

(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2005-500011

(P2005-500011A)

(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 |
| A O 1 K 67/027 | A O 1 K 67/027 | 2 G O 4 5 |
| A 6 1 K 38/00 | A 6 1 K 45/00 | 4 B O 2 4 |
| A 6 1 K 45/00 | A 6 1 K 48/00 | 4 B O 6 3 |
| A 6 1 K 48/00 | A 6 1 P 1/04 | 4 B O 6 5 |
| | | 4 C O 8 4 |
| 審査請求 未請求 予備審査請求 未請求 (全 146 頁) 最終頁に続く | | |

| | | | |
|---------------|------------------------------|----------|---------------------|
| (21) 出願番号 | 特願2002-569877 (P2002-569877) | (71) 出願人 | 502334043 |
| (86) (22) 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/GB2002/000941 | | 2 エヌユー チャーロット ストリート |
| (87) 国際公開番号 | W02002/070558 | | 6 0 |
| (87) 国際公開日 | 平成14年9月12日 (2002.9.12) | (74) 代理人 | 100059959 |
| (31) 優先権主張番号 | 0105401.4 | | 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 禎男 |
| | | (74) 代理人 | 100065189 |
| | | | 弁理士 宍戸 嘉一 |
| | | (74) 代理人 | 100074228 |
| | | | 弁理士 今城 俊夫 |
| 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、新規なタンパク質 (BAA07892.2と称され、本明細書では核ホルモンレセプターリガンド結合ドメインと特定された)、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；

(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、前記のフラグメントであるポリペプチド；または

(iii) (i) または (ii) の機能的等価物であるポリペプチド。

【請求項 2】

LBDS1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 (ii) 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 1507 と 1742 を含むと定義され、前記フラグメントが “ LBDモチーフ ” 残基 VAL1573、LEU1574、ASP1578、LYS1579、ALA1582 および MET1583、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。

【請求項 3】

請求項 1 の (iii) 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 VAL1573、LEU1574、ASP1578、LYS1579、ALA1582 および MET1583、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。

【請求項 4】

前記機能的等価物が、LBDS1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% を越える配列同一性を有する、請求項 1 - 4 のいずれかの項に記載のフラグメントまたは機能的等価物。

【請求項 6】

配列番号：2 のいずれかに記載の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 のいずれかの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。

【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 4548 から 5255 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

10

20

30

40

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抑制するリガンド。

【請求項 1 5】

抗体である請求項 1 4 に記載のリガンド。

10

【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。

20

【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 のいずれかの項に記載のポリペプチド、請求項 8 - 1 1 のいずれかの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 のいずれかの項に記載の化合物。

【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。

30

【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 または請求項 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 または請求項 2 1 に記載の方法。

【請求項 2 3】

以下の工程を含む

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前記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；

40

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が疾患を示唆する、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

50

と前記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) 前記サンプル化された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の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。

10

【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 のいずれかの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し、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。

20

【請求項 27】

前記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するハイブリッドであり、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非ハイブリダイズ部分の有無を検出することによって患者で変異の有無を検出する請求項 25 または 26 のいずれかの方法。

【請求項 28】

前記疾患が以下である請求項 20 - 27 のいずれかの項に記載の方法：

細胞増殖性疾患（新形成、癌および骨髄増殖症候群および血管形成を含む）、代謝性障害（肥満、骨粗しょう症、非インスリン依存性糖尿病、脂質代謝異常、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、高脂血症を含む）、心脈管系疾患（高血圧、アテローム性硬化症、高インスリン血症、心不整脈を含む）、腎疾患（糸球体腎炎、腎血管性高血圧を含む）、皮膚疾患（皮膚疾患、乾癬、アクネ、加齢、湿疹、創傷治癒を含む）、炎症（炎症性大腸疾患、気道の炎症、気腫、喘息、免疫不全、自己免疫疾患、多発性硬化症を含む）、神経系疾患（不安、抑うつ、神経変性疾患、アルツハイマー症、パーキンソン病、脳損傷、卒中および痛みを含む）、感染（ウイルス感染を含む）、並びに核内ホルモンが関係するその他の病的状態。

30

【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。

40

【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する細胞-細胞接着を行わせる方法。

【請求項 32】

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 - 11 のいずれか 1 項に記

50

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医薬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 に記載の医薬組成物：

細胞増殖性疾患（新形成、癌および骨髄増殖症候群および血管形成を含む）、代謝性障害（肥満、骨粗しょう症、非インスリン依存性糖尿病、脂質代謝異常、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、高脂血症を含む）、心血管系疾患（高血圧、アテローム性硬化症、高インスリン血症、心不整脈を含む）、腎疾患（糸球体腎炎、腎血管性高血圧を含む）、皮膚疾患（皮膚疾患、乾癬、アクネ、加齢、湿疹、創傷治癒を含む）、炎症（炎症性大腸疾患、気道の炎症、気腫、喘息、免疫不全、自己免疫疾患、多発性硬化症を含む）、神経系疾患（不安、抑うつ、神経変性疾患、アルツハイマー症、パーキンソン病、脳損傷、卒中および痛みを含む）、感染（ウイルス感染を含む）、並びに核内ホルモンが関係するその他の病的状態。

10

【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、患者で疾患を治療する方法。

20

【請求項 3 6】

天然の遺伝子または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。

30

【請求項 3 8】

患者の疾患の治療をモニターする方法であって、一定期間にわたって請求項 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け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の徴候である、前記患者の疾患の治療をモニターする方法。

40

【請求項 3 9】

疾患の治療および/または診断で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 4 0】

ストリンジェントな条件下で請求項 8 - 1 1 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

50

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 4 1】

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 4 0 のキット。

【請求項 4 2】

そのうちの少なくとも1つが請求項 8 - 1 1 のいずれか1項に記載の核酸分子である核酸分子のアレイを含むキット。

【請求項 4 3】

請求項 1 - 7 のいずれかの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 4 4】

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遺伝子導入またはノックアウト非ヒト動物。

【請求項 4 5】

請求項 4 4 に記載の非ヒト遺伝子導入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】

【0 0 0 1】

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特定された、BAA07892.2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。

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。

【0 0 0 2】

背景

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、問題のタンパク質の機能をタンパク質配列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、配列データの生産速度が、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、前記ツールは通常の特徴決定技術と急速に置き換えられつつある。実際、本発明の特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精査し、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、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研究および薬剤の発見のためのターゲットとして特定し特徴を決定す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。

【0 0 0 3】

10

20

30

40

50

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、可能な限り多くのデータを、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され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見を含む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。

このデータリソースは、配列のみをベースにしてタンパク質の正確な機能を予測することを可能にした。通常の技術を用いた場合、このような予測は、同じ機能群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関連するタンパク質と配列相同性が低いタンパク質の場合正確な予測は不可能である。

本件では、その配列が公的に利用可能なデータベースにKIAA0097タンパク質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D43948および GenBankタンパク質アクセッション番号BAA07892.2）が、核内ホルモンレセプターリガンド結合ドメイン類の新規なメンバーとして関与する。

【0004】

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関連する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つのキードメイン、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、前記レセプターがモノマー、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特定することができる。

リガンド結合ドメイン（LBD）は同系ホルモンと結合しこれと反応する。LBDと結合したリガンドは、すでに結合していた“核内レセプター補助リプレッサー”を排斥する構造的変化をひき起こす。続いて前記補助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、新しい“核内レセプター補助アクチベーター”で補充される。このリガンドによってひき起こされる補助アクチベーターによる補助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、補助リプレッサー結合および補助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重要な薬剤標的であるので、それらの特定が所望される。リガンド結合ドメインは低い（約15%）同一性しか共有しないが、非常に類似する構造を有し、したがって、構造に基づく関係付けツール(structure-based relationship tool)（例えばゲノムスレッダー）のための理想的なターゲットである。

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、さらに前記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公的ドメイン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特定される、核内レセプターとして注釈されてない新規ないずれのLBDも完全に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のリガンド結合ドメイン”を含むものと注釈付けする。

【0005】

10

20

30

40

50

【表 1】

| | |
|----------------------------|---|
| 表 1 : 核内ホルモンレセプタースーパーファミリー | |
| 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 プロジェ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 |
| 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 NGFI-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 |
| ファミリー : | DAX1 |
| サブファミリー : | DAX1 |

10

20

【0006】

下表 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。第 4 位と 8 位の間、および第 9 位と 12 位の間で見出される残基の数には変動があることに留意されたい。“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 個の配列でアラインメントを実施し、残基の保存パターンを特定することによって構築された。

30

【0007】

【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

【0008】

II. 核内ホルモンレセプターおよび疾患

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た。それらの多くは疾患プロセスにおいて役割を有する(表3参照)

【0009】

表3：核内ホルモンレセプターと疾患

【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予防に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 |

【0010】

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、新規な核内ホルモンレセプターのリガンド結合ドメインの特定は、それらタンパク質が上記で特定した疾患や他の症状において役割を有する可能性があるために強く希求される。新規な核内ホルモンレセプターのリガンド結合ドメインの特定はしたがって、疾患（特に表3で特定し

たような疾患)の治療および診断と密接な関係を有する。

【0011】

発明

本発明は、BAA07892.2タンパク質が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機能するという発見を基にしている。

BAA07892.2タンパク質の場合、前記タンパク質配列の残基1507-1742を含む領域は、ヒトレチノイン酸レセプターガンマ(PDBコード3LBD)の残基4(GLN185)から残基234(MET415)と等価なフォールディングを採用することが見出された。ヒトレチノイン酸レセプターガンマは、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機能することが知られている。さらにヒトレチノイン酸レセプターガンマの“LBDモチーフ”残基、PHE251、LEU254、ASP258、GLN259、LEU262およびLEU263は、BAA07892.2ではそれぞれVAL1573、LEU1574、ASP1578、LYS1579、ALA1582およびMET1583として保存されている。前記の関係は単にヒトレチノイン酸レセプターガンマとの関係ではなく、むしろ全体としての核内ホルモンレセプターリガンド結合ドメイン類との関係である。したがって、ヒトレチノイン酸レセプターガンマ(3LBD)とBAA07892.2とのゲノムスレッダー(Genome Threader)(商標)アラインメントの参照によって、BAA07892.2のVAL1573、LEU1574、ASP1578、LYS1579、ALA1582およびMET1583は“LBDモチーフ”残基を形成すると予想される。

等価なフォールディングおよび“LBDモチーフ”残基保存が組み合わされて、BAA07892.2のこの領域、したがってこの領域を含むタンパク質が核内ホルモンレセプターの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が可能である。

【0012】

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そ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そ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。

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“LBDS1ポリペプチド”と称する。

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に必要なものと予想される、LBDS1ポリペプチドの領域(以下では“LBDS1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、またはその変種であって“LBDモチーフ”(VAL1573、LEU1574、ASP1578、LYS1579、ALA1582およびMET1583、または同等な残基)を保有する。本明細書で定義されるように、LBDS1核内ホルモンレセプターリガンド結合ドメイン領域は、LBDS1ポリペプチド配列の残基1507と残基1742の間に広がっていると考えられる。

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を含むが、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(LBDS1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(ii)に記載のポリペプチドフラグメント、好ましくはLBDS1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

グメントをコードするものであるか、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ものである。

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、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、例えば発現ベクターを提供する。

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、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抑制するリガンドを提供する。

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）させる。重要なことには、LBDS1ポリペプチドのLBDS1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それぞれ定義された領域の機能を特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特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。

【0014】

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。前記分子はまた以下を含む疾患の治療を目的とする医薬の製造においても用いることができる：細胞増殖性疾患（腫瘍、癌および骨髄増殖症候群および血管形成を含む）、代謝性障害（肥満、骨粗しょう症、非インスリン依存性糖尿病、脂質代謝異常、甲状腺機能亢進、高カルシウム血症、高コレステロール血症、高脂血症を含む）、心臓血管疾患（高血圧、アテローム性硬化症、高インスリン血症、心不整脈を含む）、腎疾患（糸球体腎炎、腎血管性高血圧を含む）、皮膚疾患（skin disorders）（皮膚疾患（dermatological disease）、乾癬、アクネ、加齢、湿疹、創傷治癒を含む）、炎症（炎症性大腸疾患、気道の炎症、気腫、喘息を含む）、免疫障害、自己免疫疾患、多発性硬化症、神経系疾患（不安、抑うつ、神経変性疾患、アルツハイマー症、パーキンソン病、脳損傷、卒中および痛みを含む）、感染（ウイルス感染を含む）、並びに核内ホルモンが関係するその他の病的状態。

【0015】

第九番目の特徴では、本発明は患者で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。同様な方法は患者での疾患治療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もしく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緩解の指標となる。

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）を生物学的サンプルとリガンド-ポリペプチド複合体の形成に適した条件下で接触させる工程；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、当業者には明らかなように種々の異なる方法、例えば短いプローブによ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が存

10

20

30

40

50

在する。同様な方法を短期または長期ベースで用いて、モニターされる疾患の治療を可能にすることができる。本発明はまた前記疾患診断方法で有用なキットを提供する。

【0016】

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活性を発揮させる方法を提供し、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一緒に含有する。 10

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（腫瘍、癌および骨髄増殖症候群および血管形成を含む）、代謝性障害（肥満、骨粗しょう症、非インスリン依存性糖尿病、脂質代謝異常、甲状腺機能亢進、高カルシウム血症、高コレステロール血症、高脂血症を含む）、心臓血管疾患（高血圧、アテローム性硬化症、高インスリン血症、心不整脈を含む）、腎疾患（糸球体腎炎、腎血管性高血圧を含む）、皮膚疾患（skin disorders）（皮膚疾患（dermatological disease）、乾癬、アクネ、加齢、湿疹、創傷治癒を含む）、炎症（炎症性大腸疾患、気道の炎症、気腫、喘息を含む）、免疫障害、自己免疫疾患、多発性硬化症、神経系疾患（不安、抑うつ、神経変性疾患、アルツハイマー症、パーキンソン病、脳損傷、卒中および痛みを含む）、感染（ウイルス感染を含む）、並びに核内ホルモンが関係するその他の病的状態である。 20

【0017】

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。 30

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なければならない。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、リガンドまたは化合物はアンタゴニストでなければならない。前記拮抗薬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。 40

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させるために、または全く発現させないために形質転換した遺伝子導入または遺伝子ノックアウト非ヒト動物を提供する。前記遺伝子導入動物は、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特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。

【0018】

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、記載した特定の方法論、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実施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 50

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。

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, NY); および Handbook of Experimental Immunology, Vols. I-IV (D.M. Weir and C.C. Blackwell eds. 1986).

