

(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2005-500015

(P2005-500015A)

(43) 公表日 平成17年1月6日(2005.1.6)

(51) Int. Cl. ⁷	F I	テーマコード (参考)
C 1 2 N 15/09	C 1 2 N 15/00 Z N A A	2 G O 4 5
A O 1 K 67/027	A O 1 K 67/027	4 B O 2 4
A 6 1 K 31/7088	A 6 1 K 31/7088	4 B O 6 3
A 6 1 K 38/00	A 6 1 K 39/00 H	4 B O 6 5
A 6 1 K 39/00	A 6 1 K 39/395 D	4 C O 8 4
審査請求 未請求 予備審査請求 未請求 (全 163 頁) 最終頁に続く		

(21) 出願番号	特願2002-570600 (P2002-570600)	(71) 出願人	502334043 インファーマティカ リミテッド イギリス ロンドン ダブリュー1ティー 2 エヌユー チャーロット ストリート 6 0
(86) (22) 出願日	平成14年3月5日 (2002.3.5)	(74) 代理人	100059959 弁理士 中村 稔
(85) 翻訳文提出日	平成15年9月5日 (2003.9.5)	(74) 代理人	100067013 弁理士 大塚 文昭
(86) 国際出願番号	PCT/GB2002/000986	(74) 代理人	100082005 弁理士 熊倉 禎男
(87) 国際公開番号	W02002/070563	(74) 代理人	100065189 弁理士 宍戸 嘉一
(87) 国際公開日	平成14年9月12日 (2002.9.12)	(74) 代理人	100074228 弁理士 今城 俊夫
(31) 優先権主張番号	0105402.2		
(32) 優先日	平成13年3月5日 (2001.3.5)		
(33) 優先権主張国	英国 (GB)		

最終頁に続く

(54) 【発明の名称】 核内ホルモンレセプターリガンド結合ドメイン

(57) 【要約】

本発明は、新規タンパク質 (BAA22563.1 と称され、本明細書では核内ホルモンレセプターリガンド結合ドメインと同定された)、並びに疾患の診断、予防及び治療における前記タンパク質およびコード遺伝子由来の核酸配列の使用に関する。

【特許請求の範囲】**【請求項 1】**

以下のポリペプチド：

(i) 配列番号：2 に記載のアミノ酸配列を含むかまたは前記アミノ酸配列からなるポリペプチド；

(i i) 核内ホルモンレセプターリガンド結合ドメインの活性を有するか、または (i) のポリペプチドと共通の抗原決定基を有する、前記のフラグメントであるポリペプチド；または

(i i i) (i) または (i i) の機能的等価物であるポリペプチド。

【請求項 2】

LBDG2 ポリペプチドの核内ホルモンレセプターリガンド結合ドメイン領域を含む、請求項 1 の (i i) に記載のフラグメントであるポリペプチドであって、前記核内ホルモンレセプターリガンド結合ドメイン領域が配列番号：2 に記載のアミノ酸配列の残基 1104 から 1309 を含むと定義され、前記フラグメントが “ LBDモチーフ ” 残基 PHE1174、VAL1177、ASP1181、ASN1182、LEU1185 および LEU1186、または等価な残基を保有し、さらに核内ホルモンレセプターリガンド結合ドメインの活性を保有する前記ポリペプチド。

10

【請求項 3】

請求項 1 の (i i i) に記載の機能的等価物であり、配列番号：2 に記載のアミノ酸配列と相同であり、触媒性残基 PHE1174、VAL1177、ASP1181、ASN1182、LEU1185 および LEU1186、または等価な残基を保有し、さらに核内ホルモンレセプターリガンド結合ドメインの活性を保有するポリペプチド。

20

【請求項 4】

機能的等価物が LBDG2 ポリペプチドの核内ホルモンレセプターリガンド結合ドメイン領域に相同である、請求項 3 に記載のポリペプチド。

【請求項 5】

BLASTバージョン 2.1.3 で NCBI (the National Center for Biotechnology Information ; <http://www.ncbi.nlm.nih.gov/>) によって特定されたデフォルトパラメーター { Blosum62 マトリックス；ギャップ開放ペナルティー = 11 およびギャップ伸長ペナルティー = 1 } を用いて決定したとき、配列番号：2 に記載のアミノ酸配列、または核内ホルモンレセプターリガンド結合ドメイン活性を保有するそのフラグメントと 80% を超える配列同一性、好ましくは 85%、90%、95%、98% または 99% を超える配列同一性を有する、請求項 1 - 4 のいずれか 1 項に記載のフラグメントまたは機能的等価物。

30

【請求項 6】

配列番号：2 のいずれか一つによって与えられるアミノ酸配列を有するポリペプチド、または核内ホルモンレセプターリガンド結合ドメイン活性を保有するそのフラグメントと顕著な構造相同性を示す請求項 1 - 5 のいずれか 1 項に記載の機能的等価物。

【請求項 7】

配列番号：2 の配列に由来する 7 つもしくはそれより多い (例えば 8、10、12、14、16、18、20 またはそれより多い) アミノ酸残基からなる、請求項 1 の (i) のポリペプチドと共通の抗原決定基を有する請求項 1、2 または 5 に記載のフラグメント。

40

【請求項 8】

請求項 1 - 7 のいずれか 1 項に記載のポリペプチドをコードする精製核酸分子。

【請求項 9】

配列番号：1 に記載の核酸配列を有するか、または重複等価物もしくはそのフラグメントである請求項 8 に記載の精製核酸分子。

【請求項 10】

配列番号：1 のヌクレオチド 3351 から 3968 を含むか、またはその重複等価物である請求項 8 または請求項 9 に記載の精製核酸分子のフラグメント。

【請求項 11】

高いストリンジェンシー条件下で請求項 8 - 10 のいずれか 1 項に記載の核酸分子とハイ

50

ブリダイズする精製核酸分子。

【請求項 1 2】

請求項 8 - 1 1 のいずれか 1 項に記載の核酸分子を含むベクター。

【請求項 1 3】

請求項 1 2 に記載のベクターで形質転換された宿主細胞。

【請求項 1 4】

請求項 1 - 7 のいずれか 1 項に記載のポリペプチドと特異的に結合し、さらに好ましくは前記ポリペプチドの核内ホルモンレセプターリガンド結合ドメイン活性を阻害するリガンド。

【請求項 1 5】

抗体である請求項 1 4 に記載のリガンド。

【請求項 1 6】

請求項 1 - 7 のいずれか 1 項に記載のポリペプチドの発現レベルまたは活性レベルを増加または低下させる化合物。

【請求項 1 7】

請求項 1 - 7 のいずれか 1 項に記載のポリペプチドと結合し、前記ポリペプチドの生物学的作用のいずれをも誘発することがない請求項 1 6 に記載の化合物。

【請求項 1 8】

天然または改変された基質、リガンド、酵素、レセプターまたは構造的もしくは機能的模倣体である請求項 1 6 または 1 7 に記載の化合物。

【請求項 1 9】

疾患の治療または診断で使用することを目的とする請求項 1 - 7 のいずれか 1 項に記載のポリペプチド、請求項 8 - 1 1 のいずれか 1 項に記載の核酸分子、請求項 1 2 に記載のベクター、請求項 1 4 または 1 5 に記載のリガンド、または請求項 1 6 - 1 8 のいずれか 1 項に記載の化合物。

【請求項 2 0】

患者の疾患を診断する方法であって、前記患者の組織で請求項 1 - 7 のいずれか 1 項に記載のポリペプチドをコードする天然の遺伝子の発現レベルを評価するか、または請求項 1 - 7 のいずれか 1 項に記載のポリペプチドの活性を評価し、さらに前記発現または活性レベルをコントロールレベルと比較することを含み、前記コントロールレベルと異なるレベルが疾患を示唆する前記疾患の診断方法。

【請求項 2 1】

in vitro で実施される請求項 2 0 に記載の方法。

【請求項 2 2】

(a) 請求項 1 4 または 1 5 に記載のリガンドを、リガンド - ポリペプチド複合体の形成に適した条件下で生物学的サンプルと接触させ、さらに (b) 前記複合体を検出する工程；を含む、請求項 2 0 または 2 1 に記載の方法。

【請求項 2 3】

a) 患者由来の組織サンプルを、請求項 8 - 1 1 のいずれか 1 項に記載の核酸分子とプローブとの間でハイブリッド複合体を形成し得るストリンジェントな条件下で核酸プローブと接触させる工程；

b) コントロールサンプルと前記プローブを工程 a) で用いた条件と同じ条件下で接触させる工程；さらに

c) 前記サンプル中のハイブリッド複合体の存在を検出する工程；

を含み、コントロールサンプル中のハイブリッド複合体レベルと異なるハイブリッド複合体レベルが患者サンプルで検出されることは疾患を示唆する、請求項 2 0 または請求項 2 1 に記載の方法。

【請求項 2 4】

a) 患者の組織由来の核酸サンプルを、請求項 8 - 1 1 のいずれか 1 項に記載の核酸分子とプライマーとの間でハイブリッド複合体を形成し得るストリンジェントな条件下で核酸

10

20

30

40

50

プライマーと接触させる工程；

b) コントロールサンプルと前記プライマーを工程 a) で用いた条件と同じ条件下で接触させる工程；

c) サンプル化された前記核酸を増幅する工程；および、

d) 患者サンプルおよびコントロールサンプルの両方から、増幅された核酸レベルを検出する工程；

を含み、コントロールサンプル中の増幅核酸レベルと顕著に異なる増幅核酸レベルが患者サンプルで検出されることは疾患を示唆する、請求項 20 または請求項 21 に記載の方法。

【請求項 25】

以下の工程を含む請求項 20 または請求項 21 に記載の方法：

a) 疾患について検査しようとする患者から組織サンプルを得る工程；

b) 請求項 8 - 11 のいずれか 1 項に記載の核酸分子を前記組織サンプルから単離する工程；および、

c) 核酸分子中で疾患に関連する変異の存在を疾患の徴候として検出することによって患者を疾患について診断する工程。

【請求項 26】

核酸分子を増幅して増幅生成物を形成させ、前記増幅生成物中で変異の有無を検出することをさらに含む請求項 25 の方法。

【請求項 27】

核酸分子をストリンジェントな条件下で前記核酸分子とハイブリダイズする核酸プローブと接触させてハイブリッド二本鎖分子を形成させ、前記ハイブリッド二本鎖分子は疾患に関連する変異に対応するいずれかの部分でハイブリダイズしない核酸プローブ鎖部分を有し、および、疾患関連変異の有無を表示するものとして前記プローブ鎖のハイブリダイズしない部分の有無を検出することによって、患者において変異の有無を検出する請求項 25 または 26 の方法。

【請求項 28】

疾患が以下から選択される請求項 20 - 27 のいずれか 1 項に記載の方法：

細胞増殖性疾患（新生物、メラノーマ、肺、結腸直腸、乳房、膵、頭部および頸部の腫瘍並びに他の固形腫瘍を含む）、骨髄増殖性疾患（例えば白血病、非ホジキンリンパ腫）、白血球減少症、血小板減少症、血管形成疾患、カポジ肉腫、自己免疫/炎症性疾患（アレルギー、炎症性腸疾患、関節炎、乾癬および気道の炎症、喘息および器官の移植拒絶を含む）、心脈管系疾患（高血圧、浮腫、アングナ、アテローム性硬化症、血栓症、敗血症、ショック、再灌流障害、心不整脈および虚血を含む）、神経学的疾患（中枢神経系疾患、アルツハイマー病、脳損傷、脳卒中、筋萎縮性側索硬化症、不安、抑うつおよび痛みを含む）、発達障害、代謝性障害（真性糖尿病、骨粗しょう症、脂質代謝障害、甲状腺機能亢進、上皮小体機能亢進、高カルシウム血症、高コレステロール血症、高脂血症および肥満を含む）、腎疾患（糸球体腎炎、腎血管性高血圧を含む）、皮膚疾患（アクネ、湿疹および創傷治癒を含む）、加齢の負の作用、エイズ、感染（ウイルス感染、細菌感染、真菌感染および寄生虫感染を含む）、並びに他の病的状態（特に核内ホルモンレセプターが関与するもの）。

【請求項 29】

核内ホルモンレセプターリガンド結合ドメインとしての請求項 1 - 7 のいずれか 1 項に記載のポリペプチドの使用。

【請求項 30】

核内ホルモンレセプターリガンド結合ドメイン活性を保有するタンパク質の発現を目的とする請求項 8 - 11 のいずれか 1 項に記載の核酸分子の使用。

【請求項 31】

請求項 1 - 7 のいずれか 1 項に記載のポリペプチドを利用して細胞 - 細胞接着を行わせる方法。

10

20

30

40

50

【請求項 3 2】

請求項 1 - 7 のいずれか 1 項に記載のポリペプチド、請求項 8 - 1 1 のいずれか 1 項に記載の核酸分子、請求項 1 2 に記載のベクター、請求項 1 4 または 1 5 に記載のリガンド、または請求項 1 6 - 1 8 のいずれか 1 項に記載の化合物を含む医薬組成物。

【請求項 3 3】

請求項 1 - 7 のいずれか 1 項に記載のポリペプチドまたは請求項 8 - 1 1 のいずれか 1 項に記載の核酸分子を含むワクチン組成物。

【請求項 3 4】

以下の疾患の治療用薬物の製造で使用される、請求項 1 - 7 のいずれか 1 項に記載のポリペプチド、請求項 8 - 1 1 のいずれか 1 項に記載の核酸分子、請求項 1 2 に記載のベクター、請求項 1 4 または 1 5 に記載のリガンド、請求項 1 6 - 1 8 のいずれか 1 項に記載の化合物、または請求項 3 2 記載の医薬組成物：

細胞増殖性疾患（新生物、メラノーマ、肺、結腸直腸、乳房、膵、頭部および頸部の腫瘍並びに他の固形腫瘍）、骨髄増殖性疾患（例えば白血病、非ホジキンリンパ腫）、白血球減少症、血小板減少症、血管形成疾患、カポジ肉腫、自己免疫/炎症性疾患（アレルギー、炎症性腸疾患、関節炎、乾癬および気道の炎症、喘息および器官の移植拒絶を含む）、心臓系疾患（高血圧、浮腫、アングINA、アテローム性硬化症、血栓症、敗血症、ショック、再灌流障害、心不整脈および虚血を含む）、神経学的疾患（中枢神経系疾患、アルツハイマー病、脳損傷、脳卒中、筋萎縮性側索硬化症、不安、抑うつおよび痛みを含む）、発達障害、代謝性障害（真性糖尿病、骨粗しょう症、脂質代謝障害、甲状腺機能亢進、上皮小体機能亢進、高カルシウム血症、高コレステロール血症、高脂血症および肥満を含む）、腎疾患（糸球体腎炎、腎血管性高血圧を含む）、皮膚疾患（アクネ、湿疹および創傷治癒を含む）、加齢の負の作用、エイズ、感染（ウイルス感染、細菌感染、真菌感染および寄生虫感染を含む）、並びに他の病的状態（特に核内ホルモンレセプターが関与するもの）。

【請求項 3 5】

請求項 1 - 7 のいずれか 1 項に記載のポリペプチド、請求項 8 - 1 1 のいずれか 1 項に記載の核酸分子、請求項 1 2 に記載のベクター、請求項 1 4 または 1 5 に記載のリガンド、請求項 1 6 - 1 8 のいずれか 1 項に記載の化合物、または請求項 3 2 に記載の医薬組成物を患者に投与することを含む、患者の疾患を治療する方法。

【請求項 3 6】

天然の遺伝子の発現またはポリペプチドの活性が、健常な対象者における発現または活性レベルと比較したとき罹患対象者で低下する疾患に対して、前記対象者に投与される前記ポリペプチド、核酸分子、ベクター、リガンド、化合物、または組成物がアゴニストである請求項 3 5 に記載の方法。

【請求項 3 7】

天然の遺伝子の発現またはポリペプチドの活性が、健常な対象者における発現または活性レベルと比較したとき罹患対象者で上昇する疾患に対して、前記対象者に投与される前記ポリペプチド、核酸分子、ベクター、リガンド、化合物または組成物がアンタゴニストである請求項 3 5 に記載の方法。

【請求項 3 8】

患者の疾患の治療的処置をモニターする方法であって、請求項 1 - 7 のいずれか 1 項に記載のポリペプチドの発現もしくは活性レベル、請求項 8 - 1 1 のいずれか 1 項に記載の核酸分子の発現レベルを前記患者由来の組織で一定期間にわたってモニターすることを含み、前記期間にわたってコントロールレベルに向かって前記発現または活性レベルが変化することが前記疾患の軽減を示す、前記患者の疾患の治療的処置をモニターする方法。

【請求項 3 9】

疾患の治療および/または診断において有効な化合物の特定方法であって、前記方法が、請求項 1 - 7 のいずれか 1 項に記載のポリペプチド、請求項 8 - 1 1 のいずれか 1 項に記載の核酸分子、または請求項 1 3 に記載の宿主細胞を、前記ポリペプチドまたは核酸に対

10

20

30

40

50

する結合親和性を保有すると思われる1つまたは2つ以上の化合物と接触させ、さらに前記核酸分子またはポリペプチドと特異的に結合する化合物を選別することを含む前記有効な化合物の特定方法。

【請求項40】

ストリンジェントな条件下で請求項8-11のいずれか1項に記載の核酸分子とハイブリダイズする核酸プローブを含む第一の容器；前記核酸分子の増幅に有用なプライマーを含む第二の容器；および疾患の診断を容易にすることを目的とする前記プローブおよびプライマーの使用のための指示書を含む、疾患の診断に有用なキット。

【請求項41】

ハイブリダイズしていないRNAを消化するための薬剤を保持する第三の容器をさらに含む請求項40のキット。 10

【請求項42】

核酸分子のアレイを含むキットであって、前記核酸分子の少なくとも一つが請求項8-11のいずれか1項に記載の核酸分子である前記キット。

【請求項43】

請求項1-7のいずれか1項に記載のポリペプチドと結合する1つまたは2つ以上の抗体；および前記抗体と前記ポリペプチドとの間の結合反応の検出に有用な試薬を含むキット。

【請求項44】

請求項1-7のいずれか1項に記載のポリペプチドをより高レベルまたはより低レベルで発現するように、または前記ポリペプチドを発現しないように形質転換したトランスジェニックまたはノックアウト非ヒト動物。 20

【請求項45】

請求項44に記載の非ヒトトランスジェニック動物を候補化合物と接触させ、さらに前記動物の疾患に対する前記化合物の影響を決定することによって、疾患の治療に有効な化合物をスクリーニングする方法。

【発明の詳細な説明】

【0001】

本発明は、本明細書において核内ホルモンレセプターのリガンド結合ドメインとして同定された、BAA22563.1と称される、新規なタンパク質並びに疾患の診断、予防および治療における前記タンパク質およびそのコード遺伝子に由来する核酸配列の使用に関する。 30
本明細書に引用した全ての刊行物、特許および特許出願は、引用により完全に本明細書に含まれるものとする。

(背景技術)

薬剤の発見プロセスにおいて、機能ゲノム学(functional genomics)の時代の到来につれて根幹的な革命が現在進行している。“機能ゲノム学”という用語は、対象のタンパク質配列に機能を帰属させるためにバイオインフォマティクスツールを利用するアプローチに適用される。そのようなツールは、配列データの生成速度が、これらタンパク質配列に機能を割り当てる研究室の能力をはるかに越えるために、ますます必要性を増している。バイオインフォマティクスツールの潜在能力および精度が高まっているために、前記ツールは通常 40
の生化学的特徴付け技術と急速に置き換えられつつある。実際、本発明の同定に用いた高度なバイオインフォマティクスツールは、今や、高い信頼性をもつ結果を産出する能力を有する。

配列データが利用可能になるにつれ、種々の研究機関および企業の組織がそれらを調査し、重要な発見が絶え間なく達成され続けている。しかしながら、研究および薬剤の発見のための標的として更に新たな遺伝子およびそれらがコードするポリペプチドを同定し特徴付ける必要性は引き続き存在している。

最近、未知の機能をもつ配列を評価するための注目すべきツールが本発明の出願人によって開発された。このツールは、同時係属国際特許出願第PCT/GB01/01105号の主題であるデータベースシステム(バイオペンジウム(Biopendium)検索データベースと称される)で 50

ある。このデータベースシステムは、独占的技術を用いて作製され、利用可能な全てのタンパク質または核酸配列の完全な比較から作製された情報を含む集積データリソースから成る。

【0002】

別個のデータリソースからこれら配列データを一体化させたその背後の目的は、可能なかぎり多くのデータを、配列それ自体と各配列に関連する情報の両方に関して1つの完全なリソースにまとめることである。各配列と関係を有する全ての利用可能なデータ(入手可能な場合はコードするタンパク質の三次元構造に関するデータを含む)を一緒に統合し、各配列について知られている情報の最大限の利用を可能にし、したがって最も多くの知識に基づく予測がこれら配列の比較から入手することができる。前記データベースで作製され、各配列のエントリーに付随する注釈は、生物学的に関連がある事柄を前記配列情報に付与する。

10

このデータリソースは、配列のみからタンパク質機能の正確な予測を可能にした。通常の場合、このような予測は、同じ機能ファミリーに属する他のタンパク質に対して高度な配列同一性(約20% - 30%の同一性)を示すタンパク質についてのみ可能である。既知の機能を有する他の近縁なタンパク質と低度な配列相同性を示すタンパク質に対しては、正確な予測が不可能である。

本件では、その配列が公的に利用可能なデータベースにBAA22563.1として記録されているタンパク質(NCBI GenBankヌクレオチドアクセッション番号AB007510.1およびGenBankタンパク質アクセッション番号BAA22563.1)が、核内ホルモンレセプターリガンド結合ドメインファミリーの新規なメンバーとして意味付けられる。

20

【0003】

I. 核内ホルモンレセプターリガンド結合ドメインに関する導入部

核内ホルモンレセプター遺伝子スーパーファミリー(表1参照)は、標的遺伝子の転写を調節する構造的に近縁なタンパク質をコードする。前記タンパク質には、ステロイドホルモンおよび甲状腺ホルモン、ビタミン並びにリガンドが未だ発見されていない他のタンパク質に対するレセプターが含まれる。核内レセプターは2つの主要なドメイン、DNA結合ドメイン(DBD)およびリガンド結合ドメイン(LBD)から構成されている。DBDは、モノマー、ホモダイマーまたはヘテロダイマーとして特異的なDNA配列と結合するように前記レセプターに指令する。DBDは、核内レセプターでのみ見出される特殊なタイプのジンクフィンガーである。DBDを有する核内レセプターは、PROSITEコンセンサス配列(PS00031)とのマッチングについて検索することによって容易に配列レベルで同定することができる。

30

リガンド結合ドメイン(LBD)は同系のホルモンと結合しこれに応答する。LBDと結合したリガンドは、すでに結合していた“核内レセプターコリプレッサー”を排斥する構造的変化をひき起こす。続いて前記コリプレッサーによって以前に占有されていた部位は空席になり、“核内レセプターコアクチベーター”で補充される。このリガンドによってひき起こされるコアクチベーターによるコリプレッサーの交換が、リガンド結合が標的遺伝子の転写活性化をもたらすメカニズムである。全てのリガンド結合ドメインはコンセンサス配列、“LBDモチーフ”(表2参照)を含み、前記モチーフは、コリプレッサー結合およびコアクチベーター結合を仲介する。LBDは、今日まで全ての核内ホルモンレセプター標的薬剤のための結合部位であり、したがって新規なリガンド結合ドメインは魅力的な薬剤標的であるので、それらの同定が所望される。リガンド結合ドメインは低い配列同一性(約15%)しか共有しないが、非常に類似する構造を有し、したがって、ゲノムスレッダー(Genome Threader)のような構造に基づく関係付けツール(structure-based relationship tool)のための理想的なターゲットである。

40

PROSITEのような基本的な検索ツールを用いて、DBDが存在することによって、さらにそれをもとに推測されたLBDによって、多くのタンパク質配列が核内ホルモンレセプターとしてパブリックドメインですでに注釈付けされている。このために、ゲノムスレッダーによって同定される、核内レセプターとして注釈されてないどの新規LBDも完

50

全にDBDを欠くであろうということが予期される。LBDを有するがDBDを欠くタンパク質の先例は、DAX1によって提供される。したがって、我々は、これらDBDのないヒットを“核内ホルモンレセプター”ではなく“核内ホルモンレセプターリガンド結合ドメイン”を含むものと注釈付けする。

【0004】

【表1】

ファミリー：ステロイドホルモンレセプター	
サブファミリー：グルココルチコイドレセプター プロゲステロンレセプター アンドロゲンレセプター エストロゲンレセプター	10
ファミリー：甲状腺ホルモンレセプター様因子	
サブファミリー：レチノイン酸レセプター (RAR) レチノイドXレセプター (RXR) 甲状腺ホルモンレセプター ビタミンDレセプター NGF1-B FTZ-F1 ペルオキシソーム増殖因子活性化レセプター (PPAR) クジソンレセプター レチノイドオーファンレセプター (ROR) タイレス (Tailless) / COUP HNF-4 CF1 Knirps	20
ファミリー：DAX1	
サブファミリー：DAX1	

表1：核内ホルモンレセプタースーパーファミリー

【0005】

30

【表2】

1	2	3	4	5	6	7	8	9	10	11	12	13
L	任意の2 残基		L	任意の3残基 (又は、2残基、 または4残基)			D	Q	任意の2残基 (又は、1 残基、 または3残 基)		L	L
I			I				E	N			I	I
A			A					R			A	A
V			V					H			V	V
M			M					K			M	M
F			F					S			F	F
Y			Y					T			Y	Y
W			W								W	W

10

20

【0006】

表2：“LBDモチーフ”。最上段の数字はモチーフ内の残基の位置を示す。文字は1文字コードによるアミノ酸を示す。1つの縦の欄内の文字は全てモチーフ内のその位置について許容される。例えば、L、I、A、V、M、F、YまたはWは“LBDモチーフ”の最初の位置を占めることができる。位置4と8の間、および位置9と12の間で見出される残基の数には変動が観察されることを述べる。“LBDモチーフ”は、核内ホルモンレセプターリガンド結合ドメインの681の配列を整列化し、残基の保存パターンを同定することによって構築された。

【0007】

II. 核内ホルモンレセプターおよび疾患

30

核内ホルモンレセプターは、多様な生理学的機能において役割を果たすことが示されている。それらの多くは疾患プロセスにおいて役割を果たし得る(表3参照)

【0008】

【表3】

核内ホルモンレセプター	疾患	
アンドロゲンレセプター	<p>アンドロゲン非感受性症候群 (Lubahn et al. 1989, Proc. Natl. Acad. Sci. USA 86 : 9534-9538)</p> <p>ライフエンスタイン症候群 (Wooster et al. 1992, Nat. Genet. 2 : 132-134)</p> <p>X連鎖劣性脊髄および延髄筋萎縮 (MacLean et al. 1995 Mol. Cell. Endocrinol. 112 : 133-141)</p> <p>男性乳癌 (Wooster et al. 1992, Nat. Genet. 2 : 132-134)</p>	10
グルココルチコイドレセプター	<p>ネルソン症候群 (Karl et al. 1996, J. Clin. Endocrinol. Metab. 81 : 124-129)</p> <p>グルココルチコイド耐性急性T細胞白血病 (Hala et al. 1996, Int. J. Cancer 68 : 663-668)</p>	
鉱質コルチコイドレセプター	<p>仮性低アルドステロン症 (Chung et al. 1995, J. Clin. Endocrinol. Metab. 80 : 3341-3345)</p>	20
エストロゲンレセプター α	<p>ERα発現はヒト乳癌サブセットで上昇する。タモキシフェンの適用は乳癌進行を予防する主要な治療法である。残念ながら、ERα陽性乳癌の35%はタモキシフェン耐性である (Petrangeli et al. 1994, J. Steroid Biochem. Mol. Biol. 49 : 327-331)</p>	
ビタミンD3レセプター	<p>ビタミンD3レセプターの変異は、ビタミンD3欠乏に対する表現型 (Rickets) に類似する遺伝性疾患をもたらす (Hughes et al. 1988, Science 242 : 1702-1725)</p>	30
レチノイン酸レセプター α	<p>急性骨髄性白血病 (Lavau & Dejean 1994, Leukemia 8 : 9-15)</p>	
甲状腺ホルモンレセプター β	<p>“甲状腺ホルモンに対する普遍的耐性” (G R T H) (Refetoff 1994, Thyroid 4 : 345-349)</p>	
DAX1	<p>X連鎖先天性副腎形成不全 (AHC) および性機能不全 (Ito et al. 1997, Mol. Cell. Biol. 17 : 1476-1483)</p>	40

表3 : 核内ホルモンレセプターと疾患

【0009】

したがって、核内ホルモンレセプターのLBDと結合するリガンドによる核内ホルモンレセプターの変化は、疾患の表現型を変化させる手段を提供する。したがって、新規核内ホルモンレセプターのリガンド結合ドメインの同定は、それらタンパク質が上で同定した疾患や他の症状において役割を果たし得るために強く希求される。したがって新規核内ホルモ

ンレセプターのリガンド結合ドメインの同定は、疾患（特に表3で同定したような疾患）の治療および診断と密接な関連を有する。

【0010】

（発明の詳細な説明）

本発明は、BAA22563.1タンパク質が核内ホルモンレセプターリガンド結合ドメインとして機能するという発見を基にしている。

BAA22563.1タンパク質の場合、前記タンパク質配列の残基1104 - 1309を含む領域は、甲状腺ホルモンレセプター（PDBコード1BSX:A）の残基6(Asp216)から残基209(Asp427)と等価なフォールドを採用することが見出された。甲状腺ホルモンレセプターは、核内ホルモンレセプターリガンド結合ドメインとして機能することが知られている。さらに甲状腺ホルモンレセプターの“LBDモチーフ”残基、PHE293、LEU296、ASP300、GLN301、LEU304およびLEU305は、BAA22563.1ではそれぞれPHE1174、VAL1177、ASP1181、ASN1182、LEU1185およびLEU1186として保存されている。前記の関係は単に甲状腺ホルモンレセプターとの関係だけではなく、むしろ全体として核内ホルモンレセプターリガンド結合ドメインファミリー全体との関係である。したがって、BAA22563.1と甲状腺ホルモンレセプター（1BSX:A）とのゲノムスレッダー（商標）アラインメントの参照によって、BAA22563.1のPHE1174、VAL1177、ASP1181、ASN1182、LEU1185およびLEU1186は“LBDモチーフ”残基を形成すると予測される。

10

等価なフォールドおよび“LBDモチーフ”残基の保存が組み合わされて、BAA22563.1の前記領域の機能的注釈付けを可能とし、したがって前記領域を含むタンパク質が核内ホルモンレセプターリガンド結合ドメイン活性を有するという機能的注釈付けを可能とする。

20

【0011】

第一の特徴では、本発明はポリペプチドを提供し、前記ポリペプチドは、

（i）配列番号：2に記載のアミノ酸配列を含む；

（ii）核内ホルモンレセプターリガンド結合ドメインの活性を有するか、または（i）のポリペプチドと共通の抗原決定基を有する、配列番号：2のフラグメントである；または

（iii）（i）または（ii）の機能的等価物である。

好ましくは、前記ポリペプチドは、

（i）配列番号：2に記載のアミノ酸配列から成る；

（ii）核内ホルモンレセプターリガンド結合ドメインの活性を有するか、または（i）のポリペプチドと共通の抗原決定基を有する、配列番号：2のフラグメントである；または

30

（iii）（i）または（ii）の機能的等価物である。

配列番号：2に記載の配列を有するポリペプチドは、以下では“LBDG2ポリペプチド”と称する。

【0012】

本発明の前記特徴にしたがえば、上記（ii）に記載の好ましいポリペプチドフラグメントは、核内ホルモンレセプターリガンド結合ドメインの活性を担うと予測される、LBDG2ポリペプチドの領域（以下では“LBDG2核内ホルモンレセプターリガンド結合ドメイン領域”と称する）を含むか、または“LBDモチーフ”（PHE1174、VAL1177、ASP1181、ASN1182、LEU1185およびLEU1186、または等価な残基）を保有する変種である。本明細書で明確にされるように、LBDG2核内ホルモンレセプターリガンド結合ドメイン領域は、LBDG2ポリペプチド配列の残基1104と残基1309との間に広がっていると考えられる。

40

本発明の前記特徴はまた、ポリペプチドのフラグメントおよび上記で定義したこれらポリペプチドフラグメントの変種を取り込んだ融合タンパク質も含むが、ただし前記融合タンパク質は核内ホルモンレセプターリガンド結合ドメインとしての活性を保有することを条件とする。

【0013】

第二の特徴では、本発明は、本発明の第一の特徴に記載のポリペプチドをコードする精製

50

核酸分子を提供する。好ましくは前記精製核酸分子は、配列番号：1（LBDG2ポリペプチドをコードする）に記載の核酸配列を有するか、または前記配列の重複等価物またはフラグメントである。好ましい核酸フラグメントは、上記（i i）に記載のポリペプチドフラグメントをコードするか、好ましくはLBDG2核内ホルモンレセプターリガンド結合ドメイン領域を含むか、上記に定義したこれらフラグメントの変種をコードするポリペプチドフラグメントである。

第三の特徴では、本発明は、本発明の第二の特徴の核酸分子と高ストリンジェンシー条件下でハイブリダイズする精製核酸分子を提供する。

第四の特徴では、本発明は、発現ベクターのように本発明の第二または第三の特徴の核酸分子を含むベクターを提供する。

10

第五の特徴では、本発明は、本発明の第四の特徴のベクターで形質転換された宿主細胞を提供する。

第六の特徴では、本発明は、本発明の第一の特徴のポリペプチドと特異的に結合し、好ましくは核内ホルモンレセプターリガンド結合ドメインの活性を阻害するリガンドを提供する。

【0014】

第七の特徴では、本発明は、本発明の第一の特徴のポリペプチドをコードする天然の遺伝子の発現を変化させるか、または本発明の第一の特徴のポリペプチドの活性を調節するために有効な化合物を提供する。

本発明の第七の特徴の化合物は、前記ポリペプチドの遺伝子の発現レベルまたは活性レベルを増加（作働）させるか、または低下（拮抗）させ得る。重要なことには、それぞれLBDG2ポリペプチドのLBDG2核内ホルモンレセプターリガンド結合ドメイン領域として本明細書で明確にされた領域の機能を同定することによって、核内ホルモンレセプターリガンド結合ドメインが関与する疾患の治療および/または診断に有効な化合物を同定することが可能なスクリーニング方法をデザインすることができる。

20

第八の特徴では、本発明は、診断または治療で使用するための本発明の第一の特徴のポリペプチド、または本発明の第二もしくは第三の特徴の核酸分子、または本発明の第四の特徴のベクター、または本発明の第五の特徴のリガンド、または本発明の第六の特徴の化合物を提供する。

【0015】

本発明者は、種々のヒト組織由来の抽出物においてLBDG2のmRNAが見出されることを発見した。そのヒト脾臓における高レベルの転写の発見は、免疫系、特にリンパ球およびB細胞の発生および機能におけるLBDG2の役割と一致している。したがってLBDG2のアゴニストおよびアンタゴニストの開発は、自己免疫、アレルギーおよび免疫グロブリン機能不全に関連する疾患を含む種々のヒト免疫系疾患での治療的処置において同定の役割を有し得ることが考えられる。それらの疾患は、I型真性糖尿病、慢性関節リウマチ、多発性硬化症、乾癬、糸球体症に起因する腎不全、強皮症、炎症性腸疾患（クローン病および潰瘍性大腸炎）、移植拒絶、喘息、アトピー性皮膚炎、湿疹、骨髄腫および抗体産生を必要とする感染症（例えば、ウイルス感染細胞、結核、リステリアのような細胞内病原体）を含む。LBDG2のメッセンジャーRNAは、Daudi、IM9およびRaji細胞のようなヒトB細胞株において

30

40

【0016】

さらに副腎、卵巣、前立腺および精巣の組織において、有意なレベルのLBDG2が見出された。このことは、LBDG2に対するアゴニストおよびアンタゴニストの開発が良性前立腺肥大、前立腺癌、卵巣癌および精巣癌のような疾患において価値のあり得ることを示してい

50

る。さらにLBDG2のアゴニストまたはアンタゴニストは以下の疾患を含む疾患の治療のために開発され得る：高血圧、ストレスに対する応答（感染症のストレスを含む）、塩類および水分恒常性の調節、排卵調節を通じての受胎率制御（不妊症および避妊）、着床調節（不妊症および避妊）並びに精子形成の調節（不妊症および避妊）（しかし、前記疾患はこれだけに限られない）。

【0017】

LBDG2のmRNAがヒト脳内において有意なレベルで発現されることも見出された。このことは、ヒト疾患の状態に潜在的な関連を提供することから注目すべきことであり、LBDG2のリガンド結合ドメインに対するアゴニストおよびアンタゴニストの開発は、以下に挙げるものを含む種々のヒト疾患において治療的処置の可能性を提供する：細胞増殖性疾患（新生物、メラノーマ、肺、結腸直腸、乳房、膵、頭部および頸部の腫瘍並びに他の固形腫瘍を含む）；骨髄増殖性疾患（例えば白血病、非ホジキンリンパ腫、白血球減少症、血小板減少症、血管形成疾患、カポジ肉腫）；自己免疫/炎症性疾患（自己免疫、アレルギーおよび免疫グロブリン機能不全に関連する疾患、炎症性腸疾患、関節炎、乾癬および気道の炎症、喘息および器官の移植拒絶を含む）；心脈管系疾患（高血圧、浮腫、アンギナ、アテローム性動脈硬化症、血栓症、敗血症、ショック、再灌流障害、心不整脈および虚血を含む）；神経学的疾患（中枢神経系疾患、アルツハイマー病、脳損傷、脳卒中、筋萎縮性側索硬化症、不安、抑うつおよび痛みを含む）；発達障害；代謝性疾患（真性糖尿病、骨粗しょう症、脂質代謝障害、甲状腺機能亢進、上皮小体機能亢進、高カルシウム血症、高コレステロール血症、高脂血症および肥満を含む）；腎疾患（糸球体腎炎、腎血管性高血圧を含む）；皮膚疾患（アクネ、湿疹および創傷治癒を含む）；加齢の負の作用；エイズ；感染（ウイルス感染、細菌感染、真菌感染および寄生虫感染を含む）；並びに他の病的状態（特に核内ホルモンレセプターが関与するもの）。

【0018】

DNA結合ドメインの存在しないリガンド結合ドメインを含むLBDG2のような“非古典的”核内ホルモンレセプターの発見は、脳におけるステロイド（神経ステロイドとして知られる）の広範な効果、およびそのような効果は一般的には、転写活性化を必要とする既知の古典的ステロイドホルモン核内レセプターを介さずに媒介されることを一貫して報告している既知文献と一致している。一例を挙げると、神経ステロイドは、特にGABAおよびNMDAのレセプター並びにシグマ(Sigma)レセプターのようなレセプターの域において神経伝達に影響することが示されている。神経ステロイドは神経保護的役割を果たすことが示されている。従ってLBDG2に対するアゴニスト（またはアンタゴニスト）の開発を介する治療的処置は、脳血管疾患（例えば梗塞形成または出血（脳卒中））並びに中枢神経系および脊髄に対する損傷に続く神経変性状態（例えば痴呆、パーキンソン病および神経変性）の治療において役割を有し得る。さらに神経ステロイドは、認知処理、空間学習および記憶、不安、並びに常習行為パターンを引き起こす欲求のような行為に影響することが示された。従ってLBDG2に対するアゴニストおよびアンタゴニストの開発は、痴呆、学習困難、不安、常習行為（例えばアルコール中毒、摂食障害および薬物依存であるが、これだけに限られない）を治療するための治療的処置をもたらし得る。

【0019】

第九番目の特徴では、本発明は患者の疾患を診断する以下の工程を含む方法を提供する：本発明の第一の特徴のポリペプチドをコードする天然の遺伝子の発現レベル、または本発明の第一の特徴のポリペプチドの活性レベルを前記患者由来の組織で評価し、さらに前記発現または活性レベルをコントロールレベルと比較する工程であって、この場合前記コントロールレベルと異なるレベルは疾患を示唆する。前記の方法は好ましくはin vitroで実施されるであろう。同様な方法は患者における疾患の治療的処置のモニタリングに使用することができる。この場合、時間の経過にしたがってポリペプチドまたは核酸分子の発現または活性レベルがコントロールレベルに向かって変化するのは疾患の軽減の指標となる。

本発明の第一の特徴のポリペプチドを検出する好ましい方法は以下の工程を含む：（a）

本発明の第六の特徴のリガンド（例えば抗体）をリガンド - ポリペプチド複合体の形成に適した条件下で生物学的サンプルと接触させる工程；および、（b）前記複合体を検出する工程。

本発明の第九番目の特徴に記載の方法には、例えば短いプローブを用いる核酸ハイブリダイゼーション法、点変異分析、ポリメラーゼ連鎖反応（PCR）増幅、および異常なタンパク質レベルを検出する抗体を用いる方法といった、種々の異なる方法が存在することが当業者には明らかであろう。同様な方法を短期または長期ベースで用いて、疾患の治療的処置を患者においてモニターすることを可能にし得る。本発明はまた前記疾患診断方法に有用なキットも提供する。

【0020】

第十番目の特徴では、本発明は、核内ホルモンレセプターリガンド結合ドメインとしての本発明の第一の特徴のポリペプチドの使用を提供する。本発明はまた、核内ホルモンレセプターリガンド結合ドメインの活性をもつタンパク質の発現のために本発明の第二または第三の特徴に記載の核酸分子の使用も提供する。本発明はまた核内ホルモンレセプターリガンド結合ドメイン活性を作用させる方法も提供し、その方法は本発明の第一の特徴のポリペプチドを利用する。

第十一番目の特徴では本発明は医薬組成物を提供し、前記医薬組成物は、本発明の第一の特徴のポリペプチド、または本発明の第二もしくは第三の特徴の核酸分子、または本発明の第四の特徴のベクター、または本発明の第六の特徴のリガンド、または本発明の第七の特徴の化合物を医薬的に許容できる担体と組合せて含有する。

【0021】

第十二番目の特徴では、本発明は、疾患の診断または治療を目的とする医薬品の製造で使用するために、本発明の第一の特徴のポリペプチド、または本発明の第二のもしくは第三の特徴の核酸分子、または本発明の第四の特徴のベクター、本発明の第六の特徴のリガンド、または本発明の第七の特徴の化合物を提供する。前記診断または治療される疾患は、例えば、細胞増殖性疾患（新生物、メラノーマ、肺、結腸直腸、乳房、膵、頭部および頸部の腫瘍並びに他の固形腫瘍を含む）、骨髄増殖性疾患（例えば白血病、非ホジキンリンパ腫、白血球減少症、血小板減少症、血管形成疾患、カポジ肉腫）、自己免疫/炎症性疾患（アレルギー、炎症性腸疾患、関節炎、乾癬および気道の炎症、喘息および器官の移植拒絶を含む）、心脈管系疾患（高血圧、浮腫、アングィナ、アテローム性動脈硬化症、血栓症、敗血症、ショック、再灌流障害、心不整脈および虚血を含む）、神経学的疾患（中枢神経系疾患、アルツハイマー病、脳損傷、脳卒中、筋萎縮性側索硬化症、不安、抑うつおよび痛みを含む）、発達障害、代謝性障害（真性糖尿病、骨粗しょう症、脂質代謝障害、甲状腺機能亢進、上皮小体機能亢進、高カルシウム血症、高コレステロール血症、高脂血症および肥満を含む）、腎疾患（糸球体腎炎、腎血管性高血圧を含む）、皮膚疾患（アクネ、湿疹および創傷治癒を含む）、加齢の負の作用、エイズ、感染（ウイルス感染、細菌感染、真菌感染および寄生虫感染を含む）、他の病的状態（特に核内ホルモンレセプターが関与するもの）、並びに本発明の第八の特徴に関連して上述した他のより特異的な疾患および状態である。

【0022】

第十三番目の特徴では、本発明は患者の疾患を治療する方法を提供し、前記方法は、本発明の第一の特徴のポリペプチド、または本発明の第二もしくは第三の特徴の核酸分子、または本発明の第四の特徴のベクター、または本発明の第六の特徴のリガンド、または本発明の第七の特徴の化合物を患者に投与することを含む。

本発明の第一の特徴のポリペプチドをコードする天然の遺伝子の発現、または本発明の第一の特徴のポリペプチドの活性が、健常な対象者の発現または活性レベルと比較したとき罹患対象者で低下する疾患の場合、前記患者に投与される前記ポリペプチド、核酸分子、リガンドまたは化合物はアゴニストであろう。逆に、前記天然の遺伝子の発現、または前記ポリペプチドの活性が、健常な対象者の発現または活性レベルと比較したとき罹患対象者で上昇する疾患の場合、前記患者に投与される前記ポリペプチド、核酸分子、リガンド

10

20

30

40

50

または化合物はアンタゴニストであろう。前記アンタゴニストの例にはアンチセンス核酸分子、リボザイムおよびリガンド（例えば抗体）が含まれる。

【0023】

第十四番目の特徴では、本発明は、本発明の第一の特徴のポリペプチドを高レベルで、または低レベルで発現するように、または全く発現しないように形質転換したトランスジェニックまたはノックアウト非ヒト動物を提供する。前記トランスジェニック動物は、疾患の研究用モデルとして非常に有用であり、さらに前記疾患の治療または診断に有効な化合物の同定を目的とするスクリーニング方法で用いることができる。

本発明を利用するために用いることができる標準的な技術および方法の要旨は下記で提供される。本発明は、記載した同定の方法論、プロトコル、細胞株、ベクターおよび試薬に限定されないことは理解されよう。本明細書で用いられる専門用語は単に個々の態様を説明するためのものであり、前記用語によって本発明の範囲を限定しようとするものではないこともまた理解されよう。本発明の範囲は添付の請求の範囲の用語によってのみ限定される。

本明細書では、ヌクレオチドおよびアミノ酸についての標準的な略語が用いられる。

本発明の実施では別に指示がなければ、分子生物学、微生物学、リコンビナントDNA技術および免疫学の通常の技術が用いられるであろう。前記技術は当業者の技術範囲内である。

【0024】

前記のような技術は文献で完全に説明されている。特に適切な解説書の例には以下が含まれる：Sambrook Molecular Cloning; A Laboratory Manual, Second Edition (1989); DNA Cloning, Vol. I and II (D.N. Glover ed. 1985); Oligonucleotide Synthesis (M.J. Gait ed. 1984); Nucleic Acid Hybridization (B.D. Hames & S.J. Higgins eds. 1984); Transcription and Translation (B.D. Hames & S.J. Higgins eds. 1984); Animal Cell Culture (R.I. Freshney ed. 1986); Immobilized Cells and Enzymes (IRL Press, 1986); B. Perbal, A Practical Guide to Molecular Cloning (1984); the Methods in Enzymology series (Academic Press, Inc.)特にVol. 154 & 155; Gene Transfer Vectors for Mammalian Cells (J.H. Miller and M.P. Calos eds. 1987, Cold Spring Harbor Laboratory); Immunochemical Methods in Cell and Molecular Biology (Mayer and Walker, eds. 1987, Academic Press, London); Scopes, (1987) Protein Purification: Principles and Practice, Second Edition (Springer Verlag, N.Y.); および Handbook of Experimental Immunology, Vols. I - IV (D.M. Weir and C.C. Blackwell eds. 1986)。

【0025】

本明細書において用いる“ポリペプチド”という用語は、ペプチド結合または改変ペプチド結合によって互いに連結した2つまたは3つ以上のアミノ酸を含む任意のペプチドまたはタンパク質（すなわちペプチドイソスター）が含まれる。この用語は、短鎖（ペプチドおよびオリゴペプチド）および長鎖（タンパク質）の両方を指す。

本発明のポリペプチドは成熟タンパク質の形態を有するものでもよく、またプレ-、プロ-またはプレプロ-タンパク質であってプレ-、プロ-またはプレプロ-部分の切断によって活性化され、活性成熟ポリペプチドを生じるタンパク質でもよい。そのようなポリペプチドでは、プレ-、プロ-またはプレプロ-配列はリーダー配列もしくは分泌配列であっても、または成熟ポリペプチド配列の精製のために用いられる配列であってもよい。

本発明第一の特徴のポリペプチドは融合タンパク質の一部を形成することができる。例えば、1つまたは2つ以上の付加アミノ酸配列を含むことがしばしば有利である。前記付加アミノ酸配列は、例えばリコンビナント形成時に、分泌もしくはリーダー配列、プロ-配列、精製を促進する配列、またはより高いタンパク質安定性を付与する配列を含んでもよい。あるいは、または前記に加えて、前記成熟ポリペプチドを別の化合物、例えば前記ポリペプチドの半減期を増加させるような化合物（例えばポリエチレングリコール）を融合させることができる。

