUC-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring ADGUC. 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 ADGUC. 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 ADGUC-encoding transcript.

XII. Expression of ADGUC

Expression and purification of ADGUC is achieved using bacterial or virus-based expression systems. For expression of ADGUC 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 ADGUC upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of ADGUC 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 ADGUC 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.)

WO 02/02757

PCT/US01/20491

In most expression systems, ADGUC 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 ADGUC at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, *supra*, ch. 10 and 16). Purified ADGUC obtained by these methods can be used directly in the assays shown in Examples XVI, XVII, and XVIII where applicable.

XIII. Functional Assays

ADGUC function is assessed by expressing the sequences encoding ADGUC 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

WO 02/02757

PCT/US01/20491

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 ADGUC on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding ADGUC and either CD64 or CD64-GFP.

5 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 ADGUC and other genes of interest can be analyzed by northern
10 analysis or microarray techniques.

XIV. Production of ADGUC Specific Antibodies

ADGUC 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.

15 Alternatively, the ADGUC 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.)

20 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
25 antipeptide and anti-ADGUC activity by, for example, binding the peptide or ADGUC 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 ADGUC Using Specific Antibodies

Naturally occurring or recombinant ADGUC is substantially purified by immunoaffinity
30 chromatography using antibodies specific for ADGUC. An immunoaffinity column is constructed by covalently coupling anti-ADGUC 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.

WO 02/02757

PCT/US01/20491

Media containing ADGUC are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of ADGUC (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/ADGUC 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 ADGUC is collected.

XVI. Identification of Molecules Which Interact with ADGUC

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

Alternatively, molecules interacting with ADGUC 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).

ADGUC 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 ADGUC Activity

Adenylyl cyclase activity of ADGUC is demonstrated by the ability to convert ATP to cAMP (Mittal, C.K. (1986) *Meth. Enzymol.* 132:422-428). In this assay ADGUC is incubated with the substrate [α -³²P]ATP, following which the excess substrate is separated from the product cyclic [³²P] AMP. ADGUC activity is determined in 12 x 75 mm disposable culture tubes containing 5 μ l of 0.6 M Tris-HCl, pH 7.5, 5 μ l of 0.2 M MgCl₂, 5 μ l of 150 mM creatine phosphate containing 3 units of creatine phosphokinase, 5 μ l of 4.0 mM 1-methyl-3-isobutylxanthine, 5 μ l of 20 mM cAMP, 5 μ l 20 mM dithiothreitol, 5 μ l of 10 mM ATP, 10 μ l [α -³²P]ATP ($2-4 \times 10^6$ cpm), and water in a total volume of 100 μ l. The reaction mixture is prewarmed to 30°C. The reaction is initiated by adding ADGUC to the prewarmed reaction mixture. After 10-15 minutes of incubation at 30°C, the reaction is terminated by adding 25 μ l of 30% ice-cold trichloroacetic acid (TCA). Zero-time incubations and reactions incubated in the absence of ADGUC are used as negative controls. Products are separated by ion exchange chromatography, and cyclic [³²P] AMP is quantified using a β -radioisotope counter. The ADGUC activity is proportional to the amount of cyclic [³²P] AMP formed in the reaction.

WO 02/02757

PCT/US01/20491

Guanylyl cyclase activity of ADGUC is demonstrated by the ability to convert GTP to cGMP (Mittal, supra). In this assay ADGUC is incubated with the substrate [α - 32 P]GTP, following which the excess substrate is separated from the product cyclic [32 P]GMP. A reaction mixture contains 5 μ l of 1 M Tris-HCl, pH 7.5, 5 μ l 80 mM MnCl₂ or MgCl₂, 25 μ l of 40 mM theophylline or 2.0 mM 1-methyl-3-isobutylxanthine, 5 μ l 150 mM creatine phosphate containing 20 μ g creatine phosphokinase (120-135 units/mg protein), 5 μ l 20 mM cGMP, 10 μ l 10 mM GTP, 10 μ l [α - 32 P] GTP (containing 2-4 x 10⁶ cpm), and water in a total volume of 100 μ l. The reaction is initiated by the addition of ADGUC. After 10-15 minutes of incubation at 37°C, the reaction is terminated by adding 20 μ l of 40% ice-cold trichloroacetic acid. Zero-time incubations and reactions incubated in the absence of ADGUC are used as negative controls. Products are separated by ion exchange chromatography, and cyclic [32 P]GMP is quantified using a β -radioisotope counter. The ADGUC activity is proportional to the amount of cyclic [32 P]GMP formed in the reaction.

XVIII. Identification of ADGUC Agonists and Antagonists

Agonists or antagonists of ADGUC activation or inhibition may be tested using the assays described in section XVII, or with the use of assay technologies which allow high throughput readout in multi-well plate format, such as the Cyclic AMP FlashPlate Assay and Cyclic GMP FlashPlate Assay (NEN Life Sciences Products). Agonists cause an increase in ADGUC activity and antagonists cause a decrease in ADGUC activity.

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

Table 1

Incyte Protein ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
3298552	1	7478215	9	768924261
7478215	2	7478215	7	7478215
1505146	3	1505146CD1	8	1505146CB1
1577526	4	1577526CD1	9	1577526CB1
7478215	5	7478215CD1	10	7478215CB1

WO 02/02757

PCT/US01/20491

Table 2

Polypeptide SEQ ID NO.	Incyte Polypeptide ID	GenBank ID NO:	Probability score	GenBank Homolog
1	3685652CD1	9202676	0	[Rattus norvegicus] adenyl cyclase type IV Gao, B. and Gilman, A.G. (1991) Cloning and expression of a widely distributed (type IV) adenyl cyclase. Proc. Natl. Acad. Sci. U.S.A. 88:10178-10182.
2	7478216CD1	91758332	0	[Rattus norvegicus] adenyl cyclase type V Scholich, K. et al. (1999) Science 283:1328-1331
3	1505146CD1	9868200	2.20E-21	[Caenorhabditis elegans] similar to adenylate cyclase Wilson, R. et al. (1994) Nature 368: 32-38
4	1577526CD1	91504042	7.50E-254	[Homo sapiens] similar to yeast adenylate cyclase (S56715)
5	7478215CD1	9162613	0	[Bos taurus] adenyl cyclase Type I

Table 3 (cont.)

SPO ID No.	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
2	7478216CD1	837	S124 S128 S150 S159 S171 S190 S202 S264 S283 S298 S376 S455 S473 S519 S806 S817 S858 S878 T698 T795 Y573	M313 N519	Transmembrane Domain: S315-W331, Score 14.27 Y355-L371, Score 11.82 Adenylylate and Guanylylate catalytic domain guanylylate.cyc: T512-R696 Score 286.9; E=2.6 e-82 Guanylylate cyclases signature guanylylate.cyclases.prf: E605-G668 score 30.97 Guanylylate cyclases BL00452: G524-I666, E476-K497, I512-D554, C562-L578, P<4e-10 GUANYLYLATE CYCLASES DM00173 A45195 276-552: E406-Y683, P value 2.2e-148 ADENYLYLATE CYCLASE PM030262: M76-E180 P value 9.9e-50 PM0003877: S376-Y513 P value 4.7e-43 PM0003560: Q515-A653 P value 1.3e-24 PM002814: K754-Q837 P value 1.1e-21 Leucine Zipper: L542-L563 Guanylylate Cyclases: G525-P648 signal_classifier.MI-S64 Leucine-rich repeat: signature PR00019A:V20-I33 PR00019B:C40-L53	HMMER HMMER_Pfam PROFILES SCAN BLAST_BLOCKS BLAST_DOMO BLAST_PRODOM MOOTIPS MOOTIPS SIFOS3 BLASTP_PRIMIS
3	1505146CD1	104	S10 S14 S34 T99 Y89			

Table 3 (cont.)