10

20

【0019】

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結合した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（すなわちペプチド同族体）が含まれる。前記用語は、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いが、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、活性を有する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。

30

本発明の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、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（例えばリコンビナント形成時に）を付与する配列を含むことができる。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延長するための化合物（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。

【0020】

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公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、タンパク分解反応、リン酸化、プレニル化、ラセミ化、セレノイル化、タンパク質への転移RNA仲介アミノ酸付加（例えばアルギニル化）およびユビキチン付加が含まれる。

40

50

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、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共有結合改変によるブロッキングは、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てよい。

【0021】

ポリペプチド内に生じる改変はしばしばポリペプチドが生成される方法の関数であろう。遺伝子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種々の宿主細胞タイプ間で異なるであろう。

10

本発明のポリペプチドは任意の適切な様式で製造することができる。そのようなポリペプチドには、天然に存在する単離されたポリペプチド（例えば細胞培養から精製）、遺伝子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。

本発明の第一の特徴の機能的に等価なポリペプチドは、LBDS1ポリペプチドと相同なポリペプチドであろう。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十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ように“相同である”と称される。“同一性”とは、アラインメントを実施した配列のどの特定の場所においても、アミノ酸残基が配列間で同一であることを示す。“類似性”は、アラインメントを実施した配列のいずれの特定の場所においても、アミノ酸残基が配列間で類似のタイプであることを示す。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）。

20

【0022】

したがって、相同なポリペプチドには、LBDS1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）で置換されてあるポリペプチドを含むことができ、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された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。

30

40

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が含まれる。

【0023】

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、LBDS1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、LBDS1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。

50

本明細書の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。

本件では、LBDS1の好ましい活性フラグメントは、LBDS1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もの、および残基VAL1573、LEU1574、ASP1578、LYS1579、ALA1582およびMET1583、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、VAL1573はLEU、ILE、ALA、MET、PHE、TYRまたはTRPによって置換することができる。例えば、LEU1574はILE、ALA、VAL、MET、PHE、TYRまたはTRPによって置換することができる。例えば、ASP1578はGLUと置換することができる。例えば、LYS1579は、GLN、ASN、ARG、HIS、SERまたはTHRと置換できる。例えば、ALA1582は、LEU、ILE、VAL、MET、PHE、TYRまたはTRPと置換することができる。例えば、MET1583はLEU、ILE、ALA、VAL、PHE、TYRまたはTRPによって置換することができる。したがって本発明のこの特徴は、LBDS1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、VAL1573、LEU1574、ASP1578、LYS1579、ALA1582およびMET1583、または等価な残基をもつ“LBDモチーフ”を保有するものを含む。上記で考察したように、LBDS1の核内ホルモンレセプターリガンド結合ドメイン領域は、LBDS1ポリペプチド配列の残基1507と残基1742の間に広がっていると考えられる。

10

20

30

40

【0024】

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、構造的アラインメント技術の1つまたは2つ以上を用いて特定されたポリペプチドであろう。例えば、バイオペンジウム(Biopendium)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)、LBDS1ポリペプチドと比較したとき、配列同一性は低い、LBDS1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想される、現在はまだ知られていない機能を有するポリペプチドを特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先ず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)を用いて実施した。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実施された。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ために必要であった基準にしたがってトレーニングした。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公知であるので、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。

40

【0025】

LBDS1の構造的相同体は、LBDS1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構

50

造的相同性を共有し、“LBDモチーフ”残基VAL1573、LEU1574、ASP1578、LYS1579、ALA1582およびMET1583、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想される。

本発明の第一の特徴のポリペプチドにはまた、LBDS1ポリペプチドのフラグメント、LBDS1ポリペプチドのフラグメントの機能的等価物、およびLBDS1ポリペプチドのフラグメント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LBDS1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。

10

【0026】

本明細書において用いる、“フラグメント”という用語は、LBDS1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まなければならない。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。

本発明のこの態様の好ましいポリペプチドフラグメントは、LBDS1ポリペプチドのそれぞれLBDS1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定義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。

20

LBDS1ポリペプチドの場合、この領域は残基1507と残基1742の間に広がっていると考えられる。

このフラグメントの変種は、本発明のこの特徴の実施態様として包含されるが、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。

ある場合には、“変種”という用語は本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。

【0027】

伸長型変種の場合、LBDS1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、LBDS1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、LBDS1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アミノ酸残基が付加されて、前記ポリペプチドフラグメントが正確に折り畳まれる能力を損なうことなく、LBDS1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末端のいずれか一方または両方に包含され、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すことができる。

30

LBDS1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LBDS1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（VAL1573、LEU1574、ASP1578、LYS1579、ALA1582およびMET1583）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。

40

【0028】

第二の場合には、“変種”という用語は、上記で述べたポリペプチドフラグメントの相同体を含み、前記相同体はLBDS1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、さらに前記は“LBDモチーフ”残基（VAL1573、LEU1574、ASP1578、LYS1579、ALA1582およびMET1583）または等価な残基を保有するが、た

50

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、LBDS1ポリペプチドのそれぞれLBDS1核内ホルモンレセプターリガンド結合ドメイン領域と30%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相同体変種は、LBDS1ポリペプチドのLBDS1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、LBDS1ポリペプチドのLBDS1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。

10

【0029】

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)によって特定される、LBDS1ポリペプチド配列のLBDS1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定義されるポリペプチドフラグメントの構造と構造的相同性を示すポリペプチドフラグメントであろう。したがって、LBDS1ポリペプチド配列のLBDS1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定義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、上記で定義されたように、前記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ディングと同じフォールディングを採用するはずである。

20

LBDS1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定義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VAL1573、LEU1574、ASP1578、LYS1579、ALA1582およびMET1583、または等価な残基を保持するはずである。

【0030】

そのようなフラグメントは“独立的存在”、すなわち、他のアミノ酸もしくはポリペプチドの部分でもなく、融合も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極めて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実施態様は、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。

30

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)を用いてリガンド、例えばポリクローナルまたはモノクローナル抗体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の単離もしくは特定に、または前記ポリペプチドのアフィニティークロマトグラフィーによる精製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または治療補助としても用いられるであろう。

40

【0031】

“免疫特異的”という用語は、前記抗体が、従来技術の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。

50

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、タイログロブリンおよびキーホールリンペットの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。免疫動物から血清を採集し、公知の方法（例えばイムノアフィニティークロマトグラフィー）で処理する。

【0032】

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また当業者は容易に製造できる。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); Vole 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作製されたモノクローナル抗体群を種々の特性、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、それらを誘導させた個々のポリペプチドの精製に特に有用である。また別には、問題のモノクローナル抗体をコードする遺伝子を、例えば当分野で公知の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。

キメラ抗体[非ヒト可変領域がヒトの定常領域と結合または融合されている（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。

【0033】

抗体は、例えばヒト化によって（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)）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。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。

また別の選択肢では、前記抗体は“二重特異性”抗体であってもよい。前記は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。

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のリンパ球のPCR増幅V-遺伝子、または未感作ライブラリーのいずれかから選別することができる（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また鎖のシャッフリングによって改善することができる（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、放射能免疫アッセイ（RIA）または酵素結合免疫吸着アッセイ（ELISA）の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。

【0034】

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、LBDS1ポリペプチドの活性化フラグメント、例えばLBDS1ポリペプチド

10

20

30

40

50

配列のLBDS1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。

これら配列ストレッチを包含する核酸分子は本発明のこの特徴の好ましい実施態様を構成する。

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の連続する少なくともn個の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また上記で述べた核酸分子と相補的な配列を含む（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）。

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、例を挙げれば、例えば固相ホスホルアミダイト化学合成を用いる化学合成によって、ゲノムまたはcDNAライブラリーから、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転写によって生成することができる。

10

【0035】

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（センス鎖とも称される）でも、非コード鎖（アンチセンス鎖とも称される）でもよい。

20

“核酸分子”という用語にはまたDNAおよびRNAの類似体（例えば改変骨格を含むもの）、並びにペプチド核酸（PNA）が含まれる。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アンチジーン作用因子を指し、長さが少なくとも5ヌクレオチドのオリゴヌクレオチドを含み、前記オリゴヌクレオチドがアミノ酸残基のペプチド骨格と結合されたものをいう。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）]。

配列番号：2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。これらの分子はまた、遺伝コードの縮退の結果として、配列番号：2またはLBDS1ポリペプチドの活性フラグメント（例えばLBDS1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ができる。LBDS1核内ホルモンレセプターリガンド結合ドメイン領域は、LBDS1ポリペプチド配列の残基1507と残基1742の間に広がっていると考えられる。配列番号：1では、LBDS1核内ホルモンレセプターリガンド結合ドメイン領域は、したがってヌクレオチド4548から5255を含む核酸分子によってコードされる。この配列ストレッチ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実施態様を構成する。

30

【0036】

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）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安定性で役割を有するものである。前記核酸分子はまた、付加アミノ酸、例えば更なる官能性を提供するアミノ酸をコードする付加配列を含むことができる。

40

【0037】

50

本発明の第二および第三の特徴の核酸分子にはまた、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が含まれる。

上記で考察したように、LBDS1ポリペプチドの好ましいフラグメントは、LBDS1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4548から5255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。

10

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アミノ酸置換、欠失または挿入をもたらすであろう。

【0038】

本発明の核酸分子はまた、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多様な理由で当分野で一般的に知られている方法を用いて作製することができる。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再アッセムブリーは、ヌクレオチド配列の作製に用いる技術に含まれる。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改変、コドンの優先性の変化、スプライシング変種の生成、変異の導入など、その他を実施することができる。

20

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し、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抑制物質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。

30

【0039】

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、当業者には理解されるように、本発明のポリペプチドをコードする標的核酸を認識し、前記と特異的に結合して前記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)）。

40

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：溶媒のタイプおよび体積；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阻害する薬剤（デンハルト試薬、またはBLOTTO）；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の後の洗浄条件のストリンジェンシー度（Sambrook et al.（上掲書）を参照されたい）。

50

【0040】

完全に相補的な分子と標的とのハイブリダイゼーションの抑制は、当業者に公知の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エンシー条件下で競合および抑制する。

“ストリンジエンシー”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。高ストリンジエ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（pH7.6）、5倍のデンハルト溶液、10%の硫酸デキストラン、および20 μ g/mLの変性せん断サケ精子DNA）中で42 $^{\circ}$ で一晩インキュベーションし、続いてる液を0.1倍のSSCで約65 $^{\circ}$ で洗浄すると定義される。低ストリンジエンシー条件は、ハイブリダイゼーション反応が35 $^{\circ}$ 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エンシーを構成するものである。

10

【0041】

本発明のこの特徴の好ましい実施態様は、LBDS1ポリペプチド（配列番号：2）をコードする核酸の全長に対して少なくとも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、LBDS1ポリペプチド配列のそれぞれLBDS1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、LBDS1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。

20

本明細書に示す通り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。

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4548-5255を含む領域の全長に対して少なくとも80%同一の領域を含むもの、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である。この場合、前記の全長に対して少なくとも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、好ましい実施態様は、LBDS1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。

30

【0042】

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、LBDS1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ルソログ遺伝子のcDNAまたはゲノムクローンを単離することができる。これに関しては、当分野で公知の他の技術の中で特に以下の技術を用いることができる。これらの技術は例示として下記で考察される。DNA配列決定および分析のための方法は周知で、当分野では一般的に利用可能であり、実際に本明細書で考察される本発明の実施態様の多くを実施するため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のクレノーフラグメント、シークエナーゼ（US Biochemical Corp., Cleveland, OH）、Taqポリメラーゼ（Prekin Elmer）、熱耐性T7ポリメラーゼ（Amersham, Chicago, IL）、またはポリメラーゼと読み取り保証エキソヌクラーゼ（例えば市販（Gibco/BRL, G

40

50

aithersburg, 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、配列決定プロセスは、例えばハミルトンマイクロラボ (Hamilton Micro Lab) 2200 (Hamilton, Reno, NV)、ペルチアサーマルサイクラー (Peltier Thermal Cycler) PTC200 (MJ Research, Watertown, MA) および A B I カタリスト 373 および 377 DNA 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。

【0043】

LBDS1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LBDS1ポリペプチドのLBDS1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4548-5255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な少なくとも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、蛍光染料、および検出可能な生成物の生成を触媒することができ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から問題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、関連配列について、例えば前記のファミリー、タイプおよびノ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。

【0044】

多くの事例で、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。完全長の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、さらに上流の配列 (例えばプロモーターおよび調節エレメント) を検出することを目的とする当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、利用可能なある方法は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) によるものである。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.) を例示することができる) 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、既知の領域にもと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)。使用できるまた別の方法は捕捉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)。さらにまたゲノムDNA探索のために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つける場合に有用である。

【0045】

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別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ムライブラリーが好ましい。ランダムにプライミ

10

20

30

40

50

ングされ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。

本発明のある実施態様では、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、核酸分子は特異的に個々のヒト染色体上の特定の位置に対して標的化され、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 (ジョンズホプキンス大学、ウェルヒメディカルライブラリーからオンラインで入手できる)。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が続いてリンケージ分析(物理的に近接する遺伝子の同時遺伝)によって特定する。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者に貴重な情報が提供される。いったん疾患または症状の位置が遺伝的リンケージによって個々のゲノム領域でおおざっぱに決定されたら、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、キャリアー個体または罹患個体間で転座、逆位などによる染色体上の位置の相違を検出するために用いることができる。

10

20

【0046】

本発明の核酸分子はまた組織分布解析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に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、位置的または量的特性を有するかもしれない。

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。本発明の宿主細胞(前記は本発明のベクターによる形質転換、トランスフェクトまたは形質導入が可能である)は原核細胞でも真核細胞でもよい。

30

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として調製することができる。前記のような発現方法は当分野では周知であり、以下の文献に多くが記載されている：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, Tronto)。

【0047】

一般的には、要求されるホストでポリペプチドを製造す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、シュードレービー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

40

50

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）。ヒトの人工染色体（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フラグメントをデリバリーするために用いることができる。

【0048】

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製造に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、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、トランスフェクション、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、擦過ローディング、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)）。

【0049】

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コントロール配列を（例えば翻訳ポリペプチドの小胞体内、細胞質周囲間隙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宿主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、宿主細胞の増殖に対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ことがある。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対して反応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反応して、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したがって、多くの適切な転写および翻訳エレメント（構成的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）（Gibco BRL）プラスミド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）はベクターで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。マルチコピー配列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別可能マーカーとともに用いることができる。

【0050】

発現ベクターは、個々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対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

10

20

30

40

50

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と結合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。

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。

長期に及ぶ高収量のリコンビナントポリペプチド生成のためには安定な発現が好ましい。例えば、問題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別マーカー(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、選別培地に切り替える前に細胞を富裕培地で1-2日間増殖させることができる。選別マーカーの目的は、選別に対する耐性を付与することで、選別マーカーの存在によって、導入された配列を発現することができ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を細胞タイプ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。

【0051】

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。

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、特にインビトロゲン(Invitrogen, San Diego, CA)からキットとして("MaxBac"キット)市販ルートで入手できる。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、例えばショウジョウバエS2およびスポドプテラ(Spodoptera) Sf9細胞が含まれる。

当分野で公知の多くの植物細胞培養および植物体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また別の例も文献に記載されている(Zenk (1991) *Phytochemistry* 30:3861-3863)。

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、それによって導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)全ての植物を培養細胞または組織から再生させることができる。

【0052】

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*)、ストレプトミセスおよびバチルス・ズブチリス(*Bacillus subtilis*)細胞が含まれる。

真菌での発現に特に適切な宿主細胞の例には酵母細胞(例えばビール酵母菌(*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当分野で公知の多くの選別系が存在する。例えば、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別基準として用いてもよい。例えばジヒドロフォレートレダクターゼ(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およ

10

20

30

40

50

びホスフィノトリシンアセチルトランスフェラーゼに対する耐性を付与する。さらに別の選別可能な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。

【0053】

マーカー遺伝子の発現の有無は問題の遺伝子もまた存在することを示唆するが、問題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。また別には、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タンデムに配置することができる。通常、誘発または選別に反応するマーカー遺伝子の発現は、タンデム遺伝子の発現も意味する。

10

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特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)) が含まれるが、ただしこれらに限定されない。イムノアッセイ技術は、核酸またはタンパク質の検出および/または定量のためにメンブレン、溶液またはチップを用いる技術を含む (例えば以下を参照されたい: R. Hampton et al. (1990) Serological Methods, A Laboratory Manual, AP S Press, St Paul, MN; および D. E. Maddox et al. (1983) J. Exp. Med. 158:1211-1216)。

20

【0054】

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分野で公知で、市販ルートで入手可能であり、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ため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。

30

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原物質の他に補助因子、抑制物質、磁性粒子などが含まれる。

【0055】

本発明の核酸分子はまた遺伝子導入動物 (特にげっ歯類動物) の作製に用いることができる。そのような遺伝子導入動物は本発明のまた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実施することができる。前記のような遺伝子導入動物は、本発明のポリペプチドの調製物質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。

40

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の培養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、酸抽出、陰イオンまたは陽イオン交換クロマトグラフィー、ホスホセルロースクロマトグラフィー、疎水性反応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。高速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タンパク質が変性した場合には、タンパク質の再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な構造を再生することができる。

【0056】

50

所望の場合には、可溶性タンパク質の精製を促進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また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FLAGS伸長/アフィニティー精製系（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X A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、精製の促進に用いてもよい。そのような発現ベクターは、チオレドキシ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）による精製を促進し、一方、チオレドキシ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）。

10

【0057】

ポリペプチドをスクリーニングアッセイで使用するために発現させる場合は、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分類（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採集することができる。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。

20

本発明のポリペプチドを用いて、化合物ライブラリーを多様な任意の薬剤スクリーニング技術で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の発現レベルまたは活性レベルを活性化（作働）させ、または抑制（拮抗）することができる。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ために有効であるか、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ために有効である。

30

アゴニスト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造物質であってよい。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。

【0058】

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抑制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。このような態様で、前記ポリペプチドと正常な細胞の結合分子との結合が抑制され、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る。

40

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、または細胞内に存在していてもよい。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反応の刺激もしくは抑制を観察する。続いて前記テスト化合物と接触させた細胞の機能的反応を、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