10

20

30

40

50

【0026】

ポリペプチドは、天然のプロセス（例えば翻訳後プロセッシング）によって、または当業者に周知の化学的改変技術によって改変された、20の遺伝子コードアミノ酸以外のアミノ酸を含んでいてもよい。本発明のポリペプチドに一般的に存在する既知の改変にはグリコシル化、脂質付加、硫化、 α -カルボキシル化（例えばグルタミン酸残基の）、ヒドロキシル化およびADP-リボシル化がある。他の可能な改変には、アセチル化、アシル化、アミド化、フラビンの共有結合付加、ヘム部分の共有結合付加、ヌクレオチドまたはヌクレオチド誘導体の共有結合付加、脂質誘導体の共有結合付加、ホスファチジルイノシトールの共有結合付加、架橋、環状化、ジスルフィド結合形成、脱メチル化、共有結合架橋の形成、システインの形成、ピログルタメートの形成、ホルミル化、GPIアンカー形成、ヨード化、メチル化、ミリストイル化、酸化、タンパク分解性プロセッシング、リン酸化、プレニル化、ラセミ化、セレノイル化、タンパク質へのトランスファーRNA媒介性アミノ酸付加（例えばアルギニル化）およびユビキチン結合が含まれる。

【0027】

改変は、ペプチド骨格、アミノ酸側鎖およびアミノまたはカルボキシ末端を含むポリペプチド内のいずれの場所に生じてもよい。実際、共有結合改変によるポリペプチドのアミノまたはカルボキシ末端またはその両端の妨害(blockage)は、天然に存在するポリペプチドおよび合成ポリペプチドで一般的であり、そのような改変は本発明のポリペプチドにも存在し得る。

ポリペプチド内に生じる改変は多くの場合ポリペプチドが生成される方法の機能であろう。組換えによって生成されるポリペプチドの場合、大部分の改変の性質および程度は、個々の宿主細胞の改変能力および問題のポリペプチドのアミノ酸配列に存在する改変シグナルによって決定されるであろう。例えば、グリコシル化パターンは異なる種類の宿主細胞間で変動するであろう。

本発明のポリペプチドは任意の適切な様式で調製することができる。そのようなポリペプチドには、単離された天然に存在するポリペプチド（例えば細胞培養から精製）、組換えにより生成されたポリペプチド（融合タンパク質を含む）、合成により生成されたポリペプチド、または前記方法を併用して生成されたポリペプチドが含まれる。

【0028】

本発明の第一の特徴の機能的に等価なポリペプチドは、LBDG2ポリペプチドと相同なポリペプチドであり得る。2つのポリペプチドは、前記ポリペプチドの一方の配列が他方のポリペプチドの配列に対して充分な同一性または類似性を有する場合、本明細書で用いられる用語のように“相同である”と称される。“同一性”とは、整列化した配列のどの特定の場所においても、アミノ酸残基が前記配列間で同一であることを示す。“類似性”は、整列化した配列のいずれの特定の場所においても、アミノ酸残基が前記配列間で類似の種類であることを示す。同一性および類似性の度合いは容易に計算できる（Computational Molecular Biology, A.M. Lesk ed., Oxford University Press, New York, 1988; Biocomputing. Informatics and Genome Projects, D.W. Smith ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part 1, A.M. Griffin and H.G. Griffin eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, G. von Heinje, Academic Press, 1987; および Sequence Analysis Primer, M. Gribskov and J. Devereux eds., M. Stockton Press, New York, 1991）。

【0029】

したがって、相同なポリペプチドには、LBDG2ポリペプチドの天然の生物学的変種（例えば前記ポリペプチドが由来した種における対立形質変種または地理的変種）および変異体（例えばアミノ酸置換、挿入または欠失を含む変異体）が含まれる。前記変異体は、1つまたは2つ以上のアミノ酸残基が保存的または非保存的アミノ酸残基（好ましくは保存的アミノ酸残基）で置換されているポリペプチドを含んでもよく、さらにそのような置換アミノ酸残基は遺伝コードでコードされたものでもそうでなくてもよい。典型的な前記の置換は、Ala、Val、LeuおよびIle間で；SerとThr間で；酸性残基AspとGlu間で；AsnとGln間

で；塩基性残基LysとArg間で；または芳香族残基PheとTyr間で生じる。特に好ましいものは、いくつか（すなわち5から10、1から5、1から3、1から2、または単に1つ）のアミノ酸が任意の組合せで置換または欠失または付加された変種である。特に好ましいものは、タンパク質の特性および活性を変化させないサイレント置換、付加および欠失である。さらにこれに関して特に好ましいものは保存的置換である。

前記変異体にはまた、1つまたは2つ以上のアミノ酸残基が置換基を含むポリペプチドも含まれる。

典型的には、2つのポリペプチド間で80%を越える同一性（好ましくは特定の領域で）が機能的等価物を示すと考えられる。好ましくは、本発明の第一の特徴の機能的に等価なポリペプチドは、LBDG2ポリペプチドに関して、またはその活性なフラグメントに関して80%を越える配列同一性を有する。より好ましいポリペプチドは、LBDG2ポリペプチドまたはその活性なフラグメントに関してそれぞれ85%、90%、95%、98%または99%を越える同一性を有する。

【0030】

本明細書で言及される同一性のパーセンテージは、BLASTバージョン2.1.3でNCBI(the National Center for Biotechnology Information; <http://www.ncbi.nlm.nih.gov/>)によって特定されたデフォルトパラメーター{Blosum62マトリックス；ギャップ開放(open)ペナルティー=11およびギャップ伸長(extension)ペナルティー=1}を用いて決定されるとおりである。

本件では、LBDG2ポリペプチドの好ましい活性フラグメントは、LBDG2核内ホルモンレセプターリガンド結合ドメイン領域を含むもの、および残基PHE1174、VAL1177、ASP1181、ASN1182、LEU1185およびLEU1186、または等価な残基を有する“LBDモチーフ”を保有するものである。“等価な残基”とは、“LBDモチーフ”残基と等価である残基を意味するが、ただし核内ホルモンレセプターリガンド結合ドメイン領域が核内ホルモンレセプターリガンド結合ドメインとしての活性を保持していることを条件とする。例えば、PHE1174はLEU、ILE、ALA、VAL、MET、TYRまたはTRPによって置換することができる。例えば、VAL1177はLEU、ILE、ALA、MET、PHE、TYRまたはTRPと置換することができる。例えば、ASP1181はGLUと置換することができる。例えば、ASN1182は、GLN、ARG、HIS、LYS、SER、THRと置換することができる。例えば、LEU1185は、ILE、ALA、VAL、MET、PHE、TYRまたはTRPと置換することができる。例えば、LEU1186は、ILE、ALA、VAL、MET、PHE、TYRまたはTRPと置換することができる。したがって本発明のこの特徴は、LBDG2ポリペプチドの核内ホルモンレセプターリガンド結合ドメイン領域に関して80%を越える同一性、好ましくは85%、90%、95%、98%または99%を越える同一性をそれぞれ有するポリペプチド、および、PHE1174、VAL1177、ASP1181、ASN1182、LEU1185およびLEU1186、または等価な残基をもつ“LBDモチーフ”を保有するポリペプチドを含む。上記で考察したように、LBDG2核内ホルモンレセプターリガンド結合ドメイン領域は、LBDG2ポリペプチド配列の残基1104と残基1309の間に広がっていると考えられる。

【0031】

本発明の第一の特徴の機能的に等価なポリペプチドはまた、1つまたは2つ以上の構造的アラインメント技術を用いて同定されたポリペプチドであってもよい。例えば、バイオペンジウム検索データベースを作製するために用いられた検索ツールの1つの特徴を形成するインファーマチカゲノムスレッダー（商標）(Inpharmatica Genome Treader)技術を用いて（同時係属国際特許出願PCT/GB01/01105を参照されたい）、LBDG2ポリペプチドと比較したとき、低い配列同一性を有するが、LBDG2ポリペプチド配列と顕著な構造的相同性を共有するがゆえに、核内ホルモンレセプターリガンド結合ドメイン活性をもつと予測される、現在は未知の機能を有するポリペプチドを同定することができる。

“顕著な構造的相同性”とは、インファーマチカゲノムスレッダー（商標）が、2つのタンパク質またはタンパク質領域が、少なくとも10%、より好ましくは少なくとも20%、30%、40%、50%、60%、70%、80%、90%および前記を越える確実性で構造的相同性を共有することを予測することを意味する。前記インファーマチカゲノム

10

20

30

40

50

スレッダー（登録商標）の前記確実性の値は以下のように計算される。既知の構造を有する配列をもっぱら使用して、初めに一組の比較をインファーマチカゲノムスレッダー（登録商標）を用いて行った。いくつかの比較は（構造を基準にして）関連することが判明しているタンパク質間で行った。続いてニューラルネットワークを、CATH構造分類（www.biochem.ucl.ac.uk/bsm/cath）から得られる既知の関係と既知の非関係とを最も良く識別することを必要とすることに基づいて調整(train)した。これによって0と1の間のニューラルネットワークスコアが得られた。しかしながら、一方で関連するタンパク質の数および無関係のタンパク質の数は既知であるので、前記ニューラルネットワークの結果を小群に分配し、正確な結果のパーセンテージを経験的に計算することが可能であった。このようにして、バイオペンジウム検索データベースにおける全ての真正の予測はニューラルネットワークスコアが付随しており、信頼百分率は、インファーマチカゲノムスレッダー（登録商標）がいかにか良好なトレーニング/テストセットであるかを反映したものである。

10

【0032】

LBDG2の構造的相同体は、LBDG2核内ホルモンレセプターリガンド結合ドメイン領域と構造的相同性を共有し、“LBDモチーフ”残基PHE1174、VAL1177、ASP1181、ASN1182、LEU1185およびLEU1186、または等価な残基を保有するはずである。そのような構造的相同体は、前記ポリペプチド配列と顕著な構造的相同性を共有し、さらに“LBDモチーフ”残基を保有するがゆえに、核内ホルモンレセプターリガンド結合ドメイン活性を有すると予測される。

20

本発明の第一の特徴のポリペプチドにはまた、LBDG2ポリペプチドのフラグメント、LBDG2ポリペプチドのフラグメントの機能的等価物、およびLBDG2ポリペプチドの機能的等価物のフラグメントが含まれるが、ただし前記機能的等価物およびフラグメントは核内ホルモンレセプターリガンド結合ドメイン活性を保持するか、またはLBDG2ポリペプチドと共通の抗原決定基を有することを条件とする。

本明細書において用いる、“フラグメント”という用語は、LBDG2ポリペプチド、またはその機能的等価物のアミノ酸配列の一部（全部ではなく）と同じアミノ酸配列を有するポリペプチドを指す。フラグメントは、前記配列に由来する少なくともn個の連続するアミノ酸を含むべきである。さらに、個々の配列に応じて、nは7またはそれより大きい（例えば8、10、12、14、16、18、20またはそれより大きい）。小さなフラグメントは抗原決定基を構成することができる。

30

本発明のこの特徴の好ましいポリペプチドフラグメントは、LBDG2ポリペプチドのそれぞれLBDG2核内ホルモンレセプターリガンド結合ドメイン領域と本明細書で明確にする領域を含むフラグメントである。これらの領域は、核内ホルモンレセプターリガンド結合ドメインと注釈付けされた領域である。

【0033】

LBDG2ポリペプチドの場合、前記領域は残基1104と残基1309の間に広がっていると考えられる。

前記フラグメントの変種は、本発明のこの特徴の態様として含まれるが、ただしこれら変種は核内ホルモンレセプターリガンド結合ドメインとしての活性を保有することを条件とする。

40

ある観点では、“変種”という用語は前記ポリペプチドフラグメントの伸長型または短縮型を含むことが意図される。

伸長型変種の場合、LBDG2ポリペプチド配列内のこれら境界のC末端および/またはN末端にさらに付加された残基が前記ポリペプチドフラグメントに含まれる場合、LBDG2ポリペプチドの核内ホルモンレセプターリガンド結合ドメイン領域が正確に折り畳まれ、核内ホルモンレセプターリガンド結合ドメインの活性を示すであろうということは想像に難くない。例えば、LBDG2ポリペプチド配列、または相同な配列に由来する5、10、20、30、40または50またはそれより多い付加アミノ酸残基はLBDG2ポリペプチドの核内ホルモンレセプターリガンド結合ドメイン領域の境界のC末端および/またはN末端のい

50

ずれか一方または両方に含まれ、前記ポリペプチドフラグメントは、正確に折り畳まれる能力を損なうことなく核内ホルモンレセプターリガンド結合ドメイン活性を示し得る。LBDG2ポリペプチドの短縮型変種の場合、1つまたは2つ以上のアミノ酸残基をLBDG2ポリペプチドの核内ホルモンレセプターリガンド結合ドメイン領域のC末端またはN末端の一方または両方で欠失させることができるが、ただし“LBDモチーフ”残基(PHE1174、VAL1177、ASP1181、ASN1182、LEU1185およびLEU1186)または等価な残基は無傷のまま維持され、欠失は、前記残基のいずれかが欠失するほど前記ポリペプチド配列内に深く伸長することはない。

【0034】

第二の観点では、“変種”という用語は、LBDG2ポリペプチドの核内ホルモンレセプターリガンド結合ドメイン領域と顕著な配列相同性を保有し、かつ“LBDモチーフ”残基(PHE1174、VAL1177、ASP1181、ASN1182、LEU1185およびLEU1186)または等価な残基を保有する上記で述べたポリペプチドフラグメントの相同体を含むが、ただし前記変種は核内ホルモンレセプターリガンド結合ドメインとしての活性を保持することを条件とする。相同体には、LBDG2ポリペプチドのうち、それぞれLBDG2核内ホルモンレセプターリガンド結合ドメイン領域と80%を越える同一性を保有するポリペプチド分子が含まれる。同一性パーセンテージは、BLASTバージョン2.1.3でNCBI(the National Center for Biotechnology Information; <http://www.ncbi.nlm.nih.gov/>)によって特定されるデフォルトパラメーター{Blosum62マトリックス; ギャップ開放ペナルティー=11およびギャップ伸長ペナルティー=1}を用いて決定されるとおりである。好ましくは、本発明のこの特徴のポリペプチドフラグメントの変種相同体は、LBDG2ポリペプチドのLBDG2核内ホルモンレセプターリガンド結合ドメイン領域に関してそれぞれ80%を越える同一性を有する。より好ましくは、変種ポリペプチドは、LBDG2ポリペプチドのLBDG2核内ホルモンレセプターリガンド結合ドメイン領域に関して85%、90%、95%、98%または99%を越える同一性をそれぞれ有するが、ただし前記変種は核内ホルモンレセプターリガンド結合ドメインとしての活性を保持することを条件とする。変種ポリペプチドはまた、上記で考察したポリペプチドフラグメントの短縮型相同体も含むが、ただし前記変種は核内ホルモンレセプターリガンド結合ドメインとしての活性を保持することを条件とする。

【0035】

本発明の第一の特徴のポリペプチドフラグメントは、例えばインファーマチカゲノムスレッダー(商標)によって同定される、LBDG2ポリペプチド配列のLBDG2核内ホルモンレセプターリガンド結合ドメイン領域によって規定されるポリペプチドフラグメントの構造と顕著な構造的相同性を示すポリペプチドフラグメントであり得る。したがって、LBDG2ポリペプチド配列のLBDG2核内ホルモンレセプターリガンド結合ドメイン領域によって規定されるポリペプチドフラグメントの構造的相同体であるポリペプチドフラグメントは、上記で定義されたフォールドのような、LBDG2ポリペプチドフラグメントによって採用されるフォールドと同じフォールドを採用するはずである。

LBDG2核内ホルモンレセプターリガンド結合ドメイン領域によって規定されるポリペプチドフラグメントの構造的相同体はまた、“LBDモチーフ”残基PHE1174、VAL1177、ASP1181、ASN1182、LEU1185およびLEU1186、または等価な残基も保持するはずである。

そのようなフラグメントは、“独立的存在(free-standing)”すなわち、他のアミノ酸もしくはポリペプチドの部分でなくてもよく、他のアミノ酸もしくはポリペプチドに融合されていなくてもよく、それらフラグメントがその部分または領域を構成するより大きなポリペプチドの内部に含まれていてもよい。より大きなポリペプチドの内部に含まれている場合は、本発明のフラグメントは最も好ましくは連続するただ1つの領域を形成する。例えばある種の好ましい態様は、前記フラグメントのアミノ末端に融合したプレ-および/またはプロ-ポリペプチド領域、および/または前記フラグメントのカルボキシ末端に融合した付加的領域を有するフラグメントに関する。しかしながら、いくつかのフラグメントが単一のより大きなポリペプチドの内部に含まれてもよい。

【0036】

10

20

30

40

50

本発明のポリペプチドまたはその免疫原性フラグメント（少なくとも1つの抗原決定基を含む）を用いて、例えばポリクローナルまたはモノクローナル抗体といった、前記ポリペプチドに免疫特異的なリガンドを作製することができる。そのような抗体は、本発明のポリペプチドを発現しているクローンを単離することもしくは同定すること、またはアフィニティークロマトグラフィーによって前記ポリペプチドを精製することに用いることができる。前記抗体はまた、当業者には明らかなように他の利用の中で特に診断的または治療的補助としても用いられ得る。

“免疫特異的”という用語は、前記抗体が、従来技術において他の関連ポリペプチドに対する親和性よりも本発明のポリペプチドに対して実質的に強い親和性を有することを意味する。本明細書で用いる“抗体”という用語は、完全な分子だけでなく問題の抗原決定基と結合することができるそのフラグメント、例えばFab、F(ab')₂およびFvも意味する。したがって、そのような抗体は本発明の第一の特徴のポリペプチドと結合する。

ポリクローナル抗体が所望される場合は、選択される哺乳類（例えばマウス、ウサギ、ヤギまたはウマ）は、本発明の第一の特徴のポリペプチドで免疫することができる。動物を免疫するために用いられるポリペプチドは、リコンビナントDNA技術によって誘導するか、または化学的に合成することができる。所望する場合には、前記ポリペプチドは担体タンパク質と結合させることができる。前記ポリペプチドと化学的に結合させることができる一般的に用いられる担体には、ウシ血清アルブミン、サイログロブリンおよびキーホールリンペットヘモシアニンが含まれる。続いて前記担体結合ポリペプチドを用いて動物を免疫することができる。血清は免疫した動物から採集され、既知の方法（例えばイムノアフィニティークロマトグラフィー）にしたがって処理される。

【0037】

本発明の第一の特徴のポリペプチドに対するモノクローナル抗体も、当業者は容易に生成できる。ハイブリドーマ技術を用いてモノクローナル抗体を作製する一般的な方法論は周知である（例えば以下を参照されたい：G. Kohler & C. Milstein, *Nature* 256: 495-497(1975); Kozbor et al., *Immunology Today* 4: 72(1983); Cole et al., 77-96 “Monoclonal Antibodies and Cancer Therapy”, Alan R. Liss, Inc. (1985)）。

本発明の第一の特徴のポリペプチドに対して生成されたモノクローナル抗体のパネル(panel)を種々の特性、すなわちアイソタイプ、エピトープ、親和性などについてスクリーニングすることができる。モノクローナル抗体は、それらを作らせた個々のポリペプチドの精製に特に有用である。あるいは、対象のモノクローナル抗体をコードする遺伝子を、例えば当技術分野で知られるPCR技術によってハイブリドーマから単離し、さらにクローニングし適切なベクターで発現させることができる。

非ヒト可変領域がヒト定常領域と結合または融合されているキメラ抗体（例えば以下を参照されたい：Proc. Natl. Acad. Sci. USA, 84: 3439(1987)）もまた有用であろう。

抗体は、例えばヒト化によって改変して各個体での免疫原性を減少させることができる（例えば以下を参照されたい：Jones et al., *Nature*, 321: 522(1986); Verhoeyen et al., *Science*, 239: 1534(1988); Kabat et al., *J. Immunol.*, 147: 1709(1991); Queen et al., *Proc. Natl. Acad. Sci. USA*, 86: 10029(1989); Gorman et al., *Proc. Natl. Acad. Sci. USA*, 88: 34181(1991); Hodgson et al., *Bio/Technology* 9: 421(1991)）

。本明細書で用いられる“ヒト化抗体”という用語は、非ヒトドナー抗体の重鎖および/または軽鎖の可変ドメイン中のCDRアミノ酸および選択した他のアミノ酸がヒト抗体の等価なアミノ酸に代えて置換されている抗体分子を指す。したがって、ヒト化抗体はヒトの抗体と密接に類似するがドナー抗体の結合能力を有する。

【0038】

また別の選択肢では、前記抗体は、2つの異なる抗原結合ドメインを有し、各ドメインは異なるエピトープに向けられている“二重特異性”抗体であってもよい。

ファージディスプレイ技術を用いて、本発明のポリペプチドに対する結合活性をもつ抗体をコードする遺伝子を、関連抗体の保有についてスクリーニングされたヒト由来のリンパ球のPCR増幅V-遺伝子のレパートリー、または未感作ライブラリーのいずれかから選

10

20

30

40

50

択することができる (J. McCafferty et al., (1990) Nature 348 : 552 - 554 ; J. Marks et al., (1992) Biotechnology 10 : 779 - 783)。前記抗体の親和性は、鎖のシャッフリングによって改善することもできる (T. Clackson et al., (1991) Nature 352 : 624 - 628)。

上記の技術によって作製された抗体は (ポリクローナルであれモノクローナルであれ)、免疫アッセイ、ラジオイムノアッセイ (RIA) または酵素結合免疫吸着アッセイ (ELISA) で試薬として用いることができるというさらに別の有用性を有する。前記の利用では、これら抗体は分析的に検出可能な試薬 (例えば放射性同位元素、蛍光分子または酵素) で標識することができる。

本発明の第二および第三の特徴の好ましい核酸分子は、配列番号 : 2 に記載のポリペプチド配列および機能的に等価なポリペプチドをコードするものである。前記機能的に等価なポリペプチドには、LBDG2ポリペプチドの活性なフラグメント、例えばLBDG2ポリペプチド配列のLBDG2核内ホルモンレセプターリガンド結合ドメイン領域を含むフラグメントフラグメント、またはその相同体が含まれる。

これら配列の一続きを包含する核酸分子は本発明のこの特徴の好ましい態様を形成する。これらの核酸分子は、本明細書に記載する方法および応用で用いることができる。好ましくは本発明の核酸分子は、本明細書に開示する配列に由来する少なくとも n 個の連続するヌクレオチドを含む。 n は、個々の配列に応じて 10 またはそれより大きい (例えば 12、14、15、18、20、25、30、35、40 またはそれより大きい)。

本発明の核酸分子は、上記で述べた核酸分子に相補的な配列も含む (例えばアンチセンスまたはプローブとしての目的のために)。

【0039】

本発明の核酸分子は、RNA (例えば mRNA)、または DNA (例えば cDNA、合成 DNA またはゲノム DNA を含む) の形態をとることができる。そのような核酸分子は、クローニングによって、化学合成によって、またはそれらを併用して得ることができる。前記核酸分子は、固相ホスホルアミダイト化学合成のような技術を用いるゲノムまたは cDNA ライブラリーからの化学合成によって、または生物体から分離することによって調製することができる。RNA 分子は一般的には DNA 配列の *in vitro* または *in vivo* 転写によって作製することができる。

核酸分子は二本鎖でも一本鎖でもよい。一本鎖 DNA はコード鎖 (センス鎖としても知られる) でも、非コード鎖 (アンチセンス鎖とも称される) でもよい。

“核酸分子” という用語には、DNA および RNA のアナログ (例えば改変骨格を含むもの)、並びにペプチド核酸 (PNA) も含まれる。本明細書で用いられる “PNA” という用語はアンチセンス分子または抗遺伝子 (anti-gene) 作用因子を指し、長さが少なくとも 5 ヌクレオチドであってアミノ酸残基のペプチド骨格と結合したオリゴヌクレオチドを含む。前記ペプチド骨格は好ましくはリジンで終わり、前記末端リジンは当該組成物に可溶性を付与する。PNA は PEG 化 (pegylated) されて細胞内での寿命が延長されてもよい {細胞内では、PNA は優先的に相補性一本鎖 DNA および RNA と結合して転写物伸長を停止させる (P.E. Nielsen et al. (1993) Anticancer Drug Des. 8 : 53 - 63) }。

【0040】

配列番号 : 2 のポリペプチドまたはその活性フラグメントをコードする核酸分子は、配列番号 : 1 に示した核酸分子のコード配列と同一でもよい。これらの分子は、遺伝コードの縮退の結果として、配列番号 : 2 のポリペプチドまたは LBDG2 ポリペプチドの活性フラグメント (例えば LBDG2 核内ホルモンレセプターリガンド結合ドメイン領域を含むフラグメント)、またはその相同体をコードする多様な配列を有することもできる。LBDG2 核内ホルモンレセプターリガンド結合ドメイン領域は、LBDG2 ポリペプチド配列の残基 1104 と残基 1309 の間に広がっていると考えられる。したがって配列番号 : 1 では、LBDG2 核内ホルモンレセプターリガンド結合ドメイン領域は、ヌクレオチド 3351 から 3968 を含む核酸分子によってコードされる。前記配列の一続きを包含する核酸分子、および前記配列の相同体は、本発明のこの特徴の好ましい態様を形成する。

配列番号：2のポリペプチドをコードする前記核酸分子には、それ自体で成熟なポリペプチドのコード配列；成熟ポリペプチドおよび付加コード配列（例えばリーダー配列または分泌配列、例えばプロ-、プリ-またはプレプロ-ポリペプチド配列をコードするもの）のためのコード配列；前述の付加的コード配列を伴う、または伴わないが、さらに付加的な非コード配列（非コード5'および3'配列を含む）を伴う成熟ポリペプチドのコード配列が含まれるが、ただしこれらに限定されない。前記の非コード5'および3'配列は、例えば転写される非翻訳配列で、転写（終止シグナルを含む）、リボソーム結合およびmRNA安定性で役割を果たすものである。前記核酸分子は、更なる官能性を提供するアミノ酸のような付加アミノ酸をコードする付加配列を含むこともできる。

【0041】

本発明の第二および第三の特徴の核酸分子は、本発明の第一の特徴のポリペプチドのフラグメントまたは機能的等価物およびそのフラグメントもコードし得る。

上記で考察したように、LBDG2ポリペプチドの好ましいフラグメントは、LBDG2核内ホルモンレセプターリガンド結合ドメイン領域を含むフラグメント、またはその相同体である。前記核内ホルモンレセプターリガンド結合ドメイン領域は、配列番号：1のヌクレオチド3351から3968を含む核酸分子によってコードされる。

本発明の機能的に等価な核酸分子は、天然に存在する変種（例えば天然に存在する対立形質変種）であっても、または前記分子は天然に存在することが知られていない変種であってもよい。前記のような天然に存在しない核酸分子の変種は、突然変異誘発技術（核酸分子、細胞または生物に対して適用される技術が含まれる）によって達成できる。

このような変種の中では、特にヌクレオチドの置換、欠失または挿入によって前述の核酸分子と異なる変種が挙げられる。置換、欠失または挿入は1つまたは2つ以上のヌクレオチドを含むことができる。変種はコード領域または非コード領域またはその両方が変化していてもよい。コード領域における変化は、保存的または非保存的なアミノ酸置換、欠失または挿入をもたらすことができる。

本発明の核酸分子はまた、多様な理由で、遺伝子生成物（ポリペプチド）のクローニング、プロセッシングおよび/または発現の改変を含む当技術分野で一般的に知られている方法を用いて操作され得る。ランダムフラグメント化によるDNAシャッフリングおよび遺伝子フラグメントおよび合成オリゴヌクレオチドのPCRリアッセムブリーは、ヌクレオチド配列の操作に用いられ得る技術に含まれる。位置特異的突然変異誘発を用いて、新規な制限部位の挿入、グリコシル化パターンの変更、コドンの優先性の変化、スプライシング変種の生成、変異の導入その他を行うことができる。

【0042】

本発明の第一の特徴のポリペプチドをコードする核酸分子は異種配列に連結され、それによって結合核酸分子が融合タンパク質をコードすることができるようにしてもよい。前記のような結合核酸分子は本発明の第二または第三の特徴に包含される。例えば、本発明のポリペプチドの阻害剤のためのペプチドライブラリーをスクリーニングするために、前記のような結合核酸分子を用いて、市販の抗体によって認識される融合タンパク質を発現させることは有用であろう。融合タンパク質はまた、本発明のポリペプチド配列と異種タンパク質配列との間に位置する切断部位を含むように操作し、それによって前記ポリペプチドを異種タンパク質から切り離して精製することができるようにしてもよい。

本発明の核酸分子にはまた、本発明のポリペプチドをコードする核酸分子と部分的に相補的であり、したがってコード核酸分子とハイブリダイズする（ハイブリダイゼーション）アンチセンス分子が含まれる。そのようなアンチセンス分子（例えばオリゴヌクレオチド）は、当業者にはよく知られるように、本発明のポリペプチドをコードする標的核酸を認識し、その標的核酸と特異的に結合してその転写を妨げるようにデザインすることができる（例えば以下の文献を参照されたい：J.S. Cohen, Trends in Pharm. Sci., 10: 435(1989)； J. Okano, Neurochem. 56: 560(1991)； J. O' Connor, Neurochem. 56: 560(1991)； Lee et al., Nucleic Acids Res. 6: 3073(1979)； Cooney et al., Science 241: 456(1988)； Dervan et al., Science 251: 1360(1991)）。

10

20

30

40

50

【0043】

本明細書で用いられる“ハイブリダイゼーション”という用語は、2つの核酸分子が水素結合によって互いに結合することを指す。典型的には、1つの分子が固相支持体に固定され、他方は溶液中で遊離しているであろう。続いて2つの分子を水素結合に適した条件下で互いに接触させる。前記結合に影響する因子には以下が含まれる：溶媒の種類および体積；反応温度；ハイブリダイゼーションの時間；攪拌；液相分子の固相支持体への非特異的結合を妨害する薬剤（デンハルト試薬、またはB L O T T O）；分子の濃度；分子の結合速度を増加させる化合物の使用（硫酸デキストランまたはポリエチレングリコール）；およびハイブリダイゼーションに続く洗浄条件のストリンジェンシー（Sambrook et al.（上掲書）を参照されたい）。

10

完全に相補的な分子と標的分子とのハイブリダイゼーションの阻害は、当業者に知られるハイブリダイゼーションアッセイを用いて調べることができる（例えばSambrook et al.（上掲書）を参照されたい）。したがって、実質的に相同な分子は、文献（G.M. Wahl and S.L. Berger, 1987, Methods Enzymol. 152: 399 - 407；A.R. Kimmel, 1987, Methods Enzymol. 152: 507 - 511）に開示されたように、完全に相同な分子と標的分子との結合を種々のストリンジェンシー条件下で競合して阻害するであろう。

【0044】

“ストリンジェンシー”とは、異なる分子の結合よりも非常に類似した分子の結合に適したハイブリダイゼーション反応の条件を指す。高ストリンジェンシーのハイブリダイゼーション条件は、以下を含む溶液（50%のホルムアミド、5倍のSSC（150mM NaCl、15mM クエン酸三ナトリウム）、50mMリン酸ナトリウム（pH 7.6）、5倍のデンハルト溶液、10%の硫酸デキストラン、および20μg/mLの変性せん断サケ精子DNA）中で42で一晩インキュベーションし、続いてフィルターを0.1倍のSSCで約65で洗浄すると定義される。低ストリンジェンシー条件は、ハイブリダイゼーション反応が35で実施されることを含む（Sambrook et al.（上掲書）を参照されたい）。好ましくは、ハイブリダイゼーションに用いられる条件は高ストリンジェンシーを構成するものである。

20

本発明のこの特徴の好ましい態様は、LBDG2ポリペプチド（配列番号：2）をコードする核酸分子の全長にわたって少なくとも80%同一である核酸分子、および前記のような核酸分子に対して実質的に相補的な核酸分子である。好ましい活性フラグメントは、LBDG2ポリペプチド配列のそれぞれLBDG2核内ホルモンレセプターリガンド結合ドメイン領域を含むフラグメントである。したがって、好ましい核酸分子は、LBDG2ポリペプチド配列の核内ホルモンレセプターリガンド結合ドメイン領域をコードする核酸分子の全長にわたって少なくとも80%同一であるものを含む。

30

本明細書で言及される同一性パーセンテージは、BLASTバージョン2.1.3でNCBI（the National Center for Biotechnology Information；<http://www.ncbi.nlm.nih.gov/>）によって特定されたデフォルトパラメーターを用いて決定されるとおりである。

【0045】

好ましくは、本発明のこの特徴の核酸分子は、配列番号：1に示された配列を有する核酸分子、この配列のヌクレオチド3351 - 3968を含む領域の全長にわたって少なくとも80%同一の領域を含むか、またはこれら核酸領域のいずれかと相補的である核酸分子を含む。この場合、前記の全長にわたって少なくとも90%、好ましくは少なくとも95%、より好ましくは少なくとも98%または99%同一の核酸分子が特に好ましい。これに関して、好ましい態様は、LBDG2ポリペプチドと実質的に同じ生物学的機能または活性を保持するポリペプチドをコードする核酸分子である。

40

本発明はまた、以下の工程を含む、本発明の核酸分子を検出する方法を提供する：（a）二重鎖を形成するハイブリダイゼーション条件下で本発明の核酸プローブを生物学的サンプルと接触させる工程；および（b）形成された前記の全ての二重鎖を検出する工程。

本発明にしたがって利用することができるアッセイに関連して下記でさらに考察するように、上記で述べた核酸分子をRNA、cDNAまたはゲノムDNAのためのハイブリダイゼーションプローブとして用い、LBDG2ポリペプチドをコードする完全長のcDNAおよびゲノムクロー

50

ンを単離し、さらに前記ポリペプチドをコードする遺伝子と高い配列類似性を有する相同遺伝子またはオーソログ遺伝子のcDNAまたはゲノムクローンを単離することができる。

【0046】

これに関しては、当技術分野で公知の他の技術の中で特に以下の技術を利用することができる。これらの技術は例示として下記で考察される。DNAのシーケンシングおよび解析のための方法は周知で、当技術分野では一般的に利用可能であり、本明細書で考察される本発明の態様の多くを実施するために実際に用いることができる。そのような方法では、DNAポリメラーゼIのKlenowフラグメント、シークエナーゼ(US Biochemical Corp., Cleveland, OH)、Taqポリメラーゼ(Perkin Elmer)、耐熱性T7ポリメラーゼ(Amersham, Chicago, IL)、またはポリメラーゼと校正エキソヌクレアーゼの組合せ(例えば市販(Gibco/BRL, Gaithersburg, MD)のELONGASE増幅キットで見出されるようなもの)のような酵素を利用することができる。好ましくは、シーケンシングプロセスは、例えばハミルトンマイクロラボ(Hamilton Micro Lab) 2200(Hamilton, Reno, NV)、ペルティエサーマルサイクラー(Peltier Thermal Cycler) PTC200(MJ Research, Watertown, MA)、ABIカタリスト並びに373および377DNAシークエンサー(Perkin Elmer)のような機器を用いて自動化することができる。

LBDG2ポリペプチドの機能と等価な機能を有するポリペプチド、特にLBDG2ポリペプチドのLBDG2核内ホルモンレセプターリガンド結合ドメイン領域と同等の機能をもつポリペプチドをコードする核酸分子を単離する方法の1つは、当技術分野で知られている標準的な方法を用いる、天然のプロープまたは人工的にデザインしたプロープによるゲノムライブラリーまたはcDNAライブラリーの探索である(例えば以下の文献を参照されたい: "Current Protocols in Molecular Biology", Ausubel et al.(eds). Greene Publishing Association and John Wiley Interscience, New York, 1989, 1992)。特に有用なプロープは、適切なコード遺伝子(配列番号: 1)、特に配列番号: 1のヌクレオチド3351-3968の領域に由来する核酸配列に対応する、または前記配列と相補的である、少なくとも15、好ましくは少なくとも30、さらに好ましくは少なくとも50の連続塩基を含むプロープである。

前記のようなプロープは分析的な検出が可能な試薬で標識して前記プロープの識別を容易にすることができる。有用な試薬には、放射性同位元素、蛍光色素、および検出可能な生成物の形成を触媒し得る酵素が含まれるが、ただしこれらに限定されない。これらのプロープを用いて、当業者は、ヒト、哺乳類または他の動物供給源から対象のタンパク質をコードするゲノムDNA、cDNAまたはRNAポリヌクレオチドの相補的なコピーを単離し、近縁配列、例えば前記のファミリー、タイプおよび/またはサブタイプに属するまた別のメンバーについて、前記の供給源をスクリーニングすることができるであろう。

【0047】

多くの場合、単離されるcDNA配列は不完全で、ポリペプチドをコードする領域は短く(通常は5'末端で)切断されているであろう。完全長cDNAを得るために、または短いcDNAを伸長させるために、いくつかの方法が利用可能である。そのような配列は、部分的なヌクレオチド配列を用い、上流の配列(例えばプロモーターおよび調節エレメント)を検出するための当技術分野で公知の種々の方法を用いて伸長させることができる。例えば、使用され得るある方法は、cDNA末端迅速増幅法(RACE; 例えば以下を参照されたい: Frohman et al., Proc. Natl. Acad. Sci. USA(1988) 85: 8998-9002)に基づく。前記技術の最近の改変(例えばマラソン(Marathon)(商標)技術(Clontech Laboratories Inc.)により例示される)は、より長いcDNAの検索を顕著に単純化した。わずかに異なる技術("制限部位"PCRと称される)では、普遍的プライマーを用いて既知の遺伝子座に近接する未知の核酸配列が検索される(G. Sarkar (1993) PCR Methods Applic. 2: 318-322)。逆PCRもまた、既知の領域に基づく多様なプライマーを用いた配列の増幅または伸長に用いることができる(T. Triglia et al. (1988) Nucleic Acids Res. 16: 8186)。使用できる別の方法は捕捉PCRで、前記は、ヒトおよび酵母の人工染色体DNAで既知配列に近接するDNAフラグメントのPCR増幅を含む(M. Lagerstrom et al. (1991) PCR

Methods Applic. 1: 111 - 119)。未知の配列を検索するために利用可能なまた別の方法はパーカーの方法である (J.D. Parker et al.(1991) Nucleic Acids Res. 19: 3055 - 3060)。さらに、ゲノムDNAを歩行させるためにPCR、入れ子(nested)プライマーおよびプロモーターファインダー(PromoterFinder)(商標)ライブラリー(Clontech, Palo Alto, CA)を用いてもよい。この方法ではライブラリーのスクリーニングが不要で、イントロン/エクソン結合部を見出すことに有用である。

【0048】

完全長cDNAをスクリーニングする場合、より大きなcDNAを包含するためにサイズ選択を実施したライブラリーを用いることが好ましい。さらにまた、遺伝子の5'領域を含む配列をより多く含むという点でランダムプライミングした(random-primed)ライブラリーが好ましい。ランダムプライミングしたライブラリーの使用は、オリゴd(T)ライブラリーが完全長cDNAを生成できない状況で特に好ましいであろう。ゲノムライブラリーは、5'非転写調節領域に配列を伸長させるために有用であろう。

本発明のある態様では、染色体局在化のために本発明の核酸分子を用いることができる。この技術では、核酸分子は個々のヒト染色体上の特定の位置に対して特異的に標的化され、個々のヒト染色体上の特定の位置にハイブリダイズさせることができる。本発明の関連配列の染色体上へのマッピングは、遺伝子関連疾患に関する配列の相関性確認において重要な工程である。いったん染色体の正確な位置に配列がマッピングされたら、前記配列の染色体上の物理的な位置を遺伝子地図データと相関させることができる。そのようなデータは、例えば以下で見出すことができる: V. McKusick, Mendelian Inheritance in Man (ジョーンズホプキンス大学、ウェルチ医学図書館を通じてオンラインで入手可能である)。同じ染色体領域にマッピングされた遺伝子と疾患との関係を、次に連鎖解析(物理的に近接する遺伝子の同時遺伝(coinheritance))によって同定する。これにより、ポジショナルクローニングまたは他の遺伝子発見技術を用いて疾患遺伝子を検索する研究者に貴重な情報が提供される。いったん疾患または症候群の位置が遺伝連鎖によって特定のゲノム領域で大まかに局在化されたら、前記領域にマッピングされるいずれの配列も、更なる解析のための関連または調節遺伝子となることができる。前記核酸分子はまた、正常な個体、キャリア個体または罹患個体間で転座、逆位などによる染色体位置上の相違を検出するために用いることができる。

【0049】

本発明の核酸分子はまた、組織分布同定(tissue localisation)のために貴重である。そのような技術は、ポリペプチドをコードするmRNAの検出によって組織中の前記ポリペプチドの発現パターンの決定を可能にする。これらの技術にはin situハイブリダイゼーション技術およびヌクレオチド増幅技術(例えばPCR)が含まれる。これらの研究から得られる結果は、生物内での前記ポリペプチドの正常な機能を示唆する。さらに、変異遺伝子によってコードされるmRNAの発現パターンと正常mRNA発現パターンとの比較研究によって、変異ポリペプチドの疾患における役割に対する貴重な洞察が提供される。そのような不適切な発現は時間的、位置的または量的性質を有する場合もある。

本発明のベクターは本発明の核酸分子を含み、クローニングベクターでも発現ベクターでもよい。本発明のベクターで形質転換、トランスフェクトまたは形質導入され得る本発明の宿主細胞は、原核細胞でも真核細胞でもよい。

本発明のポリペプチドは、宿主細胞内に含まれるベクター中の前記ポリペプチドをコードする核酸分子の発現によってリコンビナント形態で調製され得る。前記のような発現方法は当業者によく知られており、以下の文献により詳細に記述され得る: Sambrook et al. (上掲書)およびFernandez & Hoeffler(1998, eds. "Gene expression systems. Using nature for the art of expression", Academic Press, San Diego, London, Boston, New York, Sydney, Tokyo, Toronto)。

【0050】

一般的には、要求されるホストでポリペプチドを生成させるために核酸分子の維持、増殖または発現に適したいずれの系またはベクターも用いることができる。周知であり日常的

10

20

30

40

50

である種々の技術のいずれによっても（例えば前掲書（Sambrook et al.）に記載されたようなもの）、適切なヌクレオチド配列を発現系に挿入することができる。一般的には、コード遺伝子は制御エレメント（例えばプロモーター、リボソーム結合部位（細菌での発現の場合）、および場合によってオペレーター）の制御下に置かれ、それによって所望のポリペプチドをコードするDNA配列を形質転換宿主細胞でRNAに転写させることができる。適切な発現系の例には、例えば染色体系、エピソーム系およびウイルス由来系、例えば以下に由来するベクターが含まれる：細菌プラスミド、バクテリオファージ、トランスポゾン、酵母エピソーム、挿入エレメント、酵母染色体エレメント、ウイルス、例えばバキュロウイルス、パポバウイルス（例えばSV40）、ワクシニアウイルス、アデノウイルス、鶏痘ウイルス、仮性狂犬病ウイルスおよびレトロウイルス、または上記の組み合わせ、例えばプラスミドとバクテリオファージの遺伝子エレメントに由来するもの（例えばコスミドおよびファージミドを含む）。ヒト人工染色体（HAC）もまた、プラスミドに包含させ発現させるには大きいDNAフラグメントを搬送するために用いることができる。

10

【0051】

特に適切な発現系には、リコンビナントバクテリオファージ、プラスミドまたはコスミドDNA発現ベクターで形質転換された微生物（例えば細菌）；酵母発現ベクターで形質転換された酵母；ウイルス発現ベクター（例えばバキュロウイルス）を感染させた昆虫細胞系；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；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）で形質転換した植物細胞系；または動物細胞系が含まれる。無細胞翻訳系もまた本発明のポリペプチドの生成に用いることができる。

20

本発明のポリペプチドをコードする核酸分子の宿主細胞への導入は、多くの標準的な実験室マニュアル（例えば、Davis et al., Basic Methods in Molecular Biology (1986)および上掲書（Sambrook et al.））に記載された方法によって達成できる。特に適切な方法には、リン酸カルシウムトランスフェクション、DEAEデキストラン仲介トランスフェクション、トランスベクション、マイクロインジェクション、陽イオン脂質仲介トランスフェクション、エレクトロポレーション、トランスダクション、スクレイプローディング(scrape loading)、弾道導入または感染が含まれる（以下を参照されたい：Sambrook et al. (1989) 上掲書；Ausbel et al. (1991) 上掲書；Spector, Goldman & Leinward, (1998)）。

30

【0052】

コード核酸分子は、所望であれば、例えばシグナルペプチドまたはリーダー配列のような制御配列をコードする配列を（例えば翻訳されたポリペプチドの小胞体内、細胞周辺腔または細胞外環境への分泌のために）含んでいても含んでいなくてもよい。このシグナルは前記ポリペプチドにとって内因性であっても異種シグナルであってもよい。リーダー配列は、翻訳後プロセッシングで細菌ホストによって取り除くことができる。

コントロール配列の他に、宿主細胞の増殖に関連して前記ポリペプチドの発現の調節を可能にする調節配列を付加することが望ましい場合がある。調節配列の例は、化学的または物理的刺激（調節化合物の存在を含む）または多様な温度もしくは代謝条件にตอบสนองして遺伝子の発現を増加させたり低下させたりする配列である。調節配列は、ベクターの非翻訳領域、例えばエンハンサー、プロモーター並びに5'および3'非翻訳領域である。これらは、宿主細胞タンパク質と相互作用して、転写および翻訳を実行する。そのような調節配列は、その強さおよび特異性を変化させることができる。用いられるベクター系および宿主に依存して、多くの適切な転写および翻訳エレメント（構成性および誘発性プロモーターを含む）を用いることができる。例えば、細菌系でクローニングするときは、誘発性プロモーター、例えばBluescriptファージミド（Stratagene, La Jolla, CA）またはpSport1（商標）プラスミド（Gibco BRL）などのハイブリッドlacZプロモーターを用いることができる。バキュロウイルスポリヘドリン(polyhedrin)プロモーターは昆虫細胞で用いることができる。植物細胞ゲノムに由来するプロモーターまたはエンハンサー（例えば熱ショック、RUBISCOおよび貯蔵タンパク質遺伝子）または植物ウイルスに由来するプロ

40

50

モーターまたはエンハンサー（例えばウイルスプロモーターまたはリーダー配列）はベクターへクローニングすることができる。哺乳類細胞系では、哺乳類遺伝子由来または哺乳類ウイルス由来のプロモーターが好ましい。配列の多数コピーを含む細胞株の作製が必要な場合、SV40またはEBVをベースにしたベクターを適切な選択マーカーとともに用いることができる。

【0053】

発現ベクターは、特定の核酸コード配列を適切な調節配列とともにベクター内に配置させることができるように構築される。前記コード配列の調節配列に関する位置および向きは、前記コード配列が調節配列の“制御下”で転写されるような位置および向きである（すなわちコントロール配列にてDNA分子と結合するRNAポリメラーゼは前記コード配列を転写する）。いくつかの事例では、前記配列を適切な向きで制御配列に付属させることができるように（すなわちリーディングフレームを維持するために）、前記配列を改変する必要があるであろう。

10

コントロール配列および他の調節配列は、ベクターへの挿入の前に核酸コード配列に連結させることができる。あるいは、コード配列は、コントロール配列および適切な制限部位を既に含む発現ベクターへ直接クローニングすることができる。

長期的なりコンビナントポリペプチドの高収量の生成のためには、安定な発現が好ましい。例えば、対象のポリペプチドを安定に発現する細胞株は、ウイルスの複製起点および/または内因性発現エレメント並びに選択マーカー（同じベクターまたは別個のベクターに存在する）を含む発現ベクターを用いて形質転換させることができる。ベクターの導入に続き、選択培地に切り替える前に細胞を栄養(enriched)培地で1-2日間増殖させることができる。選択マーカーの目的は、選択に対する耐性を付与することで、選択マーカーの存在によって、導入された配列をうまく発現する細胞の増殖および回収が可能になる。安定に形質転換された細胞の耐性クローンは、細胞の種類に適した組織培養技術を用いて増殖させることができる。

20

【0054】

発現のためのホストとして利用可能な哺乳類細胞株は当技術分野で公知であり、米国菌培養収集所(American Type Culture Collection, ATCC)から入手可能な多くの不朽化細胞株が含まれる。前記には、例えばチャイニーズハムスター卵巣細胞(CHO)、HeLa、ベイビーハムスター腎(BHK)、サル腎(COS)、C127、3T3、BHK、HEK293、ボウズ(Bowes)メラノーマおよびヒト肝細胞癌(例えばHepG2)細胞および多数の細胞株が含まれるが、これだけに限られない。

30

バキュロウイルス系では、バキュロウイルス/昆虫細胞発現系のための材料は、特にインビトロジェン(Invitrogen, San Diego, CA)からキットの形態で(“MaxBac”キット)商業的に入手可能である。前記の技術は一般的に当業者に知られており、文献には完全に記載されている(Summers & Smith, Texas Agricultural Experiment Station Bulletin No.1555(1987))。この系での使用に特に適切な宿主細胞には昆虫細胞、例えばドロソフィラ(Drosophila)S2およびスポドプテラ(Spodoptera)Sf9細胞が含まれる。