SBO ID	Inactive Polypeptide ID	Amino Acid Residues	Peptidyl Phosphorylation Sites	Relevant Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Searches
4	1577526CD1	763	S112, S113, S143, S167, S190, S273, S307, S347, S351, S622, S69, S738, T131, T138, T406, T588, T717, T759, T764	N600, N715, N821, N845, N860, N715	Signal Peptide: Y224-Y244 Transmembrane domain: M623-T645, K646-S668, M669-K691, K692-S714, N715-Q737, M430-E452, N454-K476, N477-K504, N505-P527, N552-N574, N575-Q592, N600-S622 Leucine Rich Repeat: M623-T645, K646-S668, M669-K691, K692-S714, N715-Q737, M430-E452, N454-K476, N477-K504, N505-P527, N552-N574, N575-Q592, N600-S622 Transmembrane domains: T610-F631, I638-F655 Adenylylate and guanylylate cyclase catalytic domain: I294-P478, I859-G1056 Guanylylate cyclase motifs: G407-E430, G987-D1010 Guanylylate cyclases signatures: Q968-R1030, ProfileScan E387-G450 Guanylylate cyclase protein EL00452: Q968-R1030, R1038-F1063, R1076-L1078, R1086-L1092, R1098-F1099, R1100-L1102, R1104-L1106, R1108-L1110, R1112-L1114, R1116-L1118, R1120-L1122, R1124-L1126, R1128-L1130, R1132-L1134, R1136-L1138, R1140-L1142, R1144-L1146, R1148-L1150, R1152-L1154, R1156-L1158, R1160-L1162, R1164-L1166, R1168-L1170, R1172-L1174, R1176-L1178, R1180-L1182, R1184-L1186, R1188-L1190, R1192-L1194, R1196-L1198, R1200-L1202, R1204-L1206, R1208-L1210, R1212-L1214, R1216-L1218, R1220-L1222, R1224-L1226, R1228-L1230, R1232-L1234, R1236-L1238, R1240-L1242, R1244-L1246, R1248-L1250, R1252-L1254, R1256-L1258, R1260-L1262, R1264-L1266, R1268-L1270, R1272-L1274, R1276-L1278, R1280-L1282, R1284-L1286, R1288-L1290, R1292-L1294, R1296-L1298, R1300-L1302, R1304-L1306, R1308-L1310, R1312-L1314, R1316-L1318, R1320-L1322, R1324-L1326, R1328-L1330, R1332-L1334, R1336-L1338, R1340-L1342, R1344-L1346, R1348-L1350, R1352-L1354, R1356-L1358, R1360-L1362, R1364-L1366, R1368-L1370, R1372-L1374, R1376-L1378, R1380-L1382, R1384-L1386, R1388-L1390, R1392-L1394, R1396-L1398, R1400-L1402, R1404-L1406, R1408-L1410, R1412-L1414, R1416-L1418, R1420-L1422, R1424-L1426, R1428-L1430, R1432-L1434, R1436-L1438, R1440-L1442, R1444-L1446, R1448-L1450, R1452-L1454, R1456-L1458, R1460-L1462, R1464-L1466, R1468-L1470, R1472-L1474, R1476-L1478, R1480-L1482, R1484-L1486, R1488-L1490, R1492-L1494, R1496-L1498, R1500-L1502, R1504-L1506, R1508-L1510, R1512-L1514, R1516-L1518, R1520-L1522, R1524-L1526, R1528-L1530, R1532-L1534, R1536-L1538, R1540-L1542, R1544-L1546, R1548-L1550, R1552-L1554, R1556-L1558, R1560-L1562, R1564-L1566, R1568-L1570, R1572-L1574, R1576-L1578, R1580-L1582, R1584-L1586, R1588-L1590, R1592-L1594, R1596-L1598, R1600-L1602, R1604-L1606, R1608-L1610, R1612-L1614, R1616-L1618, R1620-L1622, R1624-L1626, R1628-L1630, R1632-L1634, R1636-L1638, R1640-L1642, R1644-L1646, R1648-L1650, R1652-L1654, R1656-L1658, R1660-L1662, R1664-L1666, R1668-L1670, R1672-L1674, R1676-L1678, R1680-L1682, R1684-L1686, R1688-L1690, R1692-L1694, R1696-L1698, R1700-L1702, R1704-L1706, R1708-L1710, R1712-L1714, R1716-L1718, R1720-L1722, R1724-L1726, R1728-L1730, R1732-L1734, R1736-L1738, R1740-L1742, R1744-L1746, R1748-L1750, R1752-L1754, R1756-L1758, R1760-L1762, R1764-L1766, R1768-L1770, R1772-L1774, R1776-L1778, R1780-L1782, R1784-L1786, R1788-L1790, R1792-L1794, R1796-L1798, R1800-L1802, R1804-L1806, R1808-L1810, R1812-L1814, R1816-L1818, R1820-L1822, R1824-L1826, R1828-L1830, R1832-L1834, R1836-L1838, R1840-L1842, R1844-L1846, R1848-L1850, R1852-L1854, R1856-L1858, R1860-L1862, R1864-L1866, R1868-L1870, R1872-L1874, R1876-L1878, R1880-L1882, R1884-L1886, R1888-L1890, R1892-L1894, R1896-L1898, R1900-L1902, R1904-L1906, R1908-L1910, R1912-L1914, R1916-L1918, R1920-L1922, R1924-L1926, R1928-L1930, R1932-L1934, R1936-L1938, R1940-L1942, R1944-L1946, R1948-L1950, R1952-L1954, R1956-L1958, R1960-L1962, R1964-L1966, R1968-L1970, R1972-L1974, R1976-L1978, R1980-L1982, R1984-L1986, R1988-L1990, R1992-L1994, R1996-L1998, R2000-L2002, R2004-L2006, R2008-L2010, R2012-L2014, R2016-L2018, R2020-L2022, R2024-L2026, R2028-L2030, R2032-L2034, R2036-L2038, R2040-L2042, R2044-L2046, R2048-L2050, R2052-L2054, R2056-L2058, R2060-L2062, R2064-L2066, R2068-L2070, R2072-L2074, R2076-L2078, R2080-L2082, R2084-L2086, R2088-L2090, R2092-L2094, R2096-L2098, R2100-L2102, R2104-L2106, R2108-L2110, R2112-L2114, R2116-L2118, R2120-L2122, R2124-L2126, R2128-L2130, R2132-L2134, R2136-L2138, R2140-L2142, R2144-L2146, R2148-L2150, R2152-L2154, R2156-L2158, R2160-L2162, R2164-L2166, R2168-L2170, R2172-L2174, R2176-L2178, R2180-L2182, R2184-L2186, R2188-L2190, R2192-L2194, R2196-L2198, R2200-L2202, R2204-L2206, R2208-L2210, R2212-L2214, R2216-L2218, R2220-L2222, R2224-L2226, R2228-L2230, R2232-L2234, R2236-L2238, R2240-L2242, R2244-L2246, R2248-L2250, R2252-L2254, R2256-L2258, R2260-L2262, R2264-L2266, R2268-L2270, R2272-L2274, R2276-L2278, R2280-L2282, R2284-L2286, R2288-L2290, R2292-L2294, R2296-L2298, R2300-L2302, R2304-L2306, R2308-L2310, R2312-L2314, R2316-L2318, R2320-L2322, R2324-L2326, R2328-L2330, R2332-L2334, R2336-L2338, R2340-L2342, R2344-L2346, R2348-L2350, R2352-L2354, R2356-L2358, R2360-L2362, R2364-L2366, R2368-L2370, R2372-L2374, R2376-L2378, R2380-L2382, R2384-L2386, R2388-L2390, R2392-L2394, R2396-L2398, R2400-L2402, R2404-L2406, R2408-L2410, R2412-L2414, R2416-L2418, R2420-L2422, R2424-L2426, R2428-L2430, R2432-L2434, R2436-L2438, R2440-L2442, R2444-L2446, R2448-L2450, R2452-L2454, R2456-L2458, R2460-L2462, R2464-L2466, R2468-L2470, R2472-L2474, R2476-L2478, R2480-L2482, R2484-L2486, R2488-L2490, R2492-L2494, R2496-L2498, R2500-L2502, R2504-L2506, R2508-L2510, R2512-L2514, R2516-L2518, R2520-L2522, R2524-L2526, R2528-L2530, R2532-L2534, R2536-L2538, R2540-L2542, R2544-L2546, R2548-L2550, R2552-L2554, R2556-L2558, R2560-L2562, R2564-L2566, R2568-L2570, R2572-L2574, R2576-L2578, R2580-L2582, R2584-L2586, R2588-L2590, R2592-L2594, R2596-L2598, R2600-L2602, R2604-L2606, R2608-L2610, R2612-L2614, R2616-L2618, R2620-L2622, R2624-L2626, R2628-L2630, R2632-L2634, R2636-L2638, R2640-L2642, R2644-L2646, R2648-L2650, R2652-L2654, R2656-L2658, R2660-L2662, R2664-L2666, R2668-L2670, R2672-L2674, R2676-L2678, R2680-L2682, R2684-L2686, R2688-L2690, R2692-L2694, R2696-L2698, R2700-L2702, R2704-L2706, R2708-L2710, R2712-L2714, R2716-L2718, R2720-L2722, R2724-L2726, R2728-L2730, R2732-L2734, R2736-L2738, R2740-L2742, R2744-L2746, R2748-L2750, R2752-L2754, R2756-L2758, R2760-L2762, R2764-L2766, R2768-L2770, R2772-L2774, R2776-L2778, R2780-L2782, R2784-L2786, R2788-L2790, R2792-L2794, R2796-L2798, R2800-L2802, R2804-L2806, R2808-L2810, R2812-L2814, R2816-L2818, R2820-L2822, R2824-L2826, R2828-L2830, R2832-L2834, R2836-L2838, R2840-L2842, R2844-L2846, R2848-L2850, R2852-L2854, R2856-L2858, R2860-L2862, R2864-L2866, R2868-L2870, R2872-L2874, R2876-L2878, R2880-L2882, R2884-L2886, R2888-L2890, R2892-L2894, R2896-L2898, R2900-L2902, R2904-L2906, R2908-L2910, R2912-L2914, R2916-L2918, R2920-L2922, R2924-L2926, R2928-L2930, R2932-L2934, R2936-L2938, R2940-L2942, R2944-L2946, R2948-L2950, R2952-L2954, R2956-L2958, R2960-L2962, R2964-L2966, R2968-L2970, R2972-L2974, R2976-L2978, R2980-L2982, R2984-L2986, R2988-L2990, R2992-L2994, R2996-L2998, R3000-L3002, R3004-L3006, R3008-L3010, R3012-L3014, R3016-L3018, R3020-L3022, R3024-L3026, R3028-L3030, R3032-L3034, R3036-L3038, R3040-L3042, R3044-L3046, R3048-L3050, R3052-L3054, R3056-L3058, R3060-L3062, R3064-L3066, R3068-L3070, R3072-L3074, R3076-L3078, R3080-L3082, R3084-L3086, R3088-L3090, R3092-L3094, R3096-L3098, R3100-L3102, R3104-L3106, R3108-L3110, R3112-L3114, R3116-L3118, R3120-L3122, R3124-L3126, R3128-L3130, R3132-L3134, R3136-L3138, R3140-L3142, R3144-L3146, R3148-L3150, R3152-L3154, R3156-L3158, R3160-L3162, R3164-L3166, R3168-L3170, R3172-L3174, R3176-L3178, R3180-L3182, R3184-L3186, R3188-L3190, R3192-L3194, R3196-L3198, R3200-L3202, R3204-L3206, R3208-L3210, R3212-L3214, R3216-L3218, R3220-L3222, R3224-L3226, R3228-L3230, R3232-L3234, R3236-L3238, R3240-L3242, R3244-L3246, R3248-L3250, R3252-L3254, R3256-L3258, R3260-L3262, R3264-L3266, R3268-L3270, R3272-L3274, R3276-L3278, R3280-L3282, R3284-L3286, R3288-L3290, R3292-L3294, R3296-L3298, R3300-L3302, R3304-L3306, R3308-L3310, R3312-L3314, R3316-L3318, R3320-L3322, R3324-L3326, R3328-L3330, R3332-L3334, R3336-L3338, R3340-L3342, R3344-L3346, R3348-L3350, R3352-L3354, R3356-L3358, R3360-L3362, R3364-L3366, R3368-L3370, R3372-L3374, R3376-L3378, R3380-L3382, R3384-L3386, R3388-L3390, R3392-L3394, R3396-L3398, R3400-L3402, R3404-L3406, R3408-L3410, R3412-L3414, R3416-L3418, R3420-L3422, R3424-L3426, R3428-L3430, R3432-L3434, R3436-L3438, R3440-L3442, R3444-L3446, R3448-L3450, R3452-L3454, R3456-L3458, R3460-L3462, R3464-L3466, R3468-L3470, R3472-L3474, R3476-L3478, R3480-L3482, R3484-L3486, R3488-L3490, R3492-L3494, R3496-L3498, R3500-L3502, R3504-L3506, R3508-L3510, R3512-L3514, R3516-L3518, R3520-L3522, R3524-L3526, R3528-L3530, R3532-L3534, R3536-L3538, R3540-L3542, R3544-L3546, R3548-L3550, R3552-L3554, R3556-L3558, R3560-L3562, R3564-L3566, R3568-L3570, R3572-L3574, R3576-L3578, R3580-L3582, R3584-L3586, R3588-L3590, R3592-L3594, R3596-L3598, R3600-L3602, R3604-L3606, R3608-L3610, R3612-L3614, R3616-L3618, R3620-L3622, R3624-L3626, R3628-L3630, R3632-L3634, R3636-L3638, R3640-L3642, R3644-L3646, R3648-L3650, R3652-L3654, R3656-L3658, R3660-L3662, R3664-L3666, R3668-L3670, R3672-L3674, R3676-L3678, R3680-L3682, R3684-L3686, R3688-L3690, R3692-L3694, R3696-L3698, R3700-L3702, R3704-L3706, R3708-L3710, R3712-L3714, R3716-L3718, R3720-L3722, R3724-L3726, R3728-L3730, R3732-L3734, R3736-L3738, R3740-L3742, R3744-L3746, R3748-L3750, R3752-L3754, R3756-L3758, R3760-L3762, R3764-L3766, R3768-L3770, R3772-L3774, R3776-L3778, R3780-L3782, R3784-L3786, R3788-L3790, R3792-L3794, R3796-L3798, R3800-L3802, R3804-L3806, R3808-L3810, R3812-L3814, R3816-L3818, R3820-L3822, R3824-L3826, R3828-L3830, R3832-L3834, R3836-L3838, R3840-L3842, R3844-L3846, R3848-L3850, R3852-L3854, R3856-L3858, R3860-L3862, R3864-L3866, R3868-L3870, R3872-L3874, R3876-L3878, R3880-L3882, R3884-L3886, R3888-L3890, R3892-L3894, R3896-L3898, R3900-L3902, R3904-L3906, R3908-L3910, R3912-L3914, R3916-L3918, R3920-L3922, R3924-L3926, R3928-L3930, R3932-L3934, R3936-L3938, R3940-L3942, R3944-L3946, R3948-L3950, R3952-L3954, R3956-L3958, R3960-L3962, R3964-L3966, R3968-L3970, R3972-L3974, R3976-L3978, R3980-L3982, R3984-L3986, R3988-L3990, R3992-L3994, R3996-L3998, R4000-L4002, R4004-L4006, R4008-L4010, R4012-L4014, R4016-L4018, R4020-L4022, R4024-L4026, R4028-L4030, R4032-L4034, R4036-L4038, R4040-L4042, R4044-L4046, R4048-L4050, R4052-L4054, R4056-L4058, R4060-L4062, R4064-L4066, R4068-L4070, R4072-L4074, R4076-L4078, R4080-L4082, R4084-L4086, R4088-L4090, R4092-L4094, R4096-L4098, R4100-L4102, R4104-L4106, R4108-L4110, R4112-L4114, R4116-L4118, R4120-L4122, R4124-L4126, R4128-L4130, R4132-L4134, R4136-L4138, R4140-L4142, R4144-L4146, R4148-L4150, R4152-L4154, R4156-L4158, R4160-L4162, R4164-L4166, R4168-L4170, R4172-L4174, R4176-L4178, R4180-L4182, R4184-L4186, R4188-L4190, R4192-L4194, R4196-L4198, R4200-L4202, R4204-L4206, R4208-L4210, R4212-L4214, R4216-L4218, R4220-L4222, R4224-L4226, R4228-L4230, R4232-L4234, R4236-L4238, R4240-L4242, R4244-L4246, R4248-L4250, R4252-L4254, R4256-L4258, R4260-L4262, R4264-L4266, R4268-L4270, R4272-L4274, R4276-L4278, R4280-L4282, R4284-L4286, R4288-L4290, R4292-L4294, R4296-L4298, R4300-L4302, R4304-L4306, R4308-L4310, R4312-L4314, R4316-L4318, R4320-L4322, R4324-L4326, R4328-L4330, R4332-L4334, R4336-L4338, R4340-L4342, R4344-L4346, R4348-L4350, R4352-L4354, R4356-L4358, R4360-L4362, R4364-L4366, R4368-L4370, R4372-L4374, R4376-L4378, R4380-L4382, R4384-L4386, R4388-L4390, R4392-L4394, R4396-L4398, R4400-L4402, R4404-L4406, R4408-L4410, R4412-L4414, R4416-L4418, R4420-L4422, R4424-L4426, R4428-L4430, R4432-L4434, R4436-L4438, R4440-L4442, R4444-L4446, R4448-L4450, R4452-L4454, R4456-L4458, R4460-L4462, R4464-L4466, R4468-L4470, R4472-L4474, R4476-L4478, R4480-L4482, R4484-L4486, R4488-L4490, R4492-L4494, R4496-L4498, R4500-L4502, R4504-L4506, R4508-L4510, R4512-L4514, R4516-L4518, R4520-L4522, R4524-L4526, R4528-L4530, R4532-L4534, R4536-L4538, R4540-L4542, R4544-L4546, R4548-L4550, R4552-L4554, R4556-L4558, R4560-L4562, R4564-L4566, R4568-L4570, R4572-L4574, R4576-L4578, R4580-L4582, R4584-L4586, R4588-L4590, R4592-L4594, R4596-L4598, R4600-L4602, R4604-L4606, R4608-L4610, R4612-L4614, R4616-L4618, R4620-L4622, R4624-L4626, R4628-L4630, R4632-L4634, R4636-L4638, R4640-L4642, R4644-L4646, R4648-L4650, R4652-L4654, R4656-L4658, R4660-L4662, R4664-L4666, R4668-L4670, R4672-L4674, R4676-L4678, R4680-L4682, R4684-L4686, R4688-L4690, R4692-L4694, R4696-L4698, R4700-L4702, R4704-L4706, R4708-L4710, R4712-L4714, R4716-L4718, R4720-L4722, R4724-L4726, R4728-L4730, R4732-L4734, R4736-L4738, R4740-L4742, R4744-L4746, R4748-L4750, R4752-L4754, R4756-L4758, R4760-L4762, R4764-L4766, R4768-L4770, R4772-L4774, R4776-L4778, R4780-L4782, R4784-L4786, R4788-L4790, R4792-L4794, R4796-L4798, R4800-L4802, R4804-L4806, R4808-L4810, R4812-L4814, R4816-L4818, R4820-L4822, R4824-L4826, R4828-L4830, R4832-L4834, R4836-L4838, R4840-L4842, R4844-L4846, R4848-L4850, R4852-L4854, R4856-L4858, R4860-L4862, R4864-L4866, R4868-L4870, R4872-L4874, R4876-L4878, R4880-L4882, R4884-L4886, R4888-L4890, R4892-L4894, R4896-L4898, R4900-L4902, R4904-L4906, R4908-L4910, R4912-L4914, R4916-L4918, R4920-L4922, R4924-L4926, R4928-L4930, R4932-L4934, R4936-L4938, R4940-L4942, R4944-L4946, R4948-L4950, R4952-L4954, R4956-L4958, R4960-L4962, R4964-L4966, R4968-L4970, R4972-L4974, R4976-L4978, R4980-L4982, R4984-L4986, R4988-L4990, R4992-L4994, R4996-L4998, R5000-L5002, R5004-L5006, R5008-L5010, R5012-L5014, R5016-L5018, R5020-L5022, R5024-L5026, R5028-L5030, R5032-L5034, R5036-L5038, R5040-L5042, R5044-L5046, R5048-L5050, R5052-L5054, R5056-L5058, R5060-L5062, R5064-L5066, R5068-L5070, R5072-L5074, R5076-L5078, R5080-L5082, R5084-L5086, R5088-L5090, R5092-L5094, R5096-L5098, R5100-L5102, R5104-L5106, R5108-L5110, R5112-L5114, R5116-L5118, R5120-L5122, R5124-L5126, R5128-L5130, R5132-L5134, R5136-L5138, R5140-L5142, R5144-L5146, R5148-L5150, R5152-L5154, R5156-L5158, R5160-L5162, R5164-L5166, R5168-L5170, R5172-L5174, R5176-L5178, R5180-L5182, R5184-L5186, R5188-L5190, R5192-L5194, R5196-L5198, R5200-L5202, R5204-L5206, R5208-L5210, R5212-L5214, R5216-L5218, R5220-L5222, R5224-L5226, R5228-L5230, R5232-L5234, R5236-L5238, R5240-L5242, R5244-L5246, R5248-L5250, R5252-L5	