50

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調べる。活性化の抑制物質は、一般的には既知のアゴニストの存在下でアッセイし、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を観察する。

【0059】

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。別の実施態様では、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。

10

前記ポリペプチドをコードする細胞内のmRNAの産生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、当分野で公知の標準的な方法でモノクローナルまたはポリクローナル抗体を用いて分泌または細胞結合ポリペプチドレベルを測定するELISAを構築し、これを用いて適切に操作作製した細胞または組織におけるポリペプチド生成を抑制または強化する化合物の検索を実施できる。続いて、前記ポリペプチドと被検化合物との結合複合体の生成を測定することができる。

【0060】

使用可能な薬剤スクリーニングのためのまた別の技術は、問題のポリペプチドと適切な結合親和性を有する化合物の高処理スクリーニングを提供する（国際特許出願W084/03564を参照されたい）。前記方法では、多数の異なる小さなテスト化合物が固相基質上で合成される。続いて、前記を本発明のポリペプチドと反応させ洗浄することができる。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されるプレートに直接被覆させることができる。

20

本発明のポリペプチドを用いて、膜結合または可溶性レセプターを当分野で公知の標準的なレセプター結合技術により特定することができる。標準的な技術は、例えばリガンド結合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元素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、仮想的レセプター供給源（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測定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、さらに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作動薬および拮抗薬を特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分野では周知である。

30

【0061】

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/または酵素を特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。

40

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、治療用もしくは診断用試薬として、ワクチンとして、または他の免疫誘発性組成物として下記で詳細に説明するように適切であろう。

本明細書で用いられる用語にしたがえば、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（Yは不純物）の合計の少なくとも85質量%がXである場合は不純物を“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重量%、98質量%または99質

50

量%を構成する。

【0062】

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。本明細書で用いられるように“治療的に有効な量”とは、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（例えば新形成細胞培養アッセイ）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。動物モデルはまた、適切な濃度範囲および投与ルート決定にも用いることができる。続いて前記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

10

。ヒト患者に対する正確な有効量は、症状の重篤度、対象者の健康状態、年齢、体重および対象者の性別、食事、投与時間および回数、併用薬剤、反応の感受性および治療に対する耐性/反応性にしたがって左右される。前記の量は日常的検査および臨床医の判断で決定できる。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、医薬またはホルモンと一緒に患者に投与することができる。

【0063】

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で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ろう。

20

それらの中には医薬的に許容できる塩、例えば鉍酸塩、例えば塩酸塩、臭化水素酸塩、リン酸塩、硫酸塩など；および有機酸の塩、例えば酢酸塩、プロピオン酸塩、マロン酸塩、安息香酸塩など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完全な考察は以下のテキストで見出すことができる：Remington's Pharmaceutical Sciences (Mack Pub. Co., N.J. 1991)。

30

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、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。さらにまた、補助物質（例えば湿潤剤または乳化剤、pH緩衝物質など）も前記組成物中に含まれていてもよい。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、リキッド、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。

【0064】

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、特にヒトの患者が治療されるであろう。

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（hypospray）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もまた調製できる。

40

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デリバリーされるであろう。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マルチ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症状で過剰である場合には、いくつかのアプローチ

50

を利用できる。あるアプローチは、上記のように対象者に抑制化合物（アンタゴニスト）を医薬的に許容できる担体とともに、前記ポリペプチドの機能抑制に有効な量で投与することを含む（前記担体による前記機能抑制は、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抑制し、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、前記のような拮抗物質は抗体である。極めて好ましくは、そのような抗体は、その免疫原性を最少にするために先に記載したようにキメラおよび/またはヒト化抗体である。

【0065】

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形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、例えば上記で述べたように（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抑制することができる。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、抑制は“三重らせん”塩基対形成法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十分に開く能力を抑制するので有用である。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）。さらにまた相補的配列またはアンチセンス分子をデザインし、リボソームの結合を妨げて転写を阻害することによってmRNAの翻訳を遮断することができる。そのようなオリゴヌクレオチドを投与してもよいし、また *in vivo* で発現させることによって *in situ* で生成させることができる。

【0066】

さらにまた、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、天然または合成の触媒活性を有する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から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）から合成してリボヌクレアーゼによる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高め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添加、または分子の骨格内でホスホジエステル結合に代わって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、さらに非慣用塩基（例えばイノシン、ケオシンおよびプトシン（アセチル -、メチル -、チオ - も同様に）、および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は広げることができる。

【0067】

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がまた利用可能である。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併用して投与し、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、患者の関連する細胞による本ポリペプチドの内在性生成を行わせるこ

10

20

30

40

50

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再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“パッケージに封入”されている。遺伝子デリバリー担体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) ベクターであろう。例えば、本発明のポリペプチドをコードする核酸分子を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。この発現構築物を続いて単離し、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問題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* で細胞を高等操作するために、さらに *in vivo* でポリペプチドを発現させるために患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), In “*Human Molecular Genetics*” (1996) T. Strachan & A.P. Read, BIOS Scientific Publishers Ltd.)。 10 20

【0068】

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直接血流または筋肉組織に注射される。

本発明のポリペプチドまたは核酸分子が疾患をひき起こす原因物質である場合には、本発明は、前記疾患をひ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) であろう。そのようなワクチンは免疫を付与する抗原、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一緒に含む。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任意の担体が含まれる。さらにまた、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。 30

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) および水性および非水性滅菌懸濁剤 (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。 40

【0069】

本発明のワクチン製剤は単位用量またはマルチ用量容器で提供されてもよい。例えば封入アンプルおよびバイアルは凍結乾燥状態で保存でき、使用直前に滅菌された液状担体を添加することだけが必要である。投与量はワクチンの比活性に左右され、日常的な検査で容易に決定される。

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。本発明の核酸分子の特徴を有する、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、またはそれらを限定することができる診断ツールを提供する。前記遺伝子に変異を保有する個体は、多様な技術によって DNA レベルで検 50

出すことができる。

診断のための核酸分子は患者の細胞、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))。

【0070】

ある実施態様では、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、さらにコントロールの前記発現レベルと比較することを含む、患者で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を形成させることができ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;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;

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。

本発明のさらに別の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べき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) 疾患に関連する核酸分子内の変異の存在を検出することによって患者の疾患を診断する工程。

【0071】

上記に記載した方法で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特定することができる。完全にマッチした配列は、リボヌクレアーゼによる消化によって、または溶解温度における差異を評価することによって mismatches を有する二本鎖と区別することができる。患者に存在する変異の有無は、ストリンジェントな条件下でDNAを前記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し(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場所で前記核酸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存在する疾患関連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検出することができる。

前記のような診断は特に出生前検査および新生児検査においてすら有用である。

【0072】

点変異およびリファレンス遺伝子と“変異”遺伝子間のその他の配列の相違は、他の周知の技術、例えば直接DNA配列決定または一本鎖構造多型性(Orita et al., *Genomics*, 5:874-879(1989))によって特定できる。例えば、配列決定プライマーを二本鎖PCR生成物または改変PCRによって生成された一本鎖鋳型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て自動配列決定方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強化される。さらに、点変異および他の配列の変動(例えば多形性)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。

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電気泳動におけるDNAフラグメントの移動度変化によって、または直接DNA配列決定(例えば、Myers et al., Science (1985) 230:1242)によって検出することができる。特定の場所における配列の変化はまた、ヌクレアーゼ保護アッセイ、例えばリボヌクレアーゼおよびS1保護または化学切断法(Cotton et al., Proc. Natl. Acad. Sci. USA (1985) 85:4397-4401)によって明らかにすることができる。

【0073】

通常ゲル電気泳動およびDNA配列決定の他に、例えばミクロ欠損、異数性、転座、逆位のような変異はまた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、それらをメンブレン上に単離および/または固定しないで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、現在のところもっと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., Genet. 7:149-154(1991))。

本発明のまた別の実施態様では、本発明の核酸分子を含むオリゴヌクレオチドプローブアレーを構築し、遺伝的変種、変異および多形性の効率的スクリーニングを実施することができる。アレー分析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、遺伝的連関および遺伝的多様性を含む分子遺伝学における多様な疑問に答える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)。

【0074】

ある実施態様では、前記アレーは以下の文献に記載された方法にしたがって調製し使用される(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でよい。前記オリゴマーは、光誘導化学反応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任意の適切な固相支持体でよい。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(スロット)プロットに類似する“格子化(gridDED)”アレーを用いてcDNAフラグメントまたはオリゴヌクレオチドを配置し、さらに真空系、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結合させることができる。アレー(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(任意の適切な固相支持体)および機械(ロボットを含む))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任意の数を含むことができる(このことはアレー自体を商業的に入手可能な計測手段の有効利用に向くものとしている)。

【0075】

上記で考察した方法の他に、対象者由来サンプルから、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分野で周知の方法のいずれか、例えば核酸増幅(例えばPCR、RT-PCR、リボヌクレアーゼ保護、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。

。ホスト由来のサンプルで本発明のポリペプチドレベルを決定するために用いることができるアッセイ技術は当分野でよく知られており、さらに上記でいくらか詳細に考察されてい

る（放射能免疫アッセイ、競合結合アッセイ、ウェスタンブロットおよび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生物学的サンプルと、リガンド-ポリペプチド複合体の形成に適した条件下で接触させる工程；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レベルの変化または異常の診断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、前記ポリペプチドと複合体形成に適した条件下で混合することによって確立することができる。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。

10

【0076】

本発明のポリペプチドと特異的に結合する抗体を前記ポリペプチドの発現を特徴とす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で治療されている患者をモニターするアッセイで用いることができる。診断の目的で有用な抗体は治療薬として上記で述べた態様と同じ態様で調製することができる。前記ポリペプチドのための診断アッセイは、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。当分野で公知の多様なレポーター分子を用いることができる（それらのいくつかは上記に記載されている）。

対象者、コントロールおよび生検組織由来の疾患サンプルで発現されたポリペプチドの量を標準値と比較する。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、治療期間中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、動物実験に、臨床試験にまたは個々の患者の治療モニタリングで個々の治療方法の有効性評価に用いることができる。

20

【0077】

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。

30

本発明のある特徴では、診断キットは、ストリンジェ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入れた第三の容器を含むことができる。

本発明のまた別の特徴では、診断キットは核酸分子アレーを含み、前記核酸分子の1つが本発明の核酸であってよい。

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。

40

【0078】

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の診断に有用である：特に細胞増殖性疾患（腫瘍、癌および骨髄増殖症候群および血管形成を含む）、代謝性障害（肥満、骨粗しょう症、非インスリン依存性糖尿病、脂質代謝異常、甲状腺機能亢進、高カルシウム血症、高コレステロール血症、高脂血症を含む）、心臓血管疾患（高血圧、アテローム性硬化症、高インスリン血症、心不整脈を含む）、腎疾患（糸球体腎炎、腎血管性高血圧を含む）、皮膚疾患（skin disorders）（皮膚疾患（dermatological disease）、乾癬、アクネ、加齢、湿疹、創傷治癒を含む）、炎症（炎症性大腸疾患、気道の炎症、気腫、喘息を含む）、免疫障害、自己免疫疾患、多発性硬化症、神経系疾患（不安、抑

50

うつ、神経変性疾患、アルツハイマー症、パーキンソン病、脳損傷、卒中および痛みを含む)、感染(ウイルス感染を含む)、並びに核内ホルモンが関係するその他の病的状態。本発明の種々の特徴および実施態様を、特にLBDS1ポリペプチドを用いて実施例によってこれからより詳細に説明しよう。

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。

【0079】

実施例1: BAA07892.2 (LBDS1)

関係が離れている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内メンバー、ホモサピエンスレチノイン酸レセプターガンマ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

10

バンク(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、ホモサピエンスのレチノイン酸レセプターガンマのリガンド結合ドメインである(PDBコード3LBD; 図1参照)。

3LBDの類縁体についてバイオペンジウム(Biopendium)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を用いる)、4475のゲノムスレッダーの結果が得られる。この4475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レチノイン酸レセプターガンマ-1の残基182-418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、BAA07892.2(LBDS1; 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BAA07892.2(LBDS1)の領域(残基1507-1742の間)をホモサピエンスレチノイン酸レセプターガンマ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。ホモサピエンスレチノイン酸レセプターガンマ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、BAA07892.2(LBDS1)の残基1507-1742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71%の信頼度で明らかにしている。

20

【0080】

3LBDの類縁体についてのバイオペンジウム(Biopendium)の(ターゲットマイニングインターフェースを用いる)検索によって、850のフォワードPSI-BLASTの結果が得られる。フォワードPSI-BLAST(図2Cを参照されたい)は前記関係を特定できない。インファーマチカゲノムスレッダーだけがBAA07892.2(LBDS1)を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確認できる。

30

BAA07892.2(LBDS1)について公開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(図3)を調査する。重複配列ディスプレイページは、Q14008と呼ばれるBAA07892.2(LBDS1)とほぼ同一の配列を明らかにする。Q14008はBAA07892.2(LBDS1)の部分配列であり、BAA07892.2の6個のN-末端残基のみを欠いている。

BAA07892.2(LBDS1)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確認するPROSITEヒットもPRINTSヒットも得られない。PROSITEおよびPRINTSは、類似するファミリーのタンパク質を表示することを支援する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、BAA07892.2(LBDS1)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。

40

【0081】

他のいずれかの公開ドメイン注釈付け手段が核内ホルモンレセプターリガンド結合ドメインをBAA07892.2(LBDS1)が含むと注釈付けすることができるか否かを確認するために、BAA07892.2(LBDS1)タンパク質配列をインタープロ(IntroPro)ウェブサイト(PFAMデータベース(Protein Family Database of Alignment and hidden Markov models、アライ

50

ンメントと隠れたマルコフモデルに関するタンパク質ファミリーデータベース) に対して検索する(図4矢印1参照)。PFAM-Aマッチングは返さない。したがってPFAMでは、BAA07892.2(LBDS1)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矢印2参照)によって二連核局在化シグナルがBAA07892.2(LBDS1)の残基550-567に確認される。核内ホルモンレセプターの典型的な(指標ではないが)特徴は二連核局在化シグナルを有することである。

続いてThe National Center for Biotechnology Information(NCBI) Genbankタンパク質データベースを、BAA07892.2(LBDS1)に関連する公開ドメインでさらに別の公知の情報があるか否かを調べた。これは、タンパク質および遺伝子配列寄託に関する米国の公開ドメインデータベースである(図5A)。BAA07892.2(LBDS1)はホモサピエンスの配列で、そのGenbankタンパク質IDはBAA07892.2であり、長さは2038アミノ酸である。BAA07892.2(LBDS1)はKazusa DNA研究所(Kisarazu、日本)の科学者の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前記公開ドメインの情報は、前記遺伝子を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していない。

【0082】

続いてスイスパイオインフォマティクス研究所、SIB(Swiss Institute of Bioinformatics)のSWISS-PROTタンパク質データベースを、BAA07892.2(LBDS1)の部分配列Q14008に関連する公開ドメインでさらに別の公知の情報があるか否かを調べた。これは、タンパク質および遺伝子配列寄託に関するスイスの公開ドメインデータベースである(図5B)。BAA07892.2(LBDS1)の部分配列Q14008はホモサピエンスの配列で、そのSWISS-PROTタンパク質IDはQ14008であり、長さは2032アミノ酸である。Q14008はCH-TOG(結腸及び肝臓腫瘍過剰発現遺伝子)タンパク質と呼ばれる。Q14008、CH-TOGはCharrasseら(INSERM、モントペリエ、フランス)によってクローン化された。この遺伝子について前記公開ドメインの情報は、前記遺伝子を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していない。

したがって、全ての公開ドメインの注釈付けツールを用いて、BAA07892.2(LBDS1)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を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。

【0083】

ここで逆検索を実施する。BAA07892.2(LBDS1)を今度は問い合わせ配列としてバイオペンジウムで用いる(図6A参照)。インファーマチカゲノムスレッダーは、BAA07892.2(LBDS1)の残基1507-1742はホモサピエンスレチノイン酸レセプターガンマのリガンド結合ドメインと同じ構造を有すると信頼度71%で同定する(図6Bの矢印参照)。ホモサピエンスレチノイン酸レセプターガンマ(3LBD)のリガンド結合ドメインは元々の問い合わせ配列であった。PSI-BLASTの正の繰返しはこの結果を返さない(図6C)。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。