当技術分野で公知である多くの植物細胞培養および植物体(whole plant)遺伝子発現系が存在する。適切な植物細胞遺伝子発現系の例には米国特許第5,693,506号、5,659,122号および5,608,143号に記載されたものが含まれる。植物細胞培養における遺伝子発現の別の例は、文献に記載されている(Zenk (1991) Phytochemistry 30: 3861 - 3863)。

40

特に、プロトプラストを単離し、これを培養して完全な再生植物を形成することが可能な植物は全て利用することができ、それによって移入遺伝子を含む完全な植物が回収される。特に、サトウキビ、サトウダイコン、綿花、果実および他の樹木、マメ類および野菜の主要な種の全てを含む(ただしこれらに限定されない)全ての植物は、培養細胞または組織から再生させることができる。

【0055】

特に好ましい細菌宿主細胞の例には連鎖球菌、ブドウ球菌、大腸菌(E. coli)、ストレプトマイセスおよびバチルス・ズブチリス(Bacillus subtilis)細胞が含まれる。

50

真菌での発現に特に適切な宿主細胞の例には酵母細胞（例えば*S. cerevisiae*）およびアスペルギルス細胞が含まれる。

形質転換細胞株の回収に用いることができる選択系は、当技術分野で公知である。例えば、単純ヘルペスウイルスチミジンキナーゼ（M. Wigler et al. (1977) Cell 11: 223 - 32）およびアデニンホスホリボシルトランスフェラーゼ（I. Lowy et al. (1980) Cell 22: 817 - 23）の遺伝子が含まれ、前記はそれぞれtk-またはaprt[±]細胞で用いることができる。

さらにまた、抗代謝物質耐性、抗生物質耐性または除草剤耐性を選択基準として用いてもよい。例えばジヒドロ葉酸レダクターゼ（DHFR）はメトトレキセートに対する耐性を付与し（M. Wigler et al. (1980) Proc. Natl. Acad. Sci. 77: 3567 - 70）、nptはアミノグリコシド系ネオマイシンおよびG-418に耐性を付与し（F. Colbere - Garapin et al. (1981) J. Mol. Biol. 150: 1 - 14）、さらにalsまたはpatはそれぞれクロロスルフロン（chlorosulfuron）およびホスフィノトリシン（phosphinotricin）アセチルトランスフェラーゼに対する耐性を付与する。さらに別の選択遺伝子が報告されており、当業者にはそれらの例は明白であろう。

10

【0056】

マーカー遺伝子の発現の有無は対象の遺伝子も存在することを示唆するが、対象の遺伝子の存在および発現を確認する必要がある。例えば、関連する配列がマーカー遺伝子配列内に挿入されている場合、適切な配列を含む形質転換細胞はマーカー遺伝子機能が存在しないことによって識別することができる。あるいは、マーカー遺伝子は、ただ1つのプロモーターの制御下で本発明のポリペプチドをコードする配列とともに直列に配置することができる。通常、誘発または選択に応答するマーカー遺伝子の発現は、直列遺伝子の発現も示している。

20

あるいは、本発明のポリペプチドをコードする核酸配列を含み、前記ポリペプチドを発現する宿主細胞は、当業者に知られている多様な方法で同定することができる。前記方法には、DNA-DNAまたはDNA-RNAハイブリダイゼーションおよびタンパク質バイオアッセイ、例えば蛍光活性化細胞ソーティング（FACS）またはイムノアッセイ技術（例えば酵素結合免疫吸着アッセイ（ELISA）および放射性イムノアッセイ（RIA））が含まれ（ただしこれらに限定されない）、核酸またはタンパク質の検出および/または定量的ためのメンブレン、溶液またはチップをベースとする技術が含まれる（例えば以下を参照されたい：R. Hampton et al. (1990) Serological Methods, A Laboratory Manual, APS Press, St Paul, MN; およびD.E. Maddox et al. (1983) J. Exp. Med. 158: 1211 - 1216）。

30

多様な標識および結合技術が当業者に知られており、種々の核酸およびアミノ酸アッセイで用いることができる。本発明のポリペプチドをコードする核酸分子に関連する配列を検出するための標識ハイブリダイゼーションプローブまたはPCRプローブの作製手段には、標識したポリヌクレオチドを用いるオリゴ標識、ニックトランスレーション、末端標識またはPCR増幅が含まれる。あるいは、本発明のポリペプチドをコードする配列をベクターにクローニングしてmRNAプローブを作製することができる。そのようなベクターは当技術分野で公知で、商業的に入手可能であり、適切なRNAポリメラーゼ（例えばT7、T3またはSP6）および標識ヌクレオチドを添加することによりin vitroでRNAプローブを合成することに用いることができる。これらの方法は、種々の市販キット（Pharmacia & Upjohn (Kalamazoo, MI); Promega (Madison, WI); U.S. Biochemical Corp., (Cleveland, OH))を用いて実施することができる。

40

適切なレポーター分子または標識（前記は検出を容易にするために用いることができる）には、放射性核種、酵素および蛍光、化学発光または色素生産性物質の他に基質、コファクター、阻害剤、磁性粒子などが含まれる。

【0057】

本発明の核酸分子はまたトランスジェニック動物（特にげっ歯類動物）の作製に用いることができる。そのようなトランスジェニック動物は本発明の別の特徴を構成する。これは、体細胞の改変によって局部的に、または遺伝性改変を導入する生殖細胞系療法によって

50

実施することができる。前記のようなトランスジェニック動物は、本発明のポリペプチドのモジュレーターとして有効な薬剤分子のための動物モデルを作製するために特に有用であろう。

ポリペプチドは、周知の方法によってリコンビナント細胞培養物から回収し精製することができる。前記周知の方法には、硫酸またはエタノール沈澱、酸性抽出、陰イオンまたは陽イオン交換クロマトグラフィー、リン酸化セルロースクロマトグラフィー、疎水性相互作用クロマトグラフィー、アフィニティークロマトグラフィー、ヒドロキシルアパタイトクロマトグラフィーおよびレクチンクロマトグラフィーが含まれる。高性能液体クロマトグラフィーは、精製に特に有用である。単離および精製の間にポリペプチドが変性した場合には、タンパク質のリフォールディングのためによく知られている技術を用いて活性化構造を再生することができる。

10

【0058】

所望の場合には、可溶性タンパク質の精製を容易にするポリペプチドドメインをコードするヌクレオチド配列に本発明のポリペプチドをコードする配列を結合させることにより特殊化したベクター構築物も、タンパク質の精製を容易にするために用いることができる。そのような精製促進ドメインの例には、金属キレートペプチド（例えば固定化金属上での精製を可能にするヒスチジン-トリプトファンモジュール、固定化免疫グロブリン上での精製を可能にするプロテインAドメイン、およびFLAG伸長/アフィニティー精製システム（Immunex Corp., Seattle, WA）で用いられるドメイン）が含まれる。切断可能なリンカー配列（例えばXA因子またはエンテロキナーゼ（Invitrogen, San Diego, CA）に特異的なもの）を精製ドメインと本発明のポリペプチドとの間に包含させて、精製を容易にすることに用いてもよい。そのような発現ベクターの1つは、チオレドキシニンまたはエンテロキナーゼ切断部位に先行するいくつかのヒスチジン残基と融合させた本発明のポリペプチドを含む融合タンパク質の発現を提供する。ヒスチジン残基は、IMAC（固定金属イオンアフィニティークロマトグラフィー；J. Porath et al.(1992) Prot. Exp. Purif. 3: 263 - 281）により精製を容易にし、一方、チオレドキシニンまたはエンテロキナーゼ切断部位は融合タンパク質からポリペプチドを精製するための手段を提供する。融合タンパク質を含むベクターについての考察は以下で提供される：D.J. Kroll et al.(1993) DNA Cell Biol. 12: 441 - 453）。

20

【0059】

スクリーニングアッセイで使用するためにポリペプチドを発現させる場合は、前記ポリペプチドを発現する宿主細胞の表面で前記ポリペプチドを生成させることが一般には好ましい。この場合、宿主細胞はスクリーニングアッセイで使用する前に、例えば蛍光活性化細胞ソーティング（FACS）またはイムノアフィニティー技術のような技術を用いて収穫することができる。ポリペプチドが培養液中に分泌される場合は、前記培養液を回収して発現されたポリペプチドを回収および精製することができる。ポリペプチドが細胞内で生成される場合、ポリペプチドを回収する前に、先ず初めに細胞を溶解させねばならない。本発明のポリペプチドを用いて、種々の薬剤スクリーニング技術のいずれかで、化合物ライブラリーをスクリーニングすることができる。そのような化合物は、本発明のポリペプチドの遺伝子発現レベルまたは活性レベルを活性化（作働）させ、または阻害（拮抗）することができる。それらは本発明のさらなる特徴を形成する。好ましい化合物は、本発明の第一の特徴のポリペプチドをコードする天然の遺伝子の発現を変化させることに有効であるか、または本発明の第一の特徴のポリペプチドの活性を調節することに有効である。アゴニスト化合物またはアンタゴニスト化合物は、例えば細胞、無細胞調製物、化学物質ライブラリーまたは天然生成物の混合物から単離することができる。これらのアゴニストまたはアンタゴニストは、天然または改変された基質、リガンド、酵素、レセプターまたは構造的もしくは機能的模倣物質であってよい。前記のスクリーニング技術の適切な概論のためには以下を参照されたい：Coligan et al.(1991) Current Protocols in Immunology 1(2): Chapter 5。

30

40

【0060】

50

おそらく良好なアンタゴニストと考えられる化合物は、本発明のポリペプチドと結合し、結合したときに前記ポリペプチドの生物学的作用を誘発しない分子である。潜在的なアンタゴニストには、本発明のポリペプチドと結合し、それによって本発明のポリペプチドの活性を阻害または消滅させる小型有機分子、ペプチド、ポリペプチドおよび抗体が含まれる。そのようなやり方で、前記ポリペプチドと正常な細胞の結合分子との結合が阻害され、その結果前記ポリペプチドの正常な生物学的活性が阻害され得る。

このようなスクリーニング技術で用いられる本発明のポリペプチドは溶液中で遊離していても、固相支持体に固定されていても、細胞表面に保持されていても、または細胞内に位置していてもよい。一般に、このようなスクリーニングの方法は、前記のポリペプチドを発現している適切な細胞または細胞膜を用いることを含み、前記細胞または細胞膜をテスト化合物と接触させて、結合または機能的な応答の刺激もしくは阻害を観察する。続いて前記テスト化合物と接触させた細胞の機能的応答を、前記テスト化合物と接触させなかったコントロール細胞と比較する。前記のようなアッセイによって、テスト化合物が前記ポリペプチドの活性化によって発生するシグナルをもたらすか否かを、適切な検出系を用いて評価する。活性化の阻害剤は、一般的には既知のアゴニストの存在下でアッセイされ、テスト化合物の存在下でアゴニストによる活性化の影響が観察される。

或いは、単純な結合アッセイを用いてもよい。この場合、テスト化合物のポリペプチド保持表面への付着が直接または間接的にテスト化合物と結合させた標識手段によって検出されるか、または標識競合物質を用いた競合を含むアッセイで検出される。別の態様では、競合薬剤のスクリーニングアッセイを用いることができる。この場合、ポリペプチドと特異的に結合することができる中和抗体がテスト化合物と結合について競合する。このようにして、前記抗体を用いて前記ポリペプチドに対して特異的な結合親和性を保有する一切のテスト化合物の存在を検出することができる。

前記ポリペプチドをコードするmRNAの細胞内の生成に対する添加テスト化合物の影響を検出するアッセイをデザインすることもできる。例えば、当技術分野で公知の標準的な方法によりモノクローナルまたはポリクローナル抗体を用いてポリペプチドの分泌レベルまたは細胞結合レベルを測定するELISAを構築することができ、それを用いて適切に操作した細胞または組織からのポリペプチド生成を阻害または増強し得る化合物について検索することができる。続いて、前記ポリペプチドとテストされる化合物との結合複合体の生成を測定することができる。

【0061】

使用され得る薬剤スクリーニングのための別の技術は、対象のポリペプチドに対して適切な結合親和性を有する化合物の高速大量処理スクリーニングを提供する（国際特許出願WO 84/03564を参照されたい）。前記方法では、多数の異なる小型のテスト化合物が固相基質上で合成され、次に本発明のポリペプチドと反応させられ洗浄され得る。ポリペプチドを固定する方法の1つは非中和抗体を使用することである。続いて、当技術分野で周知の方法を用いて結合ポリペプチドを検出することができる。精製ポリペプチドはまた、前述の薬剤スクリーニング技術で使用するためにプレートに直接被覆させることができる。

本発明のポリペプチドを用いて、膜結合レセプターまたは可溶性レセプターを当技術分野で公知の標準的なレセプター結合技術により同定することができる。前記標準的な技術は、例えばリガンド結合アッセイおよび架橋アッセイであり、このアッセイでは、ポリペプチドは放射性同位体で標識されているか、化学的に改変されているか、またはその検出もしくは精製を容易にするペプチド配列と融合されており、推定上のレセプター供給源（例えば細胞の組成物、細胞膜、細胞上清、組織抽出物または体液）とインキュベートされる。結合有効性は、生物物理的技術、例えば表面プラズモン共鳴および分光法を用いて測定することができる。結合アッセイは、レセプターの精製およびクローニングのために用いることができるが、ポリペプチドとそのレセプターとの結合に競合する前記ポリペプチドのアゴニストおよびアンタゴニストを同定するためにも用いることができる。スクリーニングアッセイを実施する標準的な方法は当技術分野ではよく理解されている。

【0062】

10

20

30

40

50

本発明はまた、上記で述べたアゴニスト、アンタゴニスト、リガンド、レセプター、基質および酵素を同定する方法で有用なスクリーニングキットを含む。

本発明は、上記アゴニスト、アンタゴニスト、リガンド、レセプター、基質および酵素、並びに上記で述べた方法によって発見される、本発明のポリペプチドの活性または抗原性を調節する他の化合物を含む。

本発明はまた、本発明のポリペプチド、核酸、リガンドまたは化合物とともに適切な医薬担体を含む医薬組成物を提供する。これらの組成物は、下記で詳細に説明するように、治療用もしくは診断用試薬として、ワクチンとして、または他の免疫原性組成物として適切であり得る。

本明細書で用いられる専門用語にしたがえば、ポリペプチド、核酸、リガンドまたは化合物 { X } を含む組成物は、組成物中の X + Y の合計の少なくとも 85 質量 % が X である場合は不純物 (本明細書中では Y) を “実質的に含まない”。好ましくは X は、組成物中の X + Y の合計の少なくとも約 90 質量 % を、より好ましくは少なくとも約 95 質量 %、98 質量 % または 99 質量 % を構成する。

10

【0063】

医薬組成物は、好ましくは治療的に有効な量の本発明のポリペプチド、核酸分子、リガンドまたは化合物を含むべきである。本明細書で用いられる “治療的に有効な量” という用語は、標的疾患または症状を治療、緩和もしくは予防するために、または検出可能な治療効果もしくは予防効果を示すために必要な治療薬剤の量を指す。いずれの化合物についても、治療的に有効な投与量は、最初に細胞培養アッセイ (例えば新生物細胞培養アッセイ) または動物モデル (通常はマウス、ウサギ、イヌまたはブタ) のいずれかで見積もることができる。動物モデルは、適切な濃度範囲および投与経路の決定にも用いることができる。次にそのような情報を用いて、ヒトで有用な投与量および投与経路を決定することができる。

20

ヒト対象者に対する正確な有効量は、疾患状態の重篤度、対象者の全身の健康状態、対象者の年齢、体重および性別、食事、投与時間および投与回数、併用薬剤、反応感受性および治療に対する許容性 / 応答性に依存するであろう。前記の量は日常的検査により決定でき、それは臨床医の判断の範囲内である。一般には有効用量は 0.01mg/kg から 50mg/kg、好ましくは 0.05mg/kg から 10mg/kg であろう。本組成物は個別に、または他の薬剤、医薬品またはホルモンと一緒に患者に投与することができる。

30

【0064】

医薬組成物はまた、治療薬の投与のために医薬的に許容できる担体を含むことができる。そのような担体には、抗体および他のポリペプチド、遺伝子並びに他の治療薬剤 (例えばリポソーム) が含まれるが、ただし担体はそれ自体で前記組成物を投与される個体に有害な抗体の産生を誘発せず、さらに不都合な毒性をもたらすことなく投与されることを条件とする。適切な担体は大型でゆっくりと代謝される巨大分子、例えばタンパク質、多糖類、ポリ乳酸、ポリグリコール酸、重合アミノ酸、アミノ酸コポリマーおよび不活性ウイルス粒子であり得る。

それらの中には医薬的に許容できる塩、例えば鉍酸塩 (塩酸塩、臭化水素酸塩、リン酸塩、硫酸塩などのような) ; および有機酸の塩 (酢酸塩、プロピオン酸塩、マロン酸塩、安息香酸塩などのような) を用いることができる。医薬的に許容できる担体についての綿密な考察は以下のテキストで入手可能である: Remington's Pharmaceutical Sciences (Mack Pub. Co., N.J. 1991)。

40

治療用組成物中の医薬的に許容できる担体は、さらに液体、例えば水、生理食塩水、グリセロールおよびエタノールを含むことができる。さらに、助剤 (例えば湿潤剤または乳化剤、pH 緩衝物質など) が前記組成物中に存在していてもよい。そのような担体は、患者が摂取できるように前記医薬組成物を錠剤、ピル、糖衣錠、カプセル、液剤、ゲル、シロップ、スラリー、懸濁剤などとして製剤化することを可能にする。

いったん製剤化されたら、本発明の組成物は直接対象者に投与することができる。治療される対象者は動物で、特にヒト対象者が治療され得る。

50

本発明で用いられる医薬組成物は、多数の経路（経口、静脈内、筋肉内、動脈内、骨髄内、硬膜下腔内、心室内、経皮適用（例えばW098/20734を参照）、皮下、腹腔内、鼻内、腸内、局所、舌下、腔内または直腸的手段が含まれるが、ただしこれらに限定されない）によって投与できる。遺伝子銃またはハイポスプレーもまた、本発明の医薬組成物の投与に用いることができる。典型的には、本治療用組成物は、注射用物質（液体溶液または懸濁剤のいずれか）として調製できる。注射前に液体賦形剤中で溶液または懸濁剤とするために適した固体を調製することもできる。

本組成物の直接的デリバリーは一般に、皮下、腹腔内、静脈内または筋肉内に注射によって達成されるか、または組織の間隙腔に搬送されるであろう。前記組成物はまた、病巣に投与してもよい。投薬治療は単回投与スケジュールでも複数回投与スケジュールでもよい

。本発明のポリペプチドの活性が特定の疾患状態で過剰である場合には、いくつかのアプローチが利用可能である。あるアプローチは、医薬的に許容できる担体とともに上記のような阻害化合物（アンタゴニスト）を、前記ポリペプチドの機能阻害に有効な量で対象者に投与することを含む（前記化合物による前記機能阻害は、例えばリガンド、基質、酵素、レセプターの結合を遮断するか、または第二のシグナルを阻害し、それによって異常な症状を緩和することによって達成される）。好ましくは、前記アンタゴニストは抗体である。最も好ましくは、そのような抗体は、先に記載したようなその免疫原性を最少にするキメラ抗体および/またはヒト化抗体である。

また別のアプローチでは、問題のリガンド、基質、酵素、レセプターに対する結合親和性を保持するポリペプチドの可溶性を投与することができる。典型的には、前記ポリペプチドは、関連部分を保持するフラグメントの形態で投与することができる。

また別のアプローチでは、前記ポリペプチドをコードする遺伝子の発現は、発現遮断技術を用いて、例えば内部で生成されるかまたは別々に投与されるアンチセンス核酸分子を使用して（上記で述べたように）阻害することができる。遺伝子発現の改変は、ポリペプチドをコードする遺伝子の制御領域、5'領域または調節領域（シグナル配列、プロモーター、エンハンサーおよびイントロン）に対して相補的な配列またはアンチセンス分子（DNA、RNAまたはPNA）をデザインすることによって達成できる。同様に、阻害は“三重らせん”塩基対方法論を用いて達成することができる。三重らせん対形成は、ポリメラーゼ、転写因子または調節分子の結合のために二重らせんが充分に開く能力を阻害することから有用である。三重DNAを用いる最近の治療の進歩は文献に記載されている（J.E. Gee et al. (1994) In: B.E. Huber & B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, NY）。相補的配列またはアンチセンス分子をデザインし、リボソームに対する結合を妨げて転写を妨害することによってmRNAの翻訳を遮断することもできる。そのようなオリゴヌクレオチドは投与されてもよいし、またin vivoでの発現によりin situで生成させてもよい。

【0065】

さらに、本発明のポリペプチドの発現は、そのコードmRNA配列に特異的なリボザイムを用いることによって妨げることができる。リボザイムは、天然または合成であり得る触媒的活性型のRNAである（例えば以下を参照されたい：N. Usman et al., Curr. Opin. Struct. Biol. (1996) 6(4): 527-533）。合成リボザイムをデザインして、選択した位置でmRNAを特異的に切断し、それによってmRNAの機能的ポリペプチドへの翻訳を妨げることができる。リボザイムは、通常RNA分子で見出されるような天然のリン酸リボース骨格および天然の塩基から合成することができる。或いは、リボザイムは、非天然の骨格（例えば2'-O-メチルRNA）を用いて合成して、リボヌクレアーゼ分解から保護することができる。これらは改変塩基を含んでいてもよい。

RNA分子は細胞内安定性および半減期を増加させるために改変することができる。可能な改変には、RNA分子の5'および/または3'末端へのフランキング配列の付加、または分子の骨格内でホスホジエステル結合に代わるホスホロチオエートまたは2'-O-メチルの使用が含まれるが、ただしこれらに限定されない。この概念はPNAの生成にも受け継がれ

、非慣用塩基（例えばイノシン、ケオシン(queosine)およびブトシン(butosine)（アセチル -、メチル -、チオ - も同様に）、および同様に改変された形態のアデニン、シチジン、グアニン、チミンおよびウリジンで、これらは内因性エンドヌクレアーゼによって容易に認識されないようなものである）の包含によってPNA分子の全てに前記概念を広げることができる。

【0066】

本発明のポリペプチドおよびその活性の過小発現に関連する異常な状態を治療するためには、いくつかのアプローチも利用可能である。あるアプローチは、前記ポリペプチドを活性化する化合物（すなわち上記で述べたアゴニスト）の治療的に有効な量を対象者に投与し、異常な状態を緩和することを含む。あるいは、本ポリペプチドの治療量を適切な医薬担体と組合せて投与し、ポリペプチドの相対的な生理学的バランスを回復させてもよい。遺伝子治療を用い、対象者の関連する細胞による本ポリペプチドの内因性生成を行わせることができる。遺伝子治療は、不完全な遺伝子を修正した治療用遺伝子と置き換えることによって、前記ポリペプチドの不適切な生成を永久的に治療するために用いられる。

本発明の遺伝子治療は *in vivo* または *ex vivo* で実施できる。*Ex vivo* 遺伝子治療は、患者の細胞の単離および精製、治療用遺伝子の導入、および遺伝的に改変した細胞を患者に戻して導入することを必要とする。対照的に、*in vivo* 遺伝子治療は、患者の細胞の単離および精製を必要としない。

治療用遺伝子は患者に投与するために典型的には“封入されて”いる。遺伝子デリバリー賦形剤はリポソームのような非ウイルス、または、例えばK.L. Berkner (1992) *Curr. Top. Microbiol. Immunol.*, 158: 39 - 66に記載されたアデノウイルスのような複製欠損ウイルスまたはN. Muzyczka (1992) *Curr. Top. Microbiol. Immunol.*, 158: 97 - 129および米国特許第5,252,479号に記載されたアデノ付随ウイルス(AAV)ベクターであり得る。例えば、本発明のポリペプチドをコードする核酸分子は、複製欠損レトロウイルスベクターで発現させるために操作を施すことができる。次にこの発現構築物を単離し、前記ポリペプチドをコードするRNAを含有するレトロウイルスプラスミドベクターで形質導入したパッケージ細胞に導入することができる。その結果、前記パッケージ細胞は対象の遺伝子を含有する感染性ウイルス粒子を産生することができるようになる。これらのプロデューサー細胞は *in vivo* で細胞を操作するために、さらに *in vivo* でポリペプチドを発現させるために対象者に投与することができる（以下を参照されたい：*Gene Therapy and Other Molecular Genetic - based Therapeutic Approaches*, Chapter 20（およびその中に引用された文献）, “*Human Molecular Genetics*” (1996) T. Strachan & A.P. Read, BIOS Scientific Publishers Ltd.）。

別のアプローチは“裸のDNA”の投与で、この場合、治療用遺伝子は血流または筋肉組織に直接注射される。

本発明のポリペプチドまたは核酸分子が疾患を引き起こす原因物質である場合には、本発明は、前記疾患を引き起こす原因物質に対する抗体を生成するワクチンとして用いることができる。

本発明のワクチンは予防的（すなわち感染を防ぐ）または治療的（すなわち感染後の疾患を治療する）であり得る。そのようなワクチンは、免疫性を付与する抗原、免疫原、ポリペプチド、タンパク質または核酸を、通常は上記で述べた医薬的に許容できる担体と組合せて含む。前記担体には、それ自体で組成物を投与される個体に対して有害な抗体の産生を誘発しない担体のいずれもが含まれる。さらに、これらの担体は免疫刺激剤（“アジュバント”）として機能してもよい。さらにまた、前記抗原または免疫原は、細菌の類毒素（例えばジフテリア、破傷風、コレラ、H.ピロリ菌(*pyroli*)）および他の病原体と結合させることができる。

ポリペプチドは胃で分解されるので、ポリペプチドを含むワクチンは好ましくは非経口的に（例えば皮下、筋肉内、静脈内または皮内注射）投与される。非経口投与に適した製剤には、水性および非水性滅菌注射溶液（前記は抗酸化剤、緩衝剤、抗菌剤および製剤を受容者の血液に対し等張にする溶質を含むことができる）、並びに水性および非水性滅菌懸

10

20

30

40

50

濁剤（前記は分散剤または粘稠剤を含むことができる）が含まれる。

【0067】

本発明のワクチン製剤は、単位用量または複数単位用量容器で提供されてもよい。例えば封入したアンプルおよびバイアルは、使用直前に滅菌された液状担体を添加することだけを必要とする凍結乾燥状態で保存することができる。投与量はワクチンの比活性に依存し、日常的な検査で容易に決定することができる。

本発明はまた、診断薬としての本発明の核酸分子の使用に関する。本発明の核酸分子により特徴付けられ、機能不全に付随する遺伝子の変異型の検出は、前記遺伝子の過小発現、過剰発現または位置的もしくは時間的発現の変化から生じる疾患の診断、または疾患に対する感受性の診断に付け加えることができるか、またはそれらを明確にすることができる

10

診断のための核酸分子は対象者の細胞、例えば血液、尿、唾液、組織生検または剖検材料から入手できる。ゲノムDNAを直接検出に用いてもよいし、PCR、リガーゼ連鎖反応（LCR）、鎖置換増幅（SDA）または他の増幅技術を分析前に用いることによって、ゲノムDNAを酵素的に増幅してもよい（以下の文献を参照されたい：Saiki et al., Nature 324: 163 - 166(1986); Bej et al., Crit. Rev. Biochem. Molec. Biol., 26: 301 - 334 (1991); Birkenmeyer et al., J. Virol. Meth., 35: 117 - 126(1991); Van Brunt, J., Bio/Technology, 8: 291 - 294(1990)）。

【0068】

20

ある態様では、本発明のこの特徴は、本発明のポリペプチドをコードする天然の遺伝子の発現レベルを評価し、さらに前記発現レベルをコントロールのレベルと比較することを含む、患者における疾患を診断する方法を提供する（ここで前記コントロールレベルと異なるレベルは疾患を示唆する）。前記方法は以下の工程を含む：

a) 本発明の核酸分子と核酸プローブとの間でハイブリッド複合体の形成を可能にするストリンジェント条件下で、患者由来の組織サンプルを前記核酸プローブと接触させる工程；

b) コントロールサンプルを工程 a) で用いた条件と同じ条件下で前記プローブと接触させる工程；および、

c) 前記サンプル中のハイブリッド複合体の存在を検出する工程；

30

この場合、コントロールサンプル中のハイブリッド複合体レベルと異なるハイブリッド複合体レベルが患者サンプルで検出されることは疾患を示唆する。

本発明のさらなる特徴は、以下の工程を含む診断方法を含む：

a) 疾患について検査される患者から組織サンプルを入手する工程；

b) 前記組織サンプルから本発明の核酸分子を単離する工程；および、

c) 前記核酸分子内の疾患に付随する変異の存在を検出することによって患者を疾患について診断する工程。

上記に記載した方法における核酸分子の検出を補助するために、増幅工程、例えばPCRの使用を含むことができる。

正常な遺伝子型と比較すると、増幅生成物におけるサイズの変化によって、欠失および挿入が検出される。点変異は、増幅DNAを本発明の標識RNAとハイブリダイズさせるか、あるいは、本発明の標識アンチセンスDNA配列とハイブリダイズさせることによって同定することができる。完全にマッチした配列は、RNase消化によって、または溶融温度における差異を評価することによって、ミスマッチを有する二重鎖と区別することができる。DNAをストリンジェントな条件下で前記DNAとハイブリダイズする核酸プローブと接触させてハイブリッド二重鎖分子を形成させ（前記ハイブリッド二重鎖は、疾患に付随する変異に対応するいずれかの部分で前記核酸プローブ鎖のハイブリダイズしていない部分を有する）、DNA鎖の対応する部分における疾患付随変異の有無を示すものとして、前記プローブ鎖のハイブリダイズしていない部分の有無を検出することによって、患者における変異の有無を検出することができる。

40

50

前記のような診断は特に出生前検査で有用であり、新生児検査においてすら有用である。

【0069】

点変異、および参照遺伝子と“変異”遺伝子間で異なる他の配列は、他の周知の技術、例えば直接DNAシーケンシングまたは一本鎖構造多型性(Orita et al., Genomics, 5: 874 - 879(1989)を参照)によって同定できる。例えば、シーケンシングプライマーを二本鎖PCR生成物または改変PCRによって作製された一本鎖テンプレート分子とともに用いることができる。配列決定は、放射能標識ヌクレオチドを用いる通常の方法によって、または蛍光タグを用いる自動シーケンシング法によって実施される。クローン化DNAセグメントを、特異的DNAセグメントを検出するためのプローブとして用いることもできる。この方法の感受性は、PCRと併用したとき極めて増強される。さらに、点変異および他の配列の変動(例えば多型性)は、例えば対立遺伝子特異的オリゴヌクレオチドをただ1つのヌクレオチドが異なる配列のPCR増幅に用いることによって、上記のように検出することができる。

10

DNA配列の相違はまた、変性剤の存在下または非存在下でのゲル内のDNAフラグメントの電気泳動移動度における変化によって、または直接DNAシーケンシング(例えば、Myers et al., Science (1985) 230: 1242)によって検出することができる。特定の位置における配列の変化はまた、ヌクレアーゼ保護アッセイ(例えばRNaseおよびS1保護)または化学切断法(Cotton et al., Proc. Natl. Acad. Sci. USA (1985) 85: 4397 - 4401を参照)によって明らかにすることができる。

通常ゲル電気泳動およびDNAシーケンシングの他に、ミクロ欠損、異数性、転座、逆位のような変異は、in situ分析によっても検出できる(例えば以下を参照されたい: Keller et al., DNA Probes, 2nd Ed., Stockton Press, New York, N.Y., USA(1993))。すなわち、細胞内のDNAまたはRNA配列は、それらを単離および/またはメンブレン上に固定する必要なしに変異について分析することができる。蛍光in situハイブリダイゼーション(FISH)は、現在のところ最も一般的に用いられている方法で、FISHに関する多数の概論が存在する(例えば以下を参照されたい: Trachuck et al., Science, 250: 559 - 562(1990); およびTrask et al., Trends Genet. 7: 149 - 154(1991))。

20

本発明の別の態様では、本発明の核酸分子を含むオリゴヌクレオチドプローブのアレイを構築し、遺伝的変種、変異および多型性の効率的スクリーニングを実施することができる。アレイ技術方法はよく知られており汎用的に応用することができ、遺伝子発現、遺伝連鎖および遺伝的多様性を含む分子遺伝学における種々の疑問に取り組むために用いることができる(例えば以下を参照されたい: M. Chee et al., Science (1996) 274: 610 - 613)。

30

【0070】

ある態様では、前記アレイは以下の文献に記載された方法にしたがって調製され使用される(PCT出願W095/11995(Chee et al.); D.J. Lockhart et al.(1996) Nat. Biotech. 14: 1675 - 1680; M. Schena et al.(1996) Proc. Natl. Acad. Sci. 93: 10614 - 10619)。オリゴヌクレオチド対は2つから100万個を越える範囲を変動し得る。前記オリゴマーは、光誘導化学法を用いて基板上の指定領域で合成される。基板は紙、ナイロンまたは他のタイプのメンブレン、フィルター、チップ、ガラススライドもしくは他の適切な任意の固相支持体でよい。別の特徴では、オリゴヌクレオチドは、PCT特許出願(W095/251116, Baldeschweiler et al.)に記載されたように、化学的結合方法およびインクジェット適用装置を用いて基板表面で合成することができる。別の特徴では、ドット(またはスロット)プロットに類似する“格子化(gridDED)”アレイを用い、真空系、熱結合方法、UV結合方法、機械的または化学的結合方法を用いて基質表面にcDNAフラグメントまたはオリゴヌクレオチドを配置し、連結させることができる。アレイ(例えば上記で述べたようなもの)は手動で、または利用可能な装置(スロットプロットまたはドットプロット装置)、材料(適切な固相支持体すべて)および機械(ロボット機器を含む)を用いて作製することができ、8、24、96、384、1536または6144個のオリゴヌクレオチド、または2つから100万個を越える範囲の他のいずれの数をも含むことができ

40

50

る（このことはアレイ自体を商業的に入手可能な計測器の有効利用に向くものとしている）。

【0071】

上記で考察した方法の他に、対象者に由来するサンプルから、ポリペプチドまたはmRNAの異常な増加または低下レベルを決定することを含む方法によって疾患を診断することができる。発現低下または発現増加は、ポリヌクレオチドの定量のために当技術分野で周知の方法のいずれか、例えば核酸増幅（例を挙げるとPCR、RT-PCR、RNase保護、ノザンプロット法）および他のハイブリダイゼーション方法を用いてRNAレベルで測定することができる。

ホストに由来するサンプルで本発明のポリペプチドレベルを決定することに用いることができるアッセイ技術は当業者によく知られており、さらに上記でいくらか詳細に考察されている（ラジオイムノアッセイ、競合結合アッセイ、ウェスタンブロット分析およびELISAアッセイを含む）。本発明のこの特徴では以下の工程を含む診断方法が提供される：（a）上記に記載したリガンドを、リガンド-ポリペプチド複合体の形成に適した条件下で生物学的サンプルと接触させる工程；および（b）前記複合体を検出する工程。

例えばELISA、RIAおよびFACSのようなポリペプチドレベルを測定するためのプロトコルは、さらにポリペプチド発現の変化レベルまたは異常レベルを診断するための基礎を提供することができる。ポリペプチド発現の正常値または標準値は、正常な哺乳類対象体（好ましくはヒト）から得られた体液または細胞抽出物を、複合体形成に適した条件下で前記ポリペプチドに対する抗体と混合することによって確立される。標準的な複合体形成量は種々の方法、例えば分光測定方法によって定量することができる。

【0072】

本発明のポリペプチドと特異的に結合する抗体は、前記ポリペプチドの発現によって特徴付けられる症状または疾患の診断のために、または本発明のポリペプチド、核酸分子、リガンドおよび他の化合物を用いて治療されている患者をモニターするアッセイにおいて用いることができる。診断目的で有用な抗体は、治療薬として上記で述べたのと同じ様式で調製することができる。前記ポリペプチドについての診断アッセイは、前記抗体および標識を用いてヒトの体液または細胞もしくは組織の抽出物中のポリペプチドを検出する方法を含む。前記抗体は改変して、または改変せずに用いることができ、さらにそれらをレポーター分子と共有結合または非共有結合によって結合させることによって標識することができる（それらのいくつかは上記に記載されている）。

生検組織由来の対象者、コントロールおよび疾患サンプルで発現されたポリペプチドの量は、標準値と比較される。標準値と対象者の値との間の偏差は疾患診断のためのパラメータを確立する。診断アッセイを用いてポリペプチド発現の有無および過剰を識別し、治療的処置の間のポリペプチドレベルの調節をモニターすることができる。そのようなアッセイはまた、動物実験、臨床試験または個々の患者の治療モニタリングにおける個々の治療的処置方法の有効性を評価することに用いることができる。

【0073】

本発明の診断キットは以下を含むことができる：

- （a）本発明の核酸分子；
- （b）本発明のポリペプチド；または
- （c）本発明のリガンド。

本発明のある特徴では、診断キットは、ストリンジエントな条件下で本発明の核酸分子とハイブリダイズする核酸プローブを含む第一の容器；核酸分子を増幅させるために有用なプライマーを含む第二の容器；および疾患の診断を容易にするために前記プローブおよびプライマーの使用についての指示書を含むことができる。前記キットはさらに、ハイブリダイズしていないRNAを消化するための試薬を保持する第三の容器を含むことができる。本発明の別の特徴では、診断キットは核酸分子のアレイを含み、前記核酸分子の1つが本発明の核酸分子であってもよい。

10

20

30

40

50

本発明のポリペプチドを検出するために、診断キットは以下を含むことができる：本発明のポリペプチドと結合する1つまたは2つ以上の抗体；および前記抗体とポリペプチドとの間の結合反応の検出に有用な試薬。

そのようなキットは、以下のような疾患または疾患に対する感受性を診断することに有用であろう：特に細胞増殖疾患（新生物、メラノーマ、肺、結腸直腸、乳房、膵、頭部および頸部の腫瘍並びに他の固形腫瘍を含む）、骨髄増殖性疾患（例えば白血病、非ホジキンリンパ腫、白血球減少症、血小板減少症、血管形成疾患、カボジ肉腫）、自己免疫/炎症性疾患（アレルギー、炎症性腸疾患、関節炎、乾癬および気道の炎症、喘息および器官の移植拒絶を含む）、心脈管系疾患（高血圧、浮腫、アングィナ、アテローム性動脈硬化症、血栓症、敗血症、ショック、再灌流障害、心不整脈および虚血を含む）、神経学的疾患（中枢神経系疾患、アルツハイマー病、脳損傷、脳卒中、筋萎縮性側索硬化症、不安、抑うつおよび痛みを含む）、発達障害、代謝性障害（真性糖尿病、骨粗しょう症、脂質代謝障害、甲状腺機能亢進、上皮小体機能亢進、高カルシウム血症、高コレステロール血症、高脂血症および肥満を含む）、腎疾患（糸球体腎炎、腎血管性高血圧を含む）、皮膚疾患（アクネ、湿疹および創傷治癒を含む）、加齢の負の作用、エイズ、感染（ウイルス感染、細菌感染、真菌感染および寄生虫感染を含む）、並びに他の病的状態（特に核内ホルモンレセプターが関与するもの）。

10

本発明の種々の特徴および態様は、特にLBDG2ポリペプチドへの言及を有する実施例を介してこれからより詳細に説明されるであろう。

細部の改変は本発明の範囲を逸脱することなく実施できることは理解されよう。

20

【0074】

（実施例）

実施例：1

遠縁の新規な核内ホルモンレセプターリガンド結合ドメインの検索を開始するために、原型となるファミリーメンバー、ホモサピエンス甲状腺ホルモンレセプターのリガンド結合ドメインを選択する。より具体的には、前記検索は、タンパク質データベース（PDB）（Research Collaboratory for Structural Bioinformaticsによって運用されている）から得られた構造を用いて開始する。

選択した構造は、ホモサピエンス甲状腺ホルモンレセプターのリガンド結合ドメインである（PDBコード1BSX：A；図1参照）。

30

1BSX：Aの類縁体についてバイオベンジウムの検索を実施し（ターゲットマイニングインターフェース(Target Mining Interface)を用いる）、4096のゲノムスレッダーの結果が返される。この4096のゲノムスレッダーの結果には、典型的な核内ホルモンレセプターリガンド結合ドメイン（例えばホモサピエンス甲状腺ホルモンレセプターの残基211 - 461の間で見出されるもの）（図2Aの矢印参照）の例が含まれる。

核内ホルモンレセプターリガンド結合ドメインを含むことが判明しているタンパク質の中で、核内ホルモンレセプターリガンド結合ドメインを含むと注釈付けされていないタンパク質、BAA22563.1（LBDG2；図2Bの矢印参照）が現れる。インファーマチカゲノムスレッダーは、配列BAA22563.1（LBDG2）の領域（残基1104 - 1309の間）をホモサピエンス甲状腺ホルモンレセプターのリガンド結合ドメインと類似の構造を有すると同定した。ホモサピエンス甲状腺ホルモンレセプターのリガンド結合ドメインと類似する構造を保有することは、BAA22563.1（LBDG2）の残基1104 - 1309は核内ホルモンレセプターリガンド結合ドメインとして機能することを示唆する。ゲノムスレッダーはこれを86%の信頼度で同定する。

40

1BSX：Aの類縁体についてのバイオベンジウムの検索（ターゲットマイニングインターフェースを用いる）によって、852のフォワード(forward)PSI - BLASTの結果も得られる。フォワードPSI - BLAST（図2Cを参照されたい）は前記関係を同定できず、インファーマチカゲノムスレッダーだけがBAA22563.1（LBDG2）は核内ホルモンレセプターリガンド結合ドメインを含むと同定できる。

【0075】

50

BAA22563.1 (LBDG2) についてパブリックドメインデータベースで何が判明しているかを評価するために、重複配列ディスプレイページ (Redundant Sequence Display Page) (図 3) を調査する。BAA22563.1 (LBDG2) が核内ホルモンレセプターリガンド結合ドメインを含むと同定する PROSITE ヒットも PRINTS ヒットも得られない。PROSITE および PRINTS は、類似ファミリーのタンパク質を記述するのに役立つデータベースである。両データベースから核内ホルモンレセプターリガンド結合ドメインのヒットが返ってこないことは、PROSITE または PRINTS を用いた場合、BAA22563.1 (LBDG2) が核内ホルモンレセプターリガンド結合ドメインを含むものとして同定できないことを意味する。

他のいずれかのパブリックドメインの注釈付け手段が BAA22563.1 (LBDG2) は核内ホルモンレセプターリガンド結合ドメインを含むと注釈付けすることができるか否かを明らかにするために、BAA22563.1 (LBDG2) タンパク質配列をインタープロ (InterPro) ウェブサイトで PFAM データベース (アラインメントおよび隠れマルコフモデルのタンパク質ファミリーデータベース (Protein Family Database of Alignment and hidden Markov models)) に対して検索する (図 4 A 矢印 1 参照)。1 つの PFAM - A マッチングが PF000527/IPRO00148 に対して見出される。このことは、パピローマウイルス E7 タンパク質に対する関連性の指標である。パピローマウイルス E7 タンパク質 マッチングは、BAA22563.1 (LBDG2) の残基 1561 - 1570 の間に位置している。しかしながら、核内ホルモンレセプターリガンド結合ドメインを含むと BAA22563.1 (LBDG2) を注釈付けている PFAM - A マッチングは存在しない。したがって PFAM では、BAA22563.1 (LBDG2) が核内ホルモンレセプターリガンド結合ドメインを含むことが明らかにされない。

興味深いことには、PROSITE の PF スキャン (図 4 A 矢印 2 参照) によって二極性核局在化シグナルが BAA22563.1 (LBDG2) の残基 35 - 52 に同定される。核内ホルモンレセプターの典型的な (指標ではないが) 特徴は、二極性核局在化シグナルの保有である。

次に、BAA22563.1 (LBDG2) に関してパブリックドメインにおいて知られる更なる情報が存在するかを検討するために、全米バイオテクノロジー情報センター (NCBI) Genebank タンパク質データベースを調査した。これは、タンパク質配列および遺伝子配列の寄託のための US パブリックドメインデータベースである (図 5)。BAA22563.1 (LBDG2) はホモサピエンス配列であり、その Genebank タンパク質 ID は BAA22563.1 であり、長さは 2335 アミノ酸である。BAA22563.1 (LBDG2) は サッカロマイセス・セレビシエ PRP8、プレ - mRNA スプライス因子のホモサピエンス相同体と呼ばれる。BAA22563.1 (LBDG2) は大塚遺伝子研究所 (日本、徳島県河内町加賀須野) の科学者グループによってクローン化された。この遺伝子についてのパブリックドメイン情報は、核内ホルモンレセプターリガンド結合ドメインを含むと前記遺伝子を注釈付けしていない。

【 0 0 7 6 】

したがって、全てのパブリックドメインの注釈付けツールを用いて、BAA22563.1 (LBDG2) は核内ホルモンレセプターリガンド結合ドメインを含むと注釈付けされないであろうと結論することができる。インファーマチカゲノムスレッダーのみが、前記タンパク質は核内ホルモンレセプターリガンド結合ドメインを含むと注釈付けすることができる。

ここで逆検索を実施する。BAA22563.1 (LBDG2) を今度はクエリー配列としてバイオペンジウムで用いる (図 6 A 参照)。インファーマチカゲノムスレッダーは、BAA22563.1 (LBDG2) の残基 1104 - 1309 をホモサピエンス甲状腺ホルモンレセプターのリガンド結合ドメインと同じ構造を有すると信頼度 86% で同定する (図 6 B の矢印参照)。ホモサピエンス甲状腺ホルモンレセプター (1BSX: A) のリガンド結合ドメインは、元々のクエリー配列であった。PSI - BLAST の正の繰返しはこの結果を返さない (図 6 C)。この関係を明らかにすることができるのは、インファーマチカゲノムスレッダーだけである。

ホモサピエンス甲状腺ホルモンレセプターリガンド結合ドメインの配列は、BAA22563.1 (LBDG2) の配列アラインメントを調査するために前記に対して選択される。類似構造を有すると同定されたタンパク質に対するクエリータンパク質の AI Eye アラインメント (図 7) を観察することは、相同領域の可視化に役立つ。

ホモサピエンス甲状腺ホルモンレセプターリガンド結合ドメインは “ L B D モチーフ ”

10

20

30

40

50

を含み、前記モチーフは、今日までに注釈付けされた全ての核内ホルモンレセプターリガンド結合ドメインで見出されている。“LBDモチーフ”は核内ホルモンレセプターのコアクチベーターおよびコリプレッサーを補充することに関与する。6つの残基；PHE293、LEU296、ASP300、GLN301、LEU304およびLEU305は、ホモサピエンス甲状腺ホルモンレセプターリガンド結合ドメインにおいてこのモチーフを構成する（図7の四角枠を参照）。BAA22563.1（LBDG2）中の4残基（PHE1174、ASP1181、LEU1185、LEU1186）は、ホモサピエンス甲状腺ホルモンレセプターリガンド結合ドメイン中の6つの“LBDモチーフ”残基のうち4つ（PHE293、ASP300、LEU304、LEU305）と正確にマッチする。そのうえBAA22563.1（LBDG2）のVAL1177およびASN1182は、ホモサピエンス甲状腺ホルモンレセプターリガンド結合ドメインの“LBDモチーフ”では残りの2残基LEU296およびGLN301に保存的に置換されている。このことは、BAA22563.1（LBDG2）はホモサピエンス甲状腺ホルモンレセプターリガンド結合ドメインと類似する核内ホルモンレセプターリガンド結合ドメインを含むことを示している。

10

【0077】

同定されたタンパク質が実際にクエリー配列と関連することを保証するために、可視化プログラム“LigEye”（図8A）および“iRasmol”（図8B）を用いる。これらの可視化ツールは、既知の小型阻害分子が既知のタンパク質構造の活性部位で相互作用するアミノ酸を表示することによって前記既知タンパク質構造の活性部位を同定する。これらの相互作用は、直接的水素結合を介するか、または疎水性相互作用を介する。このようにして、活性部位のフォールド/構造が、同定した相同体と選択した既知の構造をもつタンパク質との間で保存されているか否かを理解することができる。ホモサピエンス甲状腺ホルモンレセプターリガンド結合ドメインのLigEyeによる調査は、3,5,3'-トリヨードチロニンと結合する5残基（図7の丸付き文字）を明らかにする。しかしながら、これら5残基のうち4残基（ILE276、LEU330、ASN331およびLEU346）のみがゲノムスレッダーアラインメントの範囲内にある。したがって、これら4残基のみが、核内ホルモンレセプターリガンド結合ドメインを含むというBAA22563.1（LBDG2）のゲノムスレッダーの注釈付けを強固にするために用いることができる。これら4残基の中に3個の疎水性残基が存在し、これらはホモサピエンス甲状腺ホルモンレセプターリガンド結合ドメイン（ILE276、LEU330およびLEU346）のポケットの輪郭を形成する。ホモサピエンス甲状腺ホルモンレセプターリガンド結合ドメインのLEU330およびLEU346はBAA22563.1（LBDG2）で完全に保存されている：LEU1216およびLEU1230（図7の丸付き文字）。ホモサピエンス甲状腺ホルモンレセプターリガンド結合ドメインのILE276は、BAA22563.1（LBDG2）ではLEU1161によって保存的に置換されている：（図7の穴あき丸付き文字）。結合ポケットの輪郭を形成する3つの疎水性残基のうち（ゲノムスレッダーアラインメント領域内で）3つで疎水性が保存されていることは、BAA22563.1（LBDG2）が疎水性ステロイド様リガンドと結合することを示している。