Table 4

Polynucleotide SEQ ID NO:	Inocyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
6	3685652CB1	3769	1-1652, 3727-3769, 2180-3019	7345578H1 (SYNODIN02)	3170	3741
				7248217G (SYNODIN01)	3003	3699
				7709325H1 (PANCNO02)	22	557
				FL2685652_g8346725_g202676	410	3670
				21379036	3282	3769
				653476H1 (CONFTR02)	53	286
				653476H1 (CONFTR02)	53	286
				7119306V1 (ZERSRO709)	1053	1874
				6766950J1 (BRANOR01)	1026	1670
				7357757H1 (HEARNOR03)	679	1042
7	7478216CB1	3137	1-669, 1813-2019, 2427-3137, 753-820	71153228V1	2559	3137
				67543228H1 (SINTFER02)	1894	2496
				GNN_g8567693_000003_002	1	1290
				71151710V1	2350	2967
				71152591V1	1815	2332
				71874115V1	1	538
				71874576V1	476	1072
				71874777V1	1451	1996
				71875066V1	542	1893
				71875066V1	542	1893
8	1505146CB1	1958	107-210, 189-188, 466, 352- 927	52092686 (MUP22DF01)	1335	1737
				6783247E9 (SINTMCG1)	2105	2660
				656837698 (MCLDAN05)	469	897
				71434255V1	1709	2404
				5302686H1 (DRGTNOR04)	413	667
				6980889H1 (BRAPTR04)	966	1552
				71441001V1	1925	2530
				5847513H1 (BRAENOT04)	757	1029
				GNN_g8176861_000002_002	1	2039
				71875066V1	542	1893
9	1577526CB1	2660	1-875	71434255V1	1709	2404
				5302686H1 (DRGTNOR04)	413	667
				6980889H1 (BRAPTR04)	966	1552
				71441001V1	1925	2530
				5847513H1 (BRAENOT04)	757	1029
				GNN_g8176861_000002_002	1	2039
				71875066V1	542	1893
				71875066V1	542	1893
				71875066V1	542	1893
				71875066V1	542	1893

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
10	7478215CB1	3811	225-285, 1-30, 830-1142, 2811-3811, 1723-2125	7738563H1 (BRAITFE01) 7252668J2 (BRAIFEE04) 6984562R8 (BRAIFER05) 7230580H1 (BRAIIDE15) 7282824H1 (BRAIFER05) 6981221F8 (BRAIFER05) 7292452F8 (BRAIFER05) 6894026H1 (BRAIIDE03) 7232668I2 (BRAIFEE04)	607 1 2035 1320 2681 2042 3255 568 1456	1357 616 2786 1855 3268 2768 3811 1184 2095

Table 5

Polynucleotide SEQ. ID. NO:	Incyte Project ID	Representative Library
6	368562CE1	PROSUS23
7	7478216CE1	BR30NOR01
8	1505146CE1	OTRABF07
9	157526CE1	SINLYMCO1
10	7478213CE1	BRALFER05

Table 6 (cont.)

Library	Vector	Library Description
PROST0523	pINCY	<p>This subtracted prostate tumor library was constructed using 10 million clones from a prostate tumor library that was subjected to 2 rounds of subtractive hybridization with 10 million clones from a prostate tumor library. The starting library for subtraction was constructed by pooling equal numbers of clones from 4 prostate tumor libraries using mRNA isolated from prostate tumor removed from Caucasian males at ages 58 (A), 61 (B), 65 (C), and 68 (D) during prostatectomy with lymph node excision. Pathology indicated adenocarcinoma in all donors. History included elevated PSA, induration and tobacco abuse in donor A; elevated PSA, induration, prostate hyperplasia, renal failure, osteoarthritis, renal artery stenosis, benign HFN, thrombocytopenia, hyperlipidemia, tobacco/alcohol abuse and hepatitis C (carrier) in donor B; elevated PSA, induration, osteoarthritis, renal artery stenosis, benign HFN, induration, hypercholesterolemia, and kidney calculus in donor C; and elevated PSA, induration, hypercholesterolemia, and kidney calculus in donor D. The starting prostate tissue libraries were constructed by pooling equal numbers of cDNA clones from 3 prostate tissue libraries derived from prostate tissue, prostate epithelial cells, and fibroblasts from prostate stroma from 3 different donors. Subtractive hybridization conditions were based on the methodologies of Swarcop et al., NAR 19 (1991):1954 and Bonaldo, et al., Genome Research 6 (1996):791.</p>
SIN17MC01	pINCY	<p>This large size-fractionated library was constructed using pooled cDNA from two donors. Clones generated using this library were isolated from ileum tissue removed from a 30-year-old Caucasian female (donor A) who had a history of Crohn's disease, sigmoiditis, and permanent colostomy, and from ileum tissue removed from a 70-year-old Caucasian female (donor B) during right hemicolectomy, open liver biopsy, sigmoidoscopy, colonoscopy, and per-3) manent colostomy. Pathology for the matched tumor tissue (donor A) indicated carcinoma tumor (grade 1 neuroendocrine carcinoma) arising in the terminal ileum. The tumor per-meated through the ileal wall into the mesenteric fat and extended into the adherent cecum, where tumor extended through the bowel wall up to the mucosal surface. Multiple direct tumor metastatic foci were identified in the mesenteric lymph nodes. Pathology for the matched tumor (donor B) indicated invasive, grade 2 adenocarcinoma forming an ulcerated mass, situated distal to the ileocecal valve. The tumor invaded through the muscularis propria just into the serosal adipose tissue. One regional lymph node was positive for a microfocus of metastatic adenocarcinoma. Donor A presented with flushing and unspecified abdominal/pelvic symptoms. Patient history included endo-metriosiis, and tobacco and alcohol abuse. Donor B's history included a malignant breast cancer, type 1 diabetes, hyperlipidemia, viral hepatitis, an unspecified thyroid dis-order, osteoarthritis, and a malignant skin neoplasm. Donor B's medication included tamoxifen.</p>

WO 02/02757

PCT/US01/20491

Table 7

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	Mismatch <50%
ABI/PARACCEL/FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	
ABI Auto-Assembler	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.	EST: 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 search.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	EST: fasta E value=1.00E-6 Assembled EST: fasta identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less Full Length sequences: fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:38-105; and Altwood, 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 Nussliel, Cambridge Univ. Press, pp. 1-350.	PFAM hits: Probability value= 1.0E-3 or less Signal peptide hits: Score= 0 or greater

WO 02/02757

PCT/US01/20491

Table 7 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribshov, M. et al. (1988) CABIOS 4:61-66; Gribshov, M. et al. (1989) Methods Enzymol. 183:146-159; Birooch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality score=CCG-specified "HGR" 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. 8:175-182; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phis Revised Assembly Program including SWAT and CrossMatch programs based on efficient implementation of the Smith-Waterman algorithm, used in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:159-197; and Green, P., University of Washington, Seattle, WA.	Score=120 or greater. Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claretie, J.M. and S. Audic (1997) CABIOS 7:431-439.	Score=3.5 or greater
TMAP	A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation.	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Sonnhammer, E.L. et al. (1998) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol. Glasgow et al., eds. The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182.	
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Biaroch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

WO 02/02757

PCT/US01/20491

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:1-5,
 - b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical
to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5,
 - c) a biologically active fragment of a polypeptide having an amino acid sequence selected
from the group consisting of SEQ ID NO:1-5, and
 - 10 d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from
the group consisting of SEQ ID NO:1-5.
2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-5.
- 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:6-
20 10.
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 02/02757

PCT/US01/20491

10. An isolated antibody which specifically binds to a polypeptide of claim 1.
11. An isolated polynucleotide selected from the group consisting of:
- a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting
 - 5 of SEQ ID NO:6-10,
 - b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10,
 - c) a polynucleotide complementary to a polynucleotide of a),
 - d) a polynucleotide complementary to a polynucleotide of b), and
 - 10 e) an RNA equivalent of a)-d).
12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 11.
13. A method 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.

WO 02/02757

PCT/US01/20491

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

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

18. A method for treating a disease or condition associated with decreased expression of functional ADGUC, 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 ADGUC, 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 ADGUC, comprising administering to a patient in need of such treatment a composition of claim 23.

WO 02/02757

PCT/US01/20491

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

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

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

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

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

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

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

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

WO 02/02757

PCT/US01/20491

sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 11 or fragment thereof;

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

29. A diagnostic test for a condition or disease associated with the expression of ADGUC 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
- 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,
- b) a single chain antibody,
- c) a Fab fragment,
- d) a F(ab')₂ fragment, or
- e) a humanized antibody.

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 ADGUC in a subject, comprising administering to said subject an effective amount of the composition of claim 31.

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 ADGUC in a subject, comprising administering to said subject an effective amount of the composition of claim 33.

WO 02/02757

PCT/US01/20491

35. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 10 comprising:
- a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, or an immunogenic fragment thereof, under conditions to elicit an antibody response;
 - b) isolating antibodies from said animal; and
 - c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-5.
36. An antibody produced by a method of claim 35.
37. A composition comprising the antibody of claim 36 and a suitable carrier.
38. A method of making a monoclonal antibody with the specificity of the antibody of claim 10 comprising:
- a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, or an immunogenic fragment thereof, under conditions to elicit an antibody response;
 - b) isolating antibody producing cells from the animal;
 - c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells;
 - d) culturing the hybridoma cells; and
 - e) isolating from the culture monoclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-5.
39. A monoclonal antibody produced by a method of claim 38.
40. A composition comprising the antibody of claim 39 and a suitable carrier.
41. The antibody of claim 10, wherein the antibody is produced by screening a Fab expression library.

WO 02/02757

PCT/US01/20491

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

43. A method for detecting a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-5 in a sample, comprising the steps of:

5 a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and

b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-5 in the sample.

10

44. A method of purifying a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-5 from a sample, the method comprising:

a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and

15 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:1-5.

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

20 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.

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

25

49. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:5.

50. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:6.

30 51. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:7.

52. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:8.

WO 02/02757

PCT/US01/20491

53. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:9.

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

5

WO 02/02757

PCT/US01/20491

<110> INCYTE GENOMICS, INC.
 GANDHI, Ameena
 TRIBOULEY, Catherine
 LI, Ding
 LU, Dyung Aina M.
 LEE, Ernestine A.
 YUE, Henry
 YANG, Junming
 BAUGHN, Mariah R.
 THORNTON, Michael
 YAO, Monique G.
 WALIA, Narinder K.
 TANG, Y. Tom
 ELLIOTT, Vicki S.
 LU, Yan

<120> ADENYLYL AND GUANYLYL CYCLASES

<130> FI-0145 PCT

<140> To Be Assigned
 <141> Herewith

<150> 60215,476; 60/223,545; 60/229,876; 60/234,838; 60/236,483
 <151> 2000-06-29; 2000-08-04; 2000-08-31; 2000-09-22; 2000-09-29

<160> 10
 <170> PERL Program

<210> 1
 <211> 1086
 <212> FRF
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3685652CD1

<400> 1
 Met Ala Arg Leu Phe Ser Pro Arg Pro Pro Pro Ser Glu Asp Leu
 1 5 10 15
 Phe Tyr Glu Thr Tyr Tyr Ser Leu Ser Gln Gln Tyr Pro Leu Leu
 20 25 30
 Leu Leu Leu Leu Gly Ile Val Leu Cys Ala Leu Ala Ala Leu Leu
 35 40 45
 Ala Val Ala Trp Ala Ser Gly Arg Glu Leu Thr Ser Asp Pro Ser
 50 55 60
 Phe Leu Thr Thr Val Leu Cys Ala Leu Gly Gly Phe Ser Leu Leu
 65 70 75
 Leu Gly Leu Ala Ser Arg Glu Gln Arg Leu Gln Arg Trp Thr Arg
 80 85 90
 Pro Leu Ser Gly Leu Val Trp Val Ala Leu Leu Ala Leu Gly His
 95 100 105
 Ala Phe Leu Phe Thr Gly Gly Val Val Ser Ala Trp Asp Gln Val
 110 115 120
 Ser Tyr Phe Leu Phe Val Ile Phe Thr Ala Tyr Ala Met Leu Pro