ホモサピエンスレチノイン酸レセプターガンマリガンド結合ドメインの配列は、BAA07892.2(LBDS1)の配列アラインメントを観察するために前記に対して選択される。類似の構造を有すると確認されたタンパク質に対する問い合わせタンパク質のAIEye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。

ホモサピエンスレチノイン酸レセプターガンマリガンド結合ドメインは“LBDモチーフ”を含み、前記モチーフは、今日まで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補助アクチベーターおよび補助リプレッサーのリクルーティングに中心的に関与する。残基PHE251、LEU254、ASP258、GLN259、LEU262およびLEU263は、ホモサピエンスレチノイン酸レセプターガンマ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)。BAA07892.2(LBDS1)中の残基VAL1573、LEU1574、ASP1578、LYS1579、ALA1582およびMET1583は協調して、LBDモチーフの基準を満足する(図7の実線及び破線四角枠を参照

10

20

30

40

50

)。このことは、BAA07892.2 (LBDS1) はホモサピエンスレチノイン酸レセプターガンマ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。

【0084】

特定されたタンパク質が実際問い合わせタンパク質と関連することを担保するために、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、公知の小型の抑制性分子が公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公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、直接水素結合によるか、または疎水性相互作用による。このようにして、活性部位のフォールディング/構造が、同定された類似体と選択した公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知ることができる。ホモサピエンスレチノイン酸レセプターガンマリガンド結合ドメインのLigEyeによる観察は、9-シスレチノイン酸と結合する14残基(図7の丸付き文字)を明らかにする。これらの14残基の3つのみ(ALA1556、LEU1591、LEU1725、図7の丸付き文字)がBAA07892.2(LBDS1)に正確に保存され、BAA07892.2(LBDS1)は、場合によっては特に9-シスレチノイン酸と結合しないことを示唆する。しかしながら、11個の疎水性残基が存在し、これらはホモサピエンスレチノイン酸レセプターガンマリガンド結合ドメイン(PHE230、LEU233、ALA234、LEU268、LEU271、MET272、ILE275、PHE288、PHE304、LEU400およびILE412)のポケットに一列に並んでいる。BAA07892.2(LBDS1)において、これら11残基のうちの3つ(ALA1556、LEU1591、LEU1725、図7の丸付き文字)は完全に保存される。さらに5つの残基(ILE1552、VAL1555、ILE1592、PHE1595、LEU1739、図7の破線丸付き文字)は疎水的な性質において保存される。結合ポケットの輪郭を形成する11の疎水性残基のうち8つで疎水性が保存されていることは、BAA07892.2(LBDS1)は疎水性ステロイド様疎水性リガンドと結合することを示唆している。典型的には、核内ホルモンレセプターリガンド結合ドメインの結合ポケットは、非常に重要なアルギニン残基を含んでいる。前記アルギニンのグアニジニウム(guanidinium)基は結合する機能を果たし、ステロイド/ステロイド様リガンドに存在するヒドロキシルまたはカルボキシル基を安定化する。例えば、ホモサピエンスエストロゲンレセプターリガンド結合ドメインのARG394はエストラジオールのAリングのフェノール類のヒドロキシル基と結合し、それを安定化する(Brzozowski Nature 1997 vol.389:p753-758)。LigEyeは、ホモサピエンスレチノイン酸レセプターガンマリガンド結合ドメインのARG278がレチノイン酸のカルボキシラート基を安定化する。BAA07892.2(LBDS1)において、前記ARGは保存されないが、C-末端に向かっている1残基が置換されたARG1599(アステリスク)が存在する。BAA07892.2(LBDS1)のARG1599はステロイド様リガンドに存在するヒドロキシルまたはカルボキシル基と結合し、これを安定化する機能を果たす。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、BAA07892.2(LBDS1)はホモサピエンスレチノイン酸レセプターガンマリガンド結合ドメインと類似の態様で折り畳まれることを示唆し、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確認される。

【0085】

BAA07892.2(LBDS1)の逆最大化PSI-BLASTによって、BAA07892.2(LBDS1)と76.0%の配列同一性を有する(図6Cの矢印1参照)、CAB61894.1と呼ばれるBAA07892.2(LBDS1)のアフリカツメガエル類似体、XMAP215が特定される。また、BAA07892.2(LBDS1)の逆最大化PSI-BLASTによって、BAA07892.2(LBDS1)と45.0%の配列同一性を有する(図6Cの矢印2参照)、CAB55772.1と呼ばれるBAA07892.2(LBDS1)のキイロショウジョウバエ類似体、微小管結合蛋白質が特定される。

BAA07892.2(LBDS1)、アフリカツメガエルCAB61894.1(XMAP215)及びキイロショウジョウバエCAB55772.1(微小管結合蛋白質)でアラインメントを実施し、AIEyeで観察する(図9)。AIEyeは、前記リガンドの結合ポケット(ILE1552、VAL1555、ALA1556、LEU1591、ILE1592、PHE1595、LEU1725およびLEU1739)の輪郭を形成すると予測されるBAA07892.2(LBDS1)の8つの疎水性残基は全てアフリカツメガエルCAB61894.1、XMAP215(図9を参

10

20

30

40

50

照)で正確に保存されていることを明らかにしている。さらにまた、BAA07892.2(LBDS1)の“LBDモチーフ”とマッチングする6つの残基(VAL1573、LEU1574、ASP1578、LYS1579、ALA1582、MET1583)はすべて正確にアフリカツメガエルCAB61894.1、XMAP215で保存されている(図9参照)。さらに、予想されるリガンド結合ARG1599は正確にアフリカツメガエルCAB61894.1、XMAP215で保存されている(アスタリスクが付されている、図9参照)。タンパク質の機能に本質的である残基は前記タンパク質の類似体で保存される。したがって、アフリカツメガエル類似体CAB61894.1(XMAP215)の予想されるBAA07892.2(LBDS1)の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本質的であるすべての残基の絶対的な保存性は、核内ホルモンレセプターリガンド結合ドメインを含むBAA07892.2(LBDS1)の注釈付けを強く支持する。

10

【0086】

A1Eyeアラインメントはまた、BAA07892.2(LBDS1)とキイロシヨウジョウバエCAB55772.1(微小管結合蛋白質)との保存された関係を示す。A1Eyeは、BAA07892.2(LBDS1)の8つの疎水性残基のうち2つは(ILE1552およびVAL1555)、キイロシヨウジョウバエCAB55772.1(微小管結合蛋白質)で正確に保存されていることを示す。8つのBAA07892.2(LBDS1)の疎水性残基のうち別の4つは(LEU1591、PHE1595、LEU1725、LEU1739)は、キイロシヨウジョウバエCAB55772.1(微小管結合蛋白質)でその他の疎水性残基の代わりに保存的に用いられる。さらにまた、BAA07892.2(LBDS1)“LBDモチーフ”のLys1579は、キイロシヨウジョウバエCAB55772.1(微小管結合蛋白質)で正確に保存されている。BAA07892.2(LBDS1)“LBDモチーフ”のその他の5つの残基は(VAL1573、LEU1574、ASP1578、ALA1582、MET1583)キイロシヨウジョウバエCAB55772.1(微小管結合蛋白質)で保存的に置換されている。さらに、予測されるリガンド結合ARG1599はキイロシヨウジョウバエCAB55772.1(微小管結合蛋白質)で正確に保存されている(アスタリスクが付されている、図9参照)。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類似体において保存されるであろう。したがって、キイロシヨウジョウバエ類似体CAB55772.1(微小管結合蛋白質)におけるBAA07892.2(LBDS1)の予測される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15の残基のうち4つの正確な保存(さらに9つの保存的な置換)は、BAA07892.2(LBDS1)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。

20

【0087】

また、BAA07892.2(LBDS1)の逆最大化PSI-BLASTによって、BAA07892.2(LBDS1)と95%の配列同一性を有する(図6Cの矢印3参照)、ホモサピエンス配列CAA63212.1が特定される。CAA63212.1のBAA07892.2(LBDS1)とのA1Eyeアラインメントは、異なる2つの領域のみが存在することを明らかにする(図10参照)。第1の相違は、CAA63212.1がBAA07892.2(LBDS1)の6つのN-末端アミノ酸を欠いていることである。第2の相違は、CAA63212.1が、BAA07892.2(LBDS1)と比較して、60のアミノ酸を除去する内部欠失を有することである。前記内部欠失は、cDNAが乳癌組織から得られるために(図11の矢印参照)、スプライスバリエーション又は疾病関連欠失を示している場合がある。興味深いことに、前記60のアミノ酸の欠失は、CAA63212.1が“LBDモチーフ”、ARG1599及び3つの予測される疎水性リガンド結合残基(LEU1591、ILE1592及びPHE1595)を欠いていることを意味する。したがって、CAA63212.1は、ステロイド様リガンドとの結合及び核内ホルモンの補助アクチベーターおよび補助リプレッサーのリクルーティングでコンピテントでないことが予測されるであろう。

30

40

【0088】

Charrasseらによる関連する論文が存在する(Eur. J. Biochem. 1995 Dec 1;234(2)406-13)(図12)。この論文は、BAA07892.2(LBDS1)部分配列Q14008、CH-TOGをコードするmRNAが、健康な大腸と比較して癌性の大腸組織で過剰発現していることを示している。さらに、BAA07892.2(LBDS1)部分配列Q14008、CH-TOGをコードするmRNAは、健康な大腸と比較して癌性の肝組織で過剰発現している。このことは、BAA07892.2(LBDS1)部分配列Q14008、CH-TOGのレベルと癌の状態との間の直接的な関係を明らかにしている。BAA07892.

50

2 (LBDS1) と癌の状態とを関連付けることは、前記症状の診断および/または治療のためにBAA07892.2 (LBDS1) の遺伝子またはそのコードタンパク質を標的とする薬剤の開発のための道を作る。特に、前記遺伝子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ことを確認することは、(例えば核内ホルモンレセプターリガンド結合ドメイン作動薬または拮抗薬の使用により) BAA07892.2 (LBDS1) タンパク質に特異的な薬剤の開発を促進する。

【図面の簡単な説明】

【0089】

【図1】バイオベンジウムの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“3LBD”を用いて開始する。

10

【図2】2Aでは、3LBD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選択が示されている。矢印は、ホモサピエンスレチノイン酸レセプターガンマ-1(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)を示している。2Bでは、3LBD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選択が示されている。矢印は、BAA07892.2 (LBDS1) を示している。2Cでは、3LBD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。BAA07892.2 (LBDS1) は特定されない。

【図3】BAA07892.2 (LBDS1) に対する重複配列表示の結果のページである。Q14008と呼ばれる等価な配列が存在する(矢印を参照されたい)。

【図4】BAA07892.2 (LBDS1) 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。

20

【図5】5Aは、BAA07892.2 (LBDS1) に対するNCBIタンパク質の報告である。5Bは、Q14008と呼ばれるBAA07892.2 (LBDS1) CH-TOGタンパク質と等価な配列に対するSWISS-PROTのエントリーを示す。

【図6】6Aはバイオベ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問い合わせ配列としてBAA07892.2 (LBDS1) を用いて開始する。6Bは、問い合わせ配列としてBAA07892.2 (LBDS1) を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選択である。矢印は3LBDを指している。6Cは、問い合わせ配列としてBAA07892.2 (LBDS1) を用いて得られた逆最大化PSI-BLASTの結果の選択である。1、2の番号を付与された矢印はアフリカツメガエルおよびキイロショウジョウバエのBAA07892.2 (LBDS1) の類似体を指している。3の番号を付与された矢印はBAA07892.2 (LBDS1) のホモサピエンス変異体を指している。

30

【図7】BAA07892.2 (LBDS1) および3LBDのAIEye配列アラインメントを示す。

【図8】8Aは3LBDに関するLigEyeで、9-シスレチノイン酸(REA420)とホモサピエンスレチノイン酸レセプターガンマ、3LBDのリガンド結合ドメインとの相互作用部位を示している。8Bは、3LBD、ホモサピエンスレチノイン酸レセプターガンマのリガンド結合ドメインのiRasMolによる解析図である。

【図9】BAA07892.2 (LBDS1) とBAA07892.2 (LBDS1) のアフリカツメガエル(CAB61894.1) およびキイロショウジョウバエ(CAB55772.1) 類似体とのAIEye配列アラインメントを示す。

40

【図10】BAA07892.2 (LBDS1) とCAA63212.1、BAA07892.2 (LBDS1) のホモサピエンス変異体とのAIEye配列アラインメントを示す。

【図11】CAA63212.1、BAA07892.2 (LBDS1) のホモサピエンス変異体に対するNCBIタンパク質の報告を示す。CAA63212.1はヒト乳癌細胞ライブラリー(矢印)に由来する。

【図12】Charrasse et al., Eur. J. Biochem. 1995 Dec 1;234(2)406-13による文献の概要を示す。

【 図 10 - 2 】

AIEye 出力(2月21日, 2001 4:29 PM)

FIG. 10-2

```

720 730 740 750 760 770 780 790
BA078922 1 A E A C M L P W T A E O V Y S M A F S O K N P K N O S E T L N W L S N A I K E F G S G L N V K A F I S N V K T A A A T N P A V R T A A I T L L O G W Y L
CA463212 1 A E A C M L P W T A E O V Y S M A F S O K N P K N O S E T L N W L S N A I K E F G S G L N V K A F I S N V K T A A A T N P A V R T A A I T L L O G W Y L
BA078922 Y V G S L R M F F E D E K P A L L S O D A E F E M G O S P P A R T R G I S K H S T G T O E G E D O D E P D O G S N D V D L L P R T E I S D K I T S
CA463212 Y V G S L R M F F E D E K P A L L S O D A E F E M G O S P P A R T R G I S K H S T G T O E G E D O D E P D O G S N D V D L L P R T E I S D K I T S
BA078922 E L V S K I G D K N W K I R K E G L D E V A G I I N D A K F I Q P N I G E L P T A L K G R L N D S N K I L Y Q O T L N I L Q O L A V A M G F N I K O H V K N L
CA463212 E L V S K I G D K N W K I R K E G L D E V A G I I N D A K F I Q P N I G E L P T A L K G R L N D S N K I L Y Q O T L N I L Q O L A V A M G F N I K O H V K N L
BA078922 G I P I I T V L G D S K N N V R A A L A T V N A M A E Q T G M K E L E G D L S E E L K E R P F R E L L O W L A K I P T R S T P D L I L O V P
CA463212 G I P I I T V L G D S K N N V R A A L A T V N A M A E Q T G M K E L E G D L S E E L K E R P F R E L L O W L A K I P T R S T P D L I L O V P
BA078922 H L Y S C L E D R N D V Y K K A D A L P F E M H L L Y E M A K A T G K L K T S K D O V A M L E K A K W M P A K F A P P T K A T S P A G S A P
CA463212 H L Y S C L E D R N D V Y K K A D A L P F E M H L L Y E M A K A T G K L K T S K D O V A M L E K A K W M P A K F A P P T K A T S P A G S A P
BA078922 A K F O P A S A P A E D C I S S T E P K P D P K A K A P G L S S K A S A G G K M P S K T L K E D E K S G P I F I V P N G K E O R M K D E K L K
CA463212 A K F O P A S A P A E D C I S S T E P K P D P K A K A P G L S S K A S A G G K M P S K T L K E D E K S G P I F I V P N G K E O R M K D E K L K
BA078922 V L K W N F T T P R D E Y I E O L K T O M S S C V A K W L D E M F H D F Q H N K A L A M V D H L E S E K E G V I G C L O L I L K W L T L R F F O T M T
CA463212 V L K W N F T T P R D E Y I E O L K T O M S S C V A K W L D E M F H D F Q H N K A L A M V D H L E S E K E G V I G C L O L I L K W L T L R F F O T M T
BA078922 S V L M K A L E Y L K L L F L L S E E V H L T E N E A S S F I P Y L V V K Y G E P K D V I R K D V R A I L N R M O L Y V P K M F P F I H E G T K S K N
CA463212 S V L M K A L E Y L K L L F L L S E E V H L T E N E A S S F I P Y L V V K Y G E P K D V I R K D V R A I L N R M O L Y V P K M F P F I H E G T K S K N
BA078922 S K O R A E C L E E L G C L V E S Y G M V C O P T P G K A K E I A V H I G R D N A V R A A L T I V Y N V H G D O V K I L N L S E K D M S M L
CA463212 S K O R A E C L E E L G C L V E S Y G M V C O P T P G K A K E I A V H I G R D N A V R A A L T I V Y N V H G D O V K I L N L S E K D M S M L