20

30

【0078】

ホモサピエンス甲状腺ホルモンレセプターリガンド結合ドメインのASN331は、BAA22563.1（LBDG2）ではGLN1217によって保存的に置換されている：（図7の穴あき丸付き文字）。このことは、実際インファーマチカゲノムスレッダーによって予測されたように、BAA22563.1（LBDG2）はホモサピエンス甲状腺ホルモンレセプターリガンド結合ドメインと類似の様式で折り畳まれることを示し、したがって核内ホルモンレセプターリガンド結合ドメインを含むと同定される。

40

BAA22563.1（LBDG2）の逆最大化（reverse-maximised）PSI-BLASTによってドロソフィラ・メラノガスター相同体（AAF58573.1、図6C矢印1を参照）、シノラブディス・エレガンス相同体（P34369、図6C矢印2を参照）、オリザ・サティバ相同体（BAA78744.1、図6C矢印3を参照）およびアラビドプシス・サリアナ相同体（CAB80541.1、図6C矢印4を参照）が同定される。BAA22563.1（LBDG2）、AAF58573.1（ドロソフィラ・メラノガスター相同体）、P34369（シノラブディス・エレガンス相同体）、BAA78744.1（オリザ・サティバ相同体）およびCAB80541.1（アラビドプシス・サリアナ相同体）をAIEyeで整列化して調

50

査する(図9)。AIEyeは、BAA22563.1(LBDG2)の4つの予測リガンド結合残基(ゲノムスレッダーアラインメントに含まれる; LEU1161、LEU1216、GLN1217およびLEU1230)は全て、AAF58573.1(ドロソフィラ・メラノガスター相同体)、P34369(シノラブディス・エレガンス相同体)、BAA78744.1(オリザ・サティバ相同体)およびCAB80541.1(アラビドプシス・サリアナ相同体)において正確に保存されていることを明らかにしている。さらにBAA22563.1(LBDG2)における予測“LBDモチーフ”残基の全て(PHE1174、VAL1177、ASP1181、ASN1182、LEU1185およびLEU1186)は、AAF58573.1(ドロソフィラ・メラノガスター相同体)、P34369(シノラブディス・エレガンス相同体)、BAA78744.1(オリザ・サティバ相同体)およびCAB80541.1(アラビドプシス・サリアナ相同体)において正確に保存されている。唯一の例外は、LEU1185がMETで保存的に置換されているP34369(シノラブディス・エレガンス相同体)である。タンパク質の機能にとって必須の残基は、前記タンパク質の相同体において保存されるであろう。したがって、予測されるBAA22563.1(LBDG2)核内ホルモンレセプターリガンド結合ドメインの機能に必須であろうと思われる残基のAAF58573.1(ドロソフィラ・メラノガスター相同体)、P34369(シノラブディス・エレガンス相同体)、BAA78744.1(オリザ・サティバ相同体)およびCAB80541.1(アラビドプシス・サリアナ相同体)における保存は、BAA22563.1(LBDG2)が核内ホルモンレセプターリガンド結合ドメインを含むとの注釈付けを強く支持する。

10

【0079】

実施例：2

提唱したLBDの組織発現を定量することを目的として、タックマン(Taqman)RT-PCR定量を用いた。タックマン3'-5'エキソヌクレアーゼアッセイは、2つのプライマー認識配列間の部位で標的テンプレートとハイブリダイズする二重標識蛍光発生プローブの核酸分解を含む過程によって、PCR単位複製配列の形成にシグナルを送る(米国特許第5,876,930号を参照)。ABIプリズム(Prism)7000は、各増幅サイクルの間に、生成した単位複製配列の量に関する化学量論的な前記シグナルの検出および定量的測定を自動化して行う。PCR分析のために必要な時間および労働量の実質的な節減を提供することに加えて、この技術は反応における標的配列の単純化および非常に高精度な定量化を可能にする。

20

非罹病器官から調製したヒトRNAは、Ambion Europe(Huntingdon, UK)またはClontech(BD, Franklin Lakes, NJ)から購入した。オリゴヌクレオチドプライマーおよびプローブは、Primer Expressソフトウェア(Applied Biosystems, Foster City CA)を用いて、40~60%のGC含有量を有し、プローブの5'末端にG-ヌクレオチドが存在せず、4つより多くの連続するGが存在しないように設計した。次にBLAST(登録商標)(Basic Local Alignment Search Tool, Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.: J Mol Biol 1990 Oct 5; 215(3): 403-10)を用いて各プライマーおよびプローブを解析した。各オリゴヌクレオチドは、公的に利用可能なデータベースであるユニジーン(Unigene)およびゴールデンパス(GoldenPath)において提示された他の既知cDNA配列または既知ゲノム配列と比較した場合に3 bpより高い特異性で標的配列を認識したという結果を確認した。そのプライマーおよびプローブの配列は、次のようである：

30

BAA22563.1(LBDG2) 順方向プライマー：CAG ACA GGC CGC TGA CAT T

BAA22563.1(LBDG2) 逆方向プライマー：GCC ATC AGG AGG TCA ACA ACA

40

BAA22563.1(LBDG2) プローブ：AGT TTG GCC TCT TTC CCT CTG TCT GTG C

【0080】

18sおよびヒトリボソームタンパク質の前最適化プライマーおよびプローブは、Applied Biosystems, Foster City, CAから購入した。プローブは、最も5'の塩基および最も3'の塩基において、それぞれ蛍光レポーター色素(例えば、6カルボキシ-フルオレセイン{FAM}; λ_{em} = 518 nm)および蛍光クエンチャー色素(6カルボキシテトラム(carboxytetram)-エチル-ローダミン{TAMRA}; λ_{em} = 582 nm)と共有結合させた。全てのプライマーおよびプローブは、Applied Biosystems, Germanyから入手した。プライマー/プローブ濃度を50 nMから900 nMの範囲で滴定して、効率的なPCR反応のための最適濃度を決定した。プライマーおよびプローブの最適濃度は、増幅された標的遺伝子に依存して100 nMから90

50

0 nMの間を変動する。cDNAはApplied Biosystems, Foster City CAが提供している構成成分を用いて調製した。RNaseフリーの0.5 mlチューブに、50 μ lの反応物を調製する。その反応物は、総RNA 500 ng; 1 \times 逆転写酵素緩衝液; 5.5 mM MgCl₂; 1mM dNTP; ランダムヘキサマー 2.5 μ l; RNase阻害剤 20U; および逆転写酵素 62.5Uを含む。0.5 mlの薄壁光学グレードのPCR96ウェルプレート (Applied Biosystems, Foster City CA) に25 μ lの反応物を調製した。その反応物は、1 \times 最終濃度のタックマン・ユニバーサルマスターミックス (Universal Master Mix) (Applied Biosystems, Foster City CAが独自に開発したAmpliTaq Gold DNAポリメラーゼ、AmpEraseX UNG、UTPを有するdNTP、受動的参照色素および最適化した緩衝液構成成分の混合物); 100 nM タックマンプローブ; 300 nM 順方向プライマー; 900 nM 逆方向プライマーおよびcDNAテンプレート 15 ngを含んでいた。ABIプリズム7000の操作の標準の手順または同様の検出系を用いた。これには、反応体積を50から25 μ lに変えたことを除き、例えばABIプリズム7000での全デフォルトプログラム設定の使用が含まれていた。熱サイクリング条件は、50 で2分間、95 で10分間に続いて、95 で15秒間および60 で1分間の40サイクルで構成した。サイクル閾値(Ct)決定(すなわち、PCR産物の合成により生じるレポーター色素蛍光がバックグラウンド蛍光レベルより著しく高くなるのに必要なサイクル数の非整数計算)は、デフォルトパラメータを用い、各反応に対して機器により自動的に行った。同じcDNAサンプル中の標的配列およびリボソーム18s (参照) 配列に対するアッセイは、別々の反応チューブで行った。各実験では、典型的な組織サンプルの50 ngから0.78 ngのcDNAテンプレートで標準曲線化を行った。その標準曲線から、各テストサンプルにおける実際の標的cDNAまたは18s cDNAの出発量を決定する。

10

20

【0081】

各サンプル中の標的cDNAのレベルは、比較サンプル中の標的の発現レベルに正規化した。各サンプル中の内部標準cDNAのレベルは、比較サンプル中の内部標準の発現レベルに正規化した。次にそのデータを、任意で1に設定した比較サンプルにおける発現レベルに比例する正規化標的配列の倍率発現として表した。上述のBAA22563.1 (LBDG2) に特異的なプライマー/プローブを用いて、タックマンRT-PCRを結腸cDNAの2倍希釈物に対して行った。図10は投入量cDNAの対数に対してプロットしたCt値を示しており、前記範囲の投入量cDNA濃度にわたって線形関係が見られることを図示している。テストCt値からcDNAの出発量を算出するために、標準曲線の直線回帰分析を用いた。

30

上述の18s rRNAに特異的なプライマー/プローブを用いて、タックマンRT-PCRを結腸cDNAの2倍希釈物に対して行った。図11は投入量cDNAのlog値に対してプロットしたCt値を示しており、前記範囲の投入量cDNA濃度にわたって線形関係が見られることを図示している。テストCt値からcDNAの出発量を算出するために、標準曲線の直線回帰分析を用いた。上述のヒトリボソームタンパク質mRNAに特異的なプライマー/プローブを用いて、タックマンRT-PCRをIM9 cDNAの2倍希釈物に対して行った。図12は投入量cDNAのlog値に対してプロットしたCt値を示しており、前記範囲の投入量cDNA濃度にわたって線形関係が見られることを図示している。テストCt値からcDNAの出発量を算出するために、標準曲線の直線回帰分析を用いた。

上述のBAA22563.1 (LBDG2) および18s rRNAに対して特異的なプライマー/プローブを用い、指定したcDNA15 ngを用いてタックマンRT-PCRを行った。典型的な組織サンプルの50 ngから0.78 ngのcDNAテンプレートを用いて、標的および内部標準についての標準曲線化も行った。標準曲線の直線回帰分析を用いて、各テストサンプル中の標的cDNAまたは18s cDNAの実際の出発量を算出するためにそのCt値を使用した。

40

【0082】

各サンプル中の標的cDNAレベルは、比較サンプル(この場合は胃である)中の標的の発現レベルに正規化した。各サンプル中の18s cDNAレベルも、胃の18s発現レベルに正規化した。次に、BAA22563.1 (LBDG2) の発現レベルを18sの発現レベルに正規化した。図13は、任意で1に設定した胃のcDNAにおける発現レベルに関して正規化した標的配列の倍率発現を表している。各サンプルは、2-4の個々の実験間で定量化した。図13は、複数の実

50

験についての平均値 ± SEMを示す。

上で詳細に記述したBAA22563.1 (LBDG2) およびヒトリボソームタンパク質mRNAに特異的なプライマー/プローブを用い、指定したcDNA15 ngを用いてタックマンRT-PCRを行った。50 ngから0.78 ngのcDNAテンプレートを用い、典型的な細胞株サンプルを用いて標的および内部標準に対する標準曲線化も行った。標準曲線の直線回帰分析を用い、各テストサンプルにおける標的cDNAまたはヒトリボソームタンパク質cDNAの実際の出発量を算出するためにCt値を使用した。

各サンプルにおける標的cDNAのレベルを、比較サンプル(この場合はLAK細胞である)における標的の発現レベルに対して正規化した。各サンプルにおけるヒトリボソームタンパク質cDNAのレベルも、LAK細胞におけるヒトリボソームタンパク質cDNAの発現レベルに対して正規化した。次にBAA22563.1 (LBDG2) の発現レベルを、ヒトリボソームタンパク質の発現レベルに対して正規化した。図14は、任意で1に設定したLAKのcDNAにおける発現レベルに関する正規化標的配列の倍率発現を表している。図14は、各サンプルの2回繰返した測定についての平均値 ± SEMを示す。

LBDG2のmRNAは、種々のヒト組織由来の抽出物において見出された(図13)。ヒト脾臓におけるその高レベルの転写の発見は、免疫系、特にリンパ球およびB細胞の発生および機能におけるLBDG2の役割と一致している。従ってLBDG2のアゴニストおよびアンタゴニストの開発は、自己免疫、アレルギーおよび免疫グロブリン機能不全に関連する疾患を含む種々のヒト免疫系疾患での治療的処置において役割を有し得る。そのような疾患は、1型真性糖尿病、慢性関節リウマチ、多発性硬化症、乾癬、糸球体症に起因する腎不全、強皮症、炎症性腸疾患(クローン病および潰瘍性大腸炎)、移植拒絶、喘息、アトピー性皮膚炎、湿疹、骨髄腫、および抗体産生を必要とする感染症(例えばウイルス感染細胞、結核、リステリアのような細胞内病原体)を含む。

【0083】

Daudi、IM9およびRaji細胞のようなヒトB細胞株におけるLBDG2のmRNAの発見は、脾臓におけるmRNAの発見と一致する。U937細胞における高レベルのmRNA発現の発見は単球/マクロファージの機能におけるLBDG2の役割を示唆し、それらのアゴニストまたはアンタゴニストは慢性閉塞性肺疾患(COPD)、変形性関節炎、慢性関節リウマチ、炎症性腸疾患、肝繊維症(肝硬変)および皮膚繊維症(癬痕)のような繊維症、アテローム性動脈硬化症、痴呆、多発性硬化症、炎症性疼痛を含む炎症性疾患の治療に役立つ。

さらに副腎、卵巣、前立腺および精巣における有意なレベルのLBDG2の発見は、LBDG2に対するアゴニストおよびアンタゴニストの開発が良性前立腺肥大、前立腺癌、卵巣癌、精巣癌のような疾患において価値のあり得ることを示している。さらに、LBDG2に対するアゴニストまたはアンタゴニストは、以下に挙げる疾患(しかしこれだけに限られない)の治療のために開発され得る: 高血圧、感染症のストレスを含むストレスに対する応答、塩分および水分恒常性の調節、排卵調節を通じての受胎能制御(不妊症および避妊)、着床調節(不妊症および避妊)および精子形成調節(赴任および避妊)。

【0084】

LBDG2のmRNAがヒト脳において有意なレベルで発現されているという発見は注目すべきことであり、そのことがヒト疾患の状態に潜在的な関連を与えることから、LBDG2のリガンド結合ドメインに対するアゴニストおよびアンタゴニストの開発は以下に挙げる種々のヒト疾患における治療的処置の可能性を提供する: 細胞増殖性疾患(新生物、メラノーマ、肺、結腸直腸、乳房、膵、頭部および頸部の腫瘍、および他の固形腫瘍を含む)、骨髄増殖性疾患(例えば白血病、非ホジキンリンパ腫、白血球減少症、血小板減少症、血管形成疾患、カボジ肉腫)、自己免疫/炎症性疾患(アレルギー、炎症性腸疾患、関節炎、乾癬および気道の炎症、喘息および器官の移植拒絶を含む)、心脈管系疾患(高血圧、浮腫、アングィナ、アテローム性動脈硬化症、血栓症、敗血症、ショック、再灌流障害、心不整脈および虚血を含む)、神経学的疾患(中枢神経系疾患、アルツハイマー病、脳損傷、脳卒中、筋萎縮性側索硬化症、不安、抑うつおよび痛みを含む)、発達障害、代謝性障害(真性糖尿病、骨粗しょう症、脂質代謝障害、甲状腺機能亢進、上皮小体機能亢進、高カルシ

ウム血症、高コレステロール血症、高脂血症および肥満を含む)、腎疾患(糸球体腎炎、腎血管性高血圧を含む)、皮膚疾患(アクネ、湿疹および創傷治癒を含む)、加齢の負の作用、エイズ、感染(ウイルス感染、細菌感染、真菌感染および寄生虫感染を含む)、並びに他の病的状態(特に核内ホルモンレセプターが関与するもの)。

【0085】

DNA結合ドメインが存在しないリガンド結合ドメインを含むLBDG2のような“非古典的”核内ホルモンレセプターの発見は脳におけるステロイド(神経ステロイドとして知られる)の広範囲な効果を一貫して報告している既知文献と一致しており、それらの効果は、一般的には転写活性を必要とする既知の古典的ステロイドホルモン核内レセプターを介さずに媒介される。一例を挙げると、神経ステロイドは、特にGABA、NMDAおよびシグマレセプターのようなレセプターの域において神経伝達に影響することが示されている。神経ステロイドは、神経保護的役割を果たすことが示されている。したがってLBDG2に対するアゴニスト(またはアンタゴニスト)の開発を介する治療的処置は、梗塞または出血(脳卒中)、並びに中枢神経系および脊髄の損傷のような脳血管性疾患に続く痴呆、パーキンソン病および神経変性のような神経変性状態の治療において役割を有し得る。さらに神経ステロイドは、認識処理、空間学習および記憶、不安、並びに常習行為パターンを引き起こす欲求のような行為に影響することが示されている。したがってLBDG2に対するアゴニストおよびアンタゴニストの開発は、痴呆、学習困難、不安、常習行為(例えばアルコール症、摂食障害および薬物常習であるが、これだけに限られない)を治療するための治療的処置をもたらし得る。

10

20

【図面の簡単な説明】

【0086】

【図1】バイオペンジウムターゲットマイニングインターフェースのフロントエンドを示す。データベースの検索はPDBコード“1BSX:A”を用いて開始する。

【図2A】1BSX:Aを用いた検索に対するインファーマチカゲノムスレッダーの結果の抜粋が示されている。矢印は、典型的な核内ホルモンレセプターリガンド結合ドメインを有するホモサピエンス甲状腺ホルモンレセプター-1を示している。

【図2B】1BSX:Aを用いた検索に対するインファーマチカゲノムスレッダーの結果の抜粋が示されている。矢印は、BAA22563.1(LBDG2)を示している。

【図2C】1BSX:Aを用いた検索に対するフォワードPSI-BLASTの結果の完全なリストが示されている。BAA22563.1(LBDG2)は同定されない。

30

【図3】BAA22563.1(LBDG2)に対する重複配列表示の結果のページである。

【図4】BAA22563.1(LBDG2)に対するインタープロPFAM検索結果を示す。矢印1を参照されたい。

【図5】BAA22563.1(LBDG2)に対するNCBIタンパク質レポートである。

【図6A】バイオペンジウムデータベースのフロントエンドである。データベースの検索はBAA22563.1(LBDG2)をクエリー配列として用いて開始する。

【図6B】BAA22563.1(LBDG2)をクエリー配列として用いた検索のインファーマチカゲノムスレッダーの結果の抜粋である。矢印は1BSX:Aを指している。

【図6C】クエリー配列としてBAA22563.1(LBDG2)を用いて得られた逆最大化PSI-BLASTの結果の抜粋である。1から4の番号を付与された矢印はBAA22563.1(LBDG2)の相同体を指している。

40

【図7】BAA22563.1(LBDG2)および1BSX:AのAIEye配列アラインメントを示す。

【図8A】3,5,3'-トリヨードチロニンとホモサピエンス甲状腺ホルモンレセプターのリガンド結合ドメイン、1BSX:Aとの相互作用部位を図示している、1BSX:AについてのLigEyeである。

【図8B】ホモサピエンス甲状腺ホルモンレセプターのリガンド結合ドメイン、1BSX:AのiRasMol図である。

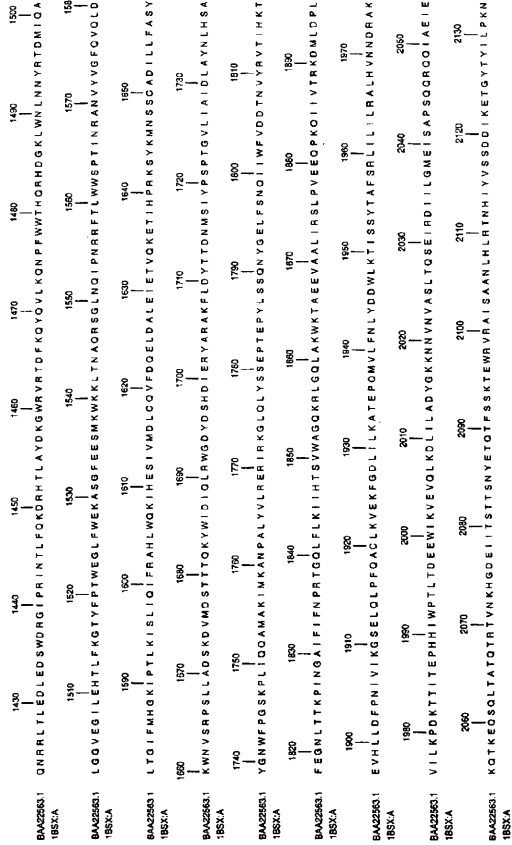
【図9】BAA22563.1(LBDG2;ホモサピエンスPRP8)と4つの相同体;AAF58573.1(ドロソフィラ・メラノガスター、1)、P34369(シノラブディス・エレガンズ、2)、BAA78744.1

50

【 7 - 3 】

AEye 出力(1月18日, 2001:5:45 PM)

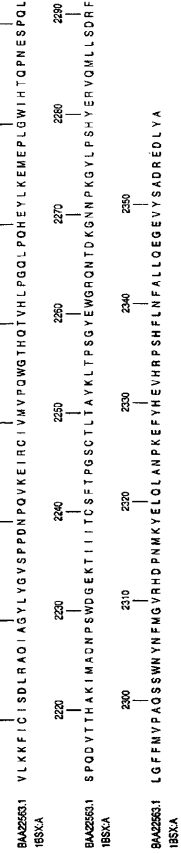
FIG. 7-3



【 7 - 4 】

AEye 出力(1月18日, 2001:5:45 PM)

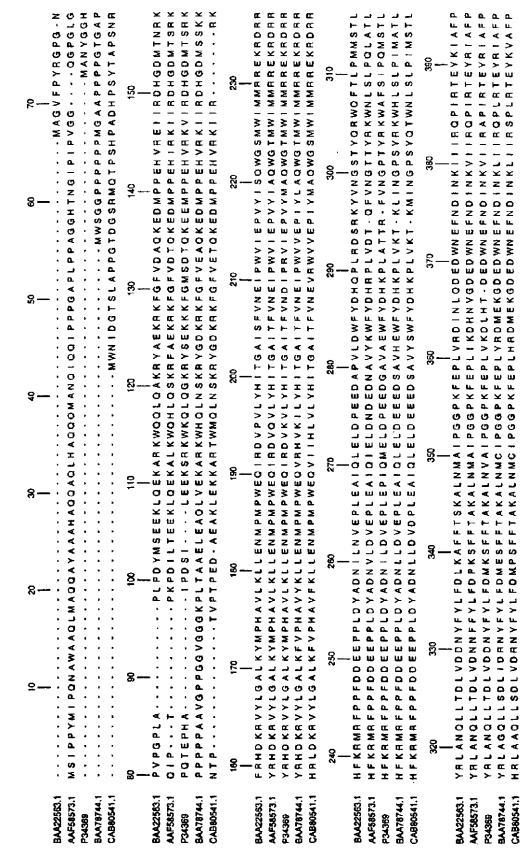
FIG. 7-4



【 9 - 1 】

AEye 出力(2月26日, 2001:9:53 AM)

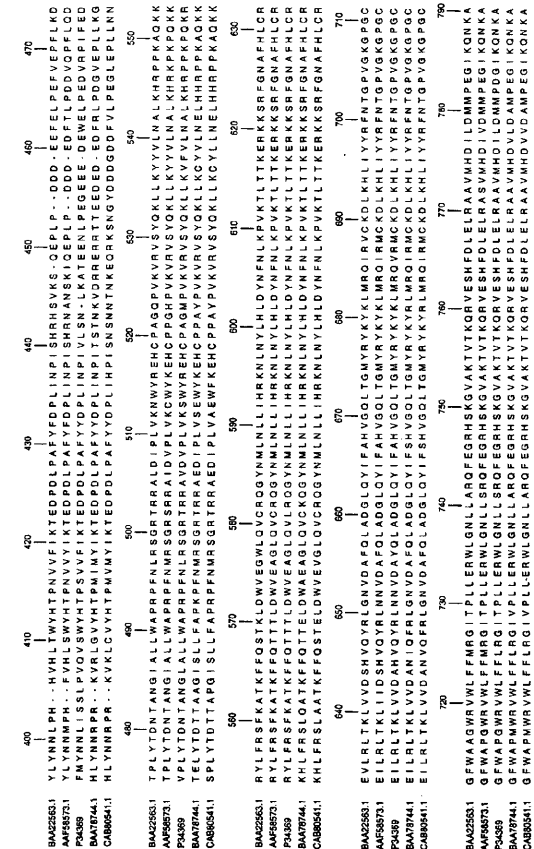
FIG. 9-1



【 9 - 2 】

AEye 出力(2月26日, 2001:9:53 AM)

FIG. 9-2



【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
12 September 2002 (12.09.2002)

PCT

(10) International Publication Number
WO 02/070563 A2

- (51) International Patent Classification: C07K 14/705 2NU (GB). POTTER, Sarah, Jane [GB/GB]; 39 Kneller Road, Brackley, London SE14 2AR (GB).
- (21) International Application Number: PCT/GB02/00986
- (22) International Filing Date: 5 March 2002 (05.03.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 0105402.2 5 March 2001 (05.03.2001) GB
- (71) Applicant (for all designated States except US): INPHARMATICA LIMITED [GB/GB]; 60 Charlotte Street, London W1T 2NU (GB).
- (72) Inventors; and
(75) Inventors/Applicants (for US only): FAGAN, Richard, Joseph [GB/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB). PHELPS, Christopher, Benjamin [GB/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB). PHILLIPS, Tom [GB/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB). PIERRON, Valerie, Nathalie [FR/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB). ALLEN, Kathryn, Elizabeth [GB/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB). ALLEN, Janet, Marjorie [GB/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB).
- (74) Agents: MERCER, Christopher, Paul et al.; Carpmaels & Ransford, 43 Bloomsbury Square, London WC1A 2RA (GB).
- (81) Designated States (national): AE, AG, AI, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KI, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, 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 (BI, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NI, SN, TD, TG).
- Published:**
without international search report and to be republished upon receipt of that report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 02/070563 A2

(54) Title: NUCLEAR HORMONE RECEPTOR LIGAND BINDING DOMAIN

(57) Abstract: This invention relates to a novel protein, termed BAA22563.1, herein identified as a Nuclear Hormone Receptor Ligand Binding Domain and to the use of this protein and nucleic acid sequence from the encoding gene in the diagnosis, prevention and treatment of disease.

WO 02/070563

PCT/GB02/00986

1

NUCLEAR HORMONE RECEPTOR LIGAND BINDING DOMAIN

This invention relates to a novel protein, termed BAA22563.1 herein identified as a Nuclear Hormone Receptor Ligand Binding Domain and to the use of this protein and nucleic acid sequence from the encoding gene in the diagnosis, prevention and treatment of disease.

All publications, patents and patent applications cited herein are incorporated in full by reference.

BACKGROUND

The process of drug discovery is presently undergoing a fundamental revolution as the era of functional genomics comes of age. The term "functional genomics" applies to an approach utilising bioinformatics tools to ascribe function to protein sequences of interest. Such tools are becoming increasingly necessary as the speed of generation of sequence data is rapidly outpacing the ability of research laboratories to assign functions to these protein sequences.

As bioinformatics tools increase in potency and in accuracy, these tools are rapidly replacing the conventional techniques of biochemical characterisation. Indeed, the advanced bioinformatics tools used in identifying the present invention are now capable of outputting results in which a high degree of confidence can be placed.

Various institutions and commercial organisations are examining sequence data as they become available and significant discoveries are being made on an on-going basis. However, there remains a continuing need to identify and characterise further genes and the polypeptides that they encode, as targets for research and for drug discovery.

Recently, a remarkable tool for the evaluation of sequences of unknown function has been developed by the Applicant for the present invention. This tool is a database system, termed the Biopendium search database, that is the subject of co-pending International Patent Application No. PCT/GB01/01105. This database system consists of an integrated data resource created using proprietary technology and containing information generated from an all-by-all comparison of all available protein or nucleic acid sequences.

The aim behind the integration of these sequence data from separate data resources is to

WO 02/070563

PCT/GB02/00986

2

combine as much data as possible, relating both to the sequences themselves and to information relevant to each sequence, into one integrated resource. All the available data relating to each sequence, including data on the three-dimensional structure of the encoded protein, if this is available, are integrated together to make best use of the information that is known about each sequence and thus to allow the most educated predictions to be made from comparisons of these sequences. The annotation that is generated in the database and which accompanies each sequence entry imparts a biologically relevant context to the sequence information.

This data resource has made possible the accurate prediction of protein function from sequence alone. Using conventional technology, this is only possible for proteins that exhibit a high degree of sequence identity (above about 20%-30% identity) to other proteins in the same functional family. Accurate predictions are not possible for proteins that exhibit a very low degree of sequence homology to other related proteins of known function.

In the present case, a protein whose sequence is recorded in a publicly available database as BAA22563.1 (NCBI Genebank nucleotide accession number AB007510.1 and a Genebank protein accession number BAA22563.1), is implicated as a novel member of the Nuclear Hormone Receptor Ligand Binding Domain family.

I. Introduction to Nuclear Hormone Receptor Ligand Binding Domains

The Nuclear Hormone Receptor gene superfamily (see Table 1) encodes structurally related proteins that regulate the transcription of target genes. These proteins include receptors for steroid and thyroid hormones, vitamins, and other proteins for which no ligands have been found. Nuclear Receptors are composed of two key domains, a DNA-Binding Domain (DBD) and a Ligand Binding Domain (LBD). The DBD directs the receptors to bind specific DNA sequences as monomers, homodimers, or heterodimers. The DBD is a particular type of zinc-finger, found only in Nuclear Receptors. Nuclear Receptors with DBDs can be readily identified at the sequence level by searching for matches to the PROSITE consensus sequence (PS00031).

WO 02/070563

PCT/GB02/00986

3

The Ligand Binding Domain (LBD) binds and responds to the cognate hormone. Ligand binding to the LBD triggers a conformational change which expels a bound "Nuclear Receptor Co-Repressor". The site previously occupied by the Co-Repressor is then free to recruit a "Nuclear Receptor Co-Activator". This Ligand-triggered swap of a Co-Repressor for a Co-Activator is the mechanism by which Ligand binding leads to the transcriptional activation of target genes. All ligand binding domains contain a consensus sequence, the "LBD motif" (see Table 2) which mediates Co-Repressor and Co-Activator binding. The LBD is the binding site for all Nuclear Hormone Receptor targeted drugs to date and it is thus desirable to identify novel Ligand Binding Domains since these will be attractive drug targets. Ligand Binding Domains share low sequence identity (~15%) but have very similar structures and so present ideal targets for a structure-based relationship tool such as Genome Threader.

Many protein sequences have already been annotated in the public domain as Nuclear Hormone Receptors by their possession of DBDs using basic search tools like PROSITE, and their LBDs inferred on the basis of this. Because of this it is anticipated that any novel LBDs identified by Genome Threader *which are not annotated as nuclear receptors* will lack the DBD entirely. A precedent for a protein which has an LBD but lacks a DBD is provided by DAX1. Thus we annotate these DBD-less hits not as "Nuclear Hormone Receptors" but rather as containing a "Nuclear Hormone Receptor Ligand Binding Domain".

Family: Steroid Hormone Receptors	
Subfamilies	Glucocorticoid Receptors
	Progesterone Receptors
	Androgen Receptors
	Estrogen Receptors
Family: Thyroid Hormone Receptor-like Factors	
Subfamilies	Retinoic Acid Receptors (RARs)

WO 02/070563

PCT/GB02/00986

4

	Retinoid X Receptors (RXRs)
	Thyroid Hormone Receptors
	Vitamin D Receptor
	NGFI-B
	FTZ-F1
	Peroxisome Proliferator Activated Receptors (PPARs)
	Ecdysone Receptors
	Retinoid Orphan Receptors (RORs)
	Tailless/COUP
	HNF-4
	CF1
	Knirps
<i>Family: DAX1</i>	
Subfamilies	DAX1

Table 1: Nuclear hormone Receptor Superfamily

WO 02/070563

PCT/GB02/00986

5

1	2	3	4	5	6	7	8	9	10	11	12	13
L	Any residues	2	L	Any 3 residues (or 2 residues or 4 residues)			D	Q	Any residues (or 1 or 3 residues)	2	L	L
I			I				E	N			I	I
A			A					R			A	A
V			V					H			V	V
M			M					K			M	M
F			F					S			F	F
Y			Y					T			Y	Y
W			W								W	W

Table 2: The "LBD motif". Numbers along the top row refer to residue position within the motif. Letters refer to amino acids by the 1-letter code. Letters within one column are all acceptable for that position within the motif. For example L, I, A, V, M, F, Y or W can occupy the first position of the "LBD motif". Note that there is observed variation in the number of residues found between position 4 and 8, and position 9 and 12. The "LBD motif" was constructed by aligning 681 sequences of Nuclear Hormone Receptor Ligand Binding Domains, and identifying conserved patterns of residues.

15

WO 02/070563

PCT/GB02/00986

6

II. Nuclear Hormone Receptors and Disease

Nuclear Hormone Receptors have been shown to play a role in diverse physiological functions, many of which can play a role in disease processes (see Table 3).

Nuclear Hormone Receptor	Disease
Androgen Receptor	Androgen Insensitivity Syndrome (Lubahn <i>et al.</i> 1989 Proc. Natl. Acad. Sci. USA 86, 9534-9538).
	Reifenstein syndrome (Wooster <i>et al.</i> 1992 Nat. Genet. 2, 132-134).
	X-linked recessive spinal and bulbar muscular atrophy (MacLean <i>et al.</i> 1995 Mol. Cell. Endocrinol. 112,133-141).
	Male breast cancer ((Wooster <i>et al.</i> 1992 Nat. Genet. 2, 132-134).
Glucocorticoid Receptor	Nelson's syndrome (Karl <i>et al.</i> 1996 J. Clin. Endocrinol. Metab. 81, 124-129).
	Glucocorticoid resistant acute T-cell leukemia (Hala <i>et al.</i> 1996 Int. J. Cancer 68, 663-668).
Mineralocorticoid Receptor	Pseudohypoaldosteronism (Chung <i>et al.</i> 1995 J. Clin. Endocrinol. Metab. 80, 3341-3345).
Estrogen Receptor alpha	ER alpha expression is elevated in a subset of human breast cancers. The application of Tamoxifen is the major therapy to prevent breast tumour progression. Unfortunately 35% of ER alpha positive breast cancers are Tamoxifen resistant (Petrangeli <i>et al.</i> 1994 J. Steroid Biochem. Mol. Biol. 49, 327-331).
Vitamin D3 Receptor	Mutations in the Vitamin D3 receptor produce a hereditary disorder similar in phenotype to Vitamin D3 deficiency (Rickets) (Hughes <i>et al.</i> 1988 Science 242, 1702-1725).
Retinoic Acid Receptor alpha	Acute Myeloid Leukemia (Lavau and Dejean 1994 Leukemia 8, 9-15).
Thyroid Hormone Receptor beta	"Generalised Resistance to Thyroid Hormones" (GRTH) (Refetoff 1994 Thyroid 4, 345-349).
DAX1	X-linked Adrenal Hypoplasia Congenita (AHC) and Hypogonadism (Ito <i>et al.</i> 1997 Mol. Cell. Biol. 17, 1476-1483).

Table 3. Nuclear Hormone Receptors and disease.

5

WO 02/070563

PCT/GB02/00986

7

Alteration of Nuclear Hormone Receptors by ligands which bind to their LBD thus provides a means to alter the disease phenotype. There is thus a great need for the identification of novel Nuclear Hormone Receptor Ligand Binding Domains, as these proteins may play a role in the diseases identified above, as well as in other disease states.

- 5 The identification of novel Nuclear Hormone Receptor Ligand Binding Domains is thus highly relevant for the treatment and diagnosis of disease, particularly those identified in Table 3.

THE INVENTION

- 10 The invention is based on the discovery that the BAA22563.1 protein functions as a Nuclear Hormone Receptor Ligand Binding Domain.

For the BAA22563.1 protein, it has been found that a region including residues 1104-1309 of this protein sequence adopts an equivalent fold to residues 6 (Asp216) to 209 (Asp427) of the Thyroid Hormone Receptor beta (PDB code 1BSX:A). Thyroid
15 Hormone Receptor beta is known to function as a Nuclear Hormone Receptor Ligand Binding Domain. Furthermore, the "LBD motif" residues PHE293, LEU296, ASP300, GLN301, LEU304 and LEU305 of the Thyroid Hormone Receptor beta are conserved as PHE1174, VAL1177, ASP1181, ASN1182, LEU1185 and LEU1186 in BAA22563.1, respectively. This relationship is not just to Thyroid Hormone Receptor beta, but rather to
20 the Nuclear Hormone Receptor Ligand Binding Domain family as a whole. Thus, by reference to the Genome Threader™ alignment of BAA22563.1 with the Thyroid Hormone Receptor beta (1BSX:A) PHE1174, VAL1177, ASP1181, ASN1182, LEU1185 and LEU1186 of BAA22563.1 are predicted to form the "LBD motif" residues.

The combination of equivalent fold and conservation of "LBD motif" residues allows the
25 functional annotation of this region of BAA22563.1, and therefore proteins that include this region, as possessing Nuclear Hormone Receptor Ligand Binding Domain activity.

In a first aspect, the invention provides a polypeptide, which polypeptide:

- (i) comprises the amino acid sequence as recited in SEQ ID NO:2;
- (ii) is a fragment thereof having Nuclear Hormone Receptor Ligand Binding Domain

WO 02/070563

PCT/GB02/00986

8

activity or having an antigenic determinant in common with the polypeptides of (i); or

(iii) is a functional equivalent of (i) or (ii).

Preferably, the polypeptide:

- 5 (i) consists of the amino acid sequence as recited in SEQ ID NO:2;
- (ii) is a fragment thereof having Nuclear Hormone Receptor Ligand Binding Domain activity or having an antigenic determinant in common with the polypeptides of (i); or
- (iii) is a functional equivalent of (i) or (ii).

10 The polypeptide having the sequence recited in SEQ ID NO:2 is referred to hereafter as "the LBDG2 polypeptide".

According to this aspect of the invention, a preferred polypeptide fragment according to part ii) above includes the region of the LBDG2 polypeptide that is predicted as that responsible for Nuclear Hormone Receptor Ligand Binding Domain activity (hereafter, 15 the "LBDG2 Nuclear Hormone Receptor Ligand Binding Domain region"), or is a variant thereof that possesses the "LBD motif" (PHE1174, VAL1177, ASP1181, ASN1182, LEU1185 and LEU1186, or equivalent residues). As defined herein, the LBDG2 Nuclear Hormone Receptor Ligand Binding Domain region is considered to extend between residue 1104 and residue 1309 of the LBDG2 polypeptide sequence.

20 This aspect of the invention also includes fusion proteins that incorporate polypeptide fragments and variants of these polypeptide fragments as defined above, provided that said fusion proteins possess activity as a Nuclear Hormone Receptor Ligand Binding Domain.

In a second aspect, the invention provides a purified nucleic acid molecule that encodes a 25 polypeptide of the first aspect of the invention. Preferably, the purified nucleic acid molecule has the nucleic acid sequence as recited in SEQ ID NO:1 (encoding the LBDG2 polypeptide), or is a redundant equivalent or fragment of this sequence. A preferred nucleic acid fragment is one that encodes a polypeptide fragment according to part ii) above, preferably a polypeptide fragment that includes the LBDG2 Nuclear Hormone

WO 02/070563

PCT/GB02/00986

9

Receptor Ligand Binding Domain region, or that encodes a variant of these fragments as this term is defined above.

In a third aspect, the invention provides a purified nucleic acid molecule which hybridizes under high stringency conditions with a nucleic acid molecule of the second aspect of the invention.

In a fourth aspect, the invention provides a vector, such as an expression vector, that contains a nucleic acid molecule of the second or third aspect of the invention.

In a fifth aspect, the invention provides a host cell transformed with a vector of the fourth aspect of the invention.

In a sixth aspect, the invention provides a ligand which binds specifically to, and which preferably inhibits the Nuclear Hormone Receptor Ligand Binding Domain activity of, a polypeptide of the first aspect of the invention.

In a seventh aspect, the invention provides a compound that is effective to alter the expression of a natural gene which encodes a polypeptide of the first aspect of the invention or to regulate the activity of a polypeptide of the first aspect of the invention.

A compound of the seventh aspect of the invention may either increase (agonise) or decrease (antagonise) the level of expression of the gene or the activity of the polypeptide. Importantly, the identification of the function of the region defined herein as the LBDG2 Nuclear Hormone Receptor Ligand Binding Domain region of the LBDG2 polypeptide, respectively, allows for the design of screening methods capable of identifying compounds that are effective in the treatment and/or diagnosis of diseases in which Nuclear Hormone Receptor Ligand Binding Domains are implicated.

In an eighth aspect, the invention provides a polypeptide of the first aspect of the invention, or a nucleic acid molecule of the second or third aspect of the invention, or a vector of the fourth aspect of the invention, or a ligand of the fifth aspect of the invention, or a compound of the sixth aspect of the invention, for use in therapy or diagnosis.

The inventors have discovered that the mRNA for LBDG2 is found in extracts from a variety of human tissues. The finding of high levels of the transcript in the human spleen is consistent with a role of LBDG2 in the immune system and in particular in lymphocyte

WO 02/070563

PCT/GB02/00986

10

development and function and in particular in B cell development and function. It is therefore considered that the development of agonists and antagonists for LBDG2 may have a particular role in the therapeutic intervention in various human diseases of the immune system including autoimmunity, allergies and diseases associated with immunoglobulin dysfunction. These diseases include type I diabetes mellitus, rheumatoid arthritis, multiple sclerosis, psoriasis, renal failure arising from glomerulopathies, scleroderma, inflammatory bowel disease (both Crohns disease and ulcerative colitis), transplant rejection, asthma, atopic dermatitis, eczema, myelomas and in infectious diseases that require production of antibodies, for example, intracellular pathogen such as virus infected cells, tuberculosis, listeria.

Messenger RNA for LBDG2 has also been found in human B cell lines such as Daudi, IM9 and Raji cells. These findings are consistent with the discovery mentioned above, that LBDG2 mRNA is found in the spleen. Finding high levels of expression of the mRNA in U937 cells suggests a role for LBDG2 in monocyte/macrophage functions and, as such, agonists or antagonists may be valuable in treating inflammatory diseases including chronic obstructive pulmonary disease (COPD), osteoarthritis, rheumatoid arthritis, inflammatory bowel disease, fibrosis such as liver fibrosis (cirrhosis) and skin fibrosis (scarring), atherosclerosis, dementia, multiple sclerosis, inflammatory pain.

In addition, significant levels of LBDG2 have been found in adrenal, ovary, prostate and testis tissue. This indicates that the development of agonists and antagonists to LBDG2 may be of value in diseases such as benign prostatic hypertrophy, prostatic cancer, ovarian cancer and testicular cancer. In addition, agonists or antagonists for LBDG2 may be developed for treatment of diseases including but not exclusive to hypertension, responses to stress including stress of infectious diseases, regulation of salt and water homeostasis, control of fertility through regulation of ovulation (infertility and contraception), regulation of implantation (infertility and contraception) and regulation of spermatogenesis (infertility and contraception).

It has also been found that the mRNA for LBDG2 is expressed at significant levels in the human brain. This is noteworthy, as this provides a potential link to human disease states and development of agonists and antagonists for the ligand binding domain of LBDG2

WO 02/070563

PCT/GB02/00986

II

offers the potential for therapeutic intervention in various human diseases including cell proliferative disorders, including neoplasm, melanoma, lung, colorectal, breast, pancreas, head and neck and other solid tumours; myeloproliferative disorders, such as leukemia, non-Hodgkin lymphoma, leukopenia, thrombocytopenia, angiogenesis disorder, Kaposi' sarcoma; autoimmune/inflammatory disorders, including autoimmunity, allergies and diseases associated with immunoglobulin dysfunction, inflammatory bowel disease, arthritis, psoriasis and respiratory tract inflammation, asthma and organ transplant rejection; cardiovascular disorders, including hypertension, oedema, angina, atherosclerosis, thrombosis, sepsis, shock, reperfusion injury, heart arrhythmia, and ischemia; neurological disorders including, central nervous system disease, Alzheimer's disease, brain injury, stroke, amyotrophic lateral sclerosis, anxiety, depression, and pain; developmental disorders; metabolic disorders including diabetes mellitus, osteoporosis, lipid metabolism disorder, hyperthyroidism, hyperparathyroidism, hypercalcemia, hypercholesterolemia, hyperlipidemia, and obesity; renal disorders, including glomerulonephritis, renovascular hypertension; dermatological disorders, including, acne, eczema, and wound healing, negative effects of aging; AIDS; infections including viral infection, bacterial infection, fungal infection and parasitic infection and other pathological conditions, particularly those in which nuclear hormone receptors are implicated.

20 The finding of a "non-classical" nuclear hormone receptor such as LBDG2 which contains a ligand binding domain in the absence of a DNA binding domain is consistent with the known literature which has consistently reported widespread effects of steroids in the brain (known as neurosteroids) and that these effects, in general, are mediated not through the known classic steroid hormone nuclear receptors which requires transcriptional activation. For instance, neurosteroids have been shown to influence neurotransmission particularly in the field of receptors such as those for GABA and NMDA and Sigma receptors. Neurosteroids have been shown to play a neuroprotective role. Therapeutic intervention through the development of agonists (or antagonists) to LBDG2 may therefore have a role in treatment of neurodegenerative conditions such as dementia, Parkinson's disease and neurodegeneration following cerebrovascular disease such as infarction or haemorrhage (stroke) and trauma to the central nervous system and

WO 02/070563

PCT/GB02/00986

12

spinal cord. In addition, neurosteroids have been shown to influence cognitive processing, spatial learning and memory, anxiety and behaviours such as craving which leads to addictive behaviour patterns. Development of agonists and antagonists to LBDG2 may therefore lead to therapeutic intervention to treat dementias, learning
5 difficulties, anxiety, addictive behaviours such as but not exclusively alcoholism, eating disorders and drug addiction.

In a ninth aspect, the invention provides a method of diagnosing a disease in a patient, comprising assessing the level of expression of a natural gene encoding a polypeptide of the first aspect of the invention or the activity of a polypeptide of the first aspect of the
10 invention in tissue from said patient and comparing said level of expression or activity to a control level, wherein a level that is different to said control level is indicative of disease. Such a method will preferably be carried out *in vitro*. Similar methods may be used for monitoring the therapeutic treatment of disease in a patient, wherein altering the level of expression or activity of a polypeptide or nucleic acid molecule over the period
15 of time towards a control level is indicative of regression of disease.

A preferred method for detecting polypeptides of the first aspect of the invention comprises the steps of: (a) contacting a ligand, such as an antibody, of the sixth aspect of the invention with a biological sample under conditions suitable for the formation of a ligand-polypeptide complex; and (b) detecting said complex.

20 A number of different such methods according to the ninth aspect of the invention exist, as the skilled reader will be aware, such as methods of nucleic acid hybridization with short probes, point mutation analysis, polymerase chain reaction (PCR) amplification and methods using antibodies to detect aberrant protein levels. Similar methods may be used on a short or long term basis to allow therapeutic treatment of a disease to be monitored
25 in a patient. The invention also provides kits that are useful in these methods for diagnosing disease.

In a tenth aspect, the invention provides for the use of a polypeptide of the first aspect of the invention as a Nuclear Hormone Receptor Ligand Binding Domain. The invention also provides for the use of a nucleic acid molecule according to the second or third
30 aspects of the invention to express a protein that possesses Nuclear Hormone Receptor

WO 02/070563

PCT/GB02/00986

13

Ligand Binding Domain activity. The invention also provides a method for effecting Nuclear Hormone Receptor Ligand Binding Domain activity, said method utilising a polypeptide of the first aspect of the invention.

5 In an eleventh aspect, the invention provides a pharmaceutical composition comprising a polypeptide of the first aspect of the invention, or a nucleic acid molecule of the second or third aspect of the invention, or a vector of the fourth aspect of the invention, or a ligand of the sixth aspect of the invention, or a compound of the seventh aspect of the invention, in conjunction with a pharmaceutically-acceptable carrier.