WO 02/02757

PCT/US01/20491

	125	130	135
Leu Gly Met Arg Asp	Ala Ala Val Ala	Gly Leu Ala Ser Ser	Leu
	140	145	150
Ser His Leu Leu Val	Leu Gly Leu Tyr	Leu Gly Pro Gln Pro	Asp
	155	160	165
Ser Arg Pro Ala Leu	Leu Pro Gln Leu	Ala Ala Asn Ala Val	Leu
	170	175	180
Phe Leu Cys Gly Asn	Val Ala Arg Arg	Ala Arg Ser Ile Ser	Ala
	185	190	195
Leu Trp Tyr Thr Pro	Ala Thr Phe Arg	Glu Ala Leu Ser Ser	Leu
	200	205	210
His Ser Arg Arg Arg	Leu Asp Thr Glu	Lys Lys His Gln Glu	His
	215	220	225
Leu Leu Leu Ser Ile	Leu Pro Ala Tyr	Leu Ala Arg Glu Met	Lys
	230	235	240
Ala Glu Ile Met Ala	Arg Leu Gln Ala	Gly Gln Gly Ser Arg	Pro
	245	250	255
Glu Ser Thr Asn Asn	Phe His Ser Leu	Tyr Val Lys Arg His	Gln
	260	265	270
Gly Val Ser Val Leu	Tyr Ala Asp Ile	Val Gly Phe Thr Arg	Leu
	275	280	285
Ala Ser Glu Cys Ser	Pro Lys Glu Leu	Val Leu Met Leu Asn	Glu
	290	295	300
Leu Phe Gly Lys Phe	Asp Gln Ile Ala	Lys Val Pro Leu His	Thr
	305	310	315
His Ile His Lys Glu	His Glu Cys Met	Arg Ile Lys Ile Leu	Gly
	320	325	330
Asp Cys Tyr Tyr Cys	Val Ser Gly Leu	Pro Leu Ser Leu Pro	Asp
	335	340	345
His Ala Ile Asn Cys	Val Arg Met Gly	Leu Asp Met Cys Arg	Ala
	350	355	360
Ile Arg Lys Leu Arg	Ala Ala Thr Gly	Val Asp Ile Asn Met	Arg
	365	370	375
Val Gly Val His Ser	Gly Ser Val Leu	Cys Gly Val Ile Gly	Leu
	380	385	390
Gln Lys Trp Gln Tyr	Asp Val Trp Ser	His Asp Val Thr Leu	Ala
	395	400	405
Asn His Met Glu Ala	Gly Gly Val Pro	Gly Arg Val His Ile	Thr
	410	415	420
Gly Ala Thr Leu Ala	Leu Leu Ala Gly	Ala Tyr Ala Val Glu	Asp
	425	430	435
Ala Gly Met Glu His	Arg Asp Pro Tyr	Leu Arg Glu Leu Gly	Glu
	440	445	450
Pro Thr Tyr Leu Val	Ile Asp Pro Arg	Ala Glu Glu Asp	Glu
	455	460	465
Lys Gly Thr Ala Gly	Gly Leu Leu Ser	Ser Leu Glu Gly Leu	Lys
	470	475	480
Met Arg Pro Ser Leu	Leu Met Thr Arg	Tyr Leu Glu Ser Trp	Gly
	485	490	495
Ala Ala Lys Pro Phe	Ala His Leu Ser	His Gly Asp Ser Pro	Val
	500	505	510
Ser Thr Ser Thr Pro	Leu Pro Glu Lys	Thr Leu Ala Ser Phe	Ser
	515	520	525
Thr Gln Trp Ser Leu	Asp Arg Ser Arg	Thr Pro Arg Gly Leu	Asp
	530	535	540
Asp Glu Leu Asp Thr	Gly Asp Ala Lys	Phe Phe Gln Val Ile	Glu

WO 02/02757

PCT/US01/20491

	545		550		555									
Gln	Leu	Asn	Ser	Gln	Lys	Gln	Trp	Lys	Gln	Ser	Lys	Asp	Phe	Asn
	560				565									570
Pro	Leu	Thr	Leu	Tyr	Phe	Arg	Glu	Lys	Glu	Met	Glu	Lys	Glu	Tyr
	575				580									585
Arg	Leu	Ser	Ala	Ile	Pro	Ala	Phe	Lys	Tyr	Tyr	Glu	Ala	Cys	Thr
	590				595									600
Phe	Leu	Val	Phe	Leu	Ser	Asn	Phe	Ile	Ile	Gln	Met	Leu	Val	Thr
	605				610									615
Asn	Arg	Pro	Pro	Ala	Leu	Ala	Ile	Thr	Tyr	Ser	Ile	Thr	Phe	Leu
	620				625									630
Leu	Phe	Leu	Leu	Ile	Leu	Phe	Val	Cys	Phe	Ser	Glu	Asp	Leu	Met
	635				640									645
Arg	Cys	Val	Leu	Lys	Gly	Pro	Lys	Met	Leu	His	Trp	Leu	Pro	Ala
	650				655									660
Leu	Ser	Gly	Leu	Val	Ala	Thr	Arg	Pro	Gly	Leu	Arg	Ile	Ala	Leu
	665				670									675
Gly	Thr	Ala	Thr	Ile	Leu	Leu	Val	Phe	Ala	Met	Ala	Ile	Thr	Ser
	680				685									690
Leu	Phe	Phe	Phe	Pro	Thr	Ser	Ser	Asp	Cys	Pro	Phe	Gln	Ala	Pro
	695				700									705
Asn	Val	Ser	Ser	Met	Ile	Ser	Asn	Leu	Ser	Trp	Glu	Leu	Pro	Gly
	710				715									720
Ser	Leu	Pro	Leu	Ile	Ser	Val	Pro	Tyr	Ser	Met	His	Cys	Cys	Thr
	725				730									735
Leu	Gly	Phe	Leu	Ser	Cys	Ser	Leu	Phe	Leu	His	Met	Ser	Phe	Glu
	740				745									750
Leu	Lys	Leu	Leu	Leu	Leu	Leu	Leu	Trp	Leu	Ala	Ala	Ser	Cys	Ser
	755				760									765
Leu	Phe	Leu	His	Ser	His	Ala	Trp	Leu	Ser	Glu	Cys	Leu	Ile	Val
	770				775									780
Arg	Leu	Tyr	Leu	Gly	Pro	Leu	Asp	Ser	Arg	Pro	Gly	Val	Leu	Lys
	785				790									795
Glu	Pro	Lys	Leu	Met	Gly	Ala	Ile	Ser	Phe	Phe	Ile	Phe	Phe	Phe
	800				805									810
Thr	Leu	Leu	Val	Leu	Ala	Arg	Gln	Asn	Glu	Tyr	Tyr	Cys	Arg	Leu
	815				820									825
Asp	Phe	Leu	Trp	Lys	Lys	Lys	Leu	Arg	Gln	Glu	Arg	Glu	Glu	Thr
	830				835									840
Glu	Thr	Met	Glu	Asn	Leu	Thr	Arg	Leu	Leu	Leu	Glu	Asn	Val	Leu
	845				850									855
Pro	Ala	His	Val	Ala	Pro	Gln	Phe	Ile	Gly	Gln	Asn	Arg	Arg	Asn
	860				865									870
Glu	Asp	Leu	Tyr	His	Gln	Ser	Tyr	Glu	Cys	Val	Cys	Val	Leu	Phe
	875				880									885
Ala	Ser	Val	Pro	Asp	Phe	Lys	Glu	Phe	Tyr	Ser	Glu	Ser	Asn	Ile
	890				895									900
Asn	His	Glu	Gly	Leu	Glu	Cys	Leu	Arg	Leu	Leu	Asn	Glu	Ile	Ile
	905				910									915
Ala	Asp	Phe	Asp	Glu	Leu	Leu	Ser	Lys	Pro	Lys	Phe	Ser	Gly	Val
	920				925									930
Glu	Lys	Ile	Lys	Thr	Ile	Gly	Ser	Thr	Tyr	Met	Ala	Ala	Thr	Gly
	935				940									945
Leu	Asn	Ala	Thr	Ser	Gly	Gln	Asp	Ala	Gln	Gln	Asp	Ala	Glu	Arg
	950				955									960
Ser	Cys	Ser	His	Leu	Gly	Thr	Met	Val	Glu	Phe	Ala	Val	Ala	Leu

WO 02/02757

PCT/US01/20491

965 970 975
 Gly Ser Lys Leu Asp Val Ile Asn Lys His Ser Phe Asn Asn Phe
 980 985 990
 Arg Leu Arg Val Gly Leu Asn His Gly Pro Val Val Ala Gly Val
 995 1000 1005
 Ile Gly Ala Gln Lys Pro Gln Tyr Asp Ile Trp Gly Asn Thr Val
 1010 1015 1020
 Asn Val Ala Ser Arg Met Glu Ser Thr Gly Val Leu Gly Lys Ile
 1025 1030 1035
 Gln Val Thr Glu Glu Thr Ala Trp Ala Leu Gln Ser Leu Gly Tyr
 1040 1045 1050
 Thr Cys Tyr Ser Arg Gly Val Ile Lys Val Lys Gly Lys Gly Gln
 1055 1060 1065
 Leu Cys Thr Tyr Phe Leu Asn Thr Asp Leu Thr Arg Thr Gly Pro
 1070 1075 1080
 Pro Ser Ala Thr Leu Gly
 1085

<210> 2

<211> 837

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7478216CD1

<400> 2

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

WO 02/02757

PCT/US01/20491

Gly Gly Gly Ser Ala Ala Ala Ala Ser Ala Gly Gly Thr Glu
 215 220 225
 Val Arg Pro Arg Ser Val Glu Val Gly Leu Glu Glu Arg Arg Gly
 230 235 240
 Lys Gly Arg Ala Ala Asp Glu Leu Glu Ala Gly Ala Val Glu Gly
 245 250 255
 Gly Glu Gly Ser Gly Asp Gly Gly Ser Ser Ala Asp Ser Gly Ser
 260 265 270
 Gly Ala Gly Pro Gly Ala Val Leu Ser Leu Gly Ala Cys Cys Leu
 275 280 285
 Ala Leu Leu Gln Ile Phe Arg Ser Lys Lys Phe Pro Ser Asp Lys
 290 295 300
 Leu Glu Arg Leu Tyr Gln Arg Tyr Phe Phe Arg Leu Asn Gln Ser
 305 310 315
 Ser Leu Thr Met Leu Met Ala Val Leu Val Leu Val Cys Leu Val
 320 325 330
 Met Leu Ala Phe His Ala Ala Arg Pro Pro Leu Gln Leu Pro Tyr
 335 340 345
 Leu Ala Asp His Met Gly Leu Ala Cys Tyr Ala Leu Ile Ala Val
 350 355 360
 Val Leu Ala Val Gln Val Val Gly Leu Leu Leu Pro Gln Pro Arg
 365 370 375
 Ser Ala Ser Glu Gly Ile Trp Trp Thr Val Phe Phe Ile Tyr Thr
 380 385 390
 Ile Tyr Thr Leu Leu Pro Val Arg Met Arg Ala Ala Val Leu Ser
 395 400 405
 Gly Val Leu Leu Ser Ala Leu His Leu Ala Ile Ala Leu Arg Thr
 410 415 420
 Asn Ala Gln Asp Gln Phe Leu Leu Lys Gln Leu Val Ser Asn Val
 425 430 435
 Leu Ile Phe Ser Cys Thr Asn Ile Val Gly Val Cys Thr His Tyr
 440 445 450
 Pro Ala Glu Val Ser Gln Arg Gln Ala Phe Gln Glu Thr Arg Glu
 455 460 465
 Cys Ile Gln Ala Arg Leu His Ser Gln Arg Glu Asn Gln Gln Gln
 470 475 480
 Glu Arg Leu Leu Leu Ser Val Leu Pro Arg His Val Ala Met Glu
 485 490 495
 Met Lys Ala Asp Ile Asn Ala Lys Gln Glu Asp Met Met Phe His
 500 505 510
 Lys Ile Tyr Ile Gln Lys His Asp Asn Val Ser Ile Leu Phe Ala
 515 520 525
 Asp Ile Glu Gly Phe Thr Ser Leu Ala Ser Gln Cys Thr Ala Gln
 530 535 540
 Glu Leu Val Met Thr Leu Asn Glu Leu Phe Ala Arg Phe Asp Lys
 545 550 555
 Leu Ala Ala Glu Asn His Cys Leu Arg Ile Lys Ile Leu Gly Asp
 560 565 570
 Cys Tyr Tyr Cys Val Ser Gly Leu Pro Glu Ala Arg Ala Asp His
 575 580 585
 Ala His Cys Cys Val Glu Met Gly Met Asp Met Ile Glu Ala Ile
 590 595 600
 Ser Leu Val Arg Glu Val Thr Gly Val Asn Val Asn Met Arg Val
 605 610 615
 Gly Ile His Ser Gly Arg Val His Cys Gly Val Leu Gly Leu Arg
 620 625 630

WO 02/02757

PCT/US01/20491

Lys Trp Gln Phe Asp Val Trp Ser Asn Asp Val Thr Leu Ala Asn
 635 640
 His Met Glu Ala Gly Gly Lys Ala Gly Arg Ile His Ile Thr Lys
 650 655
 Ala Thr Leu Asn Tyr Leu Asn Gly Asp Tyr Glu Val Glu Pro Gly
 665 670
 Cys Gly Gly Glu Arg Asn Ala Tyr Leu Lys Glu His Ser Ile Glu
 680 685
 Thr Phe Leu Ile Leu Arg Cys Thr Gln Lys Arg Lys Glu Glu Lys
 695 700
 Ala Met Ile Ala Lys Met Asn Arg Gln Arg Thr Asn Ser Ile Gly
 710 715
 His Asn Pro Pro His Trp Gly Ala Glu Arg Pro Phe Tyr Asn His
 725 730
 Leu Gly Gly Asn Gln Val Ser Lys Glu Met Lys Arg Met Gly Phe
 740 745
 Glu Asp Pro Lys Asp Lys Asn Ala Gln Glu Ser Ala Asn Pro Glu
 755 760
 Asp Glu Val Asp Glu Phe Leu Gly Arg Ala Ile Asp Ala Arg Ser
 770 775
 Ile Asp Arg Leu Arg Ser Glu His Val Arg Lys Phe Leu Leu Thr
 785 790
 Phe Arg Glu Pro Asp Leu Glu Lys Lys Tyr Ser Lys Gln Val Asp
 800 805
 Asp Arg Phe Gly Ala Tyr Val Ala Tyr Ala Ser Leu Val Phe Leu
 815 820
 Phe Ile Cys Phe Val Gln Ile Thr Ile Val Pro Gln
 830 835