```

【 図 10 - 3 】

AIEye 出力(2月21日, 2001 4:29 PM)

FIG. 10-3

```

1430 1440 1450 1460 1470 1480 1490 1500
BA078922 E E R I R R A K R P S A P I K O V E E K P O R A G N I S S N A N L R K G P A E D M S K L N G A R S M S G H P E A A G M V R R E F Q L D L E I N D N
CA463212 E E R I R R A K R P S A P I K O V E E K P O R A G N I S S N A N L R K G P A E D M S K L N G A R S M S G H P E A A G M V R R E F Q L D L E I N D N
BA078922 G Y R C E M P E L V O N K L O D I F E P V L I P E K I R A V S P H F O D M S H T A S T I N F I C O O S S D I N T S I O A L T O D E U L R O E B K A
CA463212 G Y R C E M P E L V O N K L O D I F E P V L I P E K I R A V S P H F O D M S H T A S T I N F I C O O S S D I N T S I O A L T O D E U L R O E B K A
BA078922 E E T W S G H I D O P O C A T O M O L R I L Y N T H M A D E K L E K E I I L Y S C I I G N M I S L F O I E S L A R E A S T G V L K O L M H G L I T L M L D S
CA463212 E E T W S G H I D O P O C A T O M O L R I L Y N T H M A D E K L E K E I I L Y S C I I G N M I S L F O I E S L A R E A S T G V L K O L M H G L I T L M L D S
BA078922 R I E D L E E Q O V I R S N V L V V K Y L E K S O O T N I L S A L L V L O D S L L A T A S S P F S E L V M K O L W R M V R O L P D T I N S I N L D R I
CA463212 R I E D L E E Q O V I R S N V L V V K Y L E K S O O T N I L S A L L V L O D S L L A T A S S P F S E L V M K O L W R M V R O L P D T I N S I N L D R I
BA078922 C L D I H I F M K V F P K E L K O C K S E F F I R T L L H L T L K L G P K I L D H T M I D N K M E S E L E A H C R M K M H S M D O T G S K D K
CA463212 C L D I H I F M K V F P K E L K O C K S E F F I R T L L H L T L K L G P K I L D H T M I D N K M E S E L E A H C R M K M H S M D O T G S K D K
BA078922 E T A K G A S I D A K S S K A R V D F L A E I F K I G S S E N T K E G A E L Y E K K Y S D A D I E P L N S G S F G Y E R L V I E W E
CA463212 E T A K G A S I D A K S S K A R V D F L A E I F K I G S S E N T K E G A E L Y E K K Y S D A D I E P L N S G S F G Y E R L V I E W E
BA078922 R E G K R I S T S G I S P M E V T C P T P I S T V S S I G N T N G E E V G S Y L E R L K I L R O C L O N T K Q D R P P L T S L L S P A V P
CA463212 R E G K R I S T S G I S P M E V T C P T P I S T V S S I G N T N G E E V G S Y L E R L K I L R O C L O N T K Q D R P P L T S L L S P A V P
BA078922 T V A S S T O M L H S K L S Q L R E S R E Q H O S D L S G N O T H S S O T V T S S S T A N I D D L K K L E R I K S S F K
CA463212 T V A S S T O M L H S K L S Q L R E S R E Q H O S D L S G N O T H S S O T V T S S S T A N I D D L K K L E R I K S S F K

```

【 図 1 2 】

The screenshot shows a PubMed search result for the paper: "Characterization of the cDNA and pattern of expression of a new gene over-expressed in human hepatoma and colonic tumors." by Charrier S, Meani M, Tevices S, Berni F, Chau T, Lereaux C. The abstract text is partially visible, mentioning "cDNA library" and "hepatoma and colonic tumors". The interface includes search filters, navigation buttons, and a list of related resources.

FIG. 12

【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 02/070558 A2

- (51) International Patent Classification: C07K 14/705 (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, LU, LV, MA, MD, MG, MK, MN, MW, MX, MY, 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.
- (21) International Application Number: PCT/GB02/00941 (84) Designated States (regional): ARIPO patent (GH, 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, FR, GB, GR, IE, IT, LU, MC, NL, PT, SI, TR), OAPI patent (BF, BJ, CE, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- (22) International Filing Date: 5 March 2002 (05.03.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 0105401.4 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 [US/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB); PHELPS, Christopher, Benjamin [GB/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB).
- (74) Agents: MERCER, Christopher, Paul et al.; Carpmaels & Ransford, 43 Bloomsbury Square, London WC1 2RA (GB).
- 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 02/070558 A2

(54) Title: NUCLEAR HORMONE RECEPTOR LIGAND BINDING DOMAIN

(57) Abstract: The invention relates to a novel protein, termed BAA07892.2, 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 02/070558

PCT/GB02/00941

NUCLEAR HORMONE RECEPTOR LIGAND BINDING DOMAIN

This invention relates to a novel protein, termed BAA07892.2 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
5 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
10 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.

15 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
20 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,
25 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 02/070558

PCT/GB02/00941

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 KIAA0097 protein (NCBI Genebank nucleotide accession number D43948 and a Genebank protein accession number BAA07892.2), is implicated as a novel member of the Nuclear Hormone Receptor Ligand Binding Domain family.

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).¹

WO 02/070558

PCT/GB02/00941

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".

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

WO 02/070558

PCT/GB02/00941

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 |
| | CFI |
| | Knirps |
| <i>Family: DAX1</i> | |
| Subfamilies | DAX1 |

Table 1: Nuclear hormone Receptor Superfamily

WO 02/070558

PCT/GB02/00941

5

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|-------------------|---|---|--|---|---|---|---|--|----|----|----|
| L | Any 2 residues | | L | Any 3 residues (or 2 residues or 4 residues) | | | D | Q | Any 2 residues (or 1 or 3 residues) | | 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.

WO 02/070558

PCT/GB02/00941

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).

5

Table 3. Nuclear Hormone Receptors and disease.

| 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 | Pseudohypaldosteronism (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). |

WO 02/070558

PCT/GB02/00941

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. 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.

10 THE INVENTION

The invention is based on the discovery that the BAA07892.2 protein functions as a Nuclear Hormone Receptor Ligand Binding Domain.

For the BAA07892.2 protein, it has been found that a region including residues 1507-1742 of this protein sequence adopts an equivalent fold to residues 4 (GLN185) to 234 (MET415) of the Human Retinoic Acid Receptor gamma (PDB code 3LBD). Human Retinoic Acid Receptor gamma is known to function as a Nuclear Hormone Receptor Ligand Binding Domain. Furthermore, the "LBD motif" residues PHE251, LEU254, ASP258, GLN259, LEU262 and LEU263 of the Human Retinoic Acid Receptor gamma are conserved as VAL1573, LEU1574, ASP1578, LYS1579, ALA1582 and MET1583 in BAA07892.2, respectively. This relationship is not just to Human Retinoic Acid Receptor gamma, but rather to the Nuclear Hormone Receptor Ligand Binding Domain family as a whole. Thus, by reference to the Genome Threader™ alignment of BAA07892.2 with the Human Retinoic Acid Receptor gamma (3LBD) VAL1573, LEU1574, ASP1578, LYS1579, ALA1582 and MET1583 of BAA07892.2 are predicted to form the "LBD motif" residues.

The combination of equivalent fold and conservation of "LBD motif" residues allows the functional annotation of this region of BAA07892.2, 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:

WO 02/070558

PCT/GB02/00941

8

- (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 activity or having an antigenic determinant in common with the polypeptides of (i); or
- 5 (iii) is a functional equivalent of (i) or (ii).

Preferably, the polypeptide:

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

The polypeptide having the sequence recited in SEQ ID NO:2 is referred to hereafter as "the LBDS1 polypeptide".

According to this aspect of the invention, a preferred polypeptide fragment according to

15 part ii) above includes the region of the LBDS1 polypeptide that is predicted as that responsible for Nuclear Hormone Receptor Ligand Binding Domain activity (hereafter, the "LBDS1 Nuclear Hormone Receptor Ligand Binding Domain region"), or is a variant thereof that possesses the "LBD motif" (VAL1573, LEU1574, ASP1578, LYS1579, ALA1582 and MET1583, or equivalent residues). As defined herein, the LBDS1 Nuclear

20 Hormone Receptor Ligand Binding Domain region is considered to extend between residue 1507 and residue 1742 of the LBDS1 polypeptide sequence.

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

25 Domain.

In a second aspect, the invention provides a purified nucleic acid molecule that encodes a 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 LBDS1

WO 02/070558

PCT/GB02/00941

9

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 LBDS1 Nuclear Hormone Receptor Ligand Binding Domain region, or that encodes a variant of these fragments as this term is defined above.

5 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.

10 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.

15 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.

20 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 LBDS1 Nuclear Hormone Receptor Ligand Binding Domain region of the LBDS1 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
25 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. These

WO 02/070558

PCT/GB02/00941

10

molecules may also be used in the manufacture of a medicament for the treatment of cell proliferative disorders, including neoplasm, cancer and myeloproliferative disorders and angiogenesis, metabolic disorders, including obesity, osteoporosis, non-insulin dependent diabetes, lipid metabolism disorder, hyperthyroidism, hyperparathyroidism, hypercalcemia, hypercholesterolemia, hyperlipidemia, cardiovascular disorders including, hypertension, atherosclerosis, hyperinsulinemia, heart arrhythmia, renal disorders, including glomerulonephritis, renovascular hypertension, skin disorders including dermatological disease, psoriasis, acne, aging, eczema, wound healing, inflammation, including inflammatory bowel disease, respiratory tract inflammation, emphysema, asthma, immune disorder, autoimmune disease, multiple sclerosis, nervous system disorders, including anxiety, depression, neurodegenerative disease, Alzheimer's disease Parkinson's disease, brain injury, stroke and pain, infection, including virus infection, and other conditions in which nuclear hormones are implicated.

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 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 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.

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

WO 02/070558

PCT/GB02/00941

11

on a short or long term basis to allow therapeutic treatment of a disease to be monitored 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 aspects of the invention to express a protein that possesses Nuclear Hormone Receptor 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.

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.

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, cancer and myeloproliferative disorders and angiogenesis, metabolic disorders, including obesity, osteoporosis, non-insulin dependent diabetes, lipid metabolism disorder, hyperthyroidism, hyperparathyroidism, hypercalcemia, hypercholesterolemia, hyperlipidemia, cardiovascular disorders including hypertension, atherosclerosis, hyperinsulinemia, heart arrhythmia, renal disorders, including glomerulonephritis, renovascular hypertension, skin disorders including dermatological disease, psoriasis, acne, aging, eczema, wound healing, inflammation, including inflammatory bowel disease, respiratory tract inflammation, emphysema, asthma, immune disorder, autoimmune disease, multiple sclerosis, nervous system disorders, including anxiety, depression, neurodegenerative disease, Alzheimer's disease

WO 02/070558

PCT/GB02/00941

12

Parkinson's disease, brain injury, stroke and pain, infection, including virus infection, and other conditions in which nuclear hormones are implicated.

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, 5 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 10 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, 15 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 20 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 25 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.

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

WO 02/070558

PCT/GB02/00941

13

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

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).

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.

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

WO 02/070558

PCT/GB02/00941

14

recombinant production. Alternatively or additionally, the mature polypeptide may be fused with another compound, such as a compound to increase the half-life of the 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

WO 02/070558

PCT/GB02/00941

15

polypeptides include isolated naturally-occurring polypeptides (for example purified from cell culture), recombinantly-produced polypeptides (including fusion proteins), synthetically-produced polypeptides or polypeptides that are produced by a combination of these methods.

5 The functionally-equivalent polypeptides of the first aspect of the invention may be polypeptides that are homologous to the LBDS1 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,
10 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,
20 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 LBDS1 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
25 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
30 substituted, deleted or added in any combination. Especially preferred are silent substitutions, additions and deletions, which do not alter the properties and activities of

WO 02/070558

PCT/GB02/00941

16

the protein. Also especially preferred in this regard are conservative substitutions.

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 LBDS1 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 LBDS1 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/>) [Blosun 62 matrix; gap open penalty=11 and gap extension penalty=1].

In the present case, preferred active fragments of the LBDS1 polypeptide are those that include the LBDS1 Nuclear Hormone Receptor Ligand Binding Domain region and which possess the "LBD motif" of residues VAL1573, LEU1574, ASP1578, LYS1579, ALA1582 and MET1583, 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 VAL1573 may be replaced by LEU, ILE, ALA, MET, PHE, TYR or TRP. For example LEU1574 may be replaced by ILE, ALA, VAL, MET, PHE, TYR or TRP. For example ASP1578 may be replaced by GLU. For example LYS1579 may be replaced by GLN, ASN, ARG, HIS, SER or THR. For example ALA1582 may be replaced by LEU, ILE, VAL, MET, PHE, TYR or TRP. For example MET1583 may be replaced by LEU, ILE, ALA, VAL, 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 LBDS1 polypeptide and which possess the "LBD motif" of VAL1573, LEU1574,

WO 02/070558

PCT/GB02/00941

17

ASP1578, LYS1579, ALA1582 and MET1583, or equivalent residues. As discussed above, the LBDS1 Nuclear Hormone Receptor Ligand Binding Domain region is considered to extend between residue 1507 and residue 1742 of the LBDS1 polypeptide sequence.

5 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
10 polypeptides of presently-unknown function which, while having low sequence identity as compared to the LBDS1 polypeptide, are predicted to have Nuclear Hormone Receptor Ligand Binding Domain activity, by virtue of sharing significant structural homology with the LBDS1 polypeptide sequence.

By "significant structural homology" is meant that the Inpharmatica Genome Threader™
15 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
20 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
25 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.

30 Structural homologues of LBDS1 should share structural homology with the LBDS1

WO 02/070558

PCT/GB02/00941

18

Nuclear Hormone Receptor Ligand Binding Domain region and possess the "LBD motif" residues VAL1573, LEU1574, ASP1578, LYS1579, ALA1582 and MET1583, or equivalent residues. Such structural homologues are predicted to have Nuclear Hormone Receptor Ligand Binding Domain activity by virtue of sharing significant structural
5 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 LBDS1 polypeptide, functional equivalents of the fragments of the LBDS1 polypeptide, and fragments of the functional equivalents of the LBDS1 polypeptides, provided that those functional equivalents and fragments retain Nuclear Hormone Receptor Ligand Binding
10 Domain activity or have an antigenic determinant in common with the LBDS1 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 LBDS1 polypeptides or one of its functional equivalents. The fragments should comprise at least
15 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 LBDS1 Nuclear Hormone Receptor Ligand
20 Binding Domain region of the LBDS1 polypeptides, respectively. These regions are the regions that have been annotated as Nuclear Hormone Receptor Ligand Binding Domain.

For the LBDS1 polypeptide, this region is considered to extend between residue 1507 and residue 1742.

Variants of this fragment are included as embodiments of this aspect of the invention,
25 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.

WO 02/070558

PCT/GB02/00941

19

For extended variants, it is considered highly likely that the Nuclear Hormone Receptor Ligand Binding Domain region of the LBDS1 polypeptide will fold correctly and show Nuclear Hormone Receptor Ligand Binding Domain activity if additional residues C terminal and/or N terminal of these boundaries in the LBDS1 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 LBDS1 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 LBDS1 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 LBDS1 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 LBDS1 polypeptide, although the "LBD motif" residues (VAL1573, LEU1574, ASP1578, LYS1579, ALA1582 and MET1583), 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 LBDS1 polypeptide and which possess the "LBD motif" residues (VAL1573, LEU1574, ASP1578, LYS1579, ALA1582 and MET1583), 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 30% identity with the LBDS1 Nuclear Hormone Receptor Ligand Binding Domain regions, of the LBDS1 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/>) [Blosum 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 LBDS1 Nuclear Hormone Receptor Ligand Binding Domain regions, of the LBDS1

WO 02/070558

PCT/GB02/00941

20

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 LBDS1 Nuclear Hormone Receptor Ligand Binding Domain regions of the LBDS1, polypeptides, provided that said variants retain activity as a Nuclear Hormone Receptor
5 Ligand Binding Domain. Variant polypeptides also include homologues of the truncated forms of the polypeptide fragments discussed above, provided that said variants retain 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
10 polypeptide fragment defined by the LBDS1 Nuclear Hormone Receptor Ligand Binding Domain regions, of the LBDS1 polypeptide sequence, for example, as identified by the Inpharmatica Genome Threader™. Accordingly, polypeptide fragments that are structural homologues of the polypeptide fragments defined by the LBDS1 Nuclear Hormone Receptor Ligand Binding Domain regions of the LBDS1 polypeptide sequence should
15 adopt the same fold as that adopted by this polypeptide fragment, as this fold is defined above.