10 In a twelfth aspect, the present invention provides a polypeptide of the first aspect of the invention, or a nucleic acid molecule of the second or third aspect of the invention, or a vector of the fourth aspect of the invention, or a ligand of the sixth aspect of the invention, or a compound of the seventh aspect of the invention, for use in the manufacture of a medicament for the diagnosis or treatment of a disease, such as cell proliferative disorders, including neoplasm, melanoma, lung, colorectal, breast, pancreas, head and neck and other solid tumours, myeloproliferative disorders, such as leukemia, 15 non-Hodgkin lymphoma, leukopenia, thrombocytopenia, angiogenesis disorder, Kaposis' sarcoma, autoimmune/inflammatory disorders, including allergy, inflammatory bowel disease, arthritis, psoriasis and respiratory tract inflammation, asthma, and organ transplant rejection, cardiovascular disorders, including hypertension, oedema, angina, atherosclerosis, thrombosis, sepsis, shock, reperfusion injury, heart arrhythmia, and ischemia, neurological disorders including, central nervous system disease, Alzheimer's disease, brain injury, stroke, amyotrophic lateral sclerosis, anxiety, depression, and pain, developmental disorders, metabolic disorders including diabetes mellitus, osteoporosis, lipid metabolism disorder, hyperthyroidism, hyperparathyroidism, hypercalcemia, 20 hypercholesterolemia, hyperlipidemia, and obesity, renal disorders, including glomerulonephritis, renovascular hypertension, dermatological disorders, including, acne, eczema, and wound healing, negative effects of aging, AIDS, infections including viral infection, bacterial infection, fungal infection and parasitic infection and other pathological conditions, particularly those in which nuclear hormone receptors are implicated, as well as the other more specific diseases and conditions mentioned above in 30 connection with the eighth aspect of the invention.

WO 02/070563

PCT/GB02/00986

14

In a thirteenth aspect, the invention provides a method of treating a disease in a patient comprising administering to the patient a polypeptide of the first aspect of the invention, or a nucleic acid molecule of the second or third aspect of the invention, or a vector of the fourth aspect of the invention, or a ligand of the sixth aspect of the invention, or a compound of the seventh aspect of the invention.

For diseases in which the expression of a natural gene encoding a polypeptide of the first aspect of the invention, or in which the activity of a polypeptide of the first aspect of the invention, is lower in a diseased patient when compared to the level of expression or activity in a healthy patient, the polypeptide, nucleic acid molecule, ligand or compound administered to the patient should be an agonist. Conversely, for diseases in which the expression of the natural gene or activity of the polypeptide is higher in a diseased patient when compared to the level of expression or activity in a healthy patient, the polypeptide, nucleic acid molecule, ligand or compound administered to the patient should be an antagonist. Examples of such antagonists include antisense nucleic acid molecules, ribozymes and ligands, such as antibodies.

In a fourteenth aspect, the invention provides transgenic or knockout non-human animals that have been transformed to express higher, lower or absent levels of a polypeptide of the first aspect of the invention. Such transgenic animals are very useful models for the study of disease and may also be using in screening regimes for the identification of compounds that are effective in the treatment or diagnosis of such a disease.

A summary of standard techniques and procedures which may be employed in order to utilise the invention is given below. It will be understood that this invention is not limited to the particular methodology, protocols, cell lines, vectors and reagents described. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only and it is not intended that this terminology should limit the scope of the present invention. The extent of the invention is limited only by the terms of the appended claims.

Standard abbreviations for nucleotides and amino acids are used in this specification.

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology, microbiology, recombinant DNA

WO 02/070563

PCT/GB02/00986

15

technology and immunology, which are within the skill of those working in the art.

Such techniques are explained fully in the literature. Examples of particularly suitable texts for consultation include the following: Sambrook *Molecular Cloning; A Laboratory Manual*, Second Edition (1989); DNA Cloning, Volumes I and II (D.N. Glover ed. 1985);
5 Oligonucleotide Synthesis (M.J. Gait ed. 1984); Nucleic Acid Hybridization (B.D. Hames & S.J. Higgins eds. 1984); Transcription and Translation (B.D. Hames & S.J. Higgins eds. 1984); Animal Cell Culture (R.I. Freshney ed. 1986); Immobilized Cells and Enzymes (IRL Press, 1986); B. Perbal, *A Practical Guide to Molecular Cloning* (1984); the Methods in *Enzymology* series (Academic Press, Inc.), especially volumes 154 &
10 155; Gene Transfer Vectors for Mammalian Cells (J.H. Miller and M.P. Calos eds. 1987, Cold Spring Harbor Laboratory); *Immunochemical Methods in Cell and Molecular Biology* (Mayer and Walker, eds. 1987, Academic Press, London); Scopes, (1987) *Protein Purification: Principles and Practice*, Second Edition (Springer Verlag, N.Y.); and *Handbook of Experimental Immunology*, Volumes I-IV (D.M. Weir and C. C. Blackwell eds. 1986).
15

As used herein, the term "polypeptide" includes any peptide or protein comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds, i.e. peptide isosteres. This term refers both to short chains (peptides and oligopeptides) and to longer chains (proteins).

20 The polypeptide of the present invention may be in the form of a mature protein or may be a pre-, pro- or prepro- protein that can be activated by cleavage of the pre-, pro- or prepro- portion to produce an active mature polypeptide. In such polypeptides, the pre-, pro- or prepro- sequence may be a leader or secretory sequence or may be a sequence that is employed for purification of the mature polypeptide sequence.

25 The polypeptide of the first aspect of the invention may form part of a fusion protein. For example, it is often advantageous to include one or more additional amino acid sequences which may contain secretory or leader sequences, pro-sequences, sequences which aid in purification, or sequences that confer higher protein stability, for example during recombinant production. Alternatively or additionally, the mature polypeptide may be
30 fused with another compound, such as a compound to increase the half-life of the

WO 02/070563

PCT/GB02/00986

16

polypeptide (for example, polyethylene glycol).

Polypeptides may contain amino acids other than the 20 gene-encoded amino acids, modified either by natural processes, such as by post-translational processing or by chemical modification techniques which are well known in the art. Among the known modifications which may commonly be present in polypeptides of the present invention are glycosylation, lipid attachment, sulphation, gamma-carboxylation, for instance of glutamic acid residues, hydroxylation and ADP-ribosylation. Other potential modifications include acetylation, acylation, amidation, covalent attachment of flavin, covalent attachment of a haeme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulphide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, GPI anchor formation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination.

Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. In fact, blockage of the amino or carboxyl terminus in a polypeptide, or both, by a covalent modification is common in naturally-occurring and synthetic polypeptides and such modifications may be present in polypeptides of the present invention.

The modifications that occur in a polypeptide often will be a function of how the polypeptide is made. For polypeptides that are made recombinantly, the nature and extent of the modifications in large part will be determined by the post-translational modification capacity of the particular host cell and the modification signals that are present in the amino acid sequence of the polypeptide in question. For instance, glycosylation patterns vary between different types of host cell.

The polypeptides of the present invention can be prepared in any suitable manner. Such polypeptides include isolated naturally-occurring polypeptides (for example purified from cell culture), recombinantly-produced polypeptides (including fusion proteins),

WO 02/070563

PCT/GB02/00986

17

synthetically-produced polypeptides or polypeptides that are produced by a combination of these methods.

The functionally-equivalent polypeptides of the first aspect of the invention may be polypeptides that are homologous to the LBDG2 polypeptide. Two polypeptides are said to be "homologous", as the term is used herein, if the sequence of one of the polypeptides has a high enough degree of identity or similarity to the sequence of the other polypeptide. "Identity" indicates that at any particular position in the aligned sequences, the amino acid residue is identical between the sequences. "Similarity" indicates that, at any particular position in the aligned sequences, the amino acid residue is of a similar type between the sequences. Degrees of identity and similarity can be readily calculated (Computational Molecular Biology, Lesk, A.M., ed., Oxford University Press, New York, 1988; Biocomputing. Informatics and Genome Projects, Smith, D.W., ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part 1, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; and Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991).

Homologous polypeptides therefore include natural biological variants (for example, allelic variants or geographical variations within the species from which the polypeptides are derived) and mutants (such as mutants containing amino acid substitutions, insertions or deletions) of the LBDG2 polypeptide. Such mutants may include polypeptides in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code. Typical such substitutions are among Ala, Val, Leu and Ile; among Ser and Thr; among the acidic residues Asp and Glu; among Asn and Gln; among the basic residues Lys and Arg; or among the aromatic residues Phe and Tyr. Particularly preferred are variants in which several, i.e. between 5 and 10, 1 and 5, 1 and 3, 1 and 2 or just 1 amino acids are substituted, deleted or added in any combination. Especially preferred are silent substitutions, additions and deletions, which do not alter the properties and activities of the protein. Also especially preferred in this regard are conservative substitutions.

WO 02/070563

PCT/GB02/00986

18

Such mutants also include polypeptides in which one or more of the amino acid residues includes a substituent group.

Typically, greater than 80% identity between two polypeptides (preferably, over a specified region) is considered to be an indication of functional equivalence. Preferably, functionally equivalent polypeptides of the first aspect of the invention have a degree of sequence identity with the LBDG2 polypeptide, or with active fragments thereof, of greater than 80%. More preferred polypeptides have degrees of identity of greater than 85%, 90%, 95%, 98% or 99%, respectively with the LBDG2 polypeptide, or with active fragments thereof.

Percentage identity, as referred to herein, is as determined using BLAST version 2.1.3 using the default parameters specified by the NCBI (the National Center for Biotechnology Information; <http://www.ncbi.nlm.nih.gov/>) [Blosom 62 matrix; gap open penalty=11 and gap extension penalty=1].

In the present case, preferred active fragments of the LBDG2 polypeptide are those that include the LBDG2 Nuclear Hormone Receptor Ligand Binding Domain region and which possess the "LBD motif" of residues PHE1174, VAL1177, ASP1181, ASN1182, LEU1185 and LEU1186, or equivalent residues. By "equivalent residues" is meant residues that are equivalent to the "LBD motif" residues, provided that the Nuclear Hormone Receptor Ligand Binding Domain region retains activity as a Nuclear Hormone Receptor Ligand Binding Domain. For example PHE1174 replaced by LEU, ILE, ALA, VAL, MET, TYR or TRP. For example VAL1177 may be replaced by LEU, ILE, ALA, MET, PHE, TYR or TRP. For example ASP1181 may be replaced by GLU. For example ASN1182 may be replaced by GLN, ARG, HIS, LYS, SER, THR. For example LEU1185 may be replaced by ILE, ALA, VAL, MET, PHE, TYR or TRP. For example LEU1186 may be replaced by ILE, ALA, VAL, MET, PHE, TYR or TRP. Accordingly, this aspect of the invention includes polypeptides that have degrees of identity of greater than 80%, preferably, greater than 85%, 90%, 95%, 98% or 99%, respectively, with the Nuclear Hormone Receptor Ligand Binding Domain region of the LBDG2 polypeptide and which possess the "LBD motif" of PHE1174, VAL1177, ASP1181, ASN1182, LEU1185 and LEU1186, or equivalent residues. As discussed above, the LBDG2

WO 02/070563

PCT/GB02/00986

19

Nuclear Hormone Receptor Ligand Binding Domain region is considered to extend between residue 1104 and residue 1309 of the LBDG2 polypeptide sequence.

The functionally-equivalent polypeptides of the first aspect of the invention may also be polypeptides which have been identified using one or more techniques of structural alignment. For example, the Inpharmatica Genome Threader™ technology that forms one aspect of the search tools used to generate the Biopendium search database may be used (see co-pending International patent application PCT/GB01/01105) to identify polypeptides of presently-unknown function which, while having low sequence identity as compared to the LBDG2 polypeptide, are predicted to have Nuclear Hormone Receptor Ligand Binding Domain activity, by virtue of sharing significant structural homology with the LBDG2 polypeptide sequence.

By "significant structural homology" is meant that the Inpharmatica Genome Threader™ predicts two proteins, or protein regions, to share structural homology with a certainty of at least 10% more preferably, at least 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% and above. The certainty value of the Inpharmatica Genome Threader™ is calculated as follows. A set of comparisons was initially performed using the Inpharmatica Genome Threader™ exclusively using sequences of known structure. Some of the comparisons were between proteins that were known to be related (on the basis of structure). A neural network was then trained on the basis that it needed to best distinguish between the known relationships and known not-relationships taken from the CATH structure classification (www.biochem.ucl.ac.uk/bsm/cath). This resulted in a neural network score between 0 and 1. However, again as the number of proteins that are related and the number that are unrelated were known, it was possible to partition the neural network results into packets and calculate empirically the percentage of the results that were correct. In this manner, any genuine prediction in the Biopendium search database has an attached neural network score and the percentage confidence is a reflection of how successful the Inpharmatica Genome Threader™ was in the training/testing set.

Structural homologues of LBDG2 should share structural homology with the LBDG2 Nuclear Hormone Receptor Ligand Binding Domain region and possess the "LBD motif" residues PHE1174, VAL1177, ASP1181, ASN1182, LEU1185 and LEU1186, or

WO 02/070563

PCT/GB02/00986

20

equivalent residues. Such structural homologues are predicted to have Nuclear Hormone Receptor Ligand Binding Domain activity by virtue of sharing significant structural homology with this polypeptide sequence and possessing the "LBD motif" residues.

The polypeptides of the first aspect of the invention also include fragments of the LBDG2 polypeptide, functional equivalents of the fragments of the LBDG2 polypeptide, and fragments of the functional equivalents of the LBDG2 polypeptides, provided that those functional equivalents and fragments retain Nuclear Hormone Receptor Ligand Binding Domain activity or have an antigenic determinant in common with the LBDG2 polypeptide.

As used herein, the term "fragment" refers to a polypeptide having an amino acid sequence that is the same as part, but not all, of the amino acid sequence of the LBDG2 polypeptides or one of its functional equivalents. The fragments should comprise at least n consecutive amino acids from the sequence and, depending on the particular sequence, n preferably is 7 or more (for example, 8, 10, 12, 14, 16, 18, 20 or more). Small fragments may form an antigenic determinant.

Preferred polypeptide fragments according to this aspect of the invention are fragments that include a region defined herein as the LBDG2 Nuclear Hormone Receptor Ligand Binding Domain region of the LBDG2 polypeptides, respectively. These regions are the regions that have been annotated as a Nuclear Hormone Receptor Ligand Binding Domain.

For the LBDG2 polypeptide, this region is considered to extend between residue 1104 and residue 1309.

Variants of this fragment are included as embodiments of this aspect of the invention, provided that these variants possess activity as a Nuclear Hormone Receptor Ligand Binding Domain.

In one respect, the term "variant" is meant to include extended or truncated versions of this polypeptide fragment.

For extended variants, it is considered highly likely that the Nuclear Hormone Receptor Ligand Binding Domain region of the LBDG2 polypeptide will fold correctly and show

WO 02/070563

PCT/GB02/00986

21

Nuclear Hormone Receptor Ligand Binding Domain activity if additional residues C terminal and/or N terminal of these boundaries in the LBDG2 polypeptide sequence are included in the polypeptide fragment. For example, an additional 5, 10, 20, 30, 40 or even 50 or more amino acid residues from the LBDG2 polypeptide sequence, or from a homologous sequence, may be included at either or both the C terminal and/or N terminal of the boundaries of the Nuclear Hormone Receptor Ligand Binding Domain regions of the LBDG2 polypeptide, without prejudicing the ability of the polypeptide fragment to fold correctly and exhibit Nuclear Hormone Receptor Ligand Binding Domain activity.

For truncated variants of the LBDG2 polypeptide, one or more amino acid residues may be deleted at either or both the C terminus or the N terminus of the Nuclear Hormone Receptor Ligand Binding Domain region of the LBDG2 polypeptide, although the "LBD motif" residues (PHE1174, VAL1177, ASP1181, ASN1182, LEU1185 and LEU1186), or equivalent residues should be maintained intact; deletions should not extend so far into the polypeptide sequence that any of these residues are deleted.

In a second respect, the term "variant" includes homologues of the polypeptide fragments described above, that possess significant sequence homology with the Nuclear Hormone Receptor Ligand Binding Domain region of the LBDG2 polypeptide and which possess the "LBD motif" residues (PHE1174, VAL1177, ASP1181, ASN1182, LEU1185 and LEU1186), or equivalent residues, provided that said variants retain activity as an Nuclear Hormone Receptor Ligand Binding Domain.

Homologues include those polypeptide molecules that possess greater than 80% identity with the LBDG2 Nuclear Hormone Receptor Ligand Binding Domain regions, of the LBDG2 polypeptides, respectively. Percentage identity is as determined using BLAST version 2.1.3 using the default parameters specified by the NCBI (the National Center for Biotechnology Information; <http://www.ncbi.nlm.nih.gov/>) [Blosom 62 matrix; gap open penalty=11 and gap extension penalty=1]. Preferably, variant homologues of polypeptide fragments of this aspect of the invention have a degree of sequence identity with the LBDG2 Nuclear Hormone Receptor Ligand Binding Domain regions, of the LBDG2 polypeptides, respectively, of greater than 80%. More preferred variant polypeptides have degrees of identity of greater than 85%, 90%, 95%, 98% or 99%, respectively with the

WO 02/070563

PCT/GB02/00986

22

LBDG2 Nuclear Hormone Receptor Ligand Binding Domain regions of the LBDG2, polypeptides, provided that said variants retain activity as a Nuclear Hormone Receptor Ligand Binding Domain. Variant polypeptides also include homologues of the truncated forms of the polypeptide fragments discussed above, provided that said variants retain
5 activity as a Nuclear Hormone Receptor Ligand Binding Domain.

The polypeptide fragments of the first aspect of the invention may be polypeptide fragments that exhibit significant structural homology with the structure of the polypeptide fragment defined by the LBDG2 Nuclear Hormone Receptor Ligand Binding Domain regions, of the LBDG2 polypeptide sequence, for example, as identified by the
10 Inpharmatica Genome Threader™. Accordingly, polypeptide fragments that are structural homologues of the polypeptide fragments defined by the LBDG2 Nuclear Hormone Receptor Ligand Binding Domain regions of the LBDG2 polypeptide sequence should adopt the same fold as that adopted by this polypeptide fragment, as this fold is defined above.

15 Structural homologues of the polypeptide fragment defined by the LBDG2 Nuclear Hormone Receptor Ligand Binding Domain region should also retain the "LBD motif" residues PHE1174, VAL1177, ASP1181, ASN1182, LEU1185 and LEU1186, or equivalent residues.

Such fragments may be "free-standing", i.e. not part of or fused to other amino acids or
20 polypeptides, or they may be comprised within a larger polypeptide of which they form a part or region. When comprised within a larger polypeptide, the fragment of the invention most preferably forms a single continuous region. For instance, certain preferred embodiments relate to a fragment having a pre- and/or pro- polypeptide region fused to the amino terminus of the fragment and/or an additional region fused to the carboxyl
25 terminus of the fragment. However, several fragments may be comprised within a single larger polypeptide.

The polypeptides of the present invention or their immunogenic fragments (comprising at least one antigenic determinant) can be used to generate ligands, such as polyclonal or monoclonal antibodies, that are immunospecific for the polypeptides. Such antibodies
30 may be employed to isolate or to identify clones expressing the polypeptides of the

WO 02/070563

PCT/GB02/00986

23

invention or to purify the polypeptides by affinity chromatography. The antibodies may also be employed as diagnostic or therapeutic aids, amongst other applications, as will be apparent to the skilled reader.

The term "immunospecific" means that the antibodies have substantially greater affinity for the polypeptides of the invention than their affinity for other related polypeptides in the prior art. As used herein, the term "antibody" refers to intact molecules as well as to fragments thereof, such as Fab, F(ab')₂ and Fv, which are capable of binding to the antigenic determinant in question. Such antibodies thus bind to the polypeptides of the first aspect of the invention.

If polyclonal antibodies are desired, a selected mammal, such as a mouse, rabbit, goat or horse, may be immunised with a polypeptide of the first aspect of the invention. The polypeptide used to immunise the animal can be derived by recombinant DNA technology or can be synthesized chemically. If desired, the polypeptide can be conjugated to a carrier protein. Commonly used carriers to which the polypeptides may be chemically coupled include bovine serum albumin, thyroglobulin and keyhole limpet haemocyanin. The coupled polypeptide is then used to immunise the animal. Serum from the immunised animal is collected and treated according to known procedures, for example by immunoaffinity chromatography.

Monoclonal antibodies to the polypeptides of the first aspect of the invention can also be readily produced by one skilled in the art. The general methodology for making monoclonal antibodies using hybridoma technology is well known (see, for example, Kohler, G. and Milstein, C., *Nature* 256: 495-497 (1975); Kozbor *et al.*, *Immunology Today* 4: 72 (1983); Cole *et al.*, 77-96 in *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. (1985).

Panels of monoclonal antibodies produced against the polypeptides of the first aspect of the invention can be screened for various properties, i.e., for isotype, epitope, affinity, etc. Monoclonal antibodies are particularly useful in purification of the individual polypeptides against which they are directed. Alternatively, genes encoding the monoclonal antibodies of interest may be isolated from hybridomas, for instance by PCR techniques known in the art, and cloned and expressed in appropriate vectors.

WO 02/070563

PCT/GB02/00986

24

Chimeric antibodies, in which non-human variable regions are joined or fused to human constant regions (see, for example, Liu *et al.*, Proc. Natl. Acad. Sci. USA, 84, 3439 (1987)), may also be of use.

The antibody may be modified to make it less immunogenic in an individual, for example by humanisation (see Jones *et al.*, Nature, 321, 522 (1986); Verhoeven *et al.*, Science, 239: 1534 (1988); Kabat *et al.*, J. Immunol., 147: 1709 (1991); Queen *et al.*, Proc. Natl. Acad. Sci. USA, 86, 10029 (1989); Gorman *et al.*, Proc. Natl. Acad. Sci. USA, 88: 34181 (1991); and Hodgson *et al.*, Bio/Technology 9: 421 (1991)). The term "humanised antibody", as used herein, refers to antibody molecules in which the CDR amino acids and selected other amino acids in the variable domains of the heavy and/or light chains of a non-human donor antibody have been substituted in place of the equivalent amino acids in a human antibody. The humanised antibody thus closely resembles a human antibody but has the binding ability of the donor antibody.

In a further alternative, the antibody may be a "bispecific" antibody, that is an antibody having two different antigen binding domains, each domain being directed against a different epitope.

Phage display technology may be utilised to select genes which encode antibodies with binding activities towards the polypeptides of the invention either from repertoires of PCR amplified V-genes of lymphocytes from humans screened for possessing the relevant antibodies, or from naive libraries (McCafferty, J. *et al.*, (1990), Nature 348, 552-554; Marks, J. *et al.*, (1992) Biotechnology 10, 779-783). The affinity of these antibodies can also be improved by chain shuffling (Clackson, T. *et al.*, (1991) Nature 352, 624-628).

Antibodies generated by the above techniques, whether polyclonal or monoclonal, have additional utility in that they may be employed as reagents in immunoassays, radioimmunoassays (RIA) or enzyme-linked immunosorbent assays (ELISA). In these applications, the antibodies can be labelled with an analytically-detectable reagent such as a radioisotope, a fluorescent molecule or an enzyme.

Preferred nucleic acid molecules of the second and third aspects of the invention are those which encode the polypeptide sequences recited in SEQ ID NO:2, and functionally

WO 02/070563

PCT/GB02/00986

25

equivalent polypeptides, including active fragments of the LBDG2 polypeptide, such as a fragment including the LBDG2 Nuclear Hormone Receptor Ligand Binding Domain region of the LBDG2 polypeptide sequence, or a homologue thereof.

5 Nucleic acid molecules encompassing these stretches of sequence form a preferred embodiment of this aspect of the invention.

These nucleic acid molecules may be used in the methods and applications described herein. The nucleic acid molecules of the invention preferably comprise at least n consecutive nucleotides from the sequences disclosed herein where, depending on the particular sequence, n is 10 or more (for example, 12, 14, 15, 18, 20, 25, 30, 35, 40 or
10 more).

The nucleic acid molecules of the invention also include sequences that are complementary to nucleic acid molecules described above (for example, for antisense or probing purposes).

15 Nucleic acid molecules of the present invention may be in the form of RNA, such as mRNA, or in the form of DNA, including, for instance cDNA, synthetic DNA or genomic DNA. Such nucleic acid molecules may be obtained by cloning, by chemical synthetic techniques or by a combination thereof. The nucleic acid molecules can be prepared, for example, by chemical synthesis using techniques such as solid phase phosphoramidite chemical synthesis, from genomic or cDNA libraries or by separation
20 from an organism. RNA molecules may generally be generated by the *in vitro* or *in vivo* transcription of DNA sequences.

The nucleic acid molecules may be double-stranded or single-stranded. Single-stranded DNA may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand.

25 The term "nucleic acid molecule" also includes analogues of DNA and RNA, such as those containing modified backbones, and peptide nucleic acids (PNA). The term "PNA", as used herein, refers to an antisense molecule or an anti-gene agent which comprises an oligonucleotide of at least five nucleotides in length linked to a peptide backbone of amino acid residues, which preferably ends in lysine. The terminal lysine confers

WO 02/070563

PCT/GB02/00986

26

solubility to the composition. PNAs may be pegylated to extend their lifespan in a cell, where they preferentially bind complementary single stranded DNA and RNA and stop transcript elongation (Nielsen, P.E. *et al.* (1993) *Anticancer Drug Des.* 8:53-63).

A nucleic acid molecule which encodes the polypeptide of SEQ ID NO:2, or an active fragment thereof, may be identical to the coding sequence of the nucleic acid molecule shown in SEQ ID NO:1. These molecules also may have a different sequence which, as a result of the degeneracy of the genetic code, encodes the polypeptide SEQ ID NO:2, or an active fragment of the LBDG2 polypeptide, such as a fragment including the LBDG2 Nuclear Hormone Receptor Ligand Binding Domain region, or a homologue thereof. The LBDG2 Nuclear Hormone Receptor Ligand Binding Domain region is considered to extend between residue 1104 and residue 1309 of the LBDG2 polypeptide sequence. In SEQ ID NO:1 the LBDG2 Nuclear Hormone Receptor Ligand Binding Domain region is thus encoded by a nucleic acid molecule including nucleotide 3351 to 3968. Nucleic acid molecules encompassing this stretch of sequence, and homologues of this sequence, form a preferred embodiment of this aspect of the invention.

Such nucleic acid molecules that encode the polypeptide of SEQ ID NO:2 may include, but are not limited to, the coding sequence for the mature polypeptide by itself; the coding sequence for the mature polypeptide and additional coding sequences, such as those encoding a leader or secretory sequence, such as a pro-, pre- or prepro- polypeptide sequence; the coding sequence of the mature polypeptide, with or without the aforementioned additional coding sequences, together with further additional, non-coding sequences, including non-coding 5' and 3' sequences, such as the transcribed, non-translated sequences that play a role in transcription (including termination signals), ribosome binding and mRNA stability. The nucleic acid molecules may also include additional sequences which encode additional amino acids, such as those which provide additional functionalities.

The nucleic acid molecules of the second and third aspects of the invention may also encode the fragments or the functional equivalents of the polypeptides and fragments of the first aspect of the invention.

As discussed above, a preferred fragment of the LBDG2 polypeptide is a fragment

WO 02/070563

PCT/GB02/00986

27

including the LBDG2 Nuclear Hormone Receptor Ligand Binding Domain region, or a homologue thereof. The Nuclear Hormone Receptor Ligand Binding Domain region is encoded by a nucleic acid molecule including nucleotide 3351 to 3968 of SEQ ID NO:1.

Functionally equivalent nucleic acid molecules according to the invention may be naturally-occurring variants such as a naturally-occurring allelic variant, or the molecules may be a variant that is not known to occur naturally. Such non-naturally occurring variants of the nucleic acid molecule may be made by mutagenesis techniques, including those applied to nucleic acid molecules, cells or organisms.

Among variants in this regard are variants that differ from the aforementioned nucleic acid molecules by nucleotide substitutions, deletions or insertions. The substitutions, deletions or insertions may involve one or more nucleotides. The variants may be altered in coding or non-coding regions or both. Alterations in the coding regions may produce conservative or non-conservative amino acid substitutions, deletions or insertions.

The nucleic acid molecules of the invention can also be engineered, using methods generally known in the art, for a variety of reasons, including modifying the cloning, processing, and/or expression of the gene product (the polypeptide). DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides are included as techniques which may be used to engineer the nucleotide sequences. Site-directed mutagenesis may be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, introduce mutations and so forth.

Nucleic acid molecules which encode a polypeptide of the first aspect of the invention may be ligated to a heterologous sequence so that the combined nucleic acid molecule encodes a fusion protein. Such combined nucleic acid molecules are included within the second or third aspects of the invention. For example, to screen peptide libraries for inhibitors of the activity of the polypeptide, it may be useful to express, using such a combined nucleic acid molecule, a fusion protein that can be recognised by a commercially-available antibody. A fusion protein may also be engineered to contain a cleavage site located between the sequence of the polypeptide of the invention and the sequence of a heterologous protein so that the polypeptide may be cleaved and purified

WO 02/070563

PCT/GB02/00986

28

away from the heterologous protein.

The nucleic acid molecules of the invention also include antisense molecules that are partially complementary to nucleic acid molecules encoding polypeptides of the present invention and that therefore hybridize to the encoding nucleic acid molecules (hybridization). Such antisense molecules, such as oligonucleotides, can be designed to recognise, specifically bind to and prevent transcription of a target nucleic acid encoding a polypeptide of the invention, as will be known by those of ordinary skill in the art (see, for example, Cohen, J.S., Trends in Pharm. Sci., 10, 435 (1989), Okano, J. Neurochem. 56, 560 (1991); O'Connor, J. Neurochem 56, 560 (1991); Lee *et al.*, Nucleic Acids Res 6, 3073 (1979); Cooney *et al.*, Science 241, 456 (1988); Dervan *et al.*, Science 251, 1360 (1991).

The term "hybridization" as used here refers to the association of two nucleic acid molecules with one another by hydrogen bonding. Typically, one molecule will be fixed to a solid support and the other will be free in solution. Then, the two molecules may be placed in contact with one another under conditions that favour hydrogen bonding. Factors that affect this bonding include: the type and volume of solvent; reaction temperature; time of hybridization; agitation; agents to block the non-specific attachment of the liquid phase molecule to the solid support (Denhardt's reagent or BLOTTO); the concentration of the molecules; use of compounds to increase the rate of association of molecules (dextran sulphate or polyethylene glycol); and the stringency of the washing conditions following hybridization (see Sambrook *et al.* [*supra*]).

The inhibition of hybridization of a completely complementary molecule to a target molecule may be examined using a hybridization assay, as known in the art (see, for example, Sambrook *et al.* [*supra*]). A substantially homologous molecule will then compete for and inhibit the binding of a completely homologous molecule to the target molecule under various conditions of stringency, as taught in Wahl, G.M. and S.L. Berger (1987; Methods Enzymol. 152:399-407) and Kimmel, A.R. (1987; Methods Enzymol. 152:507-511).

"Stringency" refers to conditions in a hybridization reaction that favour the association of very similar molecules over association of molecules that differ. High stringency

WO 02/070563

PCT/GB02/00986

29

hybridisation conditions are defined as overnight incubation at 42°C in a solution comprising 50% formamide, 5XSSC (150mM NaCl, 15mM trisodium citrate), 50mM sodium phosphate (pH7.6), 5x Denhardts solution, 10% dextran sulphate, and 20 microgram/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 5 0.1X SSC at approximately 65°C. Low stringency conditions involve the hybridisation reaction being carried out at 35°C (see Sambrook *et al.* [*supra*]). Preferably, the conditions used for hybridization are those of high stringency.

Preferred embodiments of this aspect of the invention are nucleic acid molecules that are at least 80% identical over their entire length to a nucleic acid molecule encoding the 10 LBDG2 polypeptide (SEQ ID NO:2), and nucleic acid molecules that are substantially complementary to such nucleic acid molecules. A preferred active fragment is a fragment that includes an LBDG2 Nuclear Hormone Receptor Ligand Binding Domain region of the LBDG2 polypeptide sequences, respectively. Accordingly, preferred nucleic acid molecules include those that are at least 80% identical over their entire length to a nucleic 15 acid molecule encoding the Nuclear Hormone Receptor Ligand Binding Domain region of the LBDG2 polypeptide sequence.

Percentage identity, as referred to herein, is as determined using BLAST version 2.1.3 using the default parameters specified by the NCBI (the National Center for Biotechnology Information; <http://www.ncbi.nlm.nih.gov/>).

20 Preferably, a nucleic acid molecule according to this aspect of the invention comprises a region that is at least 80% identical over its entire length to the nucleic acid molecule having the sequence given in SEQ ID NO:1 to a region including nucleotides 3351-3968 of this sequence, or a nucleic acid molecule that is complementary to any one of these regions of nucleic acid. In this regard, nucleic acid molecules at least 90%, preferably at 25 least 95%, more preferably at least 98% or 99% identical over their entire length to the same are particularly preferred. Preferred embodiments in this respect are nucleic acid molecules that encode polypeptides which retain substantially the same biological function or activity as the LBDG2 polypeptide.

The invention also provides a process for detecting a nucleic acid molecule of the 30 invention, comprising the steps of: (a) contacting a nucleic probe according to the

WO 02/070563

PCT/GB02/00986

30

invention with a biological sample under hybridizing conditions to form duplexes; and
(b) detecting any such duplexes that are formed.

As discussed additionally below in connection with assays that may be utilised according
to the invention, a nucleic acid molecule as described above may be used as a
5 hybridization probe for RNA, cDNA or genomic DNA, in order to isolate full-length
cDNAs and genomic clones encoding the LBDG2 polypeptide and to isolate cDNA and
genomic clones of homologous or orthologous genes that have a high sequence similarity
to the gene encoding this polypeptide.

In this regard, the following techniques, among others known in the art, may be utilised
10 and are discussed below for purposes of illustration. Methods for DNA sequencing and
analysis are well known and are generally available in the art and may, indeed, be used to
practice many of the embodiments of the invention discussed herein. Such methods may
employ such enzymes as the Klenow fragment of DNA polymerase I, Sequenase (US
Biochemical Corp, Cleveland, OH), Taq polymerase (Perkin Elmer), thermostable T7
15 polymerase (Amersham, Chicago, IL), or combinations of polymerases and proof-reading
exonucleases such as those found in the ELONGASE Amplification System marketed by
Gibco/BRL (Gaithersburg, MD). Preferably, the sequencing process may be automated
using machines such as the Hamilton Micro Lab 2200 (Hamilton, Reno, NV), the Peltier
Thermal Cycler (PTC200; MJ Research, Watertown, MA) and the ABI Catalyst and 373
20 and 377 DNA Sequencers (Perkin Elmer).

One method for isolating a nucleic acid molecule encoding a polypeptide with an
equivalent function to that of the LBDG2 polypeptide, particularly with an equivalent
function to the LBDG2 Nuclear Hormone Receptor Ligand Binding Domain region of the
LBDG2 polypeptide, is to probe a genomic or cDNA library with a natural or artificially-
25 designed probe using standard procedures that are recognised in the art (see, for example,
"Current Protocols in Molecular Biology", Ausubel *et al.* (eds). Greene Publishing
Association and John Wiley Interscience, New York, 1989,1992). Probes comprising at
least 15, preferably at least 30, and more preferably at least 50, contiguous bases that
correspond to, or are complementary to, nucleic acid sequences from the appropriate
30 encoding gene (SEQ ID NO:1), particularly a region from nucleotides 3351-3968 of SEQ

WO 02/070563

PCT/GB02/00986

31

ID NO:1, are particularly useful probes.

Such probes may be labelled with an analytically-detectable reagent to facilitate their identification. Useful reagents include, but are not limited to, radioisotopes, fluorescent dyes and enzymes that are capable of catalysing the formation of a detectable product.

- 5 Using these probes, the ordinarily skilled artisan will be capable of isolating complementary copies of genomic DNA, cDNA or RNA polynucleotides encoding proteins of interest from human, mammalian or other animal sources and screening such sources for related sequences, for example, for additional members of the family, type and/or subtype.
- 10 In many cases, isolated cDNA sequences will be incomplete, in that the region encoding the polypeptide will be cut short, normally at the 5' end. Several methods are available to obtain full length cDNAs, or to extend short cDNAs. Such sequences may be extended utilising a partial nucleotide sequence and employing various methods known in the art to detect upstream sequences such as promoters and regulatory elements. For example, one
- 15 method which may be employed is based on the method of Rapid Amplification of cDNA Ends (RACE; see, for example, Frohman *et al.*, Proc. Natl. Acad. Sci. USA (1988) 85: 8998-9002). Recent modifications of this technique, exemplified by the Marathon™ technology (Clontech Laboratories Inc.), for example, have significantly simplified the search for longer cDNAs. A slightly different technique, termed "restriction-site" PCR,
- 20 uses universal primers to retrieve unknown nucleic acid sequence adjacent a known locus (Sarkar, G. (1993) PCR Methods Applic. 2:318-322). Inverse PCR may also be used to amplify or to extend sequences using divergent primers based on a known region (Triglia, T., *et al.* (1988) Nucleic Acids Res. 16:8186). Another method which may be used is capture PCR which involves PCR amplification of DNA fragments adjacent a known
- 25 sequence in human and yeast artificial chromosome DNA (Lagerstrom, M. *et al.* (1991) PCR Methods Applic. 1: 111-119). Another method which may be used to retrieve unknown sequences is that of Parker, J.D. *et al.* (1991); Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PromoterFinder™ libraries to walk genomic DNA (Clontech, Palo Alto, CA). This process avoids the need to screen
- 30 libraries and is useful in finding intron/exon junctions.

WO 02/070563

PCT/GB02/00986

32

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

In one embodiment of the invention, the nucleic acid molecules of the present invention may be used for chromosome localisation. In this technique, a nucleic acid molecule is specifically targeted to, and can hybridize with, a particular location on an individual human chromosome. The mapping of relevant sequences to chromosomes according to the present invention is an important step in the confirmatory correlation of those sequences with the gene-associated disease. Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found in, for example, V. McKusick, Mendelian Inheritance in Man (available on-line through Johns Hopkins University Welch Medical Library). The relationships between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes). This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localised by genetic linkage to a particular genomic region, any sequences mapping to that area may represent associated or regulatory genes for further investigation. The nucleic acid molecule may also be used to detect differences in the chromosomal location due to translocation, inversion, etc. among normal, carrier, or affected individuals.

The nucleic acid molecules of the present invention are also valuable for tissue localisation. Such techniques allow the determination of expression patterns of the polypeptide in tissues by detection of the mRNAs that encode them. These techniques include in situ hybridization techniques and nucleotide amplification techniques, such as PCR. Results from these studies provide an indication of the normal functions of the polypeptide in the organism. In addition, comparative studies of the normal expression pattern of mRNAs with that of mRNAs encoded by a mutant gene provide valuable

WO 02/070563

PCT/GB02/00986

33

insights into the role of mutant polypeptides in disease. Such inappropriate expression may be of a temporal, spatial or quantitative nature.

The vectors of the present invention comprise nucleic acid molecules of the invention and may be cloning or expression vectors. The host cells of the invention, which may be transformed, transfected or transduced with the vectors of the invention may be
5 prokaryotic or eukaryotic.

The polypeptides of the invention may be prepared in recombinant form by expression of their encoding nucleic acid molecules in vectors contained within a host cell. Such expression methods are well known to those of skill in the art and many are described in detail by Sambrook *et al.* (*supra*) and Fernandez & Hoeffler (1998, eds. "Gene
10 expression systems. Using nature for the art of expression". Academic Press, San Diego, London, Boston, New York, Sydney, Tokyo, Toronto).

Generally, any system or vector that is suitable to maintain, propagate or express nucleic acid molecules to produce a polypeptide in the required host may be used. The appropriate nucleotide sequence may be inserted into an expression system by any of a
15 variety of well-known and routine techniques, such as, for example, those described in Sambrook *et al.*, (*supra*). Generally, the encoding gene can be placed under the control of a control element such as a promoter, ribosome binding site (for bacterial expression) and, optionally, an operator, so that the DNA sequence encoding the desired polypeptide is transcribed into RNA in the transformed host cell.
20

Examples of suitable expression systems include, for example, chromosomal, episomal and virus-derived systems, including, for example, vectors derived from: bacterial plasmids, bacteriophage, transposons, yeast episomes, insertion elements, yeast chromosomal elements, viruses such as baculoviruses, papova viruses such as SV40,
25 vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, or combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, including cosmids and phagemids. Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained and expressed in a plasmid.

30 Particularly suitable expression systems include microorganisms such as bacteria

WO 02/070563

PCT/GB02/00986

34

transformed with recombinant bacteriophage, plasmid or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (for example, baculovirus); plant cell systems transformed with virus expression vectors (for example, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (for example, Ti or pBR322 plasmids); or animal cell systems. Cell-free translation systems can also be employed to produce the polypeptides of the invention.

Introduction of nucleic acid molecules encoding a polypeptide of the present invention into host cells can be effected by methods described in many standard laboratory manuals, such as Davis *et al.*, Basic Methods in Molecular Biology (1986) and Sambrook *et al.*, (*supra*). Particularly suitable methods include calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction or infection (see Sambrook *et al.*, 1989 [*supra*]; Ausubel *et al.*, 1991 [*supra*]; Spector, Goldman & Leinwald, 1998). In eukaryotic cells, expression systems may either be transient (for example, episomal) or permanent (chromosomal integration) according to the needs of the system.

The encoding nucleic acid molecule may or may not include a sequence encoding a control sequence, such as a signal peptide or leader sequence, as desired, for example, for secretion of the translated polypeptide into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment. These signals may be endogenous to the polypeptide or they may be heterologous signals. Leader sequences can be removed by the bacterial host in post-translational processing.

In addition to control sequences, it may be desirable to add regulatory sequences that allow for regulation of the expression of the polypeptide relative to the growth of the host cell. Examples of regulatory sequences are those which cause the expression of a gene to be increased or decreased in response to a chemical or physical stimulus, including the presence of a regulatory compound or to various temperature or metabolic conditions. Regulatory sequences are those non-translated regions of the vector, such as enhancers, promoters and 5' and 3' untranslated regions. These interact with host cellular proteins to

WO 02/070563

PCT/GB02/00986

35

carry out transcription and translation. Such regulatory sequences may vary in their strength and specificity. Depending on the vector system and host utilised, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the Bluescript phagemid (Stratagene, LaJolla, CA) or pSportTM plasmid (Gibco BRL) and the like may be used. The baculovirus polyhedrin promoter may be used in insect cells. Promoters or enhancers derived from the genomes of plant cells (for example, heat shock, RUBISCO and storage protein genes) or from plant viruses (for example, viral promoters or leader sequences) may be cloned into the vector. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are preferable. If it is necessary to generate a cell line that contains multiple copies of the sequence, vectors based on SV40 or EBV may be used with an appropriate selectable marker.

An expression vector is constructed so that the particular nucleic acid coding sequence is located in the vector with the appropriate regulatory sequences, the positioning and orientation of the coding sequence with respect to the regulatory sequences being such that the coding sequence is transcribed under the "control" of the regulatory sequences, i.e., RNA polymerase which binds to the DNA molecule at the control sequences transcribes the coding sequence. In some cases it may be necessary to modify the sequence so that it may be attached to the control sequences with the appropriate orientation; i.e., to maintain the reading frame.

The control sequences and other regulatory sequences may be ligated to the nucleic acid coding sequence prior to insertion into a vector. Alternatively, the coding sequence can be cloned directly into an expression vector that already contains the control sequences and an appropriate restriction site.

For long-term, high-yield production of a recombinant polypeptide, stable expression is preferred. For example, cell lines which stably express the polypeptide of interest may be transformed 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

WO 02/070563

PCT/GB02/00986

36

for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells that successfully express the introduced sequences. Resistant clones of stably transformed cells may be proliferated using tissue culture techniques appropriate to the cell type.

Mammalian cell lines available as hosts for expression are known in the art and include many immortalised cell lines available from the American Type Culture Collection (ATCC) including, but not limited to, Chinese hamster ovary (CHO), HeLa, baby hamster kidney (BHK), monkey kidney (COS), C127, 3T3, BHK, HEK 293, Bowes melanoma and human hepatocellular carcinoma (for example Hep G2) cells and a number of other cell lines.

In the baculovirus system, the materials for baculovirus/insect cell expression systems are commercially available in kit form from, inter alia, Invitrogen, San Diego CA (the "MaxBac" kit). These techniques are generally known to those skilled in the art and are described fully in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987). Particularly suitable host cells for use in this system include insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells.

There are many plant cell culture and whole plant genetic expression systems known in the art. Examples of suitable plant cellular genetic expression systems include those described in US 5,693,506; US 5,659,122; and US 5,608,143. Additional examples of genetic expression in plant cell culture has been described by Zenk, (1991) *Phytochemistry* 30, 3861-3863.

In particular, all plants from which protoplasts can be isolated and cultured to give whole regenerated plants can be utilised, so that whole plants are recovered which contain the transferred gene. Practically all plants can be regenerated from cultured cells or tissues, including but not limited to all major species of sugar cane, sugar beet, cotton, fruit and other trees, legumes and vegetables.

Examples of particularly preferred bacterial host cells include streptococci, staphylococci, *E. coli*, *Streptomyces* and *Bacillus subtilis* cells.

WO 02/070563

PCT/GB02/00986

37

Examples of particularly suitable host cells for fungal expression include yeast cells (for example, *S. cerevisiae*) and *Aspergillus* cells.

Any number of selection systems are known in the art that may be used to recover transformed cell lines. Examples include the herpes simplex virus thymidine kinase (Wigler, M. *et al.* (1977) Cell 11:223-32) and adenine phosphoribosyltransferase (Lowy, I. *et al.* (1980) Cell 22:817-23) genes that can be employed in tk- or apt^r cells, respectively.

Also, antimetabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, dihydrofolate reductase (DHFR) that confers resistance to methotrexate (Wigler, M. *et al.* (1980) Proc. Natl. Acad. Sci. 77:3567-70); npt, which confers resistance to the aminoglycosides neomycin and G-418 (Colbere-Garapin, F. *et al.* (1981) J. Mol. Biol. 150:1-14) and als or pat, which confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. Additional selectable genes have been described, examples of which will be clear to those of skill in the art.

Although the presence or absence of marker gene expression suggests that the gene of interest is also present, its presence and expression may need to be confirmed. For example, if the relevant sequence is inserted within a marker gene sequence, transformed cells containing the appropriate sequences can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding a polypeptide of the invention 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.

Alternatively, host cells that contain a nucleic acid sequence encoding a polypeptide of the invention and which express said polypeptide 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 and protein bioassays, for example, fluorescence activated cell sorting (FACS) or immunoassay techniques (such as the enzyme-linked immunosorbent assay [ELISA] and radioimmunoassay [RIA]), that include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein (see Hampton, R. *et al.* (1990) Serological

WO 02/070563

PCT/GB02/00986

38

Methods, a Laboratory Manual, APS Press, St Paul, MN) and Maddox, D.E. *et al.* (1983) J. Exp. Med, 158, 1211-1216).

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
5 labelled hybridization or PCR probes for detecting sequences related to nucleic acid molecules encoding polypeptides of the present invention include oligolabelling, nick translation, end-labelling or PCR amplification using a labelled polynucleotide. Alternatively, the sequences encoding the polypeptide of the invention may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art,
10 are commercially available, and may be used to synthesise RNA probes *in vitro* by addition of an appropriate RNA polymerase such as T7, T3 or SP6 and labelled nucleotides. These procedures may be conducted using a variety of commercially available kits (Pharmacia & Upjohn, (Kalamazoo, MI); Promega (Madison WI); and U.S. Biochemical Corp., (Cleveland, OH)).

15 Suitable reporter molecules or labels, which may be used for ease of detection, include radionuclides, enzymes and fluorescent, chemiluminescent or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Nucleic acid molecules according to the present invention may also be used to create transgenic animals, particularly rodent animals. Such transgenic animals form a further
20 aspect of the present invention. This may be done locally by modification of somatic cells, or by germ line therapy to incorporate heritable modifications. Such transgenic animals may be particularly useful in the generation of animal models for drug molecules effective as modulators of the polypeptides of the present invention.

The polypeptide can be recovered and purified from recombinant cell cultures by well-
25 known methods including ammonium sulphate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. High performance liquid chromatography is particularly useful for purification. Well known techniques for refolding proteins may be
30 employed to regenerate an active conformation when the polypeptide is denatured during

WO 02/070563

PCT/GB02/00986

39

isolation and or purification.

Specialised vector constructions may also be used to facilitate purification of proteins, as desired, by joining sequences encoding the polypeptides of the invention to a nucleotide sequence encoding a polypeptide domain that will facilitate purification of soluble proteins. Examples of such purification-facilitating domains include metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilised metals, protein A domains that allow purification on immobilised immunoglobulin, and the domain utilised in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, WA). The inclusion of cleavable linker sequences such as those specific for Factor XA or enterokinase (Invitrogen, San Diego, CA) between the purification domain and the polypeptide of the invention may be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing the polypeptide of the invention fused to several histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification by IMAC (immobilised metal ion affinity chromatography as described in Porath, J. *et al.* (1992) *Prot. Exp. Purif.* 3: 263-281) while the thioredoxin or enterokinase cleavage site provides a means for purifying the polypeptide from the fusion protein. A discussion of vectors which contain fusion proteins is provided in Kroll, D.J. *et al.* (*DNA Cell Biol.* 199312:441-453).

If the polypeptide is to be expressed for use in screening assays, generally it is preferred that it be produced at the surface of the host cell in which it is expressed. In this event, the host cells may be harvested prior to use in the screening assay, for example using techniques such as fluorescence activated cell sorting (FACS) or immunoaffinity techniques. If the polypeptide is secreted into the medium, the medium can be recovered in order to recover and purify the expressed polypeptide. If polypeptide is produced intracellularly, the cells must first be lysed before the polypeptide is recovered.