<210> 3

<211> 104

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1505146CD1

<400> 3

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

<210> 4

<211> 769

WO 02/02757

PCT/US01/20491

<212> ERT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1577526CD1

```

<400> 4
Met Leu Leu Val Ala Val Leu Ala Gly Ala Leu Gln Leu Thr Gln
 1          5          10          15
Ser Arg Val Leu Cys Cys Leu Pro Cys Lys Val Glu Phe Asp Asn
 20          25          30
His Cys Ala Val Pro Trp Asp Ile Leu Lys Ala Ser Met Asn Thr
 35          40          45
Ser Ser Asn Pro Gly Thr Pro Leu Pro Leu Pro Leu Arg Ile Gln
 50          55          60
Asn Asp Leu His Arg Gln Gln Tyr Ser Tyr Ile Asp Ala Val Cys
 65          70          75
Tyr Glu Lys Gln Leu His Trp Phe Ala Lys Phe Phe Pro Tyr Leu
 80          85          90
Val Leu Leu His Thr Leu Ile Phe Ala Ala Cys Ser Asn Phe Trp
 95          100         105
Leu His Tyr Pro Ser Thr Ser Ser Arg Leu Glu His Phe Val Ala
 110         115         120
Ile Leu His Lys Cys Phe Asp Ser Pro Trp Thr Thr Arg Ala Leu
 125         130         135
Ser Glu Thr Val Ala Glu Gln Ser Val Arg Pro Leu Lys Leu Ser
 140         145         150
Lys Ser Lys Ile Leu Leu Ser Ser Ser Gly Cys Ser Ala Asp Ile
 155         160         165
Asp Ser Gly Lys Gln Ser Leu Pro Tyr Pro Gln Pro Gly Leu Glu
 170         175         180
Ser Ala Gly Ile Glu Ser Pro Thr Ser Ser Val Leu Asp Lys Lys
 185         190         195
Glu Gly Glu Gln Ala Lys Ala Ile Phe Glu Lys Val Lys Arg Phe
 200         205         210
Arg Met His Val Glu Gln Lys Asp Ile Ile Tyr Arg Val Tyr Leu
 215         220         225
Lys Gln Ile Ile Val Lys Val Ile Leu Phe Val Leu Ile Ile Thr
 230         235         240
Tyr Val Pro Tyr Phe Leu Thr His Ile Thr Leu Glu Ile Asp Cys
 245         250         255
Ser Val Asp Val Gln Ala Phe Thr Gly Tyr Lys Arg Tyr Gln Cys
 260         265         270
Val Tyr Ser Leu Ala Glu Ile Phe Lys Val Leu Ala Ser Phe Tyr
 275         280         285
Val Ile Leu Val Ile Leu Tyr Gly Leu Thr Ser Ser Tyr Ser Leu
 290         295         300
Trp Trp Met Leu Arg Ser Ser Leu Lys Gln Tyr Ser Phe Glu Ala
 305         310         315
Leu Arg Glu Lys Ser Asn Tyr Ser Asp Ile Pro Asp Val Lys Asn
 320         325         330
Asp Phe Ala Phe Ile Leu His Leu Ala Asp Gln Tyr Asp Pro Leu
 335         340         345
Tyr Ser Lys Arg Phe Ser Ile Phe Leu Ser Glu Val Ser Glu Asn
 350         355         360

```

WO 02/02757

PCT/US01/20491

Lys Leu Lys Gln Ile Asn Leu Asn Asn Glu Trp Thr Val Glu Lys 365 370 375
 Leu Lys Ser Lys Leu Val Lys Asn Ala Gln Asp Lys Ile Glu Leu 380 385 390
 His Leu Phe Met Leu Asn Gly Leu Pro Asp Asn Val Phe Glu Leu 395 400 405
 Thr Glu Met Glu Val Leu Ser Leu Glu Leu Ile Pro Glu Val Lys 410 415 420
 Leu Pro Ser Ala Val Ser Gln Leu Val Asn Leu Lys Glu Leu Arg 425 430 435
 Val Tyr His Ser Ser Leu Val Val Asp His Pro Ala Leu Ala Phe 440 445 450
 Leu Glu Glu Asn Leu Lys Ile Leu Arg Leu Lys Phe Thr Glu Met 455 460 465
 Gly Lys Ile Pro Arg Trp Val Phe His Leu Lys Asn Leu Lys Glu 470 475 480
 Leu Tyr Leu Ser Gly Cys Val Leu Pro Glu Gln Leu Ser Thr Met 485 490 495
 Gln Leu Glu Gly Phe Gln Asp Leu Lys Asn Leu Arg Thr Leu Tyr 500 505 510
 Leu Lys Ser Ser Leu Ser Arg Ile Pro Gln Val Val Thr Asp Leu 515 520 525
 Leu Pro Ser Leu Gln Lys Leu Ser Leu Asp Asn Glu Gly Ser Lys 530 535 540
 Leu Val Val Leu Asn Asn Leu Lys Lys Met Val Asn Leu Lys Ser 545 550 555
 Leu Glu Leu Ile Ser Cys Asp Leu Glu Arg Ile Pro His Ser Ile 560 565 570
 Phe Ser Leu Asn Asn Leu His Glu Leu Asp Leu Arg Glu Asn Asn 575 580 585
 Leu Lys Thr Val Glu Glu Ile Ile Ser Phe Gln His Leu Gln Asn 590 595 600
 Leu Ser Cys Leu Lys Leu Trp His Asn Asn Ile Ala Tyr Ile Pro 605 610 615
 Ala Gln Ile Gly Ala Leu Ser Asn Leu Glu Gln Leu Ser Leu Asp 620 625 630
 His Asn Asn Ile Glu Asn Leu Pro Leu Gln Leu Phe Leu Cys Thr 635 640 645
 Lys Leu His Tyr Leu Asp Leu Ser Tyr Asn His Leu Thr Phe Ile 650 655 660
 Pro Glu Glu Ile Gln Tyr Leu Ser Asn Leu Gln Tyr Phe Ala Val 665 670 675
 Thr Asn Asn Asn Ile Glu Met Leu Pro Asp Gly Leu Phe Gln Cys 680 685 690
 Lys Lys Leu Gln Cys Leu Leu Leu Gly Lys Asn Ser Leu Met Asn 695 700 705
 Leu Ser Pro His Val Gly Glu Leu Ser Asn Leu Thr His Leu Glu 710 715 720
 Leu Ile Gly Asn Tyr Leu Glu Thr Leu Pro Pro Glu Leu Glu Gly 725 730 735
 Cys Gln Ser Leu Lys Arg Asn Cys Leu Ile Val Glu Glu Asn Leu 740 745 750
 Leu Asn Thr Leu Pro Leu Pro Val Thr Glu Arg Leu Gln Thr Cys 755 760 765
 Leu Asp Lys Cys

WO 02/02757

PCT/US01/20491

<210> 5
<211> 1119
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7478215CD1

```

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

```

WO 02/02757

PCT/US01/20491

Arg Ile Lys Ile Leu Gly Asp Cys Tyr Tyr Cys Val Ser Gly Leu
 350 355 360
 Thr Gln Pro Lys Thr Asp His Ala His Cys Cys Val Glu Met Gly
 365 370 375
 Leu Asp Met Ile Asp Thr Ile Thr Ser Val Ala Glu Ala Thr Glu
 380 385 390
 Val Asp Leu Asn Met Arg Val Gly Leu His Thr Gly Arg Val Leu
 395 400 405
 Cys Gly Val Leu Gly Leu Arg Lys Trp Gln Tyr Asp Val Trp Ser
 410 415 420
 Asn Asp Val Thr Leu Ala Asn Val Met Glu Ala Ala Gly Leu Pro
 425 430 435
 Gly Lys Val His Ile Thr Lys Thr Thr Leu Ala Cys Leu Asn Gly
 440 445 450
 Asp Tyr Glu Val Glu Pro Gly Tyr Gly His Glu Arg Asn Ser Phe
 455 460 465
 Leu Lys Thr His Asn Ile Glu Thr Phe Phe Ile Val Pro Ser His
 470 475 480
 Arg Arg Lys Ile Phe Pro Gly Leu Ile Leu Ser Asp Ile Lys Pro
 485 490 495
 Ala Lys Arg Met Lys Phe Lys Thr Val Cys Tyr Leu Leu Val Gln
 500 505 510
 Leu Met His Cys Arg Lys Met Phe Lys Ala Glu Ile Pro Phe Ser
 515 520 525
 Asn Val Met Thr Cys Glu Asp Asp Asp Lys Arg Arg Ala Leu Arg
 530 535 540
 Thr Ala Ser Glu Lys Leu Arg Asn Arg Ser Ser Phe Ser Thr Asn
 545 550 555
 Val Val Tyr Thr Thr Pro Gly Thr Arg Val Asn Arg Tyr Ile Ser
 560 565 570
 Arg Leu Leu Glu Ala Arg Gln Thr Glu Leu Glu Met Ala Asp Leu
 575 580 585
 Asn Phe Phe Thr Leu Lys Tyr Lys His Val Glu Arg Glu Gln Lys
 590 595 600
 Tyr His Gln Leu Gln Asp Glu Tyr Phe Thr Ser Ala Val Val Leu
 605 610 615
 Thr Leu Ile Leu Ala Ala Leu Phe Gly Leu Val Tyr Leu Leu Ile
 620 625 630
 Phe Pro Gln Ser Val Val Val Leu Leu Leu Leu Val Phe Cys Ile
 635 640 645
 Cys Phe Leu Val Ala Cys Val Leu Tyr Leu His Ile Thr Arg Val
 650 655 660
 Gln Cys Phe Pro Gly Cys Leu Thr Ile Gln Ile Arg Thr Val Leu
 665 670 675
 Cys Ile Phe Ile Val Val Leu Ile Tyr Ser Val Ala Gln Gly Cys
 680 685 690
 Val Val Gly Cys Leu Pro Trp Ala Trp Ser Ser Lys Pro Asn Ser
 695 700 705
 Ser Leu Val Val Leu Ser Ser Gly Gly Gln Arg Thr Ala Leu Pro
 710 715 720
 Thr Leu Pro Cys Glu Ser Thr His His Ala Leu Leu Cys Cys Leu
 725 730 735
 Val Gly Thr Leu Pro Leu Ala Ile Phe Phe Arg Val Ser Ser Leu
 740 745 750
 Pro Lys Met Ile Leu Leu Ser Gly Leu Thr Thr Ser Tyr Ile Leu
 755 760 765

WO 02/02757

PCT/US01/20491

Val Leu Glu Leu Ser Gly Tyr Thr Arg Thr Gly Gly Gly Ala Val
 770 775 780
 Ser Gly Arg Ser Tyr Glu Pro Ile Val Ala Ile Leu Leu Phe Ser
 785 790 795
 Cys Ala Leu Ala Leu His Ala Arg Gln Val Asp Ile Arg Leu Arg
 800 805 810
 Leu Asp Tyr Leu Trp Ala Ala Gln Ala Glu Glu Glu Arg Glu Asp
 815 820 825
 Met Glu Lys Val Lys Leu Asp Asn Arg Arg Ile Leu Phe Asn Leu
 830 835 840
 Leu Pro Ala His Val Ala Gln His Phe Leu Met Ser Asn Pro Arg
 845 850 855
 Asn Met Asp Leu Tyr Tyr Gln Ser Tyr Ser Gln Val Gly Val Met
 860 865 870
 Phe Ala Ser Ile Pro Asn Phe Asn Asp Phe Tyr Ile Glu Leu Asp
 875 880 885
 Gly Asn Asn Met Gly Val Glu Cys Leu Arg Leu Leu Asn Glu Ile
 890 895 900
 Ile Ala Asp Phe Asp Glu Leu Met Glu Lys Asp Phe Tyr Lys Asp
 905 910 915
 Ile Glu Lys Ile Lys Thr Ile Gly Ser Thr Tyr Met Ala Ala Val
 920 925 930
 Gly Leu Ala Pro Thr Ser Gly Thr Lys Ala Lys Lys Ser Ile Ser
 935 940 945
 Ser His Leu Ser Thr Leu Ala Asp Phe Ala Ile Glu Met Phe Asp
 950 955 960
 Val Leu Asp Glu Ile Asn Tyr Gln Ser Tyr Asn Asp Phe Val Leu
 965 970 975
 Arg Val Gly Ile Asn Val Gly Pro Val Val Ala Gly Val Ile Gly
 980 985 990
 Ala Arg Arg Pro Gln Tyr Asp Ile Trp Gly Asn Thr Val Asn Val
 995 1000 1005
 Ala Ser Arg Met Asp Ser Thr Gly Val Gln Gly Arg Ile Gln Val
 1010 1015 1020
 Thr Glu Glu Val His Arg Leu Leu Arg Arg Cys Pro Tyr His Phe
 1025 1030 1035
 Val Cys Arg Gly Lys Val Ser Val Lys Gly Lys Gly Glu Met Leu
 1040 1045 1050
 Thr Tyr Phe Leu Glu Gly Arg Thr Asp Gly Asn Gly Ser Gln Ile
 1055 1060 1065
 Arg Ser Leu Gly Leu Asp Arg Lys Met Cys Pro Phe Gly Arg Ala
 1070 1075 1080
 Gly Leu Gln Gly Arg Arg Pro Pro Val Cys Pro Met Pro Gly Val
 1085 1090 1095
 Ser Val Arg Ala Gly Leu Pro Pro His Ser Pro Gly Gln Tyr Leu
 1100 1105 1110
 Pro Ser Ala Ala Ala Gly Lys Glu Ala
 1115