Structural homologues of the polypeptide fragment defined by the LBDS1 Nuclear Hormone Receptor Ligand Binding Domain region should also retain the "LBD motif" residues VAL1573, LEU1574, ASP1578, LYS1579, ALA1582 and MET1583, or
20 equivalent residues.

Such fragments may be "free-standing", i.e. not part of or fused to other amino acids or 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
25 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 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
30 least one antigenic determinant) can be used to generate ligands, such as polyclonal or

WO 02/070558

PCT/GB02/00941

21

monoclonal antibodies, that are immunospecific for the polypeptides. Such antibodies may be employed to isolate or to identify clones expressing the polypeptides of the 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
5 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
10 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
15 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
20 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).
25

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
30 polypeptides against which they are directed. Alternatively, genes encoding the

WO 02/070558

PCT/GB02/00941

22

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.

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); Quesen *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.

WO 02/070558

PCT/GB02/00941

23

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 equivalent polypeptides, including active fragments of the LBDS1 polypeptide, such as a fragment including the LBDS1 Nuclear Hormone Receptor Ligand Binding Domain region of the LBDS1 polypeptide sequence, or a homologue thereof.

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 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).

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 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.

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

WO 02/070558

PCT/GB02/00941

24

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 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 LBDS1 polypeptide, such as a fragment including the LBDS1 Nuclear Hormone Receptor Ligand Binding Domain region, or a homologue thereof. The LBDS1 Nuclear Hormone Receptor Ligand Binding Domain region is considered to extend between residue 1507 and residue 1742 of the LBDS1 polypeptide sequence. In SEQ ID NO:1 the LBDS1 Nuclear Hormone Receptor Ligand Binding Domain region is thus encoded by a nucleic acid molecule including nucleotide 4548 to 5255. 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

WO 02/070558

PCT/GB02/00941

25

the first aspect of the invention.

As discussed above, a preferred fragment of the LBDS1 polypeptide is a fragment including the LBDS1 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 4548 to 5255 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

WO 02/070558

PCT/GB02/00941

26

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 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).

WO 02/070558

PCT/GB02/00941

27

"Stringency" refers to conditions in a hybridization reaction that favour the association of very similar molecules over association of molecules that differ. High stringency 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 Denhardt's solution, 10% dextran sulphate, and 20 microgram/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 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.

10 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 LBDS1 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 LBDS1 Nuclear Hormone Receptor Ligand Binding Domain region of the LBDS1 polypeptide sequences, respectively. Accordingly, preferred nucleic acid molecules include those that are at least 80% identical over their entire length to a nucleic acid molecule encoding the Nuclear Hormone Receptor Ligand Binding Domain region of the LBDS1 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/>).

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 4548-5255 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 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 LBDS1 polypeptide.

WO 02/070558

PCT/GB02/00941

28

The invention also provides a process for detecting a nucleic acid molecule of the invention, comprising the steps of: (a) contacting a nucleic probe according to the invention with a biological sample under hybridizing conditions to form duplexes; and (b) detecting any such duplexes that are formed.

5 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 hybridization probe for RNA, cDNA or genomic DNA, in order to isolate full-length cDNAs and genomic clones encoding the LBDS1 polypeptide and to isolate cDNA and genomic clones of homologous or orthologous genes that have a high sequence similarity
10 to the gene encoding this polypeptide.

In this regard, the following techniques, among others known in the art, may be utilised 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
15 employ such enzymes as the Klenow fragment of DNA polymerase I, Sequenase (US Biochemical Corp, Cleveland, OH), Taq polymerase (Perkin Elmer), thermostable T7 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
20 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 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 LBDS1 polypeptide, particularly with an equivalent
25 function to the LBDS1 Nuclear Hormone Receptor Ligand Binding Domain region of the LBDS1 polypeptide, is to probe a genomic or cDNA library with a natural or artificially-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
30 least 15, preferably at least 30, and more preferably at least 50, contiguous bases that

WO 02/070558

PCT/GB02/00941

29

correspond to, or are complementary to, nucleic acid sequences from the appropriate encoding gene (SEQ ID NO:1), particularly a region from nucleotides 4548-5255 of SEQ 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. 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.

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 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, 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 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

WO 02/070558

PCT/GB02/00941

30

libraries and is useful in finding intron/exon junctions.

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

WO 02/070558

PCT/GB02/00941

31

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

5 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 prokaryotic or eukaryotic.

The polypeptides of the invention may be prepared in recombinant form by expression of
10 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 expression systems. Using nature for the art of expression". Academic Press, San Diego, London, Boston, New York, Sydney, Tokyo, Toronto).

15 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 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
20 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.

Examples of suitable expression systems include, for example, chromosomal, episomal and virus-derived systems, including, for example, vectors derived from: bacterial
25 plasmids, bacteriophage, transposons, yeast episomes, insertion elements, yeast chromosomal elements, viruses such as baculoviruses, papova viruses such as SV40, 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
30 also be employed to deliver larger fragments of DNA than can be contained and

WO 02/070558

PCT/GB02/00941

32

expressed in a plasmid.

Particularly suitable expression systems include microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with 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.

10 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, transfection, 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.

20 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.

WO 02/070558

PCT/GB02/00941

33

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 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

WO 02/070558

PCT/GB02/00941

34

and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for 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
5 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
10 (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
15 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.

20 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.

25 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.

30 Examples of particularly preferred bacterial host cells include streptococci,

WO 02/070558

PCT/GB02/00941

35

staphylococci, *E. coli*, Streptomyces and *Bacillus subtilis* cells.

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

5 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 aprt⁺ cells, respectively.

10 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
15 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
20 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.

25 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
30 include membrane, solution, or chip based technologies for the detection and/or

WO 02/070558

PCT/GB02/00941

36

quantification of nucleic acid or protein (see Hampton, R. *et al.* (1990) Serological 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 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, 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).

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 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-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

WO 02/070558

PCT/GB02/00941

37

employed to regenerate an active conformation when the polypeptide is denatured during 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
5 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 FLAG extension/affinity purification system (Immunex Corp.,
10 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
15 (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.*
20 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
25 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
30 (antagonise) the level of expression of the gene or the activity of the polypeptide of the

WO 02/070558

PCT/GB02/00941

38

invention and form a further aspect of the present invention. Preferred compounds are 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.

5 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).

10 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
15 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
20 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
25 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
30 indirectly associated with the test compound or in an assay involving competition with a

WO 02/070558

PCT/GB02/00941

39

labelled competitor. In another embodiment, competitive drug screening assays may be 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.

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 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 (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 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 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

WO 02/070558

PCT/GB02/00941

40

antagonists of the polypeptide, that compete with the binding of the polypeptide to its 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/or enzymes that are described above.

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 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 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 therapeutic agent needed to treat, ameliorate, or prevent a targeted 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 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,

WO 02/070558

PCT/GB02/00941

41

time and frequency of administration, drug combination(s), reaction sensitivities, and 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-

WO 02/070558

PCT/GB02/00941

42

arterial, intramedullary, intrathecal, intraventricular, transdermal or transcutaneous 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, 5 the therapeutic compositions may be prepared as injectables, either as liquid solutions or 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 10 interstitial space of a tissue. The compositions can also be administered into a lesion. 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 15 acceptable carrier in an amount effective to inhibit the function of the polypeptide, such 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.

20 In another approach, soluble forms of the polypeptide that retain binding affinity for the 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 25 inhibited using expression blocking techniques, such as the use of antisense nucleic acid 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 30 polypeptide. Similarly, inhibition can be achieved using "triple helix" base-pairing

WO 02/070558

PCT/GB02/00941

43

methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature (Gee, J.E. *et al.* (1994) In: Huber, B.E. and B.I. Carr, 5 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 10 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 15 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 20 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-, 25 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 30 the polypeptide, i.e., an agonist as described above, to alleviate the abnormal condition.

WO 02/070558

PCT/GB02/00941

44

Alternatively, a therapeutic amount of the polypeptide in combination with a suitable 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 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 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, 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, 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 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

WO 02/070558

PCT/GB02/00941

45

antibodies against the disease causing agent.

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 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.

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 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. 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 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.

Nucleic acid molecules for diagnosis may be obtained from a subject's cells, such as from

WO 02/070558

PCT/GB02/00941

46

blood, urine, saliva, tissue biopsy or autopsy material. The genomic DNA may be used 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

WO 02/070558

PCT/GB02/00941

47

amplified DNA to labelled RNA of the invention or alternatively, labelled antisense DNA 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.

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

WO 02/070558

PCT/GB02/00941

48

microdeletions, aneuploidies, translocations, inversions, can also be detected by in situ 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.

5 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
10 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
15 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
20 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
25 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
30 itself to the efficient use of commercially-available instrumentation.

WO 02/070558

PCT/GB02/00941

49

In addition to the methods discussed above, diseases may be diagnosed by methods 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
5 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
10 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.

15 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
20 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,
25 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
30 them, either covalently or non-covalently, with a reporter molecule. A wide variety of

WO 02/070558

PCT/GB02/00941

50

reporter molecules known in the art may be used, several of which are described above.

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:

- 10 (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.

25 Such kits will be of use in diagnosing a disease or susceptibility to disease, particularly cell proliferative disorders, including neoplasm, cancer and myeloproliferative disorders and angiogenesis, metabolic disorders, including obesity, osteoporosis, non-insulin dependent diabetes, lipid metabolism disorder, hyperthyroidism, hyperparathyroidism, hypercalcemia, hypercholesterolemia, hyperlipidemia, cardiovascular disorders including,

WO 02/070558

PCT/GB02/00941

51

hypertension, atherosclerosis, hyperinsulinemia, heart arrhythmia, renal disorders, including glomerulonephritis, renovascular hypertension, skin disorders including dermatological disease, psoriasis, acne, aging, eczema, wound healing, inflammation, including inflammatory bowel disease, respiratory tract inflammation, emphysema, 5 asthma, immune disorder, autoimmune disease, multiple sclerosis, nervous system disorders, including anxiety, depression, neurodegenerative disease, Alzheimer's disease Parkinson's disease, brain injury, stroke and pain, infection, including virus infection, and other conditions in which nuclear hormones are implicated.

Various aspects and embodiments of the present invention will now be described in more 10 detail by way of example, with particular reference to the LBDS1 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 15 the database is initiated using the PDB code "3LBD".

Figure 2A: A selection is shown of the Inpharmatica Genome Threader results for the search using 3LBD. The arrow indicates *Homo Sapiens* Retinoic Acid Receptor Gamma-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 20 search using 3LBD. The arrow indicates BAA07892.2 (LBDS1).

Figure 2C: Full list of forward PSI-BLAST results for the search using 3LBD. BAA07892.2 (LBDS1) is not identified.

Figure 3: The Redundant Sequence Display results page for BAA07892.2 (LBDS1). There is an equivalent sequence called Q14008 (see arrow).

25 Figure 4: InterPro PFAM search results for BAA07892.2 (LBDS1), see arrow Ⓞ.

Figure 5A: NCBI protein report for BAA07892.2 (LBDS1).

Figure 5B: SWISS-PROT entry for a sequence equivalent to BAA07892.2 (LBDS1), called Q14008, CH-TOG protein.

WO 02/070558

PCT/GB02/00941

52

Figure 6A: This is the front end of the Biopendium database. A search of the database is initiated using BAA07892.2 (LBDS1), as the query sequence.

Figure 6B: A selection of the Inpharmatica Genome Threader results of search using BAA07892.2 (LBDS1), as the query sequence. The arrow points to 3LBD.

- 5 Figure 6C: A selection of the reverse-maximised PSI-BLAST results obtained using BAA07892.2 (LBDS1), as the query sequence. The arrows numbered ① and ② point to homologues of BAA07892.2 (LBDS1) in *X.laevis* and *D.melanogaster* respectively. Arrow ③ points to a *Homo sapiens* variant of BAA07892.2 (LBDS1).

Figure 7: AIEye sequence alignment of BAA07892.2 (LBDS1) and 3LBD.

- 10 Figure 8A: LigEye for 3LBD that illustrates the sites of interaction of 9-cis Retinoic Acid (REA420) with the Ligand Binding Domain of *Homo sapiens* Retinoic Acid Receptor Gamma, 3LBD.

Figure 8B: iRasMol view of 3LBD, the Ligand Binding Domain of *Homo sapiens* Retinoic Acid Receptor Gamma.

- 15 Figure 9: AIEye sequence alignment of BAA07892.2 (LBDS1) with *X.laevis* (CAB61894.1) and *D.melanogaster* (CAB55772.1) homologues of BAA07892.2 (LBDS1).

Figure 10: AIEye sequence alignment of BAA07892.2 (LBDS1) with CAA63212.1, a *Homo sapiens* variant of BAA07892.2 (LBDS1).

- 20 Figure 11: NCBI protein report for CAA63212.1, a *Homo sapiens* variant of BAA07892.2 (LBDS1). CAA63212.1 was derived from a human breast cancer cell library (arrow).

Figure 12: Abstract of the article by by Charrasse *et al.* in the Eur. J. Biochem. 1995 Dec 1;234(2)406-13.

- 25 **Example 1: BAA07892.2 (LBDS1)**

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* Retinoic Acid Receptor Gamma. More specifically, the search is

WO 02/070558

PCT/GB02/00941

53

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* Retinoic Acid Receptor Gamma (PDB code 3LBD; see Figure 1).

- 5 A search of the Biopendium (using the Target Mining Interface) for relatives of 3LBD takes place and returns 4475 Genome Threader results. The 4475 Genome Threader results include examples of typical Nuclear Hormone Receptor Ligand Binding Domains, such as that found between residues 182-418 of the *Homo sapiens* Retinoic Acid Receptor Gamma-1 (see arrow in Figure 2A).
- 10 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, BAA07892.2 (LBDS1; see arrow in Figure 2B). The Inpharmatica Genome Threader has identified a region of the sequence BAA07892.2 (LBDS1), between residues 1507-1742, as having a structure similar to the Ligand
- 15 Binding Domain of *Homo sapiens* Retinoic Acid Receptor Gamma. The possession of a structure similar to the Ligand Binding Domain of *Homo sapiens* Retinoic Acid Receptor Gamma suggests that residues 1507-1742 of BAA07892.2 (LBDS1) function as a Nuclear Hormone Receptor Ligand Binding Domain. The Genome Threader identifies this with 71% confidence.
- 20 The search of the Biopendium (using the Target Mining Interface) for relatives of 3LBD also returns 850 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 BAA07892.2 (LBDS1) as containing a Nuclear Hormone Receptor Ligand Binding Domain.
- 25 In order to assess what is known in the public domain databases about BAA07892.2 (LBDS1) the Redundant Sequence Display Page (Figure 3) is viewed. The Redundant Sequence Display Page reveals a near-identical sequence to BAA07892.2 (LBDS1) called Q14008. Q14008 is a sub-sequence of BAA07892.2 (LBDS1), lacking only the six N-terminal residues of BAA07892.2.

WO 02/070558

PCT/GB02/00941

54

There are no PROSITE or PRINTS hits which identify BAA07892.2 (LBDS1) 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

5 BAA07892.2 (LBDS1) 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 BAA07892.2 (LBDS1) as containing a Nuclear Hormone Receptor Ligand Binding Domain, the BAA07892.2 (LBDS1) protein sequence is searched against the PFAM

10 database (Protein Family Database of Alignment and hidden Markov models) at the InterPro website (see Figure 4 arrow ①). No PFAM-A matches are returned. Thus PFAM does not identify BAA07892.2 (LBDS1) as containing a Nuclear Hormone Receptor Ligand Binding Domain.