The polypeptide of the invention can be used to screen libraries of compounds in any of a variety of drug screening techniques. Such compounds may activate (agonise) or inhibit (antagonise) the level of expression of the gene or the activity of the polypeptide of the invention and form a further aspect of the present invention. Preferred compounds are

WO 02/070563

PCT/GB02/00986

40

effective to alter the expression of a natural gene which encodes a polypeptide of the first aspect of the invention or to regulate the activity of a polypeptide of the first aspect of the invention.

Agonist or antagonist compounds may be isolated from, for example, cells, cell-free preparations, chemical libraries or natural product mixtures. These agonists or antagonists may be natural or modified substrates, ligands, enzymes, receptors or structural or functional mimetics. For a suitable review of such screening techniques, see Coligan *et al.*, Current Protocols in Immunology 1(2):Chapter 5 (1991).

Compounds that are most likely to be good antagonists are molecules that bind to the polypeptide of the invention without inducing the biological effects of the polypeptide upon binding to it. Potential antagonists include small organic molecules, peptides, polypeptides and antibodies that bind to the polypeptide of the invention and thereby inhibit or extinguish its activity. In this fashion, binding of the polypeptide to normal cellular binding molecules may be inhibited, such that the normal biological activity of the polypeptide is prevented.

The polypeptide of the invention that is employed in such a screening technique may be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. In general, such screening procedures may involve using appropriate cells or cell membranes that express the polypeptide that are contacted with a test compound to observe binding, or stimulation or inhibition of a functional response. The functional response of the cells contacted with the test compound is then compared with control cells that were not contacted with the test compound. Such an assay may assess whether the test compound results in a signal generated by activation of the polypeptide, using an appropriate detection system. Inhibitors of activation are generally assayed in the presence of a known agonist and the effect on activation by the agonist in the presence of the test compound is observed.

Alternatively, simple binding assays may be used, in which the adherence of a test compound to a surface bearing the polypeptide is detected by means of a label directly or indirectly associated with the test compound or in an assay involving competition with a labelled competitor. In another embodiment, competitive drug screening assays may be

WO 02/070563

PCT/GB02/00986

41

used, in which neutralising antibodies that are capable of binding the polypeptide specifically compete with a test compound for binding. In this manner, the antibodies can be used to detect the presence of any test compound that possesses specific binding affinity for the polypeptide.

- 5 Assays may also be designed to detect the effect of added test compounds on the production of mRNA encoding the polypeptide in cells. For example, an ELISA may be constructed that measures secreted or cell-associated levels of polypeptide using monoclonal or polyclonal antibodies by standard methods known in the art, and this can be used to search for compounds that may inhibit or enhance the production of the
- 10 polypeptide from suitably manipulated cells or tissues. The formation of binding complexes between the polypeptide and the compound being tested may then be measured.

- Another technique for drug screening which may be used provides for high throughput screening of compounds having suitable binding affinity to the polypeptide of interest
- 15 (see International patent application WO84/03564). In this method, large numbers of different small test compounds are synthesised on a solid substrate, which may then be reacted with the polypeptide of the invention and washed. One way of immobilising the polypeptide is to use non-neutralising antibodies. Bound polypeptide may then be detected using methods that are well known in the art. Purified polypeptide can also be
- 20 coated directly onto plates for use in the aforementioned drug screening techniques.

- The polypeptide of the invention may be used to identify membrane-bound or soluble receptors, through standard receptor binding techniques that are known in the art, such as ligand binding and crosslinking assays in which the polypeptide is labelled with a radioactive isotope, is chemically modified, or is fused to a peptide sequence that
- 25 facilitates its detection or purification, and incubated with a source of the putative receptor (for example, a composition of cells, cell membranes, cell supernatants, tissue extracts, or bodily fluids). The efficacy of binding may be measured using biophysical techniques such as surface plasmon resonance and spectroscopy. Binding assays may be used for the purification and cloning of the receptor, but may also identify agonists and
- 30 antagonists of the polypeptide, that compete with the binding of the polypeptide to its

WO 02/070563

PCT/GB02/00986

42

receptor. Standard methods for conducting screening assays are well understood in the art.

The invention also includes a screening kit useful in the methods for identifying agonists, antagonists, ligands, receptors, substrates, and enzymes that are described above.

5 The invention includes the agonists, antagonists, ligands, receptors, substrates and enzymes, and other compounds which modulate the activity or antigenicity of the polypeptide of the invention discovered by the methods that are described above.

The invention also provides pharmaceutical compositions comprising a polypeptide, nucleic acid, ligand or compound of the invention in combination with a suitable
10 pharmaceutical carrier. These compositions may be suitable as therapeutic or diagnostic reagents, as vaccines, or as other immunogenic compositions, as outlined in detail below.

According to the terminology used herein, a composition containing a polypeptide, nucleic acid, ligand or compound [X] is "substantially free of" impurities [herein, Y] when at least 85% by weight of the total X+Y in the composition is X. Preferably, X
15 comprises at least about 90% by weight of the total of X+Y in the composition, more preferably at least about 95%, 98% or even 99% by weight.

The pharmaceutical compositions should preferably comprise a therapeutically effective amount of the polypeptide, nucleic acid molecule, ligand, or compound of the invention. The term "therapeutically effective amount" as used herein refers to an amount of a
20 therapeutic agent needed to treat, ameliorate, or prevent a targetted disease or condition, or to exhibit a detectable therapeutic or preventative effect. For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, for example, of neoplastic cells, or in animal models, usually mice, rabbits, dogs, or pigs. The animal model may also be used to determine the appropriate concentration range and
25 route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

The precise effective amount for a human subject will depend upon the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and

WO 02/070563

PCT/GB02/00986

43

tolerance/response to therapy. This amount can be determined by routine experimentation and is within the judgement of the clinician. Generally, an effective dose will be from 0.01 mg/kg to 50 mg/kg, preferably 0.05 mg/kg to 10 mg/kg. Compositions may be administered individually to a patient or may be administered in combination with other agents, drugs or hormones.

A pharmaceutical composition may also contain a pharmaceutically acceptable carrier, for administration of a therapeutic agent. Such carriers include antibodies and other polypeptides, genes and other therapeutic agents such as liposomes, provided that the carrier does not itself induce the production of antibodies harmful to the individual receiving the composition, and which may be administered without undue toxicity. Suitable carriers may be large, slowly metabolised macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers and inactive virus particles.

Pharmaceutically acceptable salts can be used therein, for example, mineral acid salts such as hydrochlorides, hydrobromides, phosphates, sulphates, and the like; and the salts of organic acids such as acetates, propionates, malonates, benzoates, and the like. A thorough discussion of pharmaceutically acceptable carriers is available in Remington's Pharmaceutical Sciences (Mack Pub. Co., N.J. 1991).

Pharmaceutically acceptable carriers in therapeutic compositions may additionally contain liquids such as water, saline, glycerol and ethanol. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such compositions. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Once formulated, the compositions of the invention can be administered directly to the subject. The subjects to be treated can be animals; in particular, human subjects can be treated.

The pharmaceutical compositions utilised in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal or transcutaneous

WO 02/070563

PCT/GB02/00986

44

applications (for example, see WO98/20734), subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, intravaginal or rectal means. Gene guns or hyposprays may also be used to administer the pharmaceutical compositions of the invention. Typically, the therapeutic compositions may be prepared as injectables, either as liquid solutions or
5 suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection may also be prepared.

Direct delivery of the compositions will generally be accomplished by injection, subcutaneously, intraperitoneally, intravenously or intramuscularly, or delivered to the interstitial space of a tissue. The compositions can also be administered into a lesion.
10 Dosage treatment may be a single dose schedule or a multiple dose schedule.

If the activity of the polypeptide of the invention is in excess in a particular disease state, several approaches are available. One approach comprises administering to a subject an inhibitor compound (antagonist) as described above, along with a pharmaceutically acceptable carrier in an amount effective to inhibit the function of the polypeptide, such
15 as by blocking the binding of ligands, substrates, enzymes, receptors, or by inhibiting a second signal, and thereby alleviating the abnormal condition. Preferably, such antagonists are antibodies. Most preferably, such antibodies are chimeric and/or humanised to minimise their immunogenicity, as described previously.

In another approach, soluble forms of the polypeptide that retain binding affinity for the
20 ligand, substrate, enzyme, receptor, in question, may be administered. Typically, the polypeptide may be administered in the form of fragments that retain the relevant portions.

In an alternative approach, expression of the gene encoding the polypeptide can be inhibited using expression blocking techniques, such as the use of antisense nucleic acid
25 molecules (as described above), either internally generated or separately administered. Modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5' or regulatory regions (signal sequence, promoters, enhancers and introns) of the gene encoding the polypeptide. Similarly, inhibition can be achieved using "triple helix" base-pairing
30 methodology. Triple helix pairing is useful because it causes inhibition of the ability of

WO 02/070563

PCT/GB02/00986

45

the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature (Gee, J.E. *et al.* (1994) In: Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, NY). The complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes. Such oligonucleotides may be administered or may be generated in situ from expression *in vivo*.

In addition, expression of the polypeptide of the invention may be prevented by using ribozymes specific to its encoding mRNA sequence. Ribozymes are catalytically active RNAs that can be natural or synthetic (see for example Usman, N, *et al.*, Curr. Opin. Struct. Biol (1996) 6(4), 527-33). Synthetic ribozymes can be designed to specifically cleave mRNAs at selected positions thereby preventing translation of the mRNAs into functional polypeptide. Ribozymes may be synthesised with a natural ribose phosphate backbone and natural bases, as normally found in RNA molecules. Alternatively the ribozymes may be synthesised with non-natural backbones, for example, 2'-O-methyl RNA, to provide protection from ribonuclease degradation and may contain modified bases.

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 non-traditional bases such as inosine, queosine and butosine, as well as acetyl-, methyl-, thio- and similarly modified forms of adenine, cytidine, guanine, thymine and uridine which are not as easily recognised by endogenous endonucleases.

For treating abnormal conditions related to an under-expression of the polypeptide of the invention and its activity, several approaches are also available. One approach comprises administering to a subject a therapeutically effective amount of a compound that activates the polypeptide, i.e., an agonist as described above, to alleviate the abnormal condition. Alternatively, a therapeutic amount of the polypeptide in combination with a suitable

WO 02/070563

PCT/GB02/00986

46

pharmaceutical carrier may be administered to restore the relevant physiological balance of polypeptide.

Gene therapy may be employed to effect the endogenous production of the polypeptide by the relevant cells in the subject. Gene therapy is used to treat permanently the
5 inappropriate production of the polypeptide by replacing a defective gene with a corrected therapeutic gene.

Gene therapy of the present invention can occur *in vivo* or *ex vivo*. *Ex vivo* gene therapy requires the isolation and purification of patient cells, the introduction of a therapeutic gene and introduction of the genetically altered cells back into the patient. In contrast, *in*
10 *vivo* gene therapy does not require isolation and purification of a patient's cells.

The therapeutic gene is typically "packaged" for administration to a patient. Gene delivery vehicles may be non-viral, such as liposomes, or replication-deficient viruses, such as adenovirus as described by Berkner, K.L., in *Curr. Top. Microbiol. Immunol.*, 158, 39-66 (1992) or adeno-associated virus (AAV) vectors as described by Muzyczka,
15 N., in *Curr. Top. Microbiol. Immunol.*, 158, 97-129 (1992) and U.S. Patent No. 5,252,479. For example, a nucleic acid molecule encoding a polypeptide of the invention may be engineered for expression in a replication-defective retroviral vector. This expression construct may then be isolated and introduced into a packaging cell transduced with a retroviral plasmid vector containing RNA encoding the polypeptide,
20 such that the packaging cell now produces infectious viral particles containing the gene of interest. These producer cells may be administered to a subject for engineering cells *in vivo* and expression of the polypeptide *in vivo* (see Chapter 20, Gene Therapy and other Molecular Genetic-based Therapeutic Approaches, (and references cited therein) in *Human Molecular Genetics* (1996), T Strachan and A P Read, BIOS Scientific Publishers
25 Ltd).

Another approach is the administration of "naked DNA" in which the therapeutic gene is directly injected into the bloodstream or muscle tissue.

In situations in which the polypeptides or nucleic acid molecules of the invention are disease-causing agents, the invention provides that they can be used in vaccines to raise
30 antibodies against the disease causing agent.

WO 02/070563

PCT/GB02/00986

47

Vaccines according to the invention may either be prophylactic (ie. to prevent infection) or therapeutic (ie. to treat disease after infection). Such vaccines comprise immunising antigen(s), immunogen(s), polypeptide(s), protein(s) or nucleic acid, usually in combination with pharmaceutically-acceptable carriers as described above, which include
5 any carrier that does not itself induce the production of antibodies harmful to the individual receiving the composition. Additionally, these carriers may function as immunostimulating agents ("adjuvants"). Furthermore, the antigen or immunogen may be conjugated to a bacterial toxoid, such as a toxoid from diphtheria, tetanus, cholera, *H. pylori*, and other pathogens.

10 Since polypeptides may be broken down in the stomach, vaccines comprising polypeptides are preferably administered parenterally (for instance, subcutaneous, intramuscular, intravenous, or intradermal injection). Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation
15 isotonic with the blood of the recipient, and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents.

The vaccine formulations of the invention may be presented in unit-dose or multi-dose containers. For example, sealed ampoules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use.

20 The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

This invention also relates to the use of nucleic acid molecules according to the present invention as diagnostic reagents. Detection of a mutated form of the gene characterised by the nucleic acid molecules of the invention which is associated with a dysfunction will
25 provide a diagnostic tool that can add to, or define, a diagnosis of a disease, or susceptibility to a disease, which results from under-expression, over-expression or altered spatial or temporal expression of the gene. Individuals carrying mutations in the gene may be detected at the DNA level by a variety of techniques.

30 Nucleic acid molecules for diagnosis may be obtained from a subject's cells, such as from blood, urine, saliva, tissue biopsy or autopsy material. The genomic DNA may be used

WO 02/070563

PCT/GB02/00986

48

directly for detection or may be amplified enzymatically by using PCR, ligase chain reaction (LCR), strand displacement amplification (SDA), or other amplification techniques (see Saiki *et al.*, *Nature*, 324, 163-166 (1986); Bej, *et al.*, *Crit. Rev. Biochem. Molec. Biol.*, 26, 301-334 (1991); Birkenmeyer *et al.*, *J. Virol. Meth.*, 35, 117-126 (1991); Van Brunt, J., *Bio/Technology*, 8, 291-294 (1990)) prior to analysis.

In one embodiment, this aspect of the invention provides a method of diagnosing a disease in a patient, comprising assessing the level of expression of a natural gene encoding a polypeptide according to the invention and comparing said level of expression to a control level, wherein a level that is different to said control level is indicative of disease. The method may comprise the steps of:

- a) contacting a sample of tissue from the patient with a nucleic acid probe under stringent conditions that allow the formation of a hybrid complex between a nucleic acid molecule of the invention and the probe;
- b) contacting a control sample with said probe under the same conditions used in step a);
- c) and detecting the presence of hybrid complexes in said samples;

wherein detection of levels of the hybrid complex in the patient sample that differ from levels of the hybrid complex in the control sample is indicative of disease.

A further aspect of the invention comprises a diagnostic method comprising the steps of:

- a) obtaining a tissue sample from a patient being tested for disease;
- b) isolating a nucleic acid molecule according to the invention from said tissue sample; and,
- c) diagnosing the patient for disease by detecting the presence of a mutation in the nucleic acid molecule which is associated with disease.

To aid the detection of nucleic acid molecules in the above-described methods, an amplification step, for example using PCR, may be included.

Deletions and insertions can be detected by a change in the size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to labelled RNA of the invention or alternatively, labelled antisense DNA

WO 02/070563

PCT/GB02/00986

49

sequences of the invention. Perfectly-matched sequences can be distinguished from mismatched duplexes by RNase digestion or by assessing differences in melting temperatures. The presence or absence of the mutation in the patient may be detected by contacting DNA with a nucleic acid probe that hybridises to the DNA under stringent conditions to form a hybrid double-stranded molecule, the hybrid double-stranded molecule having an unhybridised portion of the nucleic acid probe strand at any portion corresponding to a mutation associated with disease; and detecting the presence or absence of an unhybridised portion of the probe strand as an indication of the presence or absence of a disease-associated mutation in the corresponding portion of the DNA strand.

10 Such diagnostics are particularly useful for prenatal and even neonatal testing.

Point mutations and other sequence differences between the reference gene and "mutant" genes can be identified by other well-known techniques, such as direct DNA sequencing or single-strand conformational polymorphism, (see Orita *et al.*, Genomics, 5, 874-879 (1989)). For example, a sequencing primer may be used with double-stranded PCR product or a single-stranded template molecule generated by a modified PCR. The sequence determination is performed by conventional procedures with radiolabelled nucleotides or by automatic sequencing procedures with fluorescent-tags. Cloned DNA segments may also be used as probes to detect specific DNA segments. The sensitivity of this method is greatly enhanced when combined with PCR. Further, point mutations and other sequence variations, such as polymorphisms, can be detected as described above, for example, through the use of allele-specific oligonucleotides for PCR amplification of sequences that differ by single nucleotides.

DNA sequence differences may also be detected by alterations in the electrophoretic mobility of DNA fragments in gels, with or without denaturing agents, or by direct DNA sequencing (for example, Myers *et al.*, Science (1985) 230:1242). Sequence changes at specific locations may also be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method (see Cotton *et al.*, Proc. Natl. Acad. Sci. USA (1985) 85: 4397-4401).

In addition to conventional gel electrophoresis and DNA sequencing, mutations such as microdeletions, aneuploidies, translocations, inversions, can also be detected by in situ

WO 02/070563

PCT/GB02/00986

50

analysis (see, for example, Keller *et al.*, DNA Probes, 2nd Ed., Stockton Press, New York, N.Y., USA (1993)), that is, DNA or RNA sequences in cells can be analysed for mutations without need for their isolation and/or immobilisation onto a membrane. Fluorescence in situ hybridization (FISH) is presently the most commonly applied method and numerous reviews of FISH have appeared (see, for example, Trachuck *et al.*, Science, 250: 559-562 (1990), and Trask *et al.*, Trends, Genet. 7:149-154 (1991)).

In another embodiment of the invention, an array of oligonucleotide probes comprising a nucleic acid molecule according to the invention can be constructed to conduct efficient screening of genetic variants, mutations and polymorphisms. Array technology methods are well known and have general applicability and can be used to address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability (see for example: M.Chee *et al.*, Science (1996) 274: 610-613).

In one embodiment, the array is prepared and used according to the methods described in PCT application WO95/11995 (Chee *et al.*); Lockhart, D. J. *et al.* (1996) Nat. Biotech. 14: 1675-1680); and Schena, M. *et al.* (1996) Proc. Natl. Acad. Sci. 93: 10614-10619). Oligonucleotide pairs may range from two to over one million. The oligomers are synthesized at designated areas on a substrate using a light-directed chemical process. The substrate may be paper, nylon or other type of membrane, filter, chip, glass slide or any other suitable solid support. In another aspect, an oligonucleotide may be synthesized on the surface of the substrate by using a chemical coupling procedure and an ink jet application apparatus, as described in PCT application WO95/251116 (Baldeschweiler *et al.*). In another aspect, a "gridded" array analogous to a dot (or slot) blot may be used to arrange and link cDNA fragments or oligonucleotides to the surface of a substrate using a vacuum system, thermal, UV, mechanical or chemical bonding procedures. An array, such as those described above, may be produced by hand or by using available devices (slot blot or dot blot apparatus), materials (any suitable solid support), and machines (including robotic instruments), and may contain 8, 24, 96, 384, 1536 or 6144 oligonucleotides, or any other number between two and over one million which lends itself to the efficient use of commercially-available instrumentation.

In addition to the methods discussed above, diseases may be diagnosed by methods

WO 02/070563

PCT/GB02/00986

51

comprising determining, from a sample derived from a subject, an abnormally decreased or increased level of polypeptide or mRNA. Decreased or increased expression can be measured at the RNA level using any of the methods well known in the art for the quantitation of polynucleotides, such as, for example, nucleic acid amplification, for instance PCR, RT-PCR, RNase protection, Northern blotting and other hybridization methods.

Assay techniques that can be used to determine levels of a polypeptide of the present invention in a sample derived from a host are well-known to those of skill in the art and are discussed in some detail above (including radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays). This aspect of the invention provides a diagnostic method which comprises the steps of: (a) contacting a ligand as described above with a biological sample under conditions suitable for the formation of a ligand-polypeptide complex; and (b) detecting said complex.

Protocols such as ELISA, RIA, and FACS for measuring polypeptide levels may additionally provide a basis for diagnosing altered or abnormal levels of polypeptide expression. Normal or standard values for polypeptide expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably humans, with antibody to the polypeptide under conditions suitable for complex formation. The amount of standard complex formation may be quantified by various methods, such as by photometric means.

Antibodies which specifically bind to a polypeptide of the invention may be used for the diagnosis of conditions or diseases characterised by expression of the polypeptide, or in assays to monitor patients being treated with the polypeptides, nucleic acid molecules, ligands and other compounds of the invention. Antibodies useful for diagnostic purposes may be prepared in the same manner as those described above for therapeutics. Diagnostic assays for the polypeptide include methods that utilise the antibody and a label to detect the polypeptide in human body fluids or extracts of cells or tissues. The antibodies may be used with or without modification, and may be labelled by joining them, either covalently or non-covalently, with a reporter molecule. A wide variety of reporter molecules known in the art may be used, several of which are described above.

WO 02/070563

PCT/GB02/00986

52

Quantities of polypeptide 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. Diagnostic assays may be used to distinguish between absence, presence, and excess expression of polypeptide and to monitor regulation of polypeptide levels during therapeutic intervention. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials or in monitoring the treatment of an individual patient.

A diagnostic kit of the present invention may comprise:

- (a) a nucleic acid molecule of the present invention;
- (b) a polypeptide of the present invention; or
- (c) a ligand of the present invention.

In one aspect of the invention, a diagnostic kit may comprise a first container containing a nucleic acid probe that hybridises under stringent conditions with a nucleic acid molecule according to the invention; a second container containing primers useful for amplifying the nucleic acid molecule; and instructions for using the probe and primers for facilitating the diagnosis of disease. The kit may further comprise a third container holding an agent for digesting unhybridised RNA.

In an alternative aspect of the invention, a diagnostic kit may comprise an array of nucleic acid molecules, at least one of which may be a nucleic acid molecule according to the invention.

To detect polypeptide according to the invention, a diagnostic kit may comprise one or more antibodies that bind to a polypeptide according to the invention; and a reagent useful for the detection of a binding reaction between the antibody and the polypeptide.

Such kits will be of use in diagnosing a disease or susceptibility to disease, particularly cell proliferative disorders, including neoplasm, melanoma, lung, colorectal, breast, pancreas, head and neck and other solid tumours, myeloproliferative disorders, such as leukemia, non-Hodgkin lymphoma, leukopenia, thrombocytopenia, angiogenesis disorder, Kaposis' sarcoma, autoimmune/inflammatory disorders, including allergy, inflammatory bowel disease, arthritis, psoriasis and respiratory tract inflammation,

WO 02/070563

PCT/GB02/00986

53

asthma, and organ transplant rejection, cardiovascular disorders, including hypertension, oedema, angina, atherosclerosis, thrombosis, sepsis, shock, reperfusion injury, heart arrhythmia, and ischemia, neurological disorders including, central nervous system disease, Alzheimer's disease, brain injury, stroke, amyotrophic lateral sclerosis, anxiety, depression, and pain, developmental disorders, metabolic disorders including diabetes mellitus, osteoporosis, lipid metabolism disorder, hyperthyroidism, hyperparathyroidism, hypercalcemia, hypercholesterolemia, hyperlipidemia, and obesity, renal disorders, including glomerulonephritis, renovascular hypertension, dermatological disorders, including, acne, eczema, and wound healing, negative effects of aging, AIDS, infections including viral infection, bacterial infection, fungal infection and parasitic infection and other pathological conditions, particularly those in which nuclear hormone receptors are implicated.

Various aspects and embodiments of the present invention will now be described in more detail by way of example, with particular reference to the LBDG2 polypeptide.

It will be appreciated that modification of detail may be made without departing from the scope of the invention.

Brief description of the Figures

Figure 1: This is the front end of the Biopendium Target Mining Interface. A search of the database is initiated using the PDB code "IBSX:A".

Figure 2A: A selection is shown of the Inpharmatica Genome Threader results for the search using IBXS:A. The arrow indicates *Homo Sapiens* Thyroid Hormone Receptor Beta-1, which has a typical Nuclear Hormone Receptor Ligand Binding Domain.

Figure 2B: A selection is shown of the Inpharmatica Genome Threader results for the search using IBXS:A. The arrow indicates BAA22563.1 (LBDG2).

Figure 2C: Full list of forward PSI-BLAST results for the search using IBXS:A. BAA22563.1 (LBDG2) is not identified.

Figure 3: The Redundant Sequence Display results page for BAA22563.1 (LBDG2).

Figure 4: InterPro PFAM search results for BAA22563.1 (LBDG2), see arrow @.

WO 02/070563

PCT/GB02/00986

54

Figure 5: NCBI Protein Report for BAA22563.1 (LBDG2).

Figure 6A: This is the front end of the Biopendium database. A search of the database is initiated using BAA22563.1 (LBDG2), as the query sequence.

Figure 6B: A selection of the Inpharmatica Genome Threader results of search using BAA22563.1 (LBDG2), as the query sequence. The arrow points to 1BSX:A.

Figure 6C: A selection of the reverse-maximised PSI-BLAST results obtained using BAA22563.1 (LBDG2), as the query sequence. The arrows numbered ① to ④ point to homologues of BAA22563.1 (LBDG2).

Figure 7: AIEye sequence alignment of BAA22563.1 (LBDG2) and 1BSX:A.

Figure 8A: LigEye for 1BSX:A that illustrates the sites of interaction of 3,5,3'-Triiodothyronine with the Ligand Binding Domain of *Homo sapiens* Thyroid Hormone Receptor Beta, 1BSX:A.

Figure 8B: iRasMol view of 1BSX:A, the Ligand Binding Domain of *Homo sapiens* Thyroid Hormone Receptor Beta.

Figure 9: AIEye sequence alignment of BAA22563.1 (LBDG2; *Homo sapiens* PRP8) with 4 homologues: AAF58573.1 (*Drosophila melanogaster* ①), P34369 (*Caenorhabditis elegans* ②), BAA78744.1 (*Oryza sativa* ③) and CAB80541.1 (*Arabidopsis thaliana* ④).

Figure 10: The linear dynamic range for target BAA22563.1 (LBDG2) reactions on colon cDNA.

Figure 11: The linear dynamic range for internal control 18s rRNA reactions on colon cDNA.

Figure 12: The linear dynamic range for internal control human ribosomal protein mRNA reactions on IM9 cell cDNA.

Figure 13: Normalised expression of BAA22563.1 (LBDG2) in 18 normal human tissues.

Figure 14: Normalised expression of BAA22563.1 (LBDG2) in a number of cell lines.

WO 02/070563

PCT/GB02/00986

55

Example 1

In order to initiate a search for novel, distantly related Nuclear Hormone Receptor Ligand Binding Domains, an archetypal family member is chosen, the Ligand Binding Domain of *Homo sapiens* Thyroid Hormone Receptor Beta. More specifically, the search is initiated using a structure from the Protein Data Bank (PDB) which is operated by the Research Collaboratory for Structural Bioinformatics.

The structure chosen is the Ligand Binding Domain of *Homo sapiens* Thyroid Hormone Receptor Beta (PDB code 1BSX:A; see Figure 1).

A search of the Biopendium (using the Target Mining Interface) for relatives of 1BSX:A takes place and returns 4096 Genome Threader results. The 4096 Genome Threader results include examples of typical Nuclear Hormone Receptor Ligand Binding Domains, such as that found between residues 211-461 of the *Homo sapiens* Thyroid Hormone Receptor Beta (see arrow in Figure 2A).

Among the proteins known to contain a Nuclear Hormone Receptor Ligand Binding Domain appears a protein which is not annotated as containing a Nuclear Hormone Receptor Ligand Binding Domain, BAA22563.1 (LBDG2; see arrow in Figure 2B). The Inpharmatica Genome Threader has identified a region of the sequence BAA22563.1 (LBDG2), between residues 1104-1309, as having a structure similar to the Ligand Binding Domain of *Homo sapiens* Thyroid Hormone Receptor Beta. The possession of a structure similar to the Ligand Binding Domain of *Homo sapiens* Thyroid Hormone Receptor Beta suggests that residues 1104-1309 of BAA22563.1 (LBDG2) function as a Nuclear Hormone Receptor Ligand Binding Domain. The Genome Threader identifies this with 86% confidence.

The search of the Biopendium (using the Target Mining Interface) for relatives of 1BSX:A also returns 852 Forward PSI-Blast results. Forward PSI-Blast (see Figure 2C) is unable to identify this relationship; only the Inpharmatica Genome Threader is able to identify BAA22563.1 (LBDG2) as containing a Nuclear Hormone Receptor Ligand Binding Domain.

WO 02/070563

PCT/GB02/00986

56

In order to assess what is known in the public domain databases about BAA22563.1 (LBDG2) the Redundant Sequence Display Page (Figure 3) is viewed. There are no PROSITE or PRINTS hits which identify BAA22563.1 (LBDG2) as containing a Nuclear Hormone Receptor Ligand Binding Domain. PROSITE and PRINTS are databases that help to describe proteins of similar families. Returning no Nuclear Hormone Receptor Ligand Binding Domain hits from both databases means that BAA22563.1 (LBDG2) is unidentifiable as containing a Nuclear Hormone Receptor Ligand Binding Domain using PROSITE or PRINTS.

In order to identify if any other public domain annotation vehicle is able to annotate BAA22563.1 (LBDG2) as containing a Nuclear Hormone Receptor Ligand Binding Domain, the BAA22563.1 (LBDG2) protein sequence is searched against the PFAM database (Protein Family Database of Alignment and hidden Markov models) at the InterPro website (see Figure 4 arrow ①). A PFAM-A match is found to PF00527/IPR000148, which is diagnostic of relatedness to Papillomavirus E7 protein. The Papillomavirus E7 protein match is located between residues 1561-1570 of BAA22563.1 (LBDG2). However, there are no PFAM-A matches annotating BAA22563.1 (LBDG2) as containing a Nuclear Hormone Receptor Ligand Binding Domain. Thus PFAM does not identify BAA22563.1 (LBDG2) as containing a Nuclear Hormone Receptor Ligand Binding Domain.

Interestingly, PROSITE PFscan (see Figure 4 arrow ②) identifies a bipartite nuclear localisation signal in BAA22563.1 (LBDG2) at residues 35-52. A typical (although non-diagnostic) feature of Nuclear Hormone Receptors is the possession of a bipartite nuclear localisation signal.

The National Center for Biotechnology Information (NCBI) Genbank protein database is then viewed to examine if there is any further information that is known in the public domain relating to BAA22563.1 (LBDG2). This is the US public domain database for protein and gene sequence deposition (Figure 5). BAA22563.1 (LBDG2) is a *Homo sapiens* sequence, its Genbank protein ID is BAA22563.1 and it is 2335 amino acids in length. BAA22563.1 (LBDG2) is called a *Homo sapiens* homologue of *Saccharomyces cerevisiae* PRP8, a Pre-mRNA splicing factor. BAA22563.1 (LBDG2) was cloned by a

WO 02/070563

PCT/GB02/00986

57

group of scientists at the Otsuka GEN Research Institute; Kagasuno, Kawauchi-cho, Tokushima, Japan. The public domain information for this gene does not annotate it as containing a Nuclear Hormone Receptor Ligand Binding Domain.

Therefore, it can be concluded that using all public domain annotation tools, BAA22563.1 (LBDG2) may not be annotated as containing a Nuclear Hormone Receptor Ligand Binding Domain. Only the Inpharmatica Genome Threader is able to annotate this protein as containing a Nuclear Hormone Receptor Ligand Binding Domain.

The reverse search is now carried out. BAA22563.1 (LBDG2) is now used as the query sequence in the Biopendium (see Figure 6A). The Inpharmatica Genome Threader identifies residues 1104-1309 of BAA22563.1 (LBDG2) as having a structure that is the same as the Ligand Binding Domain of *Homo sapiens* Thyroid Hormone Receptor Beta with 86% confidence (see arrow in Figure 6B). The Ligand Binding Domain of *Homo sapiens* Thyroid Hormone Receptor Beta (IBSX:A) was the original query sequence. Positive iterations of PSI-Blast do not return this result (Figure 6C). It is only the Inpharmatica Genome Threader that is able to identify this relationship.

The sequence of the *Homo sapiens* Thyroid Hormone Receptor Beta Ligand Binding Domain is chosen against which to view the sequence alignment of BAA22563.1 (LBDG2). Viewing the AIEye alignment (Figure 7) of the query protein against the protein identified as being of a similar structure helps to visualize the areas of homology.

The *Homo sapiens* Thyroid Hormone Receptor Beta Ligand Binding Domain contains an "LBD motif" which has been found in all annotated Nuclear Hormone Receptor Ligand Binding Domains to date. The "LBD motif" is involved in recruiting Nuclear Hormone Receptor Co-Activators and Co-Repressors. The 6 residues; PHE293, LEU296, ASP300, GLN301, LEU304 and LEU305 constitute this motif in the *Homo sapiens* Thyroid Hormone Receptor Beta Ligand Binding Domain (see square boxes Figure 7). 4 residues (PHE1174, ASP1181, LEU1185, LEU1186) in BAA22563.1 (LBDG2) precisely match 4 (PHE293, ASP300, LEU304, LEU305) out of the 6 "LBD motif" residues in the *Homo sapiens* Thyroid Hormone Receptor Beta Ligand Binding Domain. Furthermore VAL1177 and ASN1182 in BAA22563.1 (LBDG2) conservatively substitute for the remaining 2 residues LEU296 and GLN301 in the "LBD motif" of *Homo sapiens*

WO 02/070563

PCT/GB02/00986

58

Thyroid Hormone Receptor Beta Ligand Binding Domain. This indicates that BAA22563.1 (LBDG2) contains a Nuclear Hormone Receptor Ligand Binding Domain similar to The *Homo sapiens* Thyroid Hormone Receptor Beta Ligand Binding Domain.

In order to ensure that the protein identified is in fact a relative of the query sequence, the visualization programs "LigEye" (Figure 8A) and "iRasmol" (Figure 8B) are used. These visualization tools identify the active site of known protein structures by indicating the amino acids with which known small molecule inhibitors interact at the active site. These interactions are either through a direct hydrogen bond or through hydrophobic interactions. In this manner, one can see if the active site fold/structure is conserved between the identified homologue and the chosen protein of known structure. The LigEye view of the *Homo sapiens* Thyroid Hormone Receptor Beta Ligand Binding Domain reveals 5 residues which bind 3,5,3'-Triiodothyronine (circled in Figure 7). However, only 4 (ILE276, LEU330, ASN331 and LEU346) of these 5 residues lie within the Genome Threader alignment. Thus only these 4 residues can be used to consolidate the Genome Threader annotation of BAA22563.1 (LBDG2) as containing a Nuclear Hormone Receptor Ligand Binding Domain. Of these 4 residues there are 3 hydrophobic residues which line the pocket of the *Homo sapiens* Thyroid Hormone Receptor Beta Ligand Binding Domain; ILE276, LEU330 and LEU346. LEU330 and LEU346 of *Homo sapiens* Thyroid Hormone Receptor Beta Ligand Binding Domain are perfectly conserved in BAA22563.1 (LBDG2): LEU1216 and LEU1230 (circled in Figure 7). ILE276 of the *Homo sapiens* Thyroid Hormone Receptor Beta Ligand Binding Domain is conservatively substituted by LEU1161 in BAA22563.1 (LBDG2): (broken circle in Figure 7). This conservation of hydrophobicity in 3 out of the 3 hydrophobic residues (within the region of Genome Threader alignment) which line the binding pocket indicates that BAA22563.1 (LBDG2) will bind a hydrophobic steroid-like ligand.

ASN331 of the *Homo sapiens* Thyroid Hormone Receptor Beta Ligand Binding Domain is conservatively substituted by GLN1217 in BAA22563.1 (LBDG2): (broken circle in Figure 7). This indicates that indeed as predicted by the Inpharmatica Genome Threader, BAA22563.1 (LBDG2) folds in a similar manner to the *Homo sapiens* Thyroid Hormone Receptor Beta Ligand Binding Domain and as such is identified as containing a Nuclear Hormone Receptor Ligand Binding Domain.

WO 02/070563

PCT/GB02/00986

59

Reverse-maximised PSI-BLAST of BAA22563.1 (LBDG2) identifies a *Drosophila melanogaster* homologue (AAF58573.1, see Figure 6C arrow ①), a *Caenorhabditis elegans* homologue (P34369, see Figure 6C arrow ②), an *Oryza sativa* homologue (BAA78744.1, see Figure 6C arrow ③), and an *Arabidopsis thaliana* homologue (CAB80541.1, see Figure 6C arrow ④). BAA22563.1 (LBDG2), AAF58573.1 (*Drosophila melanogaster* homologue), P34369 (*Caenorhabditis elegans* homologue), BAA78744.1 (*Oryza sativa* homologue) and CAB80541.1 (*Arabidopsis thaliana* homologue) are aligned and viewed in AlEye (Figure 9). AlEye reveals that the 4 predicted ligand binding residues (within the Genome Threader alignment; LEU1161, LEU1216, GLN1217 and LEU1230) of BAA22563.1 (LBDG2) are all precisely conserved in AAF58573.1 (*Drosophila melanogaster* homologue), P34369 (*Caenorhabditis elegans* homologue), BAA78744.1 (*Oryza sativa* homologue) and CAB80541.1 (*Arabidopsis thaliana* homologue). Furthermore all of the predicted "LBD motif" residues in BAA22563.1 (LBDG2; PHE1174, VAL1177, ASP1181, ASN1182, LEU1185 and LEU1186) are precisely conserved in AAF58573.1 (*Drosophila melanogaster* homologue), P34369 (*Caenorhabditis elegans* homologue), BAA78744.1 (*Oryza sativa* homologue) and CAB80541.1 (*Arabidopsis thaliana* homologue). The only exception is P34369 (*Caenorhabditis elegans* homologue) in which LEU1185 is conservatively substituted with a MET. Residues which are essential for the function of a protein will be conserved in homologues of that protein. Thus the conservation of residues which would be essential for the function of the predicted BAA22563.1 (LBDG2) Nuclear Hormone Receptor Ligand Binding Domain in AAF58573.1 (*Drosophila melanogaster* homologue), P34369 (*Caenorhabditis elegans* homologue), BAA78744.1 (*Oryza sativa* homologue) and CAB80541.1 (*Arabidopsis thaliana* homologue) strongly supports the annotation of BAA22563.1 (LBDG2) as containing a Nuclear Hormone Receptor Ligand Binding Domain.

Example: 2

In order to determine the tissue expression of the proposed LBD, Taqman RT-PCR quantitation was used. The TaqMan 3'-5' exonuclease assay signals the formation of PCR amplicons by a process involving the nucleolytic degradation of a double-labeled fluorogenic probe that hybridises to the target template at a site between the two primer

WO 02/070563

PCT/GB02/00986

60

recognition sequences (cf. U. S. Patent 5,876,930). The ABI Prism 7000 automates the detection and quantitative measurement of these signals, which are stoichiometrically related to the quantities of amplicons produced, during each cycle of amplification. In addition to providing substantial reductions in the time and labour requirements for PCR analyses, this technology permits simplified and potentially highly accurate quantification of target sequences in the reactions.

Human RNA prepared from non-diseased organs was purchased from either Ambion Europe (Huntingdon, UK) or Clontech (BD, Franklin Lakes, NJ). Oligonucleotide primers and probes were designed using Primer Express software (Applied Biosystems, Foster City CA) with a GC-content of 40-60%, no G-nucleotide at the 5'-end of the probe, and no more than 4 contiguous Gs. Each primer and probe was then analysed using BLAST[®] (Basic Local Alignment Search Tool, Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: J Mol Biol 1990 Oct 5;215(3):403-10). Results confirmed that each oligonucleotide recognised the target sequence with a specificity >3 bp when compared to other known cDNA's or genomic sequence represented in the Unigene and GoldenPath publicly available databases. The sequence of the primers and probes were as follows:

BAA22563.1 (LBDG2) Forward primer: CAG ACA GGC CGC TGA CAT T

BAA22563.1 (LBDG2) Reverse primer: GCC ATC AGG AGG TCA ACA ACA

BAA22563.1 (LBDG2) Probe: AGT TTG GCC TCT TTC CCT CTG TCT GTG C

18s and human ribosomal protein pre-optimised primers and probe were purchased from Applied Biosystems, Foster City, CA. Probes were covalently conjugated with a fluorescent reporter dye (e.g. 6carboxy-fluorescein [FAM]; Xem = 518nm) and a fluorescent quencher dye (6carboxytetram-ethyl-rhodamine [TAMRA]; Mem = 582nm) at the most 5' and most 3' base, respectively. All primers and probes were obtained from Applied Biosystems, Germany. Primer/probe concentrations were titrated in the range of 50nM to 900nM and optimal concentrations for efficient PCR reactions were determined. Optimal primer and probe concentrations vary in between 100nM and 900nM depending on the target gene that was amplified. cDNA is prepared using components from Applied Biosystems, Foster City CA. 50µl reactions are prepared in 0.5ml RNase free tubes. Reactions contain 500ng total RNA; 1x reverse transcriptase buffer; 5.5mM MgCl₂;

WO 02/070563

PCT/GB02/00986

61

1mM dNTP's; 2.5µl random hexamers; 20U RNase inhibitor; and 62.5U reverse transcriptase. 25µl reactions were prepared in 0.5 ml thin-walled, optical grade PCR 96 well plates (Applied Biosystems, Foster City CA). Reactions contained: 1x final concentration of TaqMan Universal Master Mix (a proprietary mixture of AmpliTaq Gold DNA polymerase, AmpEraseX UNG, dNTPs with UTP, passive reference dye and optimised buffer components, Applied Biosystems, Foster City CA); 100nM Taqman probe; 300nM forward primer; 900nM reverse primer and 15ng of cDNA template.

Standard procedures for the operation of the ABI Prism 7000 or similar detection system were used. This included, for example with the ABI Prism 7000, use of all default program settings with the exception of reaction volume which was changed from 50 to 25 ul. Thermal cycling conditions consisted of two min at 50 C, 10 min at 95 C, followed by 40 cycles of 15 sec at 95 C and 1 min at 60 C. Cycle threshold (Ct) determinations, (i.e. non-integer calculations of the number of cycles required for reporter dye fluorescence resulting from the synthesis of PCR products to become significantly higher than background fluorescence levels), were automatically performed by the instrument for each reaction using default parameters. Assays for target sequences and ribosomal 18s (reference) sequences in the same cDNA samples were performed in separate reaction tubes. Within each experiment, a standard curve was carried out of a typical tissue sample, from 50ng to 0.78ng of cDNA template. From this standard curve, the amount of actual starting target or 18s cDNA in each test sample is determined.

The levels of target cDNA in each sample were normalised to the level of expression of target in a comparative sample. The levels of internal control cDNA in each sample were normalised to the level of expression of internal control in a comparative sample. The data was then represented as fold expression of normalised target sequence relative to the level of expression in the comparative sample, which is set arbitrarily to 1. Taqman RT-PCR was carried out on 2-fold dilutions of colon cDNA using primers/probes specific for BAA22563.1 (LBDG2) as described above. Figure 10 shows the Ct values plotted vs. the log input cDNA and illustrates that a linear relationship was seen over this range of input cDNA concentrations. Linear regression analysis of the standard curve was used to calculate the starting amount of cDNA from test Ct values.

WO 02/070563

PCT/GB02/00986

62

Taqman RT-PCR was carried out on 2-fold dilutions of colon cDNA using primers/probes specific for 18s rRNA as described above. Figure 11 shows the Ct values plotted vs. the log input cDNA value, and illustrates that a linear relationship was seen over this range of input cDNA concentrations. Linear regression analysis of the standard curve was used to calculate the starting amount of cDNA from test Ct values.

Taqman RT-PCR was carried out on 2-fold dilutions of IM9 cDNA using primers/probes specific for human ribosomal protein mRNA as described above. Figure 12 shows the Ct values plotted vs. the log input cDNA value, and illustrates that a linear relationship was seen over this range of input cDNA concentrations. Linear regression analysis of the standard curve was used to calculate the starting amount of cDNA from test Ct values.

Taqman RT-PCR was carried out using 15ng of the indicated cDNA using primers/probes specific for BAA22563.1 (LBDG2) and 18s rRNA as described above. A standard curve for target and internal control was also carried out, using between 50ng to 0.78ng of cDNA template of a typical tissue sample. Using linear regression analysis of the standard curves, the Ct values were used to calculate the amount of actual starting target or 18s cDNA in each test sample.

The levels of target cDNA in each sample were normalised to the level of expression of target in a comparative sample, in this case, stomach. The levels of 18s cDNA in each sample were also normalised to the level of expression of 18s in stomach. The expression levels of BAA22563.1 (LBDG2) were then normalised to the expression levels of 18s. Figure 13 represents the fold expression of normalised target sequence relative to the level of expression in stomach cDNA, which is set arbitrarily to 1. Each sample was quantitated in between 2-4 individual experiments. Figure 13 shows the mean \pm SEM for the multiple experiments.

Taqman RT-PCR was carried out using 15ng of the indicated cDNA using primers/probes specific for BAA22563.1 (LBDG2) and human ribosomal protein mRNA as described in the detailed description. A standard curve for target and internal control was also carried out using a typical cell line sample, using between 50ng to 0.78ng of cDNA template. Using linear regression analysis of the standard curves, the Ct values

WO 02/070563

PCT/GB02/00986

63

were used to calculate the amount of actual starting target or human ribosomal protein cDNA in each test sample.

The levels of target cDNA in each sample were normalised to the level of expression of target in a comparative sample, in this case LAK cells. The levels of human ribosomal protein cDNA in each sample were also normalised to the level of expression of human ribosomal protein cDNA in LAK cells. The expression levels of BAA22563.1 (LBDG2) were then normalised to the expression levels of human ribosomal protein. Figure 14 represents the fold expression of normalised target sequence relative to the level of expression in LAK cDNA, which is set arbitrarily to 1. Figure 14 shows the mean \pm SEM for duplicate measurements of each sample.

The mRNA for LBDG2 has been found in extracts from a variety of human tissues (Figure 13). The finding of high levels of the transcript in the human spleen is consistent with a role of LBDG2 in the immune system and in particular in lymphocyte development and function and in particular in B cell development and function. Development of agonists and antagonists for LBDG2 may therefore have a role in the therapeutic intervention in various human diseases of the immune system including autoimmunity, allergies and diseases associated with immunoglobulin dysfunction. These diseases include type I diabetes mellitus, rheumatoid arthritis, multiple sclerosis, psoriasis, renal failure arising from glomerulopathies, scleroderma, inflammatory bowel disease (both Crohns disease and ulcerative colitis), transplant rejection, asthma, atopic dermatitis, eczema, myelomas and in infectious diseases that require production of antibodies e.g. intracellular pathogen such as virus infected cells, tuberculosis, listeria.

The finding of mRNA for LBDG2 in human B cell lines such as Daudi, IM9 and Raji cells (Figure 14) is consistent with the finding of mRNA in the spleen. Finding high levels of expression of the mRNA in U937 cells suggests a role for LBDG2 in monocyte/macrophage functions and as such agonists or antagonists may be valuable in treating inflammatory diseases including chronic obstructive pulmonary disease (COPD), osteoarthritis, rheumatoid arthritis, inflammatory bowel disease, fibrosis such as liver fibrosis (cirrhosis) and skin fibrosis (scarring), atherosclerosis, dementia, multiple sclerosis, inflammatory pain.

WO 02/070563

PCT/GB02/00986

64

In addition, the finding of significant levels of LBDG2 in adrenal, ovary, prostate and testis indicates that development of agonists and antagonists to LBDG2 may be of value in diseases such as benign prostatic hypertrophy, prostatic cancer, ovarian cancer, testicular cancer. In addition, agonists or antagonists for LBDG2 may be developed for

5 treatment of diseases including but not exclusive to hypertension, responses to stress including stress of infectious diseases, regulation of salt and water homeostasis, control of fertility through regulation of ovulation (infertility and contraception), regulation of implantation (infertility and contraception) and regulation of spermatogenesis (infertility and contraception).