<210> 6
 <211> 3769
 <212> DNA
 <213> Homo sapiens

 <220>
 <221> misc_feature

WO 02/02757

PCT/US01/20491

<223> Incyte ID No: 3685652CB1

```

<400> 6
ggcgagcgtg actccgccat caggtccccc gtcctctccc cggacctagc ccactccgct 60
ggccacagc cgcggcagc gaggtcgggg atcaggggaa cccggggctc cctcgacca 120
accocaaatc ctgccctect ggcgatgagg cttttctagg gcaoctctac aatccgggtt 180
tgaggcaggg gaggaaagg actgagggat ccctcatcg ccaactggga gcgggctggg 240
aggcccgagg gaggccctga aaaaggagac gggattgcca cgaagttggg ggcgcgggg 300
ggtagcggct ttgagcgggt gagaaaagct caggtggggc ccgocgggoc gaaggaggta 360
accocggcgc cggccctagc cagcccccgg gctcggggct gggagatca tggcccgct 420
cttcagcccc cggccgcccc ccagogaaga cctctctac gagactact acagcctgag 480
ccagcagtac ccctctctgc tgcctctgct ggggatcgtg ctctctgctc tcggcggct 540
gctcgcagtc gectgggcca gggcaggga gctgacctca gaccagagct tcctgaccac 600
tgtctctgc gctcggggc gctctctgct gctgctgggc ctgcttccc gggagcagcg 660
actgcaagc tggacgcgtc ccctctcggc ctgggtatgg gtcgagctgc tagcctagg 720
ccacgcttc ctgttaccg gggcctgggt gagcctgct gaccaggtgt cctattttct 780
cttcctacc ttccagcgt atgcccctgt gccctgggc atgcgggacg ccgcccctgc 840
ggcctcgc tcctactct cgcactctgt gtcctcggg ctgtacttg ggcacacagc 900
ggactcacgg cctgcaactg tgcgcagtt ggcagcaaac gcaagctgtt tcctgtggc 960
gnactggcc cgcaggggcc gctccatcag gccttggg tacactcctg ccactgtcc 1020
ggagcagct agctccctgc actcaccgcg gcggctggac accgagaaga agcaccagga 1080
acactttct ttgtccatcc ttctcctca cctggccgga gagatgaagg cagagatcat 1140
ggcaggctg caggcaggac aggggtcacg gccagagagc actaacaatt tccacagcct 1200
ctatgtcaag agccaccagg gactcagcgt gctgtatgct gacatcgtgg gcttaccgct 1260
gttgccagc gactgttccc ctaaggagct ggtgtcctg ctcaatgagc tctttggcaa 1320
gttcyaccag actgccaagg ttctctctga caccacaba cacaaggagc atgaaatgca 1380
gggatcaag atctggggg actgttacta ctgtgtctct gggctggccc tctcactgcc 1440
agaccatgcc atcaactgag tgcgatagg cctggacatg tgcggggcca tccaggaaact 1500
ggggcagcc actggcgtgg acatcaacat gctgtggggc gtgcactcag gcaagctact 1560
gtgtggagtc atcggcctgc agaagtggca gtaccagctt tggtcacatg atgtccact 1620
ggtaaccac atggaggcag ggggtgtacc agggcaggtt cacatcacag gggctaccct 1680
ggcctcgtg gcaggggctt atgctgtgga ggaagcaggg atggagcacc gggaccocca 1740
ccttcggggc ctgggggagc ctacctatct ggtcatcgat ccaagggcag aggagaggga 1800
tgagaagggc actgcaggag gcttctctgc ctgcttgag ggcctcaaga tgcgtccatc 1860
actgctgat acccgttacc tggagctctg gggcgcagcc aagccttttg cccaactgag 1920
ccacgggagc agccctgtgt ccacctccac ccctctcccg gagaagacc tggcttctt 1980
cagcaccagc tggagcctgg atcggagccc taccccccgg ggaactagat atgaactgga 2040
caccgggat gccaaattct tccaggtcat tgagcagctc aactcgaga aacagtggaa 2100
gtaccgact gacttcaacc cactgacact gtacttcaga gagaaggaga tggagaaga 2160
gtaccgact tctgcaatcc ccgcttcaa atactatgaa gctgcacct tcctggtttt 2220
tcttccaac ttcatcatcc agatcctagt gacaaacagg cccccagctc tggccatcac 2280
gtatagcatc accitctctc tcttctctct catccttttt gctctgtct cagaggacct 2340
gatgaggtgt gtcctgaaag gccccaagat gctgcactgg ctgctgcaac tgcctggcct 2400
ggtggccaca cagaccagac tgagaatagc cttgggcaec gccaccatcc tcttctctt 2460
tgccatggcc attaccagcc tgttctcttt cccaacatca tcagactgcc ctttccaagc 2520
tcccaatgct tctccatga ttccaacct ctctcgggag ctccctgggt ctctgctct 2580
catcagctc ccatctcca tgaactgctg cagctggggc ttctctctct gctccctct 2640
tctgcaatg agcttcgagc tgaagctgct gctgctctg ctgtggctgg cggcatctg 2700
ctccctctc ctgcactccc atgctggct gtcggaatgc ctcatctcc gctctatct 2760
ggccctctg gactccagc ccggagtgct gaaggagccc aaactgaggg gtgctatct 2820
ctctctatc ttctcttca cctcctgtt cctggctcgc cagaatgag actactgccc 2880
cctgacttc ctgtggaaga agaagctgag gcaggagagg gaggagacag agcagatga 2940
gaactgact cggctgctct tggagaact gctccctgca cactggccc ccaagttcat 3000
tggcagaac cggcgaagc aggatctcta ccaccagtc tatgaatgcy tttgtgtct 3060
cttcctca gtcccagct tcaaggagtt ctactctgaa tccaacatca atcatgagg 3120
cctagaggt ctgaggctgc tcaatgagat aattgctgat tttgatgagc tgctctccaa 3180

```

WO 02/02757

PCT/US01/20491

```

gcccagttc agtggggtg agaagatcaa gaccatcggc agcacctaca tggcagccac 3240
aggcttaaat gccacctctg gacaggatgc acaacaggat gctgaecgga gctgcagcca 3300
ccttggcact atggttgaat ttgcctggc cctgggtctc aagctggacg tcatcaacaa 3360
gcattcaltc aacaacttcc gectgcgagt ggggttgaac catggaccgg tagtagctgg 3420
agttatgggg gcccaagaagc cgaatataga catttggggc aacacagtga acgtggccag 3480
ccgataggag agtacaggag tctctggcaa aatccaagtg actgaggaga cagcatgggc 3540
cctacagtcc ctgggtcaca cctgtctacg ccgggggtgc atcaagtgga aaggcaagg 3600
gcagctctgc acctacttcc tgaacacaga cttgacacga actggaacct cttcagctac 3660
cctaggtgta gatgcactc gccttctaag aacctcaata aagagacctt ggggtgtctg 3720
ggccccaaaa aaaaaaaaaa aaaaaaaaaa aagatgcggc gcaagctta 3769

```

<210> 7

<211> 3137

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7478216CB1

<400> 7

```

atgcccggc ggcggcagg ctggggccat ggccaggccg gggcgaggc cggcctcgg 60
eggctgacgg ggggccaag atgcccgcct gcagcgacgg gccccggcc gccccgacg 120
tgtgacctca gctgggtccc cctgctggc cgtccgccct ccccttggag acccccgcc 180
cggcttccgg gggaggagga aggagacgac gaggccgagg ggggatgtc cggctccaaa 240
agcgtgagcc ccccgggcta cggcgccgac aagactcggg cggccggccc cggggaggg 300
cccgaacacc gctctgctg gggcgaggcc gattcccgcg cgaatggcta ccccagtc 360
cccgggggct ctgcccggg ctccaccaag aaaccgggg gggcggtag cccgagcag 420
cagcagcccc tggccagccg ctggcgccgc gacgaagac acgatctcc cctgagcgg 480
gaacaccccc tggcccgggg ctcccgctc agcttccgct ccaagtccg ctggagggg 540
cggcgccggc acgactcggg tggcgccgc cgcggccgc gggggggcg gccccgggg 600
ggcagcacc gggcgccccc tggcgccgc ggcggcgct cggcgccgg cgtgctcgg 660
ggggcgagg cggaggtgct cctcgcctc gtggaggtag gctctggaga gggcggggg 720
aagggcgcg cggccacga gctggagccc ggcgcctcg agggcgcaa ggggtccgg 780
gatggcgcca gctcggcggg ctggggctcg ggcgggggg cggcgcggt gctgtcccty 840
ggcgcctgct gctcggcgtt gctgcagata ttccgctcca agaagttccc gtcggacaaa 900
ctggagcgcc tgtaccagcg ctactcttc cgcctgaacc agagcagcct caccatgctc 960
atgctgtgct tgggtgctgt gtgctggtc atgttggcct tccacgccc gggcccccg 1020
ctccagctgc cctacctgct cgaaccatg ggcctggcct gctatgcgct catcgccgtg 1080
gtcctggcgg tccaggtggt gggcctgctg ctgcccgcgc caegcagcgc ctctgagggc 1140
atcctgtgga cctgttctct catctacacc atctacacgc tgcctccgct ggcgatcggg 1200
gcccagctgc tccggggggt gctcctgtcc gccctccacc tggccatcgc cctggcacc 1260
aaegcccagg accagttcct gctgaagcag ctgtctcca atgttctcat ttctcctgc 1320
accaacatcg tgggtgtctg caccacctat cggctgagg tctccagag acaggcttcc 1380
caggagacc gagagtcgat ccaggcgcg ctccactcgc agcgggagaa ccagcagcag 1440
gaacggctcc tgetgtctgt ccttccccgt catgttgcga tggagatgaa agcagacatc 1500
aacgcgaagc agggagatg gatgttccat aagatttaca tccagaacaa tgacaacgtg 1560
agcatcctgt ttgctgacat cgaaggcttc accagcctgg cgtcccagtg cactgcacag 1620
gaactgttca tgacctcaa cgaactcttc gccgccttgg acaagctggc cgcagagaat 1680
cactgtttac gtattaagat ccttggggat tgttattact gcgctcggg gctgcctgaa 1740
gcaagggtctg accacgccc ctgctgtgtg gagatgggca tggacatgat cggagccatc 1800
tcggttgcct gggaggtgac aggggtgaac gtgaacatgc gtgtgggaat tccagcggg 1860
cgaatcact cgggtgtcct tggctcagg aagtggcagt tgcagctctg gctcaacgat 1920
gtcagctag ccaaccacat ggaagctggc ggcaaggcag gacgatcca catcacaag 1980
gtacactca actacctgaa tggggaetac gagggtggag caggctgtg gggcgagcg 2040
aacgctacc tcaaggagca cagtatcgag accttctca tctgtcctg caccagaag 2100

```

WO 02/02757

PCT/US01/20491

```

cggaaagaag agaagggcat gatcgccaag atgaaccgcc agagaaccac cccatcggg 2160
cacaaccacc cacactgggg gctgagcgc ccttctaca accactcggg tggcaaccag 2220
gtgtccaagg agatgaagcg gatgggttt gaagaccoca aggaacaaga cgcaccaggag 2280
agtgcgaacc ctgaggatga agtggatgag ttctcggccc gtgccaltga cgcaccaggag 2340
attgatagcc ttccgtctga gcaagtcgc aagttctccc tgacctcag ggagcctgac 2400
ttagagaaga agtactcaca gcaggtagac gaccgatgg gtgcctatgt ggctgatgcc 2460
tcgctcgtct tctcttcat ctgctttgtc cagatcacca tctgcoccca gtgagtatcc 2520
ccgtccctct tgggctctc cctctggctg cagggaatcc caggagtggg agggggtggt 2580
catttaggag aacagagcag ggagtctggc ctgcacttgc tcaagcaccg agagaagttt 2640
ctgagtgtgc ccttgggtgg gtgagaagcc aggggtttga cctctctggg cgggaaggtg 2700
atgtgtcctat tccaggtgg gcaactctca ctaggcctg cacagctgca tgaagatgc 2760
ttcttttttt ttttttggg gtcggctgca gggggacaca gcgtctcgt cctgggtggc 2820
ccaggctggg atgtgcagag gctatgatct gggctcactg gcaatttctg gcctcccaag 2880
ttcaggcaat tctccgctc cagcctcctg agtagctggg agaatacagg gtggcctacc 2940
accacgcccg gctaatttt ttggtattta agtagagac gggtttacc atgttgccca 3000
gggctgggtc tcaaacctcc aagctcaggc aatccaccgg cctgggcctc ccaaatgtct 3060
ggggattaca ggcgtgagcc accaggcctg gtcaagagac gccttttata tcaactaggg 3120
ctgcaggaat cggatat 3137

```

```

<210> 8
<211> 1958
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1505146CBI

```

```

<400> 8
agacgaagat agaagcgcaa ggcgagctca cgcgtgccc cgggttatag aactagtga 60
tccccacct ggtgtcttcc agcttaagga ccgaggcctg accgaggtga gcactggag 120
ggacacgacc aggaacttaa agggagaggg acggcggggg cggcaactt agtccgtgcg 180
ggattcccga cgtgtgaaac tcttctctcc agtccccgc agacttgcaag aagctgacga 240
gcaactctcag gaccatcgac ttgtccaaca acaagatcga aagcctaccg cctttgtcga 300
taggaaagt cactctgctg aagagcctct ccttgaacaa caacaaactg agtatggctc 360
tcgccaggga ggccttggct agtgcacgc tgttaccagg ggtggttttc tcttacaagc 420
tgatttttgt tttgttatg aaaccgtgca gggggctca agttgcaat gcgggaatcc 480
ttgaagcctt ctttctactt cttgtaatta cabctaaat aagccttata atgattgggt 540
gttataatga ttgtactggt gtttaattgg acttgaaga gaagaacait ttgggatgt 600
aatgtaaca aagctacctc acttggatg ggggatggat gagaagggga ggaatctcca 660
aatatcaaca aatgtctgtc ggaattgtgc ttgggcccct ttacaaaagg agcccatat 720
acacacaat gctctccggc attttgatgg cacctcacc ttctaaagac tatgttggta 780
aaatgcaata tcaaaaatg gcaatgtgg gagttttaca tataatcaac agagggcgaag 840
caggctgcca tatagtcttt ataaacttca tcaagttgt tttcagctct atcaactca 900
ggtttataaa aaatttttct ttgcagctg ttctgctgca tgagatagtc aatctgaaaa 960
aactagagac gctaagccta aacaacaatc acctagaga gctgcccgtc accttgggc 1020
aactctctgc cctcaagacc ctgagctct ctgggaacca actgggagca ttacctcccc 1080
aactttgtag cctacggcac ctggatgtga tggactctc taagaaccag altcgaagta 1140
tacctgactc agtggagag ctgcaagta tcaactcaa cctcaaccag aatcagat 1200
ctcagatctc agtgaagata tcttctgtc cacgccttaa aattcttgc ctggaagaga 1260
attgtcttga gctcagcatg cttccocaga gcatcctcag tgattccag atctgtctgc 1320
ttctgttggg aggcaactct tttgaataa agaacttgc agaactggaa ggctatgata 1380
agtacatgga ggggttcaca gccaccaga agaagtttc gtgaagttct ccaggactat 1440
ggaacotta caggatacty acttagaac tctgttggaa tctggctgag tcaaacctc 1500
ctgtgtgtg taggggtatc tacagtaagg agatgatac tcaggagatt atatttcaat 1560
caatgatctt ttctcattc agggctcttc tcaataaagc taaaagaaaa aggatcagga 1620