Interestingly, PROSITE PFscan (see Figure 4 arrow ②) identifies a bipartite nuclear

15 localisation signal in BAA07892.2 (LBDS1) at residues 550-567. 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) Genebank protein database is then viewed to examine if there is any further information that is known in the public

20 domain relating to BAA07892.2 (LBDS1). This is the US public domain database for protein and gene sequence deposition (Figure 5A). BAA07892.2 (LBDS1) is a *Homo sapiens* sequence, its Genebank protein ID is BAA07892.2 and it is 2038 amino acids in length. BAA07892.2 (LBDS1) was cloned by a group of scientists at the Kazusa DNA Research Institute, Kisarazu, Japan. The public domain information for this gene does not

25 annotate it as containing a Nuclear Hormone Receptor Ligand Binding Domain.

The Swiss Institute of Bioinformatics (SIB) SWISS-PROT protein database is then viewed to examine if there is any further information that is known in the public domain relating to the BAA07892.2 (LBDS1) subsequence Q14008. This is the Swiss public

30 domain database for protein and gene sequence deposition (Figure 5B). The BAA07892.2 (LBDS1) subsequence Q14008 is a *Homo Sapiens* sequence, its SWISS-PROT protein ID

WO 02/070558

PCT/GB02/00941

55

is Q14008 and it is 2032 amino acids in length. Q14008 is called CH-TOG protein (Colonic and Hepatic Tumour Over-Expressed Gene) protein. Q14008, CH-TOG was cloned by Charrasse *et al.* at INSERM, Montpellier, France. 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, BAA07892.2 (LBDS1) 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. BAA07892.2 (LBDS1) is now used as the query sequence in the Biopendium (see Figure 6A). The Inpharmatica Genome Threader identifies residues 1507-1742 of BAA07892.2 (LBDS1) as having a structure that is the same as the Ligand Binding Domain of *Homo sapiens* Retinoic Acid Receptor Gamma with 71% confidence (see arrow in Figure 6B). The Ligand Binding Domain of *Homo sapiens* Retinoic Acid Receptor Gamma (3LBD) 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* Retinoic Acid Receptor Gamma Ligand Binding Domain is chosen against which to view the sequence alignment of BAA07892.2 (LBDS1). Viewing the AlEye 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* Retinoic Acid Receptor Gamma 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 residues PHE251, LEU254, ASP258, GLN259, LEU262 and LEU263 constitute this motif in the *Homo sapiens* Retinoic Acid Receptor Gamma Ligand Binding Domain (see square boxes Figure 7). The residues VAL1573, LEU1574, ASP1578, LYS1579, ALA1582 and MET1583 in BAA07892.2 (LBDS1) align with and satisfy the LBD motif criteria (full and broken square boxes in Figure 7). This indicates that BAA07892.2 (LBDS1) contains a Nuclear Hormone

WO 02/070558

PCT/GB02/00941

56

Receptor Ligand Binding Domain similar to The *Homo sapiens* Retinoic Acid Receptor Gamma 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* Retinoic Acid Receptor Gamma Ligand Binding Domain reveals 14 residues which bind 9-cis Retinoic Acid (circled in Figure 7). Only 3 (ALA1556, LEU1591, LEU1725, circled in Figure 7) of these 14 residues are precisely conserved in BAA07892.2 (LBDS1), suggesting that BAA07892.2 (LBDS1) may not bind 9-cis Retinoic Acid in particular. However, there are 11 hydrophobic residues which line the pocket of the *Homo sapiens* Retinoic Acid Receptor Gamma Ligand Binding Domain; PHE230, LEU233, ALA234, LEU268, LEU271, MET272, ILE275, PHE288, PHE304, LEU400 and ILE412. In BAA07892.2 (LBDS1), 3 (ALA1556, LEU1591, LEU1725, circled in Figure 7) of these 11 residues are perfectly conserved. A further 5 residues (ILE1552, VAL1555, ILE1592, PHE1595, LEU1739, broken circles in Figure 7) are conserved in hydrophobic nature. This conservation of hydrophobicity in 8 out of the 11 hydrophobic residues which line the binding pocket indicates that BAA07892.2 (LBDS1) will bind a hydrophobic steroid-like hydrophobic ligand. Typically, the binding pocket of Nuclear Hormone Receptor Ligand Binding Domains will contain a crucial arginine residue. The guanidinium group of this arginine functions to bind and stabilise a hydroxyl or carboxyl group present on the steroid/steroid-like ligand. For example, ARG394 of the *Homo sapiens* Estrogen Receptor Ligand Binding Domain binds and stabilises the phenolic hydroxyl group of the A ring of Estradiol (Brzozowski Nature 1997 vol.389:p753-758). LigEye shows that ARG278 of the *Homo sapiens* Retinoic Acid Receptor Gamma Ligand Binding Domain stabilises the carboxylate group of Retinoic Acid. In BAA07892.2 (LBDS1) this ARG is not conserved, but there is ARG1599 (asterisk) which is displaced one residue towards the C-terminus. ARG1599 in

WO 02/070558

PCT/GB02/00941

57

BAA07892.2 (LBDS1) could function to bind and stabilise a hydroxyl or carboxyl group present on a steroid-like ligand. This indicates that indeed as predicted by the Inpharmatica Genome Threader, BAA07892.2 (LBDS1) folds in a similar manner to of the *Homo sapiens* Retinoic Acid Receptor Gamma Ligand Binding Domain and as such is identified as containing a Nuclear Hormone Receptor Ligand Binding Domain.

Reverse-maximised PSI-BLAST of BAA07892.2 (LBDS1) identifies a *X.laevis* homologue of BAA07892.2 (LBDS1) called CAB61894.1, XMAP215, which has 76.0% sequence identity to BAA07892.2 (LBDS1), see Figure 6C arrow Ⓞ. Reverse-maximised PSI-BLAST of BAA07892.2 (LBDS1) also identifies a *D.melanogaster* homologue of BAA07892.2 (LBDS1) called CAB55772.1, Microtubule Associated Protein, which has 45.0% sequence identity to BAA07892.2 (LBDS1), see Figure 6C arrow Ⓟ.

BAA07892.2 (LBDS1), *X.laevis* CAB61894.1 (XMAP215) and *D.melanogaster* CAB55772.1 (Microtubule Associated Protein) are aligned and viewed in AIEye (Figure 9). AIEye reveals that the 8 hydrophobic residues of BAA07892.2 (LBDS1) predicted to line the ligand binding pocket (ILE1552, VAL1555, ALA1556, LEU1591, ILE1592, PHE1595, LEU1725 and LEU1739) are all precisely conserved in *X.laevis* CAB61894.1, XMAP215 (see Figure 9). Furthermore, the six residues matching the "LBD motif" in BAA07892.2 (LBDS1), (VAL1573, LEU1574, ASP1578, LYS1579, ALA1582, MET1583) are all precisely conserved in *X.laevis* CAB61894.1, XMAP215 (see Figure 9). In addition the predicted ligand binding ARG1599 is precisely conserved in *X.laevis* CAB61894.1, XMAP215 (marked with an asterisk, see Figure 9). Residues which are essential for the function of a protein will be conserved in homologues of that protein. Thus the absolute conservation of all the residues which would be essential for the function of the predicted BAA07892.2 (LBDS1) Nuclear Hormone Receptor Ligand Binding Domain in the *X.laevis* homologue CAB61894.1 (XMAP215) strongly supports the annotation of BAA07892.2 (LBDS1) as containing a Nuclear Hormone Receptor Ligand Binding Domain.

The AIEye alignment also reveals conserved relationships between BAA07892.2 (LBDS1) and *D.melanogaster* CAB55772.1 (Microtubule Associated Protein). AIEye indicates that 2 (ILE1552 and VAL1555) of the 8 hydrophobic residues of BAA07892.2

WO 02/070558

PCT/GB02/00941

58

(LBDS1) are precisely conserved in *D.melanogaster* CAB55772.1 (Microtubule Associated Protein). Another 4 of the 8 BAA07892.2 (LBDS1) hydrophobic residues (LEU1591, PHE1595, LEU1725, LEU1739) are conservatively substituted for other hydrophobic residues in *D.melanogaster* CAB55772.1 (Microtubule Associated Protein).

5 Furthermore, the Lys1579 of the BAA07892.2 (LBDS1) "LBD motif" is precisely conserved in *D.melanogaster* CAB55772.1 (Microtubule Associated Protein). The other five residues of the BAA07892.2 (LBDS1) "LBD motif" (VAL1573, LEU1574, ASP1578, ALA1582, MET1583) are conservatively substituted in *D.melanogaster* CAB55772.1 (Microtubule Associated Protein). In addition the predicted ligand binding

10 ARG1599 is precisely conserved in *D.melanogaster* CAB55772.1 (Microtubule Associated Protein), marked with an asterisk, see Figure 9. Residues which are essential for the function of a protein will be conserved in homologues of that protein. Thus the precise conservation of 4 (and conservative substitution of a further 9) of the 15 residues which would be essential for the function of the predicted BAA07892.2 (LBDS1)

15 Nuclear Hormone Receptor Ligand Binding Domain in the *D.melanogaster* homologue CAB55772.1 (Microtubule Associated Protein) supports the annotation of BAA07892.2 (LBDS1) as containing a Nuclear Hormone Receptor Ligand Binding Domain.

Reverse-maximised PSI-BLAST of BAA07892.2 (LBDS1) also identifies a *Homo sapiens* sequence CAA63212.1 which has 95% sequence identity to BAA07892.2 (LBDS1), see Figure 6C arrow ③. An AlEye alignment of CAA63212.1 with

20 BAA07892.2 (LBDS1) reveals that there are only two regions of difference (see Figure 10). The first difference is that CAA63212.1 lacks the 6 N-terminal amino acids of BAA07892.2 (LBDS1). The second difference is that CAA63212.1 has an internal deletion compared to BAA07892.2 (LBDS1) which removes 60 amino acids. This

25 internal deletion may represent a splice variant or disease-linked deletion since the cDNA was obtained from breast cancer tissue (see Figure 11 arrow). Interestingly this 60 amino acid deletion means that CAA63212.1 lacks the "LBD motif", ARG1599 and three hydrophobic predicted ligand binding residues (LEU1591, ILE1592 and PHE1595). As such CAA63212.1 would be predicted to be non-competent at binding a steroid-like

30 ligand and recruiting Nuclear Hormone Co-Activators and Co-Repressors.

WO 02/070558

PCT/GB02/00941

59

There is an associated article by Charrasse *et al.* in the Eur. J. Biochem. 1995 Dec 1;234(2)406-13 (Figure 12). This article demonstrates that mRNA encoding the BAA07892.2 (LBDS1) subsequence Q14008, CH-TOG is over-expressed in cancerous colon tissue compared to healthy colon. In addition, mRNA encoding the BAA07892.2 (LBDS1) subsequence Q14008, CH-TOG is over-expressed in cancerous liver tissue compared to healthy colon. This establishes a direct correlation between levels of the BAA07892.2 (LBDS1) subsequence Q14008, CH-TOG and the cancerous state. Connecting BAA07892.2 (LBDS1) to the cancerous state paves the way for the development of agents that target the BAA07892.2 (LBDS1) gene or its encoded protein, to diagnose and/or treat these conditions. In particular, the identification of this gene as containing a Nuclear Hormone Receptor Ligand Binding Domain facilitates the development of agents that are specific to the BAA07892.2 (LBDS1) protein, for example, through the use of Nuclear Hormone Receptor Ligand Binding Domain agonists or antagonists.

WO 02/070558

PCT/GB02/00941

60

Sequence Listing

SEQ ID NO: 1 Nucleotide coding sequence for BAA07892.2 (LBDS1) protein (nucleotides 30-6146 translate to give SEQ ID NO:2)

```

5 1   ccagctgag gaaatactct taattctaag gaaaacctgg aagcacaatg ggagatgaca
61  gtgagtggtt gaaactgcca gttgatcaga aatgtgaaca caagctgtgg aaagcaaggt
121 taagtgggta tgaagaggcc ctgaagatct tccagaaaat aaaggatgaa aagagcccag
181 agtgggccaa atttttagga ttgatcaaaa aatttgtcac tgattccaat gcagtggttc
241 aatggaaggg attagaagct gcacttgttt atgttgaaaa tgcccatgta gcaggaaaaa
10 301 ccacaggaga agttgtgtca ggtgtgttaa gtaaggtgtt caatcaacct aaagctaaag
361 ccaaggagct gggcatagag atctgtctta tgtacataga gatttgaaaa ggagaggctg
421 ttcaagaaga gctcctgaaa ggcttggaaca ataagaatcc caagatcata gtggcctgta
481 tggagacact gagggaagcc ttaagtgaat ttggttccaa aatcatcttg cttaaGCCaa
541 ttatcaaaat gttgccaaaa ctctttgagt ctcgagagaa ggcgtgtcga gatgaagcca
15 601 aactaattgc tgtggagatt tacagatgga ttctgggatg cctgagaccc ccattacaaa
661 atataaactc tgttcagttg aaagaactag aagaagaatg ggtcaaaact ccaacaagt
721 ctctagacc tactcgattt ctctgttccc aacaagaact agaagctaaa ttggaacca
781 aacagctcgc tggggagat gctgaagsag gttggtgatg aggtgatgag gtgccacaaa
841 tagatgctta tgaagcttta gaagctgtag aaatcttcc caaacttccc aaagactttt
20 901 atgcaaaaat tgaagcacia aatggcaag agagaaaaa ggcctggag tctgtagaag
961 tactaataaa aaacccaaa ctggaagctg gcgattatgc agatttagta aaagcattaa
1021 agaaggttgt tggaaaggac accaatgtca tgttgggtgc ttggcagca aaatgctta
1081 ctggcctggc tgttgggcta aggaagaaat ttggacaata tgcaggacat gttgtgcaa
1141 ccactctgga gaaattcaaa gagaagaaac ctcaagtgtt acaagccctg caggaggcaa
25 1201 ttgatgcaat ctctcttact accacactac agaacatcag tgaggatgtt tttagcagtaa
1261 tggataataa aaatccaacc atcaagcagc agacatctct ttttattgca agaagtctcc
1321 gccactgcac tgcctctacc ctgocaaaaga gcttgcataa gcccttttgt gctgcactac
1381 ttaagcacat caatgttctt gctcctgaag tcagagatgc ogcatttgaa gcattgggta
1441 ctgctttgaa ggtggttggc gagaagcag taaaaccatt cctagctgat gtggacaaac
30 1501 tcaagcttga taagatcaaa gaatgttcag aaaaggtaga actgatacat ggtaaagaag
1561 cbgactagc tgcgtataag aaggaattca aacctctgco tggaaaggact gctgcttca
1621 gggctgcaag agataaggac acaaaggaca tttctgcacc caaacaggga cctcaaaaa
1681 aggcacctgc tgcctaagct ggtgggcccac caaaaagggg gaaaccagct gcaccaggag
1741 ggcagggaaa tactggaacc aagaacaaga aaggactgga gactaaagaa atagtggagc
35 1801 ctgagctctc gatagaagta tgtgaagaaa aagcttcagc tgtctctccc cctacctgta
1861 tacagcttct tgacagcagt aactggaaa gaaaggctgco ttgtatggaa gagttccaga
1921 aggcctgtta gctaatggac cgaactgaaa tgcctatgcca gccatctagt aggatgctag
1981 ccaagaaacc tggatggaaa gaaactaatt ttcaggtgat gcaaatgaag ctctcatatg
2041 ttgctttgat tgcocagaag gaaaattttt ccaaaaacgtc agctcaggtt gtatttagatg
40 2101 gccctgttga caagattgga gatgtgaaat gtgggaacaa tgcaaaagaa gctatgacag
2161 caatagocga agcctgtatg ttaccatgga ctgctgaaca ggttctgtca atggctttct