10 The finding that the mRNA for LBDG2 is expressed at significant levels in the human brain is noteworthy as this provides a potential link to human disease states and development of agonists and antagonists for the ligand binding domain of LBDG2 offers the potential for therapeutic intervention in various human diseases including cell proliferative disorders, including neoplasm, melanoma, lung, colorectal, breast, pancreas,

15 head and neck and other solid tumours, myeloproliferative disorders, such as leukemia, non-Hodgkin lymphoma, leukopenia, thrombocytopenia, angiogenesis disorder, Kaposi' sarcoma, autoimmune/inflammatory disorders, including allergy, inflammatory bowel disease, arthritis, psoriasis and respiratory tract inflammation, asthma, and organ transplant rejection, cardiovascular disorders, including hypertension, oedema, angina,

20 atherosclerosis, thrombosis, sepsis, shock, reperfusion injury, heart arrhythmia, and ischemia, neurological disorders including, central nervous system disease, Alzheimer's disease, brain injury, stroke, amyotrophic lateral sclerosis, anxiety, depression, and pain, developmental disorders, metabolic disorders including diabetes mellitus, osteoporosis, lipid metabolism disorder, hyperthyroidism, hyperparathyroidism, hypercalcemia,

25 hypercholesterolemia, hyperlipidemia, and obesity, renal disorders, including glomerulonephritis, renovascular hypertension, dermatological disorders, including, acne, eczema, and wound healing, negative effects of aging, AIDS, infections including viral infection, bacterial infection, fungal infection and parasitic infection and other pathological conditions, particularly those in which nuclear hormone receptors are

30 implicated.

WO 02/070563

PCT/GB02/00986

65

The finding of a "non-classical" nuclear hormone receptor such as LBDG2 which contains a ligand binding domain in the absence of a DNA binding domain is consistent with the known literature which has consistently reported widespread effects of steroids in the brain (known as neurosteroids) and that these effects, in general, are mediated not
5 through the known classic steroid hormone nuclear receptors which requires transcriptional activation. For instance, neurosteroids have been shown to influence neurotransmission particularly in the field of receptors such as those for GABA and NMDA and Sigma receptors. Neurosteroids have been shown to play a neuroprotective role. Therapeutic intervention through the development of agonists (or antagonists) to
10 LBDG2 may therefore have a role in treatment of neurodegenerative conditions such as dementia, Parkinson's disease and neurodegeneration following cerebrovascular disease such as infarction or haemorrhage (stroke) and trauma to the central nervous system and spinal cord. In addition, neurosteroids have been shown to influence cognitive processing, spatial learning and memory, anxiety and behaviours such as craving which
15 leads to addictive behaviour patterns. Development of agonists and antagonists to LBDG2 may therefore lead to therapeutic intervention to treat dementias, learning difficulties, anxiety, addictive behaviours such as but not exclusively alcoholism, eating disorders and drug addiction.

CLAIMS

1. A polypeptide, which polypeptide:
 - (i) comprises or consists of the amino acid sequence as recited in SEQ ID NO:2;
 - (ii) is a fragment thereof having Nuclear Hormone Receptor Ligand Binding Domain activity or having an antigenic determinant in common with the polypeptide of (i); or
 - (iii) is a functional equivalent of (i) or (ii).
2. A polypeptide which is a fragment according to claim 1(ii), which includes the Nuclear Hormone Receptor Ligand Binding Domain region of the LBDG2 polypeptide, said Nuclear Hormone Receptor Ligand Binding Domain region being defined as including residues 1104 to 1309 inclusive, of the amino acid sequence recited in SEQ ID NO:2, wherein said fragment possesses the "LBD motif" residues PHE1174, VAL1177, ASP1181, ASN1182, LEU1185 and LEU1186, or equivalent residues, and possesses Nuclear Hormone Receptor Ligand Binding Domain activity.
3. A polypeptide which is a functional equivalent according to claim 1(iii), is homologous to the amino acid sequence as recited in SEQ ID NO:2, possesses the catalytic residues PHE1174, VAL1177, ASP1181, ASN1182, LEU1185 and LEU1186, or equivalent residues, and has Nuclear Hormone Receptor Ligand Binding Domain activity.
4. A polypeptide according to claim 3, wherein said functional equivalent is homologous to the Nuclear Hormone Receptor Ligand Binding Domain region of the LBDG2 polypeptide.
5. A fragment or functional equivalent according to any one of claims 1-4, which has greater than 80% sequence identity with an amino acid sequence as recited in SEQ ID NO:2, or with a fragment thereof that possesses Nuclear Hormone Receptor Ligand Binding Domain activity, preferably greater than 85%, 90%, 95%, 98% or 99% sequence identity, as determined using BLAST version 2.1.3 using the default

parameters specified by the NCBI (the National Center for Biotechnology Information; <http://www.ncbi.nlm.nih.gov/>) [Blosum 62 matrix; gap open penalty=11 and gap extension penalty=1].

6. A functional equivalent according to any one of claims 1-5, which exhibits significant structural homology with a polypeptide having the amino acid sequence given in any one of SEQ ID NO:2, or with a fragment thereof that possesses Nuclear Hormone Receptor Ligand Binding Domain activity.
7. A fragment as recited in claim 1, 2, or 5, having an antigenic determinant in common with the polypeptide of claim 1(i), which consists of 7 or more (for example, 8, 10, 12, 14, 16, 18, 20 or more) amino acid residues from the sequence of SEQ ID NO:2.
8. A purified nucleic acid molecule which encodes a polypeptide according to any one of the preceding claims.
9. A purified nucleic acid molecule according to claim 8, which has the nucleic acid sequence as recited in SEQ ID NO:1, or is a redundant equivalent or fragment thereof.
10. A fragment of a purified nucleic acid molecule according to claim 8 or claim 9, which comprises nucleotides 3351 to 3968 of SEQ ID NO:1, or is a redundant equivalent thereof.
11. A purified nucleic acid molecule which hybridizes under high stringency conditions with a nucleic acid molecule according to any one of claims 8-10.
12. A vector comprising a nucleic acid molecule as recited in any one of claims 8-11.
13. A host cell transformed with a vector according to claim 12.
14. A ligand which binds specifically to, and which preferably inhibits the Nuclear Hormone Receptor Ligand Binding Domain activity of, a polypeptide according to any one of claims 1-7.
15. A ligand according to claim 14, which is an antibody.
16. A compound that either increases or decreases the level of expression or activity of a polypeptide according to any one of claims 1-7.

17. A compound according to claim 16 that binds to a polypeptide according to any one of claims 1-7 without inducing any of the biological effects of the polypeptide.
18. A compound according to claim 16 or claim 17, which is a natural or modified substrate, ligand, enzyme, receptor or structural or functional mimetic.
- 5 19. A polypeptide according to any one of claim 1-7, a nucleic acid molecule according to any one of claims 8-11, a vector according to claim 12, a ligand according to claim 14 or 15, or a compound according to any one of claims 16-18, for use in therapy or diagnosis of disease.
20. A method of diagnosing a disease in a patient, comprising assessing the level of
10 expression of a natural gene encoding a polypeptide according to any one of claim 1-7, or assessing the activity of a polypeptide according to any one of claim 1-7, in tissue from said patient and comparing said level of expression or activity to a control level, wherein a level that is different to said control level is indicative of disease.
21. A method according to claim 20 that is carried out *in vitro*.
- 15 22. A method according to claim 20 or claim 21, which comprises the steps of: (a) contacting a ligand according to claim 14 or claim 15 with a biological sample under conditions suitable for the formation of a ligand-polypeptide complex; and (b) detecting said complex.
23. A method according to claim 20 or claim 21, comprising the steps of:
- 20 a) contacting a sample of tissue from the patient with a nucleic acid probe under stringent conditions that allow the formation of a hybrid complex between a nucleic acid molecule according to any one of claims 8-11 and the probe;
- b) contacting a control sample with said probe under the same conditions used in step a);
and
- 25 c) detecting the presence of hybrid complexes in said samples;
wherein detection of levels of the hybrid complex in the patient sample that differ from levels of the hybrid complex in the control sample is indicative of disease.
24. A method according to claim 20 or claim 21, comprising:

- a) contacting a sample of nucleic acid from tissue of the patient with a nucleic acid primer under stringent conditions that allow the formation of a hybrid complex between a nucleic acid molecule according to any one of claims 8-11 and the primer;
- b) contacting a control sample with said primer under the same conditions used in step a); and
- c) amplifying the sampled nucleic acid; and
- d) detecting the level of amplified nucleic acid from both patient and control samples;
- wherein detection of levels of the amplified nucleic acid in the patient sample that differ significantly from levels of the amplified nucleic acid in the control sample is indicative of disease.
25. A method according to claim 20 or claim 21 comprising:
- a) obtaining a tissue sample from a patient being tested for disease;
- b) isolating a nucleic acid molecule according to any one of claims 8-11 from said tissue sample; and
- c) diagnosing the patient for disease by detecting the presence of a mutation which is associated with disease in the nucleic acid molecule as an indication of the disease.
26. The method of claim 25, further comprising amplifying the nucleic acid molecule to form an amplified product and detecting the presence or absence of a mutation in the amplified product.
27. The method of either claim 25 or 26, wherein the presence or absence of the mutation in the patient is detected by contacting said nucleic acid molecule with a nucleic acid probe that hybridises to said nucleic acid molecule under stringent conditions to form a hybrid double-stranded molecule, the hybrid double-stranded molecule having an unhybridised portion of the nucleic acid probe strand at any portion corresponding to a mutation associated with disease; and
- detecting the presence or absence of an unhybridised portion of the probe strand as an indication of the presence or absence of a disease-associated mutation.

28. A method according to any one of claims 20-27, wherein said disease is selected from cell proliferative disorders, including neoplasm, melanoma, lung, colorectal, breast, pancreas, head and neck and other solid tumours, myeloproliferative disorders, such as leukemia, non-Hodgkin lymphoma, leukopenia, thrombocytopenia, angiogenesis disorder, Kaposi's sarcoma, autoimmune/inflammatory disorders, including allergy, inflammatory bowel disease, arthritis, psoriasis and respiratory tract inflammation, asthma, and organ transplant rejection, cardiovascular disorders, including hypertension, oedema, angina, atherosclerosis, thrombosis, sepsis, shock, reperfusion injury, heart arrhythmia, and ischemia, neurological disorders including, central nervous system disease, Alzheimer's disease, brain injury, stroke, amyotrophic lateral sclerosis, anxiety, depression, and pain, developmental disorders, metabolic disorders including diabetes mellitus, osteoporosis, lipid metabolism disorder, hyperthyroidism, hyperparathyroidism, hypercalcemia, hypercholesterolemia, hyperlipidemia, and obesity, renal disorders, including glomerulonephritis, renovascular hypertension, dermatological disorders, including, acne, eczema, and wound healing, negative effects of aging, AIDS, infections including viral infection, bacterial infection, fungal infection and parasitic infection and other pathological conditions, particularly those in which nuclear hormone receptors are implicated..
29. Use of a polypeptide according to any one of claims 1-7 as a Nuclear Hormone Receptor Ligand Binding Domain.
30. Use of a nucleic acid molecule according to any one of claims 8-11 to express a protein that possesses Nuclear Hormone Receptor Ligand Binding Domain activity.
31. A method for effecting cell-cell adhesion, utilising a polypeptide according to any one of claims 1-7.
32. A pharmaceutical composition comprising a polypeptide according to any one of claims 1-7, a nucleic acid molecule according to any one of claims 8-11, a vector according to claim 12, a ligand according to claim 14 or 15, or a compound according to any one of claims 16-18.
33. A vaccine composition comprising a polypeptide according to any one of claims 1-7 or a nucleic acid molecule according to any one of claims 8-11.

34. A polypeptide according to any one of claims 1-7, a nucleic acid molecule according to any one of claims 8-11, a vector according to claim 12, a ligand according to claim 14 or 15, a compound according to any one of claims 16-18, or a pharmaceutical composition according to claim 32 for use in the manufacture of a medicament for the treatment of cell proliferative disorders, including neoplasm, melanoma, lung, colorectal, breast, pancreas, head and neck and other solid tumours, myeloproliferative disorders, such as leukemia, non-Hodgkin lymphoma, leukopenia, thrombocytopenia, angiogenesis disorder, Kaposi's sarcoma, autoimmune/inflammatory disorders, including allergy, inflammatory bowel disease, arthritis, psoriasis and respiratory tract inflammation, asthma, and organ transplant rejection, cardiovascular disorders, including hypertension, oedema, angina, atherosclerosis, thrombosis, sepsis, shock, reperfusion injury, heart arrhythmia, and ischemia, neurological disorders including, central nervous system disease, Alzheimer's disease, brain injury, stroke, amyotrophic lateral sclerosis, anxiety, depression, and pain, developmental disorders, metabolic disorders including diabetes mellitus, osteoporosis, lipid metabolism disorder, hyperthyroidism, hyperparathyroidism, hypercalcemia, hypercholesterolemia, hyperlipidemia, and obesity, renal disorders, including glomerulonephritis, renovascular hypertension, dermatological disorders, including, acne, eczema, and wound healing, negative effects of aging, AIDS, infections including viral infection, bacterial infection, fungal infection and parasitic infection and other pathological conditions, particularly those in which nuclear hormone receptors are implicated.
35. A method of treating a disease in a patient, comprising administering to the patient a polypeptide according to any one of claims 1-7, a nucleic acid molecule according to any one of claims 8-11, a vector according to claim 12, a ligand according to claim 14 or 15, a compound according to any one of claims 16-18, or a pharmaceutical composition according to claim 32.
36. A method according to claim 35, wherein, for diseases in which the expression of the natural gene or the activity of the polypeptide is lower in a diseased patient when compared to the level of expression or activity in a healthy patient, the polypeptide,

- nucleic acid molecule, vector, ligand, compound or composition administered to the patient is an agonist.
37. A method according to claim 35, wherein, for diseases in which the expression of the natural gene or activity of the polypeptide is higher in a diseased patient when compared to the level of expression or activity in a healthy patient, the polypeptide, nucleic acid molecule, vector, ligand, compound or composition administered to the patient is an antagonist.
38. A method of monitoring the therapeutic treatment of disease in a patient, comprising monitoring over a period of time the level of expression or activity of a polypeptide according to any one of claims 1-7, or the level of expression of a nucleic acid molecule according to any one of claims 8-11 in tissue from said patient, wherein altering said level of expression or activity over the period of time towards a control level is indicative of regression of said disease.
39. A method for the identification of a compound that is effective in the treatment and/or diagnosis of disease, comprising contacting a polypeptide according to any one of claims 1-7, a nucleic acid molecule according to any one of claims 8-11, or a host cell according to claim 13 with one or more compounds suspected of possessing binding affinity for said polypeptide or nucleic acid molecule, and selecting a compound that binds specifically to said nucleic acid molecule or polypeptide.
40. A kit useful for diagnosing disease comprising a first container containing a nucleic acid probe that hybridises under stringent conditions with a nucleic acid molecule according to any one of claims 8-11; a second container containing primers useful for amplifying said nucleic acid molecule; and instructions for using the probe and primers for facilitating the diagnosis of disease.
41. The kit of claim 40, further comprising a third container holding an agent for digesting unhybridised RNA.
42. A kit comprising an array of nucleic acid molecules, at least one of which is a nucleic acid molecule according to any one of claims 8-11.

43. A kit comprising one or more antibodies that bind to a polypeptide as recited in any one of claims 1-7; and a reagent useful for the detection of a binding reaction between said antibody and said polypeptide.
44. A transgenic or knockout non-human animal that has been transformed to express higher, lower or absent levels of a polypeptide according to any one of claims 1-7.
- 5 45. A method for screening for a compound effective to treat disease, by contacting a non-human transgenic animal according to claim 44 with a candidate compound and determining the effect of the compound on the disease of the animal.

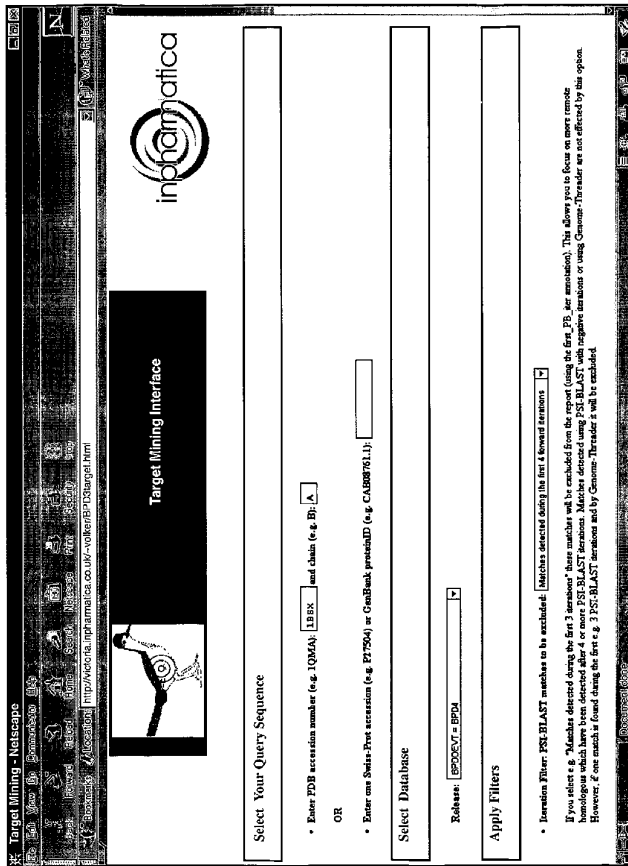


FIG. 1

SUBSTITUTE SHEET (RULE 26)

Accession	Gene	Protein	Species	Length	Score	Expect	Ident	Start	End
U01489	CD44	CD44	Human	105	16.28	2e-399	6	1	105
U01489	CD44	CD44	Human	105	15.95	2e-399	6	1	105
U01489	CD44	CD44	Human	105	15.62	2e-399	6	1	105
U01489	CD44	CD44	Human	105	15.29	2e-399	6	1	105
U01489	CD44	CD44	Human	105	14.96	2e-399	6	1	105
U01489	CD44	CD44	Human	105	14.63	2e-399	6	1	105
U01489	CD44	CD44	Human	105	14.30	2e-399	6	1	105
U01489	CD44	CD44	Human	105	13.97	2e-399	6	1	105
U01489	CD44	CD44	Human	105	13.64	2e-399	6	1	105
U01489	CD44	CD44	Human	105	13.31	2e-399	6	1	105
U01489	CD44	CD44	Human	105	12.98	2e-399	6	1	105
U01489	CD44	CD44	Human	105	12.65	2e-399	6	1	105
U01489	CD44	CD44	Human	105	12.32	2e-399	6	1	105
U01489	CD44	CD44	Human	105	11.99	2e-399	6	1	105
U01489	CD44	CD44	Human	105	11.66	2e-399	6	1	105
U01489	CD44	CD44	Human	105	11.33	2e-399	6	1	105
U01489	CD44	CD44	Human	105	11.00	2e-399	6	1	105
U01489	CD44	CD44	Human	105	10.67	2e-399	6	1	105
U01489	CD44	CD44	Human	105	10.34	2e-399	6	1	105
U01489	CD44	CD44	Human	105	10.01	2e-399	6	1	105
U01489	CD44	CD44	Human	105	9.68	2e-399	6	1	105
U01489	CD44	CD44	Human	105	9.35	2e-399	6	1	105
U01489	CD44	CD44	Human	105	9.02	2e-399	6	1	105
U01489	CD44	CD44	Human	105	8.69	2e-399	6	1	105
U01489	CD44	CD44	Human	105	8.36	2e-399	6	1	105
U01489	CD44	CD44	Human	105	8.03	2e-399	6	1	105
U01489	CD44	CD44	Human	105	7.70	2e-399	6	1	105
U01489	CD44	CD44	Human	105	7.37	2e-399	6	1	105
U01489	CD44	CD44	Human	105	7.04	2e-399	6	1	105
U01489	CD44	CD44	Human	105	6.71	2e-399	6	1	105
U01489	CD44	CD44	Human	105	6.38	2e-399	6	1	105
U01489	CD44	CD44	Human	105	6.05	2e-399	6	1	105
U01489	CD44	CD44	Human	105	5.72	2e-399	6	1	105
U01489	CD44	CD44	Human	105	5.39	2e-399	6	1	105
U01489	CD44	CD44	Human	105	5.06	2e-399	6	1	105
U01489	CD44	CD44	Human	105	4.73	2e-399	6	1	105
U01489	CD44	CD44	Human	105	4.40	2e-399	6	1	105
U01489	CD44	CD44	Human	105	4.07	2e-399	6	1	105
U01489	CD44	CD44	Human	105	3.74	2e-399	6	1	105
U01489	CD44	CD44	Human	105	3.41	2e-399	6	1	105
U01489	CD44	CD44	Human	105	3.08	2e-399	6	1	105
U01489	CD44	CD44	Human	105	2.75	2e-399	6	1	105
U01489	CD44	CD44	Human	105	2.42	2e-399	6	1	105
U01489	CD44	CD44	Human	105	2.09	2e-399	6	1	105
U01489	CD44	CD44	Human	105	1.76	2e-399	6	1	105
U01489	CD44	CD44	Human	105	1.43	2e-399	6	1	105
U01489	CD44	CD44	Human	105	1.10	2e-399	6	1	105
U01489	CD44	CD44	Human	105	0.77	2e-399	6	1	105
U01489	CD44	CD44	Human	105	0.44	2e-399	6	1	105
U01489	CD44	CD44	Human	105	0.11	2e-399	6	1	105

FIG. 2B

SUBSTITUTE SHEET (RULE 26)

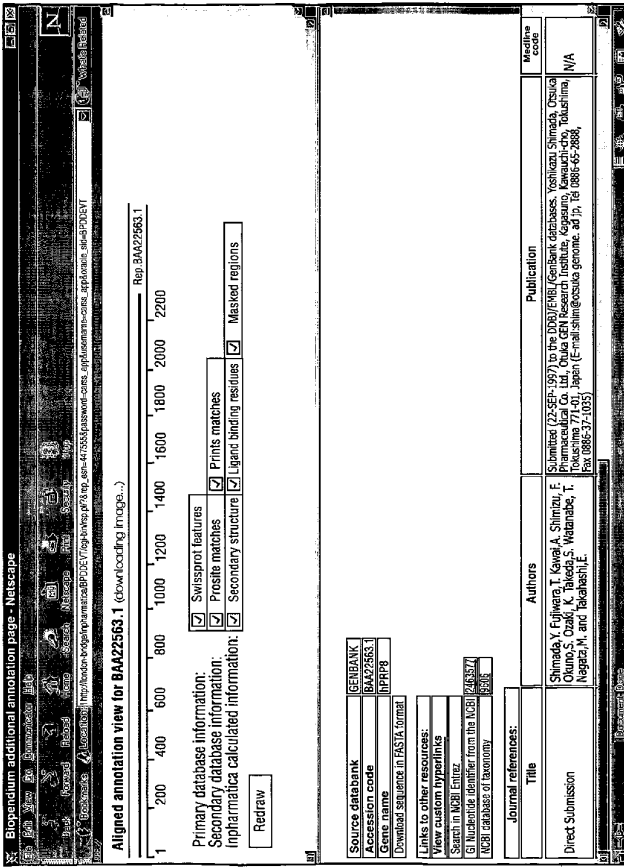


FIG. 3

InterPro search Results.

1. Query Sequence submitted Length 2335aa.

InterPro	Results of PfSearch against PROSITE	Results of FingerPRINTScan against PRINTS	Results of HMMDecipher against PFAM-A
[IPR000148] Papillomavirus E7 protein			[PF00527] [1561-1570]
[IPR0001472] Bipartite nuclear localization signal	[PS50079] [35-52]		

[IPR000148] [PF00527]
[IPR000172] [PF00579]

XML / TXT formatted

Applications furnished normally.

FIG. 4

NCBI Science Viewer - Nucleotide

BLAST Related Sequences, Nucleotide, Taxonomy, LinkOut

FASTA protein [Homo sapiens]

LOCUS BAA22563 2335 aa PRI 03-OCT-1997

DEFINITION PR80 protein [Homo sapiens].

ACCESSION U00001

VERSION 2446377

FEATURES

locus AB007510 accession AB007510.1

ORIGIN

Human.

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

Shimada, Y., Fujikawa, T., Kawai, A., Shimizu, F., Okuno, F., Ozaki, K., Otsuda, H., Kasanabe, T., Nagata, M. and Takahashi, E. Submitted (22-SEP-1997) Yoshikazu Shimada, Otsuka Pharmaceutical Tokushima, Oshimizu 772-0275, Japan; Date: Kaguno, Kawachi-cho, Tokushima, Oshimizu 772-0275, Japan. (E-mail: shimadetuoka.genose@id.jpr.tol.0886-66-2889.

2 (ORIGIN 1 to 2335)

AUTHORS Shimada, Y., Fujikawa, T., Kawai, A., Shimizu, F., Okuno, F., Ozaki, K., Otsuda, H., Kasanabe, T., Nagata, M. and Takahashi, E.

TITLE Human homologues of Saccharomyces cerevisiae PR80, pre-mRNA splicing factor

JOURNAL JOURNAL OF BIOLOGICAL CHEMISTRY

FEATURES

source

1..2335 = "Homo sapiens" /db_xref="taxon:9606"

chromosome=17

1..2335 = "Human lib-Fetal brain" /db_xref="taxon:9606"

1..2335 = "PR80 protein" /gene="hprp8" /db_xref="AB007510.1:42..7049"

ORIGIN

1 mgtatprrp gttggtggt tctctctg gctctctc tctctctc tctctctc

61 tctctctc tctctctc tctctctc tctctctc tctctctc tctctctc

121 tttgtaifvn aipaeiepy laqgmas atreairch tkatppid ceapldpy

181 tttgtaifvn aipaeiepy laqgmas atreairch tkatppid ceapldpy

241 tttgtaifvn aipaeiepy laqgmas atreairch tkatppid ceapldpy

301 klltqspit syiaefpy mlphvht wbtppvfi krapdipaf yfpllpa

361 hruvageep ipdsoefal psvpffka tplycdan qallwvrrp ftrvgtr

FIG. 5

SUBSTITUTE SHEET (RULE 26)

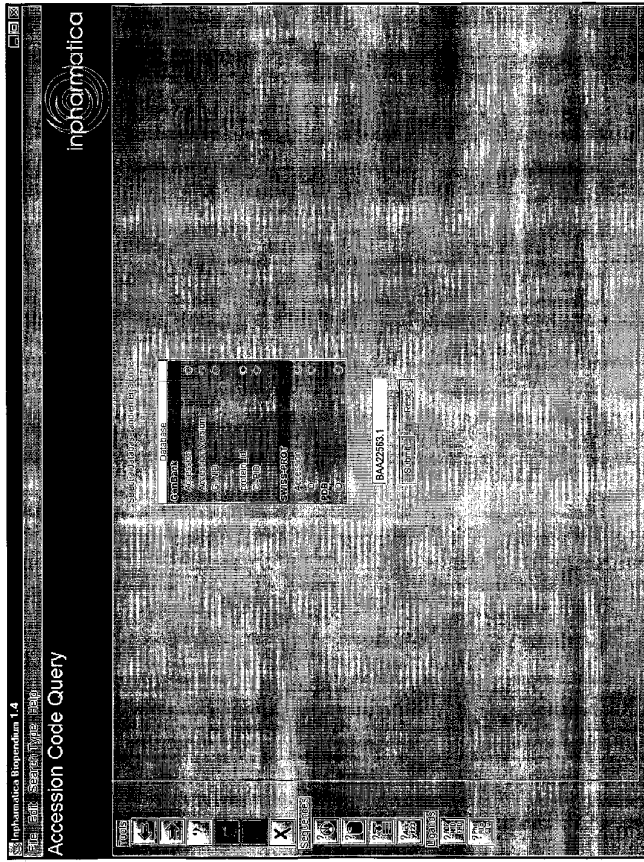


FIG. 6A

SUBSTITUTE SHEET (RULE 26)

AIEye output (January 18, 2001 5:45 PM)

FIG. 7-1

```

10 20 30 40 50 60 70
BA422553.1  MAGVFPYRGGPNVPGPLAPLDYMSSEKLEKARKWQQLQAKRYAEKRFGEVDAOKEDMPPEHVEIIRDHGDMTNR
1BSXA
80 90 100 110 120 130 140 150
BA422553.1  KFRHDKRYVGLGALKYMPHAVLKLLENMPWQIIRDVPVLYHTGATISFVNEIPWVIEPVYISOWGSMWIMMRREKDR
1BSXA
160 170 180 190 200 210 220 230
BA422553.1  RHFKRMRFPPDDEEPPLDYADNINLNVPLEAQLLEDPEDAPVLDWFDVHQLDLSRKYVNGSTYORWQFTLPMWST
1BSXA
240 250 260 270 280 290 300 310
BA422553.1  LYRLANOLLTLVDDNYFYFLDKAFFTSKALNMAIPGGPKFELVYRDINLOEDWNEFDINKIIRQPIRTEYKIAF
1BSXA
320 330 340 350 360 370 380 390
BA422553.1  PLYNNLPHHVHLTWHTPNVVFIKTEDPLPAFYFDPLINPISHRHSVKSQEPPLPDDDEFEFLPEFVEFLKDTPLYT
1BSXA
400 410 420 430 440 450 460 470
BA422553.1  DNTANGIALLWAPRFNLSGRTRALDIPLYKNWYREHCPAGOPYKRVSYOKLLKYVYLNALKHPKPAKQKRYLFR
1BSXA
480 490 500 510 520 530 540 550
BA422553.1  SFKATKFFQSTKLDWVEGVLOVCRGYNMLNLIHRKNLNYLHLDYFNFLKPVKTLTKERKSRFGNAFHLCREVLRL
1BSXA
560 570 580 590 600 610 620 630
BA422553.1  TKLVVDSHVGVRLGNVDAFOLADGLOYIFAHVGLTGMRYRYKYLKMRQIRVCKDLKHLIYYRFNTGVPVGGKGGGFMAA
1BSXA
640 650 660 670 680 690 700 710
BA422553.1  GWRVWLFMRGITPLLERWLGMLLAROFEGRHSKGVAKTVTQORVESHFDLELRAAVMHDILOMMPGIIKONKARTILQ
1BSXA

```

SUBSTITUTE SHEET (RULE 26)

FIG. 7-3 AIEye output (January 18, 2001 5:45 PM)

FIG. 7-3

BAA22563.1 1430 1440 1450 1460 1470 1480 1490 1500
 1BSX/A ONRRLLLEDSWDRGI PRINTLFOKDRHTLAYD KGWFRVTRD KGYOVL KONPFWWTHORHDKGLWNLNNYRMTDMLQA

BAA22563.1 1510 1520 1530 1540 1550 1560 1570 1580
 1BSX/A LGGVEGILEHTLFKGTFFTWECLEFWEKASGFESMKWKKLTNAQRSGLNQIPNRRFTLWWSPTINRANVYVGFQVOLD

BAA22563.1 1590 1600 1610 1620 1630 1640 1650
 1BSX/A LTGIFMHGKIPTLKISLIQIFRAHLWQKIHESIYMDLCOVFOEIDALEIETVOKETIHPKRSYKMNSSCADILLFASY

BAA22563.1 1660 1670 1680 1690 1700 1710 1720 1730
 1BSX/A KWNVSRPSELLADSKDVMDSITTKYKIDQLRWGDYDSDHIERARAKFDYTTDMMSIYPSPTGYLIIAIDLAYNLHSA

BAA22563.1 1740 1750 1760 1770 1780 1790 1800 1810
 1BSX/A YGNWFPGSKPLIQOAMAKIMKANPALYVLRERIRKGLOLYSSEPTPEYLLSSONYGELFSNOIIFWYDDTNVYRVTIHKT

BAA22563.1 1820 1830 1840 1850 1860 1870 1880 1890
 1BSX/A FEGNLTTPINGAIFIFNPRTGGLFLKIHTSVWAGOKRGLGAKWKTAEEYAALIRSLPVEEOPKQIIIVTRKMDLPL

BAA22563.1 1900 1910 1920 1930 1940 1950 1960 1970
 1BSX/A EVHLLDFPNIYIKGSELOLFFOACLKVEKFGDLILKATEPQWVLFNLYDDWLKTISSYTAFSGLIILRALHVNDRRAK

BAA22563.1 1980 1990 2000 2010 2020 2030 2040 2050
 1BSX/A VILKPKDTTITEPHHIWPPTLDEEWIKYEVOLKDLILADYKKNMNVNASLTQSEIRDIILGMEISAPSOOROOIAEIE

BAA22563.1 2060 2070 2080 2090 2100 2110 2120 2130
 1BSX/A KOTKEQSOLTATQTRTVNKHGDEIITSTTSNYETQTFSSKTEWRVRAISAANHLRTHNIYVSSDDIKETGYTYILLPKN

SUBSTITUTE SHEET (RULE 26)

FIG. 7-4

ALEye output (January 18, 2001 5:45 PM)

2140 | 2150 | 2160 | 2170 | 2180 | 2190 | 2200 | 2210
 BA4225831 | VLK K F I C I S D L R A Q I A G Y L Y G V S P D N P Q V K E I R C I V M V P O W G T H Q T V H L P G Q L P O H E Y L K E M E F L G W I H T O P N E S P O L
 1BSX:A

2220 | 2230 | 2240 | 2250 | 2260 | 2270 | 2280 | 2290
 BA4225831 | S P O D V T T H A K I M A D N P S W D G E K T I I T C S F T P S C T L T A Y K L T P S G Y E W G R Q N T D K G N N P K G Y L P S H Y E R V O M L L S D R F
 1BSX:A

2300 | 2310 | 2320 | 2330 | 2340 | 2350
 BA4225831 | L G F F M Y P A O S S W N Y N F M G V R H D P N M K Y E L O L A N P K E F Y H E V H R P S H F L N F A L L O G E G E V Y S A D R E D L Y A
 1BSX:A

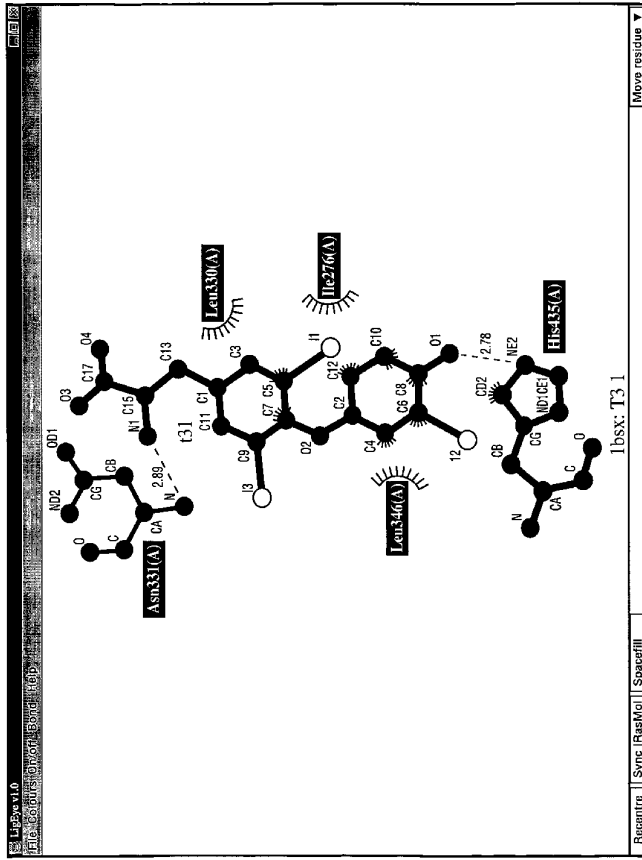


FIG. 8A

SUBSTITUTE SHEET (RULE 26)

AIEye output (February 26, 2001 9:53 AM)

FIG. 9-3

BA02563.1 RTILOHLS EAWRCWKANI PWKVPGLPTPIENMI LRYVKA KADWWTN TAHYNER IRRGATVDTKTKKKNLGRLTRLYK
 AAF5873.1 RTILOHLS EAWRCWKANI PWKVPGLPTPIENMI LRYVKA KADWWTN TAHYNER IRRGATVDTKTKKKNLGRLTRLYK
 P04369 RTILOHLS EAWRCWKANI PWKVPGLPTPIENMI LRYVKA KADWWTN SAHYNER VRRGATVDTKTKKKNLGRLTRLYK
 BA07844.1 RTILOHLS EAWRCWKANI PWKVPGLPTPIENMI LRYVKA KADWWTN VAHYNER IRRGATVDTKTKKKNLGRLTRLYK
 CAB80541.1 RTILOHLS EAWRCWKANI PWKVPGLPTPIENMI LRYVKA KADWWTN VAHYNER IRRGATVDTKTKKKNLGRLTRLYK

 BA02563.1 AEQERHNYLKDGPYI TAETVAVYTTVHMLESRFSP IFFPPLSYKHDTKLLI LALERLKEAYSVKSRLLNOSOREEL
 AAF5873.1 AEQERHNYLKDGPYI SAEVAIYTTVHMLESRFSP IFFPPLSYKHDTKLLI LALERLKEAYSVKSRLLNOSOREEL
 P04369 SEQERHNYLKDGPYI SAEVAIYTTVHMLESRFSP IFFPPLSYKHDTKLLI LALERLKEAYSVKSRLLNOSOREEL
 BA07844.1 AEQERHNYLKDGPYI TAETVAVYTTVHMLESRFSA IFFPPLSYKHDTKLLI LALERLKEAYSVAVRLNOSOREEL
 CAB80541.1 AEQERHNFQDGPYVTADEGIAIYSTVHMLESRKFSAI IFFPPLSYKHDTKLLI LALERLKEAYSVAVRLNOSOREEL

 BA02563.1 GLIEQAYDNHEALSRIKRHLLTORAFKEVGEI EFNLDLYSHLPVYDVEPLEKITDYLDOYLWYEAOKRLLFPWIKPA
 AAF5873.1 GLIEQAYDNHEALSRIKRHLLTORAFKEVGEI EFNLDLYSHLPVYDVEPLEKITDYLDOYLWYEAOKRLLFPWIKPA
 P04369 ALIEQAYDNHEALSRIKRHLLTORAFKEVGEI EFNLDLYSHLPVYDVEPLEKITDYLDOYLWYEAOKRLLFPWIKPA
 BA07844.1 GLIEQAYDNHEALSRIKRHLLTORAFKEVGEI EFNLDLYSHLPVYDVEPLEKITDYLDOYLWYEAOKRLLFPWIKPA
 CAB80541.1 GLIEQAYDNHEALSRIKRHLLTORAFKEVGEI EFNLDLYSHLPVYDVEPLEKITDYLDOYLWYEAOKRLLFPWIKPA

 BA02563.1 DTEPPPLLYKWCQGINNLDQVWETSEGEENVMLSEKFEKVEKIDLTLLNRLVLI VDHNIADYMTAKNNVVIYKDM
 AAF5873.1 DTEPPPLLYKWCQGINNLDQVWETSEGEENVMLSEKFEKVEKIDLTLLNRLVLI VDHNIADYMTAKNNVVIYKDM
 P04369 DTEPPPLLYKWCQGINNLDQVWETSEGEENVIMETKLEKIAEKMDLTLNRLVLI VDHNIADYMTAKNNVLIYKDM
 BA07844.1 DSEPPPLLYKWCQGINNLDQVWETSDGOCVVMLOTXFEKFEKIDLTLLNRLVLI VDHNIADYMTAKNNVLIYKDM
 CAB80541.1 DSEPPPLLYKWCQGINNLDQVWETSDGOCVVMLOTXFEKFEKIDLTLLNRLVLI VDHNIADYMTAKNNVLIYKDM

 BA02563.1 NHTNSYGIIRGLQFASFIVQYGLVMDLLVQLHRASEMAGPOMPNDLFSFODIATEAHP IRLFCRYIDRIHIFFRF
 AAF5873.1 NHTNSYGIIRGLQFASFIVQYGLVMDLLVQLHRASEMAGPOMPNDLFSFODIATEAHP IRLFCRYIDRIHIFFRF
 P04369 NHTNSYGIIRGLQFASFIVQYGLVMDLLVQLHRASEMAGPOMPNDLFSFODIATEAHP IRLFCRYIDRIHIFFRF
 BA07844.1 SHNTSYGLIRGLQFASFIVQYGLVMDLLVQLHRASEMAGPOMPNDLFSFODIATEAHP IRLFCRYIDRIHIFFRF
 CAB80541.1 SYTNTYGLIRGLQFASFIVQYGLVMDLLVQLHRASEMAGPOMPNDLFSFODIATEAHP IRLFCRYIDRIHIFFRF

AIEye output (February 26, 2001 9:53 AM)

FIG. 9-5

BAA22653.1 F P T W E G L F W E K A S G F E E S M K Y K L L T N A O R S G L N Q I P M R R F T L W S P T I N R A N V Y V G F O V Q L D L T G I F M H G K I P T L K I S L 1590 1600 1610 1620 1630 1640 1650
 AAF58573.1 F P T W E G L F W E K A S G F E E S M K Y K L L T N A O R S G L N Q I P M R R F T L W S P T I N R A N V Y V G F O V Q L D L T G I F M H G K I P T L K I S L 1660 1670 1680 1690 1700 1710 1720 1730
 P34389 F P T W E G L F W E K A S G F E E S M K Y K L L T N A O R S G L N Q I P M R R F T L W S P T I N R A N V Y V G F O V Q L D L T G I F M H G K I P T L K I S L 1740 1750 1760 1770 1780 1790 1800 1810
 BAA78744.1 F P T W E G L F W E K A S G F E E S M K Y K L L T N A O R S G L N Q I P M R R F T L W S P T I N R A N V Y V G F O V Q L D L T G I F M H G K I P T L K I S L 1820 1830 1840 1850 1860 1870 1880 1890
 CAB80541.1 F P T W E G L F W E K A S G F E E S M K Y K L L T N A O R S G L N Q I P M R R F T L W S P T I N R A N V Y V G F O V Q L D L T G I F M H G K I P T L K I S L 1900 1910 1920 1930 1940 1950 1960 1970
 BAA22653.1 I Q I F R A H L W O K I H E S I V M D L C O V F D Q E L D A L E I E T V K E T I H P R K S Y K M S S C A D I L L F A Y K W N V S R P S L L A D S K D V M 1740 1750 1760 1770 1780 1790 1800 1810
 AAF58573.1 I Q I F R A H L W O K I H E S I V M D L C O V F D Q E L D A L E I E T V K E T I H P R K S Y K M S S C A D I L L F A Y K W N V S R P S L L A D S K D V M 1820 1830 1840 1850 1860 1870 1880 1890
 P34389 I Q I F R A H L W O K I H E S I V M D L C O V F D Q E L D A L E I E T V K E T I H P R K S Y K M S S C A D I L L F A Y K W N V S R P S L L A D S K D V M 1900 1910 1920 1930 1940 1950 1960 1970
 BAA78744.1 I Q I F R A H L W O K I H E S I V M D L C O V F D Q E L D A L E I E T V K E T I H P R K S Y K M S S C A D I L L F A Y K W N V S R P S L L A D S K D V M 1980 1990 2000 2010 2020 2030 2040 2050
 CAB80541.1 I Q I F R A H L W O K I H E S I V M D L C O V F D Q E L D A L E I E T V K E T I H P R K S Y K M S S C A D I L L F A Y K W N V S R P S L L A D S K D V M 2060 2070 2080 2090 2100 2110 2120 2130
 BAA22653.1 D S T T T O K Y W I D I Q L R W G D Y D S H D I E R Y A R A K F L D Y T T D N M S I Y P S P T G V L I A I D L A Y N L H S A Y G N W F P G S K P L I O O A M A 1740 1750 1760 1770 1780 1790 1800 1810
 AAF58573.1 D N T T T O K Y W L D I Q L R W G D Y D S H D V E R Y A R A K F L D Y T T D N M S I Y P S P T G V L I A I D L A Y N L H S A Y G N W F P G S K P L I O O A M A 1820 1830 1840 1850 1860 1870 1880 1890
 P34389 D N T T T O K Y W L D V O L R W G D Y D S H D V E R Y A R A K F L D Y T T D N M S I Y P S P T G V L I A I D L A Y N L H S A Y G N W F P G S K P L I O O A M A 1900 1910 1920 1930 1940 1950 1960 1970
 BAA78744.1 D O K A S N K Y I D V O L R W G D Y D S H D I E R Y T R A K F M D Y T T D N M S I Y P S P T G V I I G L D L A Y N L H S A F G N W F P G S K P L L Q A M N 1980 1990 2000 2010 2020 2030 2040 2050
 CAB80541.1 D O K A S N K Y I D V O L R W G D Y D S H D I E R Y T K A K F M D Y T T D N M S I Y P S P T G V I I G L D L A Y N L H S A F G N W F P G S K P L L Q A M N 2060 2070 2080 2090 2100 2110 2120 2130
 BAA22653.1 K I M K A N P A L Y V L R E R I R K G L Q L Y S S E P T E P L S S O N Y G E I F S N Q I I W F V D D T N V Y R V T I H K T F E G N L T T K P I N G A I F I F 1740 1750 1760 1770 1780 1790 1800 1810
 AAF58573.1 K I M K A N P A L Y V L R E R I R K A L Q L Y S S E P T E P L S S O N Y G E I F S N Q I I W F V D D T N V Y R V T I H K T F E G N L T T K P I N G A I F I F 1820 1830 1840 1850 1860 1870 1880 1890
 P34389 K I I K A N P A F Y V L R E R I R K G L Q L Y S S E P T E P L S S O N Y G E I F S N Q I I W F V D D T N V Y R V T I H K T F E G N L T T K P I N G A I F I F 1900 1910 1920 1930 1940 1950 1960 1970
 BAA78744.1 K I M K S N P A L Y V L R E R I R K G L Q L Y S S E P T E P L S S O N Y G E I F S N Q I I W F V D D T N V Y R V T I H K T F E G N L T T K P I N G A I F I F 1980 1990 2000 2010 2020 2030 2040 2050
 CAB80541.1 K I M K S N P A L Y V L R E R I R K G L Q L Y S S E P T E P L S S O N Y G E I F S N Q I I W F V D D T N V Y R V T I H K T F E G N L T T K P I N G A I F I F 2060 2070 2080 2090 2100 2110 2120 2130
 BAA22653.1 N P R T G Q L F L K I I H T S Y W A G G K R L G O L A K K W T A E E V A A L I R S L P V E E O P K O I I V T R K M D L D P L E V H L L D F F N I V I K G S E L 1740 1750 1760 1770 1780 1790 1800 1810
 AAF58573.1 N P R T G Q L F L K I I H T S Y W A G G K R L G O L A K K W T A E E V A A L I R S L P V E E O P K O I I V T R K M D L D P L E V H L L D F F N I V I K G S E L 1820 1830 1840 1850 1860 1870 1880 1890
 P34389 N P R T G Q L F L K I I H T S Y W A G G K R L S O L A K K W T A E E V A A L I R S L P V E E O P K O I I V T R K M D L D P L E V H L L D F F N I V I K G S E L 1900 1910 1920 1930 1940 1950 1960 1970
 BAA78744.1 N P R T G Q L F L K V I H T S Y W A G G K R L G O L A K K W T A E E V A A L V R S L P V E E O P K O I I V T R K M D L D P L E V H L L D F F N I V I K G S E L 1980 1990 2000 2010 2020 2030 2040 2050
 CAB80541.1 N P R T G Q L F L K I I H T S Y W A G G K R L G O L A K K W T A E E V A A L V R S L P V E E O P K O I I V T R K M D L D P L E V H L L D F F N I V I K G S E L 2060 2070 2080 2090 2100 2110 2120 2130

AlEye output (February 26, 2001 9:53 AM)