```

WO 02/02757

PCT/US01/20491

```

gacaggaaaa gtcttcggtt ttgagtcatt agtagggcaa tagacaaggt tetcttcaaa 1680
accatcatta gtttggcttt aagaaaccag tagctagctg ctatttataat ggtgaggggg 1740
tgctgcctgg taacagaata gctccacacc acagcttgag attttgttta gtttcaactg 1800
gtgagctttc ataaagcttg ttgccattcc atctcttgtt taacacttca tattttatag 1860
aaattcagat aattgtgaga ggctggcatg gatctagggg tcataatcca tccagtcacc 1920
agtcagcgcg gttttcttag ctttggagtc ccacgggtg 1958

```

<210> 9

<211> 2660

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1577526CB1

<400> 9

```

atgctgctgg tggccgtgct ggccggagct ctccagctga cgcagagcag ggttctgtgc 60
tgtcttccat gcaaaagtga atttgacaat cacgtgcccg tgccctggga catcctgaaa 120
gccagcatga acacatcttc taatctgggg acaccgcttc cgtctcccct ccgaaltcag 180
aatgacctcc accgacagca gtaactctat attgatcccg tctgttaaga gaaacagctc 240
cattggtttg caaagttttt cctctatctg gtgctcttgc acacgctcat ctttgcagcc 300
tgcagcaact ttggcttca ctaccaccgt accagttcca ggcctgagca ttttggccc 360
atccttcaca agtgcctega ttctccatgg accaccgcy ccttttcaga aaacgtggct 420
ggcagctcag tgaggcctct gaaactctcc aagtccaaga ttttgcttc gtccctcagg 480
tgttccagct acatagattc cggcaaacag tcattgccct accccagacc aggtttggag 540
tcagctggca tagaaagccc aacttccagt gtccctggca agaagggagg tgaacagccc 600
aaagccatct ttgaaaaagt gaaaagattc cgcattgcat tggagcagaa ggcatacatt 660
tatagagtat atctgaaaca gataatagtc aaagtcaatt tgtttgtgct catcataact 720
tatgttccat attttttaac ccacatcact cttgaaatcg actgttcaat tgatgtgca 780
gcttttacag gatataagcg ctaccagctg gtctattcct tggcagaaat ctttaagctc 840
ctggcttcat tttatgtcat tttggttata ctttatggtc tgacctcttc ctacagcctg 900
tgggtgatgc tggaggattc cctgaagcaa tattctcttg aggccttaag agaaaaaagc 960
aactacagtg acatccctga tgtcaagaat gaacttgcct tcaccttca tctgctgat 1020
cagtatgctc ctctttatcc caaacgcttc tccatattcc tatcagaggt cagtgagaac 1080
aaactgaaac agatcaacct caataatgaa tggacagttg agaaactgaa aagttaagctt 1140
gtgaaaaagt ccagagcaaa gatagaactg catcctttta tgcctcaagg tcttccagac 1200
aatgtctttg agttaactga aalggaaagt ctaagcctgg agcttatccc agaggtgag 1260
ctgctcctct cagtctcaca gctgctcaac ctcaaggagc ttctgtgtga ccatcactct 1320
ctggtctgag accatcctgc actggccttt ctgagggaga atttaaaaat cctccgctg 1380
aaattactg aaatgggaaa aatcccagc tgggtatttc acctcaagaa tctcaaggaa 1440
ctttatcttt cgggctgtgt tetcccgtga cagttgagta ctatgcagtt ggagggcttt 1500
caggacttaa aaaaataag gaccctgac ttgaaagca gcctctccc gatcccaaaa 1560
gtgtttacag acctcctgcc ttcaattgag aaactgtccc ttgataatga gggaaacaaa 1620
ctggttgtgt tgaacnactt gaaaagatg gtcaacttga aaagcctaga actgatcagc 1680
tfgagactgg aacgcactcc acattccatt ttcagcctga ataatttga tgagttagac 1740
ctaagggaaa atcaacctaa aactgtgaaa gagatcatta gctttcagca tcttcagaat 1800
ctttctgct taaagtgtg gcacaataac attgcttata ttcttgcaca gattggggca 1860
ttactaaccc tagscagct ctctttggac cataataata ttgagaactc gcccttgcag 1920
ctttctctat gcactaaact acattatttg gatcctagct ataaccactt gaaccttact 1980
ccagaagaaa tccagtatct gagtaatttg cagtactttg ctgtgaccaa caacaatatt 2040
gagatgctac cagatgggct gtttcagtc aaaaagctgc agtgtttact ttgggggaaa 2100
aatagcttga tgaatttgc cctccatgtg ggtgagctgt caaaccttac tcatctggag 2160
ctcattggta attacctgga aacacttctc cctgaactag aaggatgca gtccctaaaa 2220
cggaaactgc tgattgttga ggagaacttg ctcaactact tctctctccc tgaacagaa 2280
cgtttacaga cgtgctttag caaatgttga cttaagaaa agagaccctg gtttcaaaat 2340

```

WO 02/02757

PCT/US01/20491

```

catttttaaa agtatgctcg gccggcggtg gtggctcagc cctataatcc cagcaactttg 2400
ggggcccaag atggcgggat tgcttgaggt caggagttcg agaccagtct ggccaacctg 2460
gtgaaaccoc atctctgcta aaactacaaa aaaattagcc aggcgtggtg gcgtgcccct 2520
gtaatcccag ctacttggga ggctgagca ggggaattgc ttgaaccagg gaggtggagg 2580
ttgcagtagc ccgagattgt gccactgtac accagcctgg gtgacagagc aagactotta 2640
tctcaaaaaa aaaaaaaaaa 2660

```

```

<210> 10
<211> 3811
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7478215C81

```

```

<400> 10
gcattggcgt gagatggcgg gggcgccggc cggcgaggcg ggcggcgagg gcggcgcggt 60
cgagcccggg ggcgcccagc gggcgccggc gacaagccgc cggcgcgagg tccggggcgt 120
cgacgaggag ttccgcttgc cagagctgga ggctgtgttc cggcgctaca cgtcggcgct 180
ggagcaggcg gccacgtgga aggcgctggc cgttctcagc ctgctggcgg gcgctcgtgc 240
gctggccgag ctgctggcgg cgcggggggc cgcggccggc ctggccaagg gctcacaccc 300
ggtgcaactgc gtcctcttcc tggcgctgct cgtgttaacc aacgtccggc cctgcaggt 360
gccccagctg cagcagctgc gccagctggc gctgctcttc agcctcaact tccgctgct 420
ctgctgtctc ttccgctggc gggcccccgc cgggggttcc gccggggccc ctggggggcc 480
agcgaccgcc gaacaagggg ttggcagctc ccttttgctc acctctgctg cctatgctt 540
gctgccctgc cgcagcctgc ttggccatagg ctttggctc gtgtggctg cgtcgcact 600
gctgttcaca gccaccttgg tcccgccaaa gcgcccaact ctctggaggc cgtcggctg 660
caatgccttg etctctgtgc gtgtgaacat gtabggggtc ttgtgogga ttctgactga 720
gcgttcaacg aggaagcgtt tctcagggc cgggagctgc attgaggacc gactgaggct 780
ggaggatgag aacgagaagc aggagcggct cctcagagc ctctgcccc ggaacgttgc 840
cattgagatg aaggagactc tctgaagcc cctgagaggc attttccaca agatttcat 900
ccagagcac gacaatgtga gcatcctgtt tgcagacac gtgggttca cgggcttggc 960
atcccagctc accgccagc agctgtgtaa actcctcaat gagctcttg gcaagttcga 1020
tgaattagcc accggaaccc actgtccgc cacaagatt ctggggact gctactactg 1080
cgtgtcggcg ctaccaccgc ccagactga ccatgcccac tgcctgttgg agatgggact 1140
cgacatgatt gataccatca catctgtggc tgaagccacc gagggtgale tgaacatgcy 1200
tgtgggtctg cacacgggca gggctcctcg tgggtcctg ggccttgoca agtggcagta 1260
cgactgtggt tccaatgatg tgaccttggc caatgtcatg gaagccgctg gcttgcagg 1320
gaagttcat atcacaaaga ccacccctagc gtgctgaaat ggggactacy aggtagaacc 1380
gggttaccga catgagagga acagtttctt gaaaaactat aacatcgaaa cctttttat 1440
tgttccatcc catgcccga agatatttcc aggcctgatt ctctcagata taaaaccggc 1500
caaaaggatg aagtccaaga ctgtctgcta cctgctgggt cagctcatgc actgcccga 1560
aatgttcaag gccgagatcc ccttctccaa tgtcatgacc tgcgaggacy atgacaagcg 1620
gagggcatta agaacagcct cggaaaaact cagaaacccg tcatcttttt ctaccaacgt 1680
tgtctacacc accccgggca ctgcgctcaa caggtacatc agccgcctct tagaagcccg 1740
ccagacagag ctggagatgg cagacctgaa cttctttacc ctgaaagta aacatgtcga 1800
acgggagcaa aagtaccacc agcttcagga cagatattc accagcggcg ttgtcctcac 1860
ctcactcctg gctgccttat ttgcccctgt ctaccttota ataltccac agagtgtggt 1920
cgtcctgctc ctgctagtat tctgcatctg cttcctgggt gcctgtgtcc tgtactctga 1980
catcaccogg gtccaggttt ttccagggct cctgacgatt cagattcgca cgtcctctgt 2040
tattttcata gtggtcttaa tctactcagt agcccaaggt tgtgtgtggc gctgctgct 2100
ttggcctggc agtccaagc ccaacagctc cctgggtgctc cttctgtctg gggccagcg 2160
caccagcctg cccaccctgc cctgcagctc tacacaccat gccctgctct gctgcttgt 2220
gggccaccctc ccgctagcca tatttttccg ggtgtcctcc ttgccaaaaa tgcctctgt 2280
ctccggctc accacgtctc acatctcgt tctggagctc agcggataca ccaggactgg 2340

```

WO 02/02757

PCT/US01/20491

```

gggtggtgcc gtctcggggc gcagctacga gccgattgtg gccatcctgc tcttctectg 2400
tgcgctggcc ctgcatgcca ggcaggtgga catcaggetg aggetggact acctctgggc 2460
cgcaacaggc gaggaggagc gagagacat ggagaaggtg aagctggaca acaggcgcat 2520
cctctcaac ctctgcccgg cccagctgc ccagcacttc ctcattgcca accctggaa 2580
catggacctc tactaacagt cctactccca ggtgggcgtc atgtttgct ccatccca 2640
cttcaatgac ttctacatcg agctggcgg caacaacatg ggggtggagt gtctgcygt 2700
tctcaacag atcatcgcc actttgacya gctcatggaa aaagactttt acaaggacat 2760
agagaagatc aagaccatcg ggagcaccla catggccgtc gtggggctag cggccacctc 2820
ggggaccaag gctaaagaagt ccatctcttc ccacctgagc acgctggcgg actttgcat 2880
tgagatgttt gactttctgg atgaaatcaa ctaccagtct tacaacgact ttgtctccg 2940
agttggcttc aatgttggcc ctgtggtggc tggagtgatt ggcgctcgca ggccccagta 3000
cgacatctgg gyaaacacag tcaacgtggc cagtctgatg gatagcacag gggcccaggg 3060
cagaatccag gtgactgagg aagtcaccg gctgctgaga aggtgcccct accactttgt 3120
gtgcccaggc aaagtctagt tcaaggycan aggcagatg ttgacatact ttctagaagg 3180
caggactgat gyaaacggtc cccaatcag gtccctgggc ttggatcgga aaatgtgtcc 3240
atttgggaga gctggccttc agggcagagc tcccccggtc tgcccctatg ctggcgtctc 3300
agttagggct gggctccctc cacactcccc aggccagtac ctgcccctcg cagcagctgg 3360
gaaggaggct tagtggagcc cacgtgggccc tctgggggtc acatgggggtg ggaatgctcc 3420
gggggtgaca caggccacgg tggctccagc caggaccagc cagaccagca gagcaggag 3480
ccacttgcca gsgtggagga ggaactattt ccaggcatgg cctgtggccc gaggcccaac 3540
cacgagcag gcacagcaca gcagtactc ggtgagggga ggacaccga ctgtgcaac 3600
ttcagggct cctggagag gtgtccactc catcctccc tccgtggcgt agggagcact 3660
gtttctttcc agcaagcctg ttcaggtttg gccaggtcgg catcaatga aggccttca 3720
gagcatccca gsgttacagc aagagccact gagggtgtgc tggcagagca actgaggagc 3780
tctcgagag tgtccctct gagacagccc t 3811

```

【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
10 January 2002 (10.01.2002)

PCT

(10) International Publication Number
WO 02/002757 A3

- (51) International Patent Classification: C12N 15/00, 15/88
- (21) International Application Number: PCT/US01/20491
- (22) International Filing Date: 26 June 2001 (26.06.2001)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/215,476 29 June 2000 (29.06.2000) US
60/223,545 4 August 2000 (04.08.2000) US
60/229,876 31 August 2000 (31.08.2000) US
60/234,838 22 September 2000 (22.09.2000) US
60/236,483 29 September 2000 (29.09.2000) US
- (71) Applicant (for all designated States except US): INCYTE GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).
- (72) Inventors; and
(75) Inventors/Applicants (for US only): GANDHI, Ameena, R. [US/US]; 837 Roble Avenue, #1, Menlo Park, CA 94025 (US). TRIBOULEY, Catherine [FR/US]; 1121 Tennessee Street #5, San Francisco, CA 94107 (US). DING, Li [CH/US]; 3353 Alma Street, #146, Palo Alto, CA 94306 (US). LU, Dyung, Aina, M. [US/US]; 233 Coy Drive, San Jose, CA 95123 (US). LEE, Ernestine, A. [US/US]; 624 Kains Street, Albany, CA 94706 (US). YUE, Henry [US/US]; 826 Loh Avenue, Sunnyvale, CA 94087 (US). YANG, Junming [CH/US]; 7125 Bank Lane, San Jose, CA 95129 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). THORNTON, Michael [US/US]; 9 Medway Road, Woodside, CA 94062 (US). YAO, Monique, G. [US/US]; 111 Frederick Court, Mountain View, CA 94043 (US). WALIA, Narinder, K. [US/US]; 890 Davis Street #205, San Leandro, CA 94577 (US). TANG, Y. Tom [US/US]; 4230 Rawick Court, San Jose, CA 95118 (US). ELLIOTT, Vicki, S. [US/US]; 3770 Palom Place Way, San Jose, CA 95121 (US). LU, Yan [—US]; 3885 Corrina Way, Palo Alto, CA 94303 (US).
- (74) Agents: HAMLET-COX, Diana et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, ES, FI, GB, GR, GT, HK, HU, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GI, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- Published:
with international search report
- (88) Date of publication of the international search report:
3 January 2002
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/002757 A3

(54) Title: ADENYLYL AND GUANYLYL CYCLASES

(57) Abstract: The invention provides human adenylyl and guanylyl cyclases (ADGUC) and polynucleotides which identify and encode ADGUC. 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 ADGUC.