FIG. 9-6

1980 1990 2000 2010 2020 2030 2040 2050
 BAA22563.1 QLPFQACLKVEKFGDLILKATEPOMVLFNLYDDWLKTISSYTAFSRLILRALHVNNDRAKVILKPKDKTITTEPHHIW
 AAF56573.1 QLPFQACLKVEKFGDLILKATEPOMVLFNLYDDWLKTISSYTAFSRLILRALHVNNDRAKVILKPKDKTITTEPHHIW
 P34369 MLPFQAIMKVEKFGDLILKATEPOMVLFNLYDDWLKTISSYTAFSRAVLIMRGMHINPKKTKVITTEPHHIW
 BAA78744.1 QLPFQACLKIEKFGDLILKATEPOMVLFNLYDDWLKTISSYTAFSRLILRALHVNNDRAKVILKPKDKTITTEPHHIW
 CAB80541.1 QLPFQACLKIEKFGDLILKATEPOMALFNLYDDWLMTVSSYTAQRLILRALHVNNDRAKVILKPKDKTITTEPHHIW
 2060 2070 2080 2090 2100 2110 2120 2130
 BAA22563.1 PTLTDEEWIKVEYGLKDLILADYGGKNNVNVASLTQSEIRDIILGMEISAPSQRQIQIAIEKQTKEQSOLTATQTRTV
 AAF56573.1 PTLTDEEWIKVEYGLKDLILADYGGKNNVNVASLTQSEIRDIILGMEISAPSQRQIQIAIEKQTKEQSOLTATQTRTV
 P34369 PTLTDEEWIKVELAKOMLILADYGGKNNVNVASLTQSEVRDIILGMEISAPSQRQIQIAIEKQTKEQSOLTATQTRTV
 CAB80541.1 PSLTDDCWMMKVEVALRDLILSDYAKKNNVNTSALTQSEIRDIILGAEIIPPSOORQIQIAIEKQTKEQSOLTATQTRTV
 2140 2150 2160 2170 2180 2190 2200 2210
 BAA22563.1 NKHGDEIITSTTSNYETOTFSSKTEWRVRAISANLHLRTHNHYVSSDDIKETGYTYILPKNVLKKFICISDLRAQIAG
 AAF56573.1 NKHGDEIITSTTSNYETOTFSSKTEWRVRAISANLHLRTHNHYVSSDDIKETGYTYILPKNVLKKFICISDLRAQIAG
 P34369 NKHGDEIITATTSNYETAFASRTEWRVRAISSTNLHLRTHNHYVSSDDIKETGYTYILPKNVLKKFICISDLRAQIAG
 BAA78744.1 NVHGDELIITTSPYEQAFASKTDWRVRAISATNLYLRVNHLYVNSDDIKETGYTYILPKNVLKKFICISDLRAQIAG
 CAB80541.1 NVHGDELIITTSIPEYEQSAFSGSKTDWRVRAISATNLYLRVNHLYVNSDDIKETGYTYILPKNVLKKFICISDLRAQIAG
 2220 2230 2240 2250 2260 2270 2280 2290
 BAA22563.1 YLYGVSPPDNPOVKEIRCIWMPQWGTHTVHLPGOLPOHEYLKMEPELGMWHTOPNEPOLSPODVTTHAKIMADNPS
 AAF56573.1 YLYGVSPPDNPOVKEIRCIWMPQWGTHTVHLPGOLPOHEYLKMEPELGMWHTOPNEPOLSPODVTTHAKIMADNPS
 P34369 FMYGVSPPDNPOVKEIRCIWLPQTSHQOYVLPALPOHELLRDEPLGSMHTOPNEPOLSPODVTTHAKILDNIS
 BAA78744.1 FLYGLSPPDNPOVKEIRCIAPQHSHTQOYVLPALPEHELNDLEPLGSMHTOPNEPOLSPODVTTHAKILDNIS
 CAB80541.1 YLYGVSPPDNPOVKEIRCIWMPQWGTHTVHLPGOLPOHEYLKMEPELGMWHTOPNEPOLSPODVTTHAKILDNIS
 2300 2310 2320 2330 2340 2350 2360 2370
 BAA22563.1 WDGEKTIITCSTFPGSCSLTAYKLTSPGVWGRQNTDGNPKGYLPSHYERVQMLLSDRFLGFFMVAQSSWVNFEM
 AAF56573.1 WDGEKTIITCSTFPGSCSLTAYKLTSPGVWGRQNTDGNPKGYLPSHYERVQMLLSDRFLGFFMVAQSSWVNFEM
 P34369 WDGEKTIITCSTFPGSCSLTAYKLTSPGVWGRQNTDGNPKGYLPSHYERVQMLLSDRFLGFFMVAQSSWVNFEM
 BAA78744.1 WDGEKTIITCSTFPGSCSLTAYKLTSPGVWGRQNTDGNPKGYLPSHYERVQMLLSDRFLGFFMVAQSSWVNFEM
 CAB80541.1 WDGEKTIITCSTFPGSCSLTAYKLTQAGYEWGRQNTDGNPKGYLPSHYERVQMLLSDRFLGFFMVAQSSWVNFEM

FIG. 10

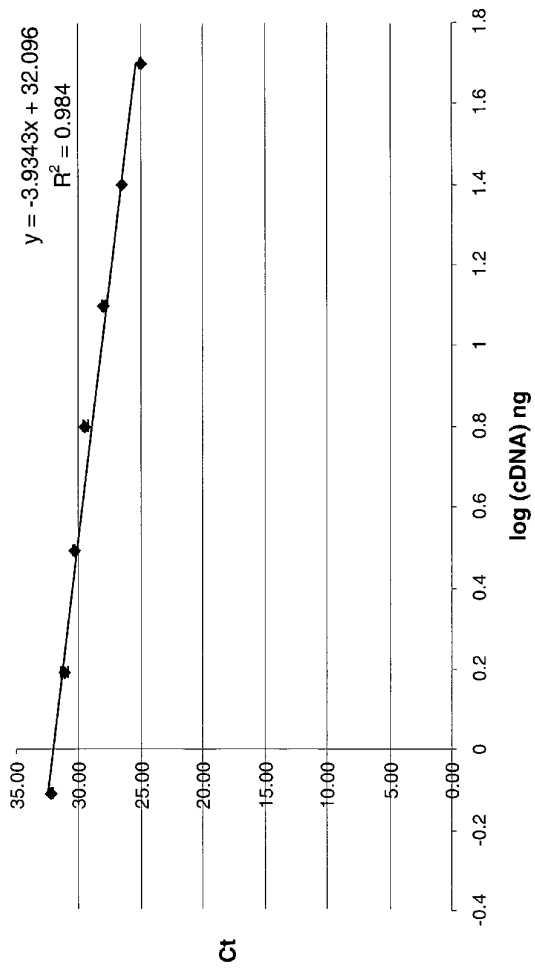


FIG. 11

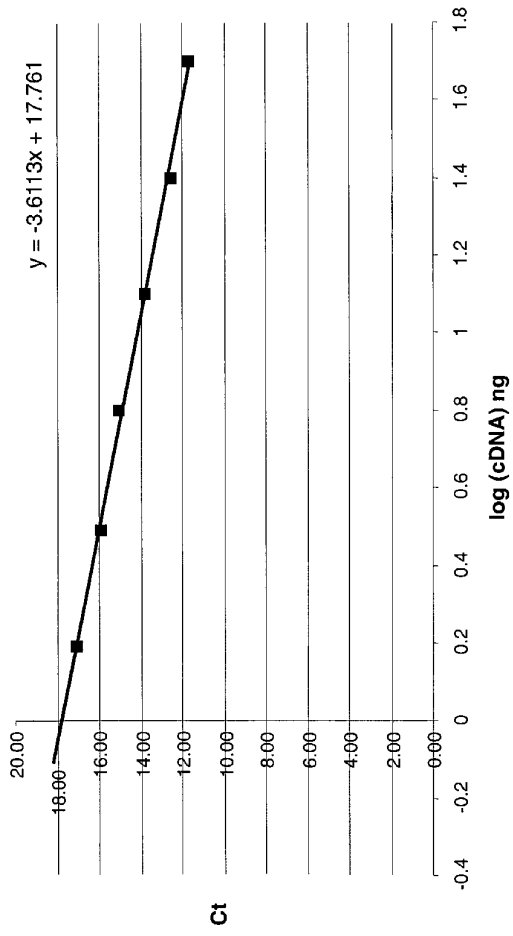


FIG. 12

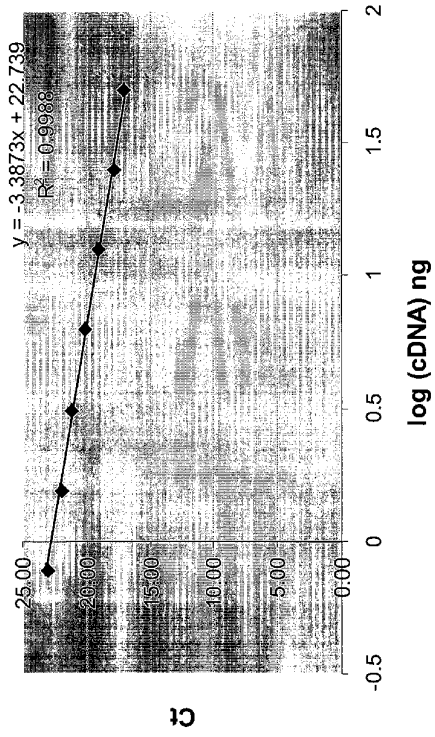


FIG. 13

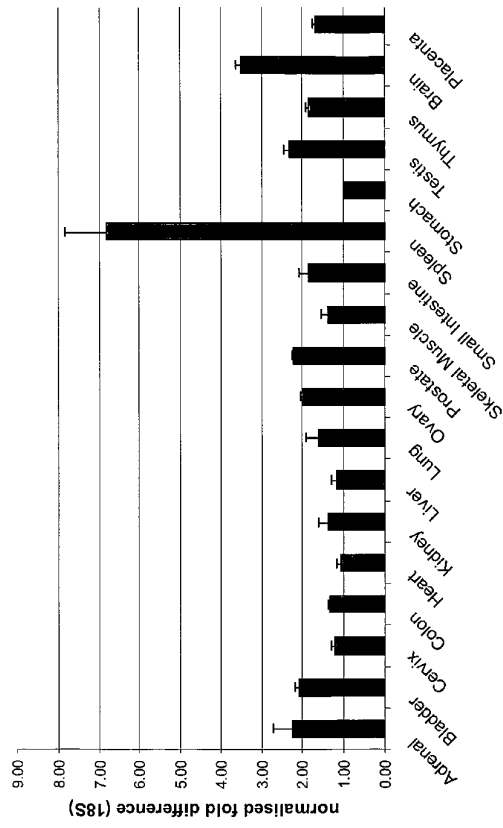
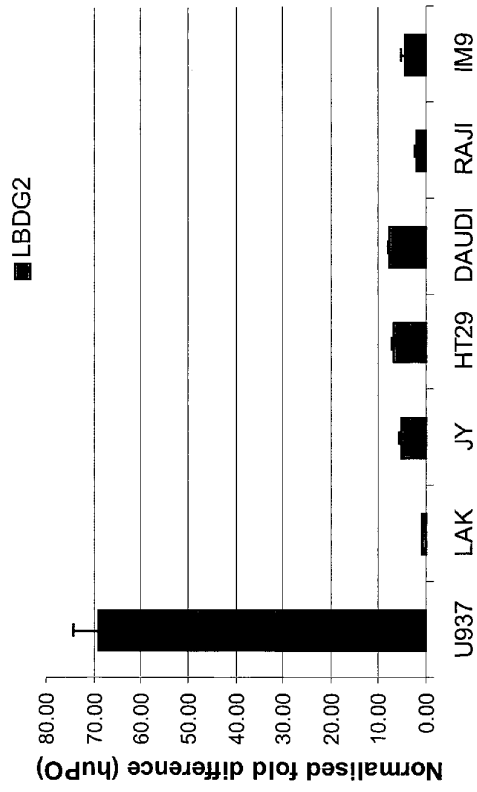


FIG. 14



WO 02/070563

1/5

PCT/GB02/00986

Sequence Listing

SEQ ID NO: 1 Nucleotide coding sequence for BAA22563.1 (LBDG2) protein

```

1  cgggcggcct cttgtgtgag ggcctgtggg attctccgga tatggccgga gtgtttccctt
5  61  atcgagggcc gggtaaccgg gtgectggcc ctctagcccc gctaccggac tacatgtcgg
121 aggagaagct gcaggagaaa gctcgaaaat ggcagcaatt gcaggccaag cgctatgcag
181 aaaagcggaa gtttgggttt gtggatgcc agaaggaaga catgcccca gaacatgca
241 gggagatcat tcgagccat ggagacatga ccaacaggaa gtcccgccat gacaaaaggg
301 tttactggg tgcctaaag tacatgcccc acgcagtct caaacctctg gagaacatgc
10 361 ctatgccttg ggagcagatt cgggatgtgc ccgtgctgta ccacatact ggagccattt
421 ccttcgtcaa tgagattccc tgggtcattg aacctgtcta catctcccag tgggggtcaa
481 tgtgattat gatgcgccga gaaaaagag ataggagca ttcaagaga atgcgttttc
541 cccctttga tgatgaggag ccgccccttg actatgctga caacatccta aatgttgagc
601 cactggaggc cattcageta gagctggacc ctgaggagga cgcctctgtg ttggactggt
15 661 tctatgacca ccagccgttg agggacagca ggaagtatgt aaatggctcc acttaccagc
721 gctgagcagtt cacactacct atgatgtcaa ctctctaccg cctggctaac cagctcctga
781 cagacttggg ggatgacaac tactcttacc tgtttgatt gaaggcttc ttacgtcca
841 aggcactcaa tatggccatt cctggaggcc ccaaatltga acotcttgtt cgagacatca
901 acctacagga tgaagactgg aatgaattca atgatattaa caagattatc atccggcagc
20 961 ctatccggac tgagtacaag attgcttttc cttacttgta caacaattt ccacaccatg
1021 tccacctcac ctggtaccat actcccaatg ttgtattcat caaaactgaa gatcctgact
1081 tgcagcttt ctactttgac cctttgatca acccaatctc ccataggcac tcagtcaaga
1141 gccaggaacc attgccgat gatgatgagg aatttgagct ccggagttt gtggagccct
1201 tctgaagga cacaccctc tatacagaca atacagccaa tggcattgcc ctgctctggg
25 1261 ccccgcgccc cttcaacctc cgtctggtc gcaccctgc ggccctggac atacccttg
1321 tcaagaactg gtatcgggag cattgtcctg cggggcagcc tgtgaaagt agggctctct
1381 accagaagct gcttaagtac tatgtctga atgcccgtaa gcctggccc ctaaggctc
1441 aaaagaagag gtatttggtc cgtccttca aagccaccaa attctttcag tccaaaagc
1501 tggactgggt ggagggttgg ctccaggttt gccgccaggg ctacaacatg ctcaacctc
30 1561 tcattcaccg caaaaacctc aactacctgc acctggacta caactcaac ctcaagcctg
1621 tgaaaacgct caccaccaag gaaagaaga aatctcgttt tgggaatgct ttccactgt

```

WO 02/070563

2/5

PCT/GB02/00986

1681 gtcgggaagt tctgctttg actaagctgg tggtagatg tcaagtgcag tctcgctgg
1741 gcaatgtgga tgccttccag ctggcagatg gattgcagta tatattggcc catgttgggc
1801 agttgacggg catgtatcga tacaataca agctgatgag acagattcgc gtgtgcaagg
1861 acctgaagca tctcatctat tatcgtttca acacaggccc tgtaggaag ggtcctggct
5 1921 gtgcttctg ggcctccggg tggcagctc ggccttttt catgcctggc attaccocct
1981 tattagacgc atggcttggc aacctcctgg cccggcagtt tgaagtcga cactcaaagg
2041 gggtagcaaa gacagtaaca aagcagcag tggagtcaca tttgacctt gagctgcggg
2101 cagctgtgat gcatgatatt ctggacatga tgcctgaggg gatcaagcag aacaaggccc
2161 ggacaatcct gcagcacctc agtgaagcct ggcctgctg gaaagccaac attcctggg
10 2221 aggtccctgg gctgcccagc cccatagaga atatgatcct tcgatacgtg aaggccaagg
2281 ctgactggtg gaccnaccct gccactaca accgagaacg gatccgcca ggggcccctg
2341 tggacaagac tgtttgtaaa aagaatctgg gccgcctcac ccgctctat ctgaagcag
2401 aacaggagcg gcagcacaac tacctgaagg acgggcctta catcacagcg gaggaaacag
2461 tggcagtata taccaccaca gtgcattggt tggaaaagcc caggtttca cccatccat
15 2521 tccccccact ctectataag catgacacca agttgctcat cttggcattg gagcggctca
2581 aggaagctta tagtgtgaag tctcggttga accagtcctca gagggaaggag ctaggtctga
2641 tcgacagggc ctacgataac ctccacgagg cgcctgccc gataaagcgt cacctcctca
2701 cacagagagc ctccaagag gtgggcattg agttcatgga tctgtatagc cacctcgttc
2761 cagtatatga tgttgagccc ctggagaaga taactgatgc ttacctggac cagtacctgt
20 2821 ggtatgaagc cgcacaagcc cgcctgttcc caccctggat taagcctgca gacacagaac
2881 cacctccact gcttgtttac aagtgggtgc aaggcatcaa taacctgcag gacgtgtggg
2941 agacagtgta aggcagctgc aatgtcatgc tggaaatccc ctttgagaag atgtatgaga
3001 agatcgactt gactctgctc aacaggctcg tgcgcctcat cgtggaccac aacatagccc
3061 actacatgac agccaagaac aacctgctca tcaactataa ggacatgaac catacgaatt
25 3121 caatgggat catcagagcc ctgcagtttg cctcattcat agtgcagtat tatggcctgg
3181 tgatggattt gcttgtattg ggaatgcacc gggccagtga gatggctggg cccctcaga
3241 tgcctaatga ctttctcagt ttcnaggaca tagccactga ggcgcccac cccatccgtc
3301 tcttctgcag atacattgat cgcatecata ttttttccag gttcacagca gatgagcctc
3361 gggacctgat tcaacgttac ctgacagagc accctgcccc caataatgaa aacatcgttg
30 3421 gctataataa caagaagtgc tggccccgag atgccccgat gcgcctcatg aaacatgatg
3481 ttaacttagg ccgggcggta ttctgggaca tcaagaaccy cttgccacgy tcaagtacta

WO 02/070563

3/5

PCT/GB02/00986

3541 cagttcagtg ggagaacagc ttcgtgtctg tglacagtaa ggacaacccc aacctgctgt
3601 tcaacatgtg tggcttcgag tgcgcacccc tgctaagtg ccgcaccagc tatgaggagt
3661 tcacccacaa ggacgggggtc tggaaacctgc agaattgaggt tactaaggag cgcacagctc
3721 agtgtttcct gcgtgtggac gatgagtc aa tgcagcctt ccacaaccgc gtgcgtcaga
5 3781 ttctcatggc ctctgggtcc accacctca ccaagattgt gaataagtgg aatacagctc
3841 tcattggcct tatgacatac ttctgggagg ctgtggtgaa caccacaagc ctcttggaat
3901 tactggtgaa gtgtgagcac aaaatccaga cacgtatcaa gattggactc aactccaaga
3961 tgccaagtgc gttccccccg gttgtgttct acaccctaa ggagttgggt ggactcggca
4021 tgcctcaaat gggccatgtg ctcatcccc aatccgacct caggtggtcc aaacagacag
10 4081 atgtaggtat cacacacttt cgttcaggaa tgagccatga agaagaccag ctcaattcca
4141 acttgtaacc ctacatacag ccatgggaga gcgagttcat tgattctcag cgggtctggg
4201 ctgagtaact actcaagaga caagaggcca ttgctcagaa cagacgcctg actttagaag
4261 acctagaaga ttcatgggat cgtggcattc ctogaatcaa tacctcttc cagaaggacc
4321 ggcacacact ggcttatgat aagggtctgc gtgtcagaac tgactttaag cagtatcagg
15 4381 ttttgaagca gaatccgttc tgggtggcac accagcggca tgatgggaaag ctctggaacc
4441 tgaacaacta ccgtacagac atgatccagg ccctgggagg tgtggaaggc attctggaac
4501 acaactctt taaggcact tacttcccta cctgggagg gctttctgg gagaaggcca
4561 gtggcttga ggaatctatg aagtggaaaga agctaactaa tgctcagcga tcaggaactga
4621 accagattcc caatctaga ttcacctctt ggtgtcccc gaccaliaat cgagccaatg
20 4681 tatatgtagg ctttcagggtg cagctagacc tgaagggtat ctctatgac ggcaagatcc
4741 caacgtgaa gatctctctc atccagatct tccgagctca cttgtggcag aagatccatg
4801 agagcattgt tatggactta tgcaggtgt ttgaccagga acttgatgca ctggaattg
4861 agacagtaca aaaggagaca atccatcccc gaaagtcata taagatgaac tcttctgtg
4921 cagatctct gctctttgcc tctataagt ggaatgtctc ccggccctca ttgctggctg
25 4981 actccaagg tglgatggac agcaccacca cccagaaata ctggattgac atccagttgc
5041 gctgggggga ctatgattcc cacgacatg agcgtacgc ccgggccaag ttctctgact
5101 acaccacoga caacatgagt atctaccctt cgcaccacag tgtactcctc gccattgacc
5161 tggcctataa cttgcacagt gccatggaa actgggtccc aggcagcaag cctctcatac
5221 aacaggccat ggccaagatc atgaaggcaa accctgcct gtatgtgta cgtgaacgga
30 5281 tccgcaagg gctacagctc tatteatctg aaccactga gccttatttg tcttctcaga
5341 actatggtga gctctctccc aaccagatta tctggttgt ggatgacacc aacgtctaca

WO 02/070563

4/5

PCT/GB02/00986

5401 gagtgactat tcacaagacc ttgaaggga acttgacaac caagcccatc aacggagcca
5461 tcttcattct caaccacgc acagggcagc tgttctctaa gataatccac acgtccgtgt
5521 gggcgggaca gaagcgtttg gggcagttgg ctaagtggaa gacagctgag gaggtggccg
5581 ccttgatccg atctctgcct gtggagggagc agcccaagca gatcattgtc accaggaagg
5 5641 acatgctgga cccactggag gtgcacttac tggacttccc caatattgtc atcaaaggat
5701 cggagctcca actcccttcc caggcgtgtc tcaagtgga aaaattcggg gatctcatcc
5761 ttaaagccac tgagcccccag atggttctct tcaacctcta tgacgactgg ctcaagacta
5821 ttctatctta cagggccttc tccgtctca tctgatctc gcgtgcctca catgtgaaca
5881 acgatcgggc aaaagtgtac ctgaagccag acaagactac tattacagaa ccacaccaca
10 5941 tctggccccc tctgactgac gaagaatgga tcaaggtcga ggtgcagctc aaggatctga
6001 tcttgctgca ctacggcaag aaaaacaatg tgaactgtgc atcactgaca caatcagaaa
6061 ttcgagacat catcctgggt atggagatct cggcacctgc acagcagcg cagcagatcg
6121 ctgagatcga gaagcagacc aaggaacaat cgcagctgac gcaacacag actcgcactg
6181 tcaacaagca tggcgtatg atcatcacct ccaccaccag caactatgag acccagactt
15 6241 tctcatccaa gactgagtg aggtcaggg ccactctctg tgccaacctg cacctaagga
6301 ccaatcacat ctatgttca tctgacgaca tcaaggagac tggctacacc tacatcttc
6361 ccaagaatgt gcttaagaag ttcactctca tatctgacct tggggcccaa attgagatg
6421 acctatatgg ggtgagccca ccagataacc cccagtgaa ggagatccgc tgcattgtga
6481 tggtagccga gtggggcact caccagaccg tgcacctgcc tggccagctg cccagcatg
20 6541 agtacctcaa ggagatggaa ccttaggtt ggcaccacac tcagcccaat gagtccccc
6601 agttatcacc ccaggatgtc accaccatg ccaagatcat ggctgacaac ccatcttggg
6661 atggcgagaa gaccattatc atcacatgca gcttcacgcc aggtcctctg acactgacgg
6721 cctacaagct gaccctagt ggctacgaat ggggcccca gaacacagac aagggaaca
6781 accccaaggy ctacctgctc tcacactatg agaggggtca gatgctgctg tcggaccgtt
25 6841 tcttggctt cttcatggtc cctgccagt cctcgtggaa ctacaacttc atgggtgttc
6901 ggcctgacc caacatgaaa tatgactac agctggcgaa ccccaagag ttctaccag
6961 aggtgcacag gccctctcac tctctcaact ttgctctct gcaggagggg gaggtttact
7021 ctggggatcg ggagacctg tatgctgac cgtttccctg cctcctgctt cagcctccc
7081 aggcggaagc ctcagccctc ccagacaggc cgtgacatt cagcagttg gcctcttcc
30 7141 ctctgtctgt gcttgtgtg ttgacctct gatggctgt catctgaat aaaatataat
7201 aataaatttt gtataaatag g

WO 02/070563

5/5

PCT/GB02/00986

SEQ ID NO: 2 Protein sequence for BAA22563.1 (LBDG2)

```

1  magvfpyrgp  gnpvpgplap  lpdymseekl  qekarkwqq  qakryaekrk  fgfvdaqked
5  61  mppshvreii  rdhgdntnrk  frhdkrvylg  alkymphavl  kllenmpmpw  eqirdvvpvly
121 hitgaisfvn  eipwviepvy  isqwgsmwim  murrekrdrh  fkrmrpppfd  deeppldyad
181 nilnveplea  iqleldpeed  apvldwfydh  qplrdsrkyv  ngstyqrwgf  tlpmmstlyr
241 lanqltldlv  ddnyfylfdl  kafttskain  maipggpkfe  plvrdinlqd  edwnsfndin
301 kiiirpirt  eykiafpyly  nnlphhvilt  wyhtpnvffi  ktedpdlpaf  yfdplinpis
10 361 hrhsvksqep  lpdddeefel  pefvepfikd  tplytdntan  giallwaprp  fnlrsgztrr
421 aldiplvknw  yrehcpagqp  vkvrsvyqkl  lkyvlnalk  hrppkagkkr  ylfrrfkatk
481 ffgstkldgw  egwlqvcrgp  ymnlhlihr  knlnylhldy  nfnltpvktl  ttkerkksrf
541 gnafhicrev  lrllklvvdv  hvqyrlgnvd  afqladglcy  ifahvgqltg  myrkykklmr
601 qirvckdlkh  liyyrfntgp  vkgpgcgfw  aagwrvwlff  mrgitpller  wlglllarqf
15 661 egrhskgvak  tvtkqrvesh  fdlelraavm  hdildmpep  ikqnkartil  qhlseawrcw
721 kanipkwvpg  lptpiemmil  ryvkakadw  tntahynrer  irrgatvdkt  vckknlgrlt
781 rlylkaequer  ghnylkdgpy  itaetvavy  tttvhwlesr  rfsppippl  sykhtkllii
841 lalerlkeay  svksrlnqsg  reelglieqa  ydnlhealsr  akrhlrtgra  fkevgiefnd
901 lyshlvpyvd  veplekitda  yldqylwea  dkrrlfpwwi  kpadtepppl  lvykwcqgin
20 961 niqdvwtsev  gecnvmlsrs  fekmyekidl  tllnrlvrli  vchniadytc  aknnvvinyk
1021 dmnhtnsygi  irglqfasfi  vqyglvmdl  lvlglrhase  magppqmpnd  flsfqdiate
1081 aahpirlfcr  yidrihiffr  ftadeardli  qryltehpdp  nnenivgynn  kkcwprdamr
1141 rlmkhdvnlq  ravfwdiknr  lprsvttvqw  ensfsvvysk  dnpnllfmc  gfecrilpkc
1201 rtsyeeftkh  dgvmnlqnev  tkertaqefl  rvddesmqrf  hnrvrqilma  sgsttftkiv
25 1261 nkwtaligl  mtyfreavvn  tqelldllvk  cehkiqtrik  iglnskmpsr  fpgvvfytpk
1321 eigglgmlsm  ghvliqgsdl  rwsqktdvgi  thfrsgmshe  edqlipnlyx  yiqpwesefi
1381 dsqrwaeys  lkrqeaiaqn  rrltledied  swdrqiprin  tlfqkdrhtl  aydkgwrvrt
1441 dfkqyqvikq  nplfwthqrh  dgklwlnny  rtdmialgg  vegilehtlf  kgtyfptweg
1501 lfwekasgfe  esmkwkklt  n  aqrsqnlqip  nrrftlwsp  tinranvyfp  fqvqldltgi
30 1561 fmhngkiptk  isliqifrah  lwqkiesiv  mdlcqvfdqe  ldaleietvq  ketihprkxy
1621 kmsscadil  lfasykwnvs  rpslladskd  vmdstttqky  widiqlrwdg  ydshdierya
1681 rakfldytt  nmsiypsptg  vliaidlayn  lhsaygnwfp  gskpliggam  akimkanpal
1741 yvlrezirkq  lqlyssepte  pylssqnyge  lfsnqiawfv  ddtvnyrvti  hktfegmltt
1801 kpingaifif  nprtgqlflk  iihstsvagg  krlgqlakwk  taeevaalir  slpveeqpkq
35 1861 iivtrkdmld  plevhllfdp  nivikgselq  lpfqacikve  kfgdlikat  epgmwlfnlly
1921 ddwlktissy  tafsrllili  rahvmndra  kvilkpdkt  itephhiwpt  ltdeewikve
1981 vqlkdilad  ygkknvnva  sltqseirdi  ilgmeisaps  qqrqiaei  kqtkeqsqit
2041 atqtrtvnkh  gdeitstts  nyetqtfsk  tewrvraisa  anhlrtnhi  yvssddiket
2101 gytvilpknv  lkkficusdl  raqiagylyg  vsppdpqv  eircivmpq  wgtqgtvhlp
40 2161 gqllpqheyk  emepglwiht  qnespqlsp  qdvtthakin  adnpswdgek  tiiitcsftp
2221 gsctltaykl  tpsgyewgrq  ntdkgnpkg  ylpshyervq  mllsdrflgf  fmwpaqsswn
2281 ynfngvrhdp  nmkyelqan  pkefyhevhr  pshfinfall  qegevsadr  edlya

```

【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
12 September 2002 (12.09.2002)

PCT

(10) International Publication Number
WO 02/070563 A3

- (51) International Patent Classification: C07K 14/705, 14/72, A61K 38/17
- (52) International Application Number: PCT/GB02/00986
- (22) International Filing Date: 5 March 2002 (05.03.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 0105402.2 5 March 2001 (05.03.2001) GB
- (71) Applicant (for all designated States except US): INPHARMATICA LIMITED [GB/GB]; 60 Charlotte Street, London W1T 2NU (GB).
- (72) Inventors; and
(75) Inventors/Applicants (for US only): FAGAN, Richard, Joseph [GB/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB). PHELPS, Christopher, Benjamin [GB/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB). PHILLIPS, Tom [GB/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB). PIERRON, Valerie, Nathalie [FR/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB). ALLEN, Kathryn, Elizabeth [GB/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB). ALLEN, Janet, Marjorie [GB/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB). POTTER, Sarah, Jane [GB/GB]; 39 Kneller Road, Brockley, London SE14 2AR (GB).
- (74) Agents: MERCER, Christopher, Paul et al.; Carpmals & Ramsford, 43 Bloomsbury Square, London WC1A 2RA (GB).
- (81) Designated States (national): AI, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GI, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LI, LU, LV, MA, MD, MG, MK, MN, MW, MX, MY, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TH, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.
- (84) Designated States (regional): ARIPO patent (GI, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LI, MC, NL, PT, SE, TR), OAPI patent (BF, BI, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:
with international search report
before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments
- (88) Date of publication of the international search report:
28 November 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 02/070563 A3

(54) Title: NUCLEAR HORMONE RECEPTOR LIGAND BINDING DOMAIN

(57) Abstract: This invention relates to a novel protein, termed BAA22563.1, herein identified as a Nuclear Hormone Receptor Ligand Binding Domain and to the use of this protein and nucleic acid sequence from the encoding gene in the diagnosis, prevention and treatment of disease.

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International Application No. PCT/GB 02/00986
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C07K14/705 C07K14/72 A61K38/17		
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 C07K		
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) BIOSIS, EPO-Internal, EMBL, SEQUENCE SEARCH, WPI Data, PAJ, MEDLINE		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE "Online! 3 October 1997 (1997-10-03) SHIMADA Y ET AL: "Homo sapiens mRNA for PRP8 protein, complete cds." retrieved from EMBL Database accession no. AB007510 XP002214548 cited in the application Sequences with 100% identity with SEQ ID Nos:1 and 2 over 7221 nucleotides and 2335 amino acids respectively the whole document --- -/--	1-13
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex		
* Special categories of cited documents: *A* document defining the general state of the art which is not considered to be of particular relevance *E* earlier document but published on or after the international filing date *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) *O* document referring to an oral disclosure, use, exhibition or other means *P* document published prior to the international filing date but later than the priority date claimed *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art *Z* document member of the same patent family		
Date of the actual completion of the international search	Date of mailing of the international search report	
24 September 2002	14/10/2002	
Name and mailing address of the ISA European Patent Office, P.O. Box 5516 Patentstr 2 NL - 2280 LV Rijswijk Tel: (+31-70) 340-2000, Tx: 31 651 epo nl Fax: (+31-70) 340-3016	Authorized officer: Nichogiannopoulou, A	

Form PCT/ISA210 (second sheet) July 1992

INTERNATIONAL SEARCH REPORT

Int'l Application No
PCT/GB 02/00986

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	LUO HONGBO R ET AL: "The human Prp8 protein is a component of both U2- and U12-dependent spliceosomes." RNA (NEW YORK), vol. 5, no. 7, July 1999 (1999-07), pages 893-908, XP001106096 ISSN: 1355-8382 figure 1	1-13
X	-& DATABASE 'Online! 29 September 1998 (1998-09-29) LUO HR ET AL: "Homo sapiens splicing factor Prp8 mRNA, complete cds." retrieved from EMBL Database accession no. AF092565 XP002214549 Sequences with 99.6% identity with SEQ ID No:1 and 2 over 7219 nucleotides and 2335 amino acids respectively the whole document	1-13
X	WO 99 51727 A (SCHMITT ARMIN ;SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN BERND) 14 October 1999 (1999-10-14) Sequence with 99.8% identity over nucleotides 4470-7720 of SEQ ID No:1 page 191 -page 192	1-13, 15, 19-45
A	SPANJAARD R A ET AL: "LIGAND-BINDING AND HETERODIMERIZATION ACTIVITIES OF A CONSERVED REGION IN THE LIGAND-BINDING DOMAIN OF THE THYROID HORMONE RECEPTOR" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES, vol. 88, no. 19, 1991, pages 8587-8591, XP002214547 1991 ISSN: 0027-8424 figure 1	1

Form: PCT/ISA/210 (continuation of second sheet) (July 1992)

International Application No. PCT/GB 02 00986

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

Although claims 20, 22-28 are directed to a diagnostic method practised on the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
Although claims 31, 35, 38 and 45 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.

Continuation of Box I.2

Claims Nos.: 14, 16-18

Present claims 14 and 16-18 relate to compounds defined by reference to a desirable characteristic or property, namely their ability to bind to a claimed polypeptide or their ability to increase or decrease the level of expression or activity of a claimed polypeptide.

The claims cover all compounds having this characteristic or property, whereas the application provides no support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for any such compound. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the compound by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, no search has been carried out for these claims.

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	International application No. PCT/GB 02/00986
Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)	
This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:	
1. <input checked="" type="checkbox"/> Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely: see FURTHER INFORMATION sheet PCT/ISA/210	
2. <input checked="" type="checkbox"/> Claims Nos.: 14, 16-18 because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically: see FURTHER INFORMATION sheet PCT/ISA/210	
3. <input type="checkbox"/> Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).	
Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)	
This International Searching Authority found multiple inventions in this international application, as follows:	
1. <input type="checkbox"/> As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.	
2. <input type="checkbox"/> As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.	
3. <input type="checkbox"/> As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:	
4. <input 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.:	
Remark on Protest	<input type="checkbox"/> The additional search fees were accompanied by the applicant's protest. <input type="checkbox"/> No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International Application No
PCT/GB 02/00986

Patent document cited in search report	Publication data	Patent family member(s)	Publication date
WO 9951727	A	14-10-1999	DE 19816395 A1 07-10-1999
			WO 9951727 A2 14-10-1999
			EP 1068231 A2 17-01-2001
			JP 2002510486 T 09-04-2002

フロントページの続き

(51)Int.Cl. ⁷	F I		テーマコード(参考)
A 6 1 K 39/395	A 6 1 K 39/395	N	4 C 0 8 5
A 6 1 K 45/00	A 6 1 K 45/00		4 C 0 8 6
A 6 1 K 48/00	A 6 1 K 48/00		4 H 0 4 5
A 6 1 P 1/00	A 6 1 P 1/00		
A 6 1 P 1/18	A 6 1 P 1/18		
A 6 1 P 3/04	A 6 1 P 3/04		
A 6 1 P 3/06	A 6 1 P 3/06		
A 6 1 P 3/10	A 6 1 P 3/10		
A 6 1 P 3/14	A 6 1 P 3/14		
A 6 1 P 5/00	A 6 1 P 5/00		
A 6 1 P 7/00	A 6 1 P 7/00		
A 6 1 P 7/02	A 6 1 P 7/02		
A 6 1 P 9/00	A 6 1 P 9/00		
A 6 1 P 9/06	A 6 1 P 9/06		
A 6 1 P 9/08	A 6 1 P 9/08		
A 6 1 P 9/10	A 6 1 P 9/10		
A 6 1 P 9/12	A 6 1 P 9/12		
A 6 1 P 11/00	A 6 1 P 11/00		
A 6 1 P 13/12	A 6 1 P 13/12		
A 6 1 P 15/00	A 6 1 P 15/00		
A 6 1 P 17/00	A 6 1 P 17/00		
A 6 1 P 17/06	A 6 1 P 17/06		
A 6 1 P 17/10	A 6 1 P 17/10		
A 6 1 P 19/02	A 6 1 P 19/02		
A 6 1 P 19/10	A 6 1 P 19/10		
A 6 1 P 21/00	A 6 1 P 21/00		
A 6 1 P 25/00	A 6 1 P 25/00		
A 6 1 P 25/04	A 6 1 P 25/04		
A 6 1 P 25/22	A 6 1 P 25/22		
A 6 1 P 25/24	A 6 1 P 25/24		
A 6 1 P 25/28	A 6 1 P 25/28		
A 6 1 P 31/04	A 6 1 P 31/04		
A 6 1 P 31/10	A 6 1 P 31/10		
A 6 1 P 31/12	A 6 1 P 31/12		
A 6 1 P 31/18	A 6 1 P 31/18		
A 6 1 P 33/00	A 6 1 P 33/00		
A 6 1 P 35/00	A 6 1 P 35/00		
A 6 1 P 35/02	A 6 1 P 35/02		
A 6 1 P 37/00	A 6 1 P 37/00		
A 6 1 P 43/00	A 6 1 P 43/00	1 1 1	
C 0 7 K 14/705	C 0 7 K 14/705		
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 5/10	C 1 2 Q 1/68	A	
C 1 2 Q 1/68	G 0 1 N 33/15	Z	
G 0 1 N 33/15	G 0 1 N 33/50	Z	

G 0 1 N 33/50	G 0 1 N 33/53	D
G 0 1 N 33/53	G 0 1 N 33/53	M
G 0 1 N 37/00	G 0 1 N 37/00	1 0 2
	C 1 2 N 5/00	A
	A 6 1 K 37/02	

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

(74)代理人 100084009

弁理士 小川 信夫

(74)代理人 100082821

弁理士 村社 厚夫

(74)代理人 100086771

弁理士 西島 孝喜

(74)代理人 100084663

弁理士 箱田 篤

(72)発明者 ファガン リチャード ジョセフ

イギリス ロンドン ダブリュ1ティー 2エヌユー チャーロット ストリート 60 インフ
ァーマティカ リミテッド

(72)発明者 フェルプス クリストファー ベンジャミン

イギリス ロンドン ダブリュ1ティー 2エヌユー チャーロット ストリート 60 インフ
ァーマティカ リミテッド

(72)発明者 フィリップス トム

イギリス ロンドン ダブリュ1ティー 2エヌユー チャーロット ストリート 60 インフ
ァーマティカ リミテッド

(72)発明者 ピアロン ヴァレリー ナサリー

イギリス ロンドン ダブリュ1ティー 2エヌユー チャーロット ストリート 60 インフ
ァーマティカ リミテッド

(72)発明者 アレン キャスリン エリザベス

イギリス ロンドン ダブリュ1ティー 2エヌユー チャーロット ストリート 60 インフ
ァーマティカ リミテッド

(72)発明者 アレン ジャネット マジョリー

イギリス ロンドン ダブリュ1ティー 2エヌユー チャーロット ストリート 60 インフ
ァーマティカ リミテッド

(72)発明者 ポッター サラ ジェイン

イギリス ロンドン エスイー4 2エイアール ブロックリー ネラー ロード 39

Fターム(参考) 2G045 AA34 DA13 DA36 FB02 FB03

4B024 AA01 AA12 AA15 CA02 CA09 HA14 HA17

4B063 QA19 QQ08 QQ43 QR55 QR62 QS34

4B065 AB01 BA02 CA45 CA46

4C084 AA01 AA02 AA07 AA13 AA14 AA17 BA01 BA08 BA22 BA23

BA35 CA18 NA01 NA13 NA14 NA15 ZA022 ZA082 ZA122 ZA162

ZA362 ZA392 ZA422 ZA452 ZA512 ZA592 ZA662 ZA702 ZA812 ZA892

ZA942 ZA962 ZA972 ZB072 ZB112 ZB132 ZB262 ZB272 ZB332 ZB352

ZB392 ZC012 ZC212 ZC332 ZC352 ZC552

4C085 AA03 AA13 AA14 BB11 BB23 BB41 BB43 CC02 CC04 CC05

	CC21	CC23	CC32	EE01						
4C086	AA01	AA02	AA03	EA16	MA01	MA04	NA01	NA13	NA14	NA15
	ZA02	ZA08	ZA12	ZA16	ZA36	ZA39	ZA42	ZA45	ZA51	ZA59
	ZA66	ZA70	ZA81	ZA89	ZA94	ZA96	ZA97	ZB07	ZB11	ZB13
	ZB26	ZB27	ZB33	ZB35	ZB39	ZC01	ZC21	ZC33	ZC35	ZC55
4H045	AA10	AA30	BA10	CA40	DA50	EA21	EA23	EA24	EA26	EA28
	EA29	EA31	EA51	EA52	EA54					

专利名称(译)	核激素受体配体结合域		
公开(公告)号	JP2005500015A	公开(公告)日	2005-01-06
申请号	JP2002570600	申请日	2002-03-05
[标]申请(专利权)人(译)	Inpharmatica公司有限公司		
申请(专利权)人(译)	Inpharmatica公司有限公司		
[标]发明人	ファガンリチャードジョセフ フェルプスクリストファーベンジャミン フィリップストム ピアロンヴァレリーナサリー アレンキャスリンエリザベス アレンジャネットマジョリー ポッターサラジェイン		
发明人	ファガン リチャード ジョセフ フェルプス クリストファー ベンジャミン フィリップストム ピアロン ヴアレリー ナサリー アレン キャスリン エリザベス アレン ジャネット マジョリー ポッター サラ ジェイン		
IPC分类号	A01K67/027 A61K31/7088 A61K35/76 A61K38/00 A61K39/00 A61K39/395 A61K45/00 A61K48/00 A61P1/00 A61P1/18 A61P3/00 A61P3/04 A61P3/06 A61P3/10 A61P3/14 A61P5/00 A61P5/14 A61P7 /00 A61P7/02 A61P9/00 A61P9/06 A61P9/08 A61P9/10 A61P9/12 A61P11/00 A61P11/06 A61P13/12 A61P15/00 A61P17/00 A61P17/02 A61P17/04 A61P17/06 A61P17/10 A61P19/00 A61P19/02 A61P19 /10 A61P21/00 A61P25/00 A61P25/02 A61P25/04 A61P25/22 A61P25/24 A61P25/28 A61P29/00 A61P31/00 A61P31/04 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P35/00 A61P35/02 A61P37 /00 A61P37/02 A61P37/06 A61P37/08 A61P43/00 C07K14/705 C07K14/72 C07K16/28 C12M1/00 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N15/09 C12Q1/68 G01N33/15 G01N33/50 G01N33/53 G01N37/00		
CPC分类号	A61K38/00 A61P1/00 A61P1/18 A61P3/00 A61P3/14 A61P11/00 A61P11/06 A61P13/12 A61P15/00 A61P17/00 A61P17/02 A61P17/04 A61P17/06 A61P17/10 A61P19/00 A61P19/02 A61P19/10 A61P21 /00 A61P25/00 A61P25/02 A61P25/04 A61P25/22 A61P25/24 A61P25/28 A61P29/00 A61P31/00 A61P31/04 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P35/00 A61P35/02 C07K14/72		
FI分类号	C12N15/00.ZNA.A A01K67/027 A61K31/7088 A61K39/00.H A61K39/395.D A61K39/395.N A61K45/00 A61K48/00 A61P1/00 A61P1/18 A61P3/04 A61P3/06 A61P3/10 A61P3/14 A61P5/00 A61P7/00 A61P7 /02 A61P9/00 A61P9/06 A61P9/08 A61P9/10 A61P9/12 A61P11/00 A61P13/12 A61P15/00 A61P17/00 A61P17/06 A61P17/10 A61P19/02 A61P19/10 A61P21/00 A61P25/00 A61P25/04 A61P25/22 A61P25 /24 A61P25/28 A61P31/04 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P35/00 A61P35/02 A61P37/00 A61P43/00.111 C07K14/705 C12N1/15 C12N1/19 C12N1/21 C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.D G01N33/53.M G01N37/00.102 C12N5/00.A A61K37/02		
F-TERM分类号	2G045/AA34 2G045/DA13 2G045/DA36 2G045/FB02 2G045/FB03 4B024/AA01 4B024/AA12 4B024 /AA15 4B024/CA02 4B024/CA09 4B024/HA14 4B024/HA17 4B063/QA19 4B063/QQ08 4B063/QQ43 4B063/QR55 4B063/QR62 4B063/QS34 4B065/AB01 4B065/BA02 4B065/CA45 4B065/CA46 4C084 /AA01 4C084/AA02 4C084/AA07 4C084/AA13 4C084/AA14 4C084/AA17 4C084/BA01 4C084/BA08 4C084/BA22 4C084/BA23 4C084/BA35 4C084/CA18 4C084/NA01 4C084/NA13 4C084/NA14 4C084 /NA15 4C084/ZA022 4C084/ZA082 4C084/ZA122 4C084/ZA162 4C084/ZA362 4C084/ZA392 4C084 /ZA422 4C084/ZA452 4C084/ZA512 4C084/ZA592 4C084/ZA662 4C084/ZA702 4C084/ZA812 4C084		

/ZA892 4C084/ZA942 4C084/ZA962 4C084/ZA972 4C084/ZB072 4C084/ZB112 4C084/ZB132 4C084/ZB262 4C084/ZB272 4C084/ZB332 4C084/ZB352 4C084/ZB392 4C084/ZC012 4C084/ZC212 4C084/ZC332 4C084/ZC352 4C084/ZC552 4C085/AA03 4C085/AA13 4C085/AA14 4C085/BB11 4C085/BB23 4C085/BB41 4C085/BB43 4C085/CC02 4C085/CC04 4C085/CC05 4C085/CC21 4C085/CC23 4C085/CC32 4C085/EE01 4C086/AA01 4C086/AA02 4C086/AA03 4C086/EA16 4C086/MA01 4C086/MA04 4C086/NA01 4C086/NA13 4C086/NA14 4C086/NA15 4C086/ZA02 4C086/ZA08 4C086/ZA12 4C086/ZA16 4C086/ZA36 4C086/ZA39 4C086/ZA42 4C086/ZA45 4C086/ZA51 4C086/ZA59 4C086/ZA66 4C086/ZA70 4C086/ZA81 4C086/ZA89 4C086/ZA94 4C086/ZA96 4C086/ZA97 4C086/ZB07 4C086/ZB11 4C086/ZB13 4C086/ZB26 4C086/ZB27 4C086/ZB33 4C086/ZB35 4C086/ZB39 4C086/ZC01 4C086/ZC21 4C086/ZC33 4C086/ZC35 4C086/ZC55 4H045/AA10 4H045/AA30 4H045/BA10 4H045/CA40 4H045/DA50 4H045/EA21 4H045/EA23 4H045/EA24 4H045/EA26 4H045/EA28 4H045/EA29 4H045/EA31 4H045/EA51 4H045/EA52 4H045/EA54

代理人(译)	中村稔 小川伸男 西岛隆义
優先権	2001005402 2001-03-05 GB
其他公开文献	JP2005500015A5
外部链接	Espacenet

摘要(译)

本发明涉及一种新型蛋白质，称为LBDG3，在本文中被鉴定为核激素受体配体结合域，并且涉及该蛋白质和来自编码基因的核酸序列在疾病的诊断，预防和治疗中的用途。

(P2005-500015A)
(43) 公表日 平成17年1月6日 (2005.1.6)

(5) Int. Cl. ⁷	F I	テーマコード (参考)
C12N 15/09	C12N 15/00 ZNAA	2G045
A01K 67/027	A01K 67/027	4B024
A61K 31/7088	A61K 31/7088	4B063
A61K 39/00	A61K 39/00 H	4B065
A61K 39/00	A61K 39/395 D	4C084
審査請求 未請求 予備審査請求 未請求 (全 163 頁) 最終頁に続く		
(21) 出願番号	特願2002-570600 (P2002-570600)	(71) 出願人 502334043
(86) (2) 出願日	平成14年3月5日 (2002.3.5)	インファーマティカ リミテッド
(85) 翻訳文提出日	平成15年9月5日 (2003.9.5)	イギリス ロンドン ダブリュー1ティー
(86) 国際出願番号	PCT/G82002/000986	2 エヌユー チャーロット ストリー
(87) 国際公開番号	W02002/070563	60
(87) 国際公開日	平成14年9月12日 (2002.9.12)	(74) 代理人 100058959
(31) 優先権主張番号	0105402.2	弁理士 中村 稔
(32) 優先日	平成13年3月5日 (2001.3.5)	(74) 代理人 100067013
(33) 優先権主張国	英国 (GB)	弁理士 大塚 文昭
		(74) 代理人 100082005
		弁理士 熊倉 慎男
		(74) 代理人 100065189
		弁理士 穴戸 善一
		(74) 代理人 100074228
		弁理士 今城 俊夫
最終頁に続く		

(F4) 「註明の名称」 精肉ホルモンレチナブターリガンド結合ドメイン