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		Interns	Application No
		PCT/us	01/20491
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/00 C12N15/88			
According to International Patent Classification (IPC) or to both national classification and IPC			
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 C12N			
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched			
Electronic data base consulted during the international search (name of data base and, where practical, search terms used)			
EPO-Internal, BIOSIS, CHEM ABS Data, SEQUENCE SEARCH, WPI Data			
C. DOCUMENTS CONSIDERED TO BE RELEVANT			
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	
X	GAO ET AL: "Cloning and expression of widely distributed (type IV) adenylyl cyclase" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, NATIONAL ACADEMY OF SCIENCE, WASHINGTON, US, vol. 88, November 1991 (1991-11), pages 10178-10182, XP002162218 ISSN: 0027-8424 compare ACC P26770 the whole document --- -/--	1-19,21, 22, 24-45,50	
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C.		<input checked="" type="checkbox"/> Patent family members are listed in annex.	
* Special categories of cited documents : *A* document defining the general state of the art which is not considered to be of particular relevance *E* earlier document but published on or after the international filing date *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) *O* document referring to an oral disclosure, use, exhibition or other means *P* document published prior to the international filing date but later than the priority date claimed ** later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. *Z* document member of the same patent family			
Date of the actual completion of the international search		Date of mailing of the international search report	
23 April 2002		12.06.02	
Name and mailing address of the ISA European Patent Office, P.O. Box 5816 Patentlaan 2 NL - 2200 HV Rijswijk Tel: (+31-70) 340-2040, Tx: 31 651 epo nl, Fax: (+31-70) 340-3016		Authorized officer Meyer, W	

Form PCT/ISA/10 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT		Interns Application No PCT/us 01/20491
C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	HURLEY JAMES H: "The adenylyl and guanylyl cyclase superfamily." CURRENT OPINION IN STRUCTURAL BIOLOGY, vol. 8, no. 6, December 1998 (1998-12), pages 770-777, XP002197060 ISSN: 0959-440X the whole document ----	
A	HURLEY JAMES H: "Structure, mechanism, and regulation of mammalian adenylyl cyclase." JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 274, no. 12, 19 March 1999 (1999-03-19), pages 7599-7602, XP002197061 ISSN: 0021-9258 the whole document ----	
A	TUCKER CHANDRA L ET AL: "Two amino acid substitutions convert a guanylyl cyclase, RetGC-1, into an adenylyl cyclase." PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES, vol. 95, no. 11, 26 May 1998 (1998-05-26), pages 5993-5997, XP002197062 May 26, 1998 ISSN: 0027-8424 the whole document ----	
A	LIU YU ET AL: "Catalytic mechanism of the adenylyl and guanylyl cyclases: Modeling and mutational analysis." PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES, vol. 94, no. 25, 9 December 1997 (1997-12-09), pages 13414-13419, XP002197063 Dec. 9, 1997 ISSN: 0027-8424 the whole document ----	
P,X	WO 01 25448 A (KAPELLER LIBERMANN ROSANA ;CHUN MIYOUNG (US); MILLENNIUM PHARM INC) 12 April 2001 (2001-04-12) the whole document ----	1-19,21, 22, 24-45,50
P,X	US 6 107 076 A (GILMAN ALFRED G ET AL) 22 August 2000 (2000-08-22) column 59-66 -----	1-19,21, 22, 24-45,50

INTERNATIONAL SEARCH REPORT	International application No. PCT/US 01/20491
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 and 23 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.:</p> <p style="padding-left: 40px;">1-19, 21-22, 24-45 and 50</p>	
<p>Remark on Protest</p> <p><input type="checkbox"/> The additional search fees were accompanied by the applicant's protest.</p> <p><input type="checkbox"/> No protest accompanied the payment of additional search fees.</p>	

International Application No. PCT/US 01 20491

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

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

1. Claims: 1-19, 21, 22, 24-45 and 50
Subject-matter relating to SEQ ID NO 1 and 6
2. Claims: 1-19, 21, 22, 24-44, 46 and 51
Subject-matter relating to SEQ ID NO 2 and 7
3. Claims: 1-19, 21, 22, 24-44, 47 and 52
Subject-matter relating to SEQ ID NO 3 and 8
4. Claims: 1-19, 21, 22, 24-44, 48 and 53
Subject-matter relating to SEQ ID NO 4 and 9
5. Claims: 1-19, 21, 22, 24-44, 49 and 54
Subject-matter relating to SEQ ID NO 5 and 10

International Application No. PCT/US 01 20491

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

Although claims 18 and 24 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Although claim 34 is 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 and 23

Present claims 20 and 23 relate to a compound defined by reference to a desirable characteristic or property, namely to act as an agonist compound

These "reach through claims" comprise a huge number of heterogeneous chemicals (e.g. sodium chloride, lead, ...) that may act as inhibitors. The application does not provide support within the meaning of Article 6 PCT or disclosure within the meaning of Article 5 PCT for such chemicals. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful examination is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the compound used by reference to a result to be achieved. Such claims are devoid of any essential technical features which would allow a meaningful examination of the patentability as set out in Article 33 PCT.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT
on patent family members

International Application No
PCT/JP 01/20491

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0125448	A	12-04-2001	US 6403358 B1 11-06-2002 AU 7860800 A 10-05-2001 EP 1222287 A1 17-07-2002 WO 0125448 A1 12-04-2001
US 6107076	A	22-08-2000	NONE

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 21/00	A 6 1 P 21/04	4 B 0 6 5
A 6 1 P 21/04	A 6 1 P 25/00	4 C 0 8 4
A 6 1 P 25/00	A 6 1 P 27/00	4 H 0 4 5
A 6 1 P 27/00	A 6 1 P 43/00	1 1 1
A 6 1 P 43/00	C 0 7 K 16/40	
C 0 7 K 16/40	C 0 7 K 16/46	
C 0 7 K 16/46	C 0 7 K 19/00	
C 0 7 K 19/00	C 1 2 N 1/15	
C 1 2 N 1/15	C 1 2 N 1/19	
C 1 2 N 1/19	C 1 2 N 1/21	
C 1 2 N 1/21	C 1 2 N 9/88	
C 1 2 N 5/10	C 1 2 Q 1/68	A
C 1 2 N 9/88	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	M
G 0 1 N 33/50	G 0 1 N 33/566	
G 0 1 N 33/53	G 0 1 N 33/573	A
G 0 1 N 33/566	C 1 2 N 5/00	A
G 0 1 N 33/573	A 6 1 K 37/56	
// C 1 2 P 21/08	C 1 2 P 21/08	

(31)優先権主張番号 60/234,838

(32)優先日 平成12年9月22日(2000.9.22)

(33)優先権主張国 米国(US)

(31)優先権主張番号 60/236,483

(32)優先日 平成12年9月29日(2000.9.29)

(33)優先権主張国 米国(US)

(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 トリボレー、キャサリーン

アメリカ合衆国カリフォルニア州94107・サンフランシスコ・#5・テネシーストリート 1
121

(72)発明者 ディング、リー

アメリカ合衆国カリフォルニア州94306・パロアルト・#146・アルマストリート 335
3

(72)発明者 リュ、デュング・アイナ・エム

アメリカ合衆国カリフォルニア州95123・サンノゼ・コイドライブ 233

(72)発明者 リー、アーンステーション・エイ

アメリカ合衆国カリフォルニア州94706・アルバニー・ケインズストリート 624

(72)発明者 ユエ、ヘンリー

アメリカ合衆国カリフォルニア州94087・サニーバイル・ルイスアベニュー 826

(72)発明者 ヤング、ジュンミン

- アメリカ合衆国カリフォルニア州 9 5 1 2 9 ・ サンノゼ ・ バンクレーン 7 1 2 5
 (72)発明者 ボーグン、マライア・アール
 アメリカ合衆国カリフォルニア州 9 4 5 7 7 ・ サンレアンドロ ・ サンティアゴロード 1 4 2 4 4
 (72)発明者 ソートン、マイケル
 アメリカ合衆国カリフォルニア州 9 4 0 6 2 ・ ウッドサイド ・ メッドウェイロード 9
 (72)発明者 ヤオ、モニック・ジー
 アメリカ合衆国カリフォルニア州 9 4 0 4 3 ・ マウンテンビュー ・ フレデリックコート 1 1 1
 (72)発明者 チョーラ、ナリンダー・ケイ
 アメリカ合衆国カリフォルニア州 9 4 5 8 7 ・ ユニオンシティ ・ # 7 1 2 ・ ユニオンスクエア 3
 3
 (72)発明者 タング、ワイ・トム
 アメリカ合衆国カリフォルニア州 9 5 1 1 8 ・ サンノゼ ・ ランウィックコート 4 2 3 0
 (72)発明者 エリオット、ビッキー・エス
 アメリカ合衆国カリフォルニア州 9 5 1 2 1 ・ サンノゼ ・ ポルトンプレイスウェイ 3 7 7 0
 (72)発明者 リュ、ヤン
 アメリカ合衆国カリフォルニア州 9 4 3 0 3 ・ パロアルト ・ コリーナウェイ 3 8 8 5
- F ターム(参考) 2G045 AA25 AA40 BB03 BB20 CB01 DA12 DA13 DA14 DA20 DA36
 DA77 FB02 FB03
 4B024 AA01 AA11 BA07 BA43 BA61 CA04 DA02 EA02 EA04 HA01
 HA04 HA14 HA15
 4B050 CC03 DD07 FF16E LL01 LL03
 4B063 QA01 QA05 QA18 QQ44 QQ53 QR32 QR55 QS25 QS34 QS36
 QX02 QX07
 4B064 AG27 CA10 CA19 CA20 CC24 DA01 DA13
 4B065 AA90X AA93Y AA95X AB01 BA02 CA27 CA44 CA46
 4C084 AA01 AA02 AA17 BA01 BA08 BA22 BA23 CA18 DC26 NA14
 ZA011 ZA331 ZA361 ZA811 ZA941 ZB351
 4H045 AA10 AA11 AA20 AA30 DA76 EA20 EA50 FA74

专利名称(译)	腺苷酸环化酶和鸟苷酸环化酶		
公开(公告)号	JP2004529603A	公开(公告)日	2004-09-30
申请号	JP2002507997	申请日	2001-06-26
[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	ガンディーアミーナアール トリボレーキャサリーン デイングリー リュデユングアイナエム リーアーンステイーンエイ ユエヘンリー ヤングジュンミング ボーグンマライアアール ソーントンマイケル ヤオモニークジー チョーラナリンダーケイ タングワイトム エリオットビッキーエス リュヤン		
发明人	ガンディー、アミーナ・アール トリボレー、キャサリーン デイング、リー リュ、デユング・アイナ・エム リー、アーンステイーン・エイ ユエ、ヘンリー ヤング、ジュンミング ボーグン、マライア・アール ソーントン、マイケル ヤオ、モニーク・ジー チョーラ、ナリンダー・ケイ タング、ワイトム エリオット、ビッキー・エス リュ、ヤン		
IPC分类号	G01N33/50 A61K38/00 A61K38/51 A61K45/00 A61P9/00 A61P15/00 A61P21/00 A61P21/04 A61P25/00 A61P27/00 A61P43/00 C07K16/40 C07K16/46 C07K19/00 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N9/88 C12N15/09 C12P21/08 C12Q1/68 G01N33/15 G01N33/53 G01N33/566 G01N33/573		
CPC分类号	A01K2217/05 A61K38/00 A61P15/00 A61P21/00 A61P21/04 A61P25/00 A61P27/00 C12N9/88 C12Y406/01001 C12Y406/01002		
FI分类号	C12N15/00.ZNA.A A61K45/00 A61P9/00 A61P15/00 A61P21/00 A61P21/04 A61P25/00 A61P27/00 A61P43/00.111 C07K16/40 C07K16/46 C07K19/00 C12N1/15 C12N1/19 C12N1/21 C12N9/88 C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.M G01N33/566 G01N33/573.A C12N5/00.A A61K37/56 C12P21/08		
F-TERM分类号	2G045/AA25 2G045/AA40 2G045/BB03 2G045/BB20 2G045/CB01 2G045/DA12 2G045/DA13 2G045/DA14 2G045/DA20 2G045/DA36 2G045/DA77 2G045/FB02 2G045/FB03 4B024/AA01 4B024/AA11 4B024/BA07 4B024/BA43 4B024/BA61 4B024/CA04 4B024/DA02 4B024/EA02 4B024/EA04 4B024/HA01 4B024/HA04 4B024/HA14 4B024/HA15 4B050/CC03 4B050/DD07 4B050/FF16E 4B050/LL01		