```

WO 02/070558

PCT/GB02/00941

61

2221 cacaaaagaa tcccaaaaat oagtcagaaa ctctgaattg gctatcaaat gccataaaag
 2281 aatttgggttt ttctgggttg aatgtcaaaag ctttcattag caatgtgaag acagctcttg
 2341 ctgcaacaaa cccagctgtg aggactgctg ccataaccct gottggcgtg atgtatctgt
 2401 atgttgggtcc ctctttggga atgttctttg aggatgagaa gccctgccctc ctatcccaga
 5 2461 tagatgcaga atttgagaag atgcagggac aaagtccacc tgcaccaacc agaggaattt
 2521 ccaagcatag cacaagtgtt acagatgaag gagaagatgg agatgaacca gatgacggga
 2581 gcaatgatgt cgttgatctt ttgccgagga cggagatcag tgataaaatc acttcagagt
 2641 tggatctcaa gattggtgat aagaattgga agattaggaa agaaggccta gatgaagtgg
 2701 caggtattat taatgacgca aaatttatcc aaccgaatat aggtgaactt ccaactgcct
 10 2761 tgaagggtcg actcaatgat tcaataaaa ttttggatca gcaaacgctg aatatcctgc
 2821 aacaactggc agtagccatg ggcaccaata ttaagcaaca tgtaaaaaat ttaggcatcc
 2881 ctatcatcac agtcccttga gacagcaaga acaatgttcg agctgctgcc ctgacgactg
 2941 tgaatgcttg ggcagacaag actggcactg aggaatggct ggaaggagaa gatctttctg
 3001 aagagctcaa aaagggaaaat cctttcttga ggcagagctc tctgggctgg ctggctgaga
 15 3061 aactacctac tcttcttcc accctacag acctatctc ttgtgttctc catctctact
 3121 cctgctaga agatcgaaat ggagatgtgc gaaagaaggc ccaagatgcc ttgccattct
 3181 tcatgatgca tttaggatat gaaaaatgg ccaaggctac tgggaaacta aagccaactt
 3241 ctaaagatca ggtattggcc atgctagaga aagccaaagt taacatgcca gccaaagctg
 3301 cccaccacc taaagcaact tctaaccaca tgggagggtc cgtccagcc aaatccagc
 20 3361 ctgcatagc acctgctgaa gattgtattt ccagcagtac agaaccacaa cctgatccaa
 3421 aaaagcccaa agctccagga ttatccteta aagcaaaagag tgcaacaagg aagaagatgc
 3481 caagcaaac cagcttaag gaggatgaag acaaatccgg gccatttttt atgttgttc
 3541 caaatggaaa agcgaagg atgaaagatg aaaaaggatt gaaggtgcta aagtgaatt
 3601 ttactacccc acgggatgaa tacattgagc aactaaagac tcaaatgtct agctgtgtg
 25 3661 ctaaatggtt acaagatgag atgtttcact cagactttca gcatacaac aaagcccttg
 3721 ctgttatggt tgatcacttg gagagtgaaa aagaaggagt tattggttgc ctggatotta
 3781 tcttaaagtg gcttaccctg aggttttttg acaccaatac aagcgtcctg atgaaagcac
 3841 tagaatattt aaaattgctc ttcaacctgc taagtgaaga agaataatcat cttactgaga
 3901 atgaagcacc ttcttcafc cctatcttg tctcaaggt tggagaacca aaggatgta
 30 3961 ttctgtaaga tgttctgccc atcctgaacc ggatgtgctc tbtctacca gctagcaaga
 4021 tgtttccctt tatcatggaa ggaaccaaat ccaaaaactc taagcagaga gcagagtgcc
 4081 tggaaagact gggatgtctg gttgagtcct atggcatgaa tgtttgcaa caaccaccag
 4141 gaaaagcctt aaaggaata gctgttcaca taggagaccg tgacaatgct gtacgcaatg
 4201 ctgcaactca caccattgta acggtgtaca atgtacatgg ggcacaggtg ttcaactga
 35 4261 ttggaatct ttctgaaaag gatatgagca tgctcgagga gaggattaag cgttcagcaa
 4321 agagaccctc tgcagacca ataaaacagg tggagagaa acctcagcgt gcacagacaa
 4381 taagctocaa tgcaacatg ttaogcaagg gaccagctga ggacatgtct tccaactca
 4441 accaagcccg aagcatgagt gggcactcctg aggcagccca gatggtccgc cgagaattcc
 4501 agctggatct agatgagatt gagaatgaca atggtacagt ccgatgtgaa atgocagaac
 40 4561 ttgttcagca caaactggat gacatttttg agccagtcct tattctgaa cccaagatcc
 4621 gggctgtttc tccacacttc gatgacatgc acagtaatac agcatccaca atcaatttca
 4681 ttatctccca agtagccagt ggtgacatca acacaagtat ccaagctctg acacagatcg
 4741 atgggtctct gagacaggaa gacaaaagctg aagccatgta cggccatatt gatcagtttc

WO 02/070558

PCT/GB02/00941

62

```

4801 tgatagccac ttttatgcag ctaagactca tctacaacac acacatggca gatgagaaat
4861 tggagaagga cgagatcatc aagtgtgata gctgtatcat tggcaacatg atttgcctgt
4921 ttccagataga gagccttgcc cgggaggcct ccactggagt actaaaagac ctaatgcatg
4981 gcoctcatcac cttaaatgctg gattctcggg ttgaagatct tggagaagga caacaggtea
5 5041 tccgctctgt gaaacctctg gtggtgaaag ttctggagaa gtcagaccag accaacatcc
5101 tgagtgcctc actgtttttg ctccaagaca gccgtctagc aacagccagt tctccaaat
5161 tctcagagct tgttatgaag tgtctctgga gaatggttcg actgttgcct gataccatca
5221 atagcattaa cctagacaga attctctcgg atatccacat tttcatgaag gtcttcccca
5281 aagagaaact gaagcaatgc aaaagtgaat ttccataaag gccctaaag accctgctac
10 5341 acaccttatg caaattaaaa gggcccaaga tcttggacca cctaacgatg atcgacaaca
5401 aaaacgagtc tgagctggag gccatctctc gccggatgat gaagcacagt atggaccaga
5461 ctgggagcaa gtctgataag gaaacagaaa agggagcatc tcgaaatgat gaaaaatcat
5521 caaagcccaa agtgaatgat ttcttagctg agatttttaa gaagattggc tctaagaaaa
5581 acactaaga gggactagca gagttataty aatataagaa gaaatactca gatgctgaca
15 5641 ttgaaccatt tctgaaaaat tctccacagt tcttccagag ctatgtcgaa agaggcctc
5701 ggttgattga gatggagggg gagggcaaa gtctatcttc cactcaaca ggcctctccc
5761 ctccagatgga agtcacatgt gtgcccacgc ccacaagcgc agtgtctctc ataggttaaca
5821 caaatgggga agaagtgggg ccactctgtc acttggaaag gctaaaagtc ctccgacagc
5881 gatgtggtct ggacaacaca aagcaagatg accgaactcc ttgacctct ctctctcca
20 5941 aaccagagtc tctactgtc gccctctcca cagacatgct ccacagcaaa ctctctcagc
6001 tccgggagtc acgggagcag caccagcatt cagacatgga ttctaaccag actcaactct
6061 caggaactgt gacctctccc tctccacag ctaacataga cgaactgaaa aaaagactgg
6121 agagaataaa gagcagtcgc aaatgaagct gccccactcc cccggcacc ctgagcttta
6181 gtttactaaa ctgaaagtc ccatagttaa aaatggcctc agcaggccta gtgtatacaa
25 6241 actggttgta tgtatcatgc cgtggagcta gggggaggag tcattgtggc acaagtattt
6301 gtacataact tgcctctctc tgtcagcgtc ctgctgctct agaagactgt ccgtggatga
6361 gtttagtgta cagacttgta aacagctgac cctctctctg tcaagtctagt tcccagatcc
6421 tttctctttt aattgctcat ttgtaaaatt gtcctaact tctctagctt ttaatagtt
6481 aatattagaa actctttaat agttttcctt tcagtttgg agctctctct tgtcgcctcg
30 6541 aagggtoact gtattctgta tgaatgcatg gcatgatata actaatttaa gagtotttta
6601 taataaagt ttgcattaac tatacctg

```

SEQ ID NO: 2 Protein sequence for BAA07892.2 (LBDS1) MET7= start methionine

```

1 gkpgstmgdd sewlklpvdq kcehklwkar lsgyeealki fqkikdeksp ewskflglik
35 61 kfvtdsnavv qlkgleaalv yvenahvagk ttgevvsqv svfnqpkak akelgieicl
121 myieiekgea vgeellkgl d nknpkii vac ietlrkalse fgskiillkp iikvlpklfe
181 srekavrdea kliaveiyrv irdalrpplq ninsvqlkel eeewklpts aprptrflrs
241 qqealeakleq qagsggdaeg ggdgddevpq idayelleav eilsklpkdf ydkieakkwq
301 erkealesve vliknplea gdyadlvkal kvvvgktnv mlvalaakcl tglavglrkk
40 361 fgqyaghvvp tilekfkckk pqvvalqlea idaiflttl qnisedvlav mdknptikq
421 qtslfiaarsf rhtastlpk sllkpfcaal lkhindsape vrdaafealg talkvvgeka
481 vkpfladvdk lkldkikecs ekvelihgkk aglaadkkef kplpgrtaas gaagdkdtkd

```

WO 02/070558

PCT/GB02/00941

63

541 isapkgpplk kapaakagpp pkkkpaapg gagntgtknk kgletkeive pelsievcee
561 kasavlpptc iqlldssnwk erlaomeefc kavelmdrte mpcqalvrml akkpgwketn
562 fqvzqmkllhi valiaqkgnf sktsaqvvlid glvdkigdvk cgnnakeamt aiaeaclpw
721 taevvsmaf sqknpknqse tinwlsnaik efgfeglrvk afisnvktal aatnpavrt
5 781 aittllgmyl yvgsplrmff edekpalleq idaefekwqg qpppaprtri sknstsqtde
841 gedgdpepddg sndvvdllpr teisdkitse lvsikigdknw kirkegldev agiindakfi
901 qpnlgeplta lkgrlndsnk ilvqgtlnil qqlavamgpn ikghvknlgi piitvlgdek
961 nrvraaalat vnawaeqtgm kewlegedls eelkknepfl rqellgwlae klptlrstpt
1021 dlilcvphly scledrngdv rkkaqdalpf fmmhlyyem akatgklkpt skdqvlamle
10 1081 kakvnpakp apptkatskp mggsapakfq pasapaedci ssttepkipd kkakapglss
1141 kaksagkkm psktslkede dksppifivv pngkeqmkd ekglkvlkn fttprdeyie
1201 qlktqmsscv akwldemfh sdfghhkal avmvdhlese kegvigoldl ilkwltrff
1261 dtntsvlmlka leyikllftl lseeeyhlte neassfipyl vvkvgpekdv irkdvrailn
1321 rmclyypask mfpfimegk sknskgraec leelgclves ygmvcqptp gkalkeiavh
15 1381 igdrnavrn aaintivty nvhgqvfkl ignlsekdms mlearikraa krpsaapikq
1441 veekpqrqn issnamlrx gpaedmskl nqarsmsghp eaaqevref qldldelend
1501 ngtvrcmpe lvghklddif epvlpepki ravsphfdm hentastinf iisqvasgdi
1561 ntsiqaltqi devlrqedka eamsghldqf liatfmqlrl iynthadek lekdeikkly
1621 sciiignmsl fqiaslarea stgvikdlmh glitlmldr iedleegqv irsvnlvvk
20 1681 vleksdqtni lsallvllqd sllatasspk fselvmkclw rnvrlpdti nsinldrill
1741 dihimkvfp keklkqckse fpirtkltl hlclckkqk ildhitmidn kneseleahl
1801 crmmkhsmdq tgsksedete kgasrideks skakvndfla eifkkgiske ntkeglaely
1861 eykkysdad iepflkssq ffqsyvergl rviemeregk griststgis pqmevtcvpt
1921 ptstsvsign tngeevgsyv ylerlkilrq rcglntkqd drppltslls kpavptvass
25 1981 tdmhsklsq lresreqhah edldsngthe sgtvtssst aniddkkrl eriksrx

WO 02/070558

PCT/GB02/00941

64

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 LBDS1 polypeptide, said Nuclear Hormone Receptor Ligand Binding Domain region being defined as including residues 1507 and 1742 inclusive, of the amino acid sequence recited in SEQ ID NO:2, wherein said fragment possesses the "LBD motif" residues VAL1573, LEU1574, ASP1578, LYS1579, ALA1582 and MET1583, 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 VAL1573, LEU1574, ASP1578, LYS1579, ALA1582 and MET1583, 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 LBDS1 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

WO 02/070558

PCT/GB02/00941

65

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 4548 to 5255 inclusive 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.

WO 02/070558

PCT/GB02/00941

66

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 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, for use in therapy or diagnosis of disease.
20. A method of diagnosing a disease in a patient, comprising assessing the level of expression of a natural gene encoding a polypeptide according to any one of claims 1-7, or assessing the activity of a polypeptide according to any one of claims 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.
- 10 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:

WO 02/070558

PCT/GB02/00941

67

- 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.

WO 02/070558

PCT/GB02/00941

68

28. A method according to any one of claims 20-27, wherein said disease is a cell proliferative disorders, including neoplasm, cancer and myeloproliferative disorders and angiogenesis, metabolic disorders, including obesity, osteoporosis, non-insulin dependent diabetes, lipid metabolism disorder, hyperthyroidism, hyperparathyroidism, hypercalcemia, hypercholesterolemia, hyperlipidemia, cardiovascular disorders including, hypertension, atherosclerosis, hyperinsulinemia, heart arrhythmia, renal disorders, including glomerulonephritis, renovascular hypertension, skin disorders including dermatological disease, psoriasis, acne, aging, eczema, wound healing, inflammation, including inflammatory bowel disease, respiratory tract inflammation, emphysema, asthma, immune disorder, autoimmune disease, multiple sclerosis, nervous system disorders, including anxiety, depression, neurodegenerative disease, Alzheimer's disease Parkinson's disease, brain injury, stroke and pain, infection, including virus infection, and other conditions in which nuclear hormones 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

WO 02/070558

PCT/GB02/00941

69

- composition according to claim 32 for use in the manufacture of a medicament for the treatment of cell proliferative disorders, including neoplasm, cancer and myeloproliferative disorders and angiogenesis, metabolic disorders, including obesity, osteoporosis, non-insulin dependent diabetes, lipid metabolism disorder,
- 5 hyperthyroidism, hyperparathyroidism, hypercalcemia, hypercholesterolemia, hyperlipidemia, cardiovascular disorders including, hypertension, atherosclerosis, hyperinsulinemia, heart arrhythmia, renal disorders, including glomerulonephritis, renovascular hypertension, skin disorders including dermatological disease, psoriasis, acne, aging, eczema, wound healing, inflammation, including inflammatory bowel
- 10 disease, respiratory tract inflammation, emphysema, asthma, immune disorder, autoimmune disease, multiple sclerosis, nervous system disorders, including anxiety, depression, neurodegenerative disease, Alzheimer's disease Parkinson's disease, brain injury, stroke and pain, infection, including virus infection, and other conditions in which nuclear hormones are implicated.
- 15 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.
- 20 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.
- 25 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.

WO 02/070558

PCT/GB02/00941

70

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
5 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
10 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
15 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.
- 20 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.
- 25 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.
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.

Target Mining Results - Metascope

File Edit View Help Information 100%

Back Forward Reload Home Stop Metascope Print Search Stop

Bookmarks Location: http://function-bridge.inpharmatics.com/kgp-jim/vic/gegal-target-p103.pl

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

| Target | Gene | Protein | Species | Accession | Score | Rank | Gene | Protein | Species | Accession | Score | Rank |
|---------------------------------|--------|---------|---------|-----------|-------|------|-------------------------------|---------|-----------|-----------|---------|------|
| <input type="checkbox"/> Rad54a | Rad54a | Rad54a | Human | AA024814 | 11.2% | 64 | KIAA0711 protein | Human | NP_055123 | 64.33 | 314.51 | 47 |
| <input type="checkbox"/> Rad54b | Rad54b | Rad54b | Human | AA024815 | 11.2% | 65 | KIAA0711 protein | Human | NP_055123 | 64.33 | 314.51 | 48 |
| <input type="checkbox"/> Rad54c | Rad54c | Rad54c | Human | AA024816 | 9.2% | 66 | CANSD ASSEMBLY PROTEIN ULT. | Human | U00000 | 9.215 | 200.39 | 43 |
| <input type="checkbox"/> Rad54d | Rad54d | Rad54d | Human | AA024817 | 8.2% | 67 | stroma antigen 3, STAG3 | Human | U00000 | 8.217 | 211.71 | 57 |
| <input type="checkbox"/> Rad54e | Rad54e | Rad54e | Human | AA024818 | 7.2% | 68 | phosphatidylinositol-4-kinase | Human | U00000 | 7.219 | 209.82 | 56 |
| <input type="checkbox"/> Rad54f | Rad54f | Rad54f | Human | AA024819 | 6.2% | 69 | Palme 4 kinase | Human | U00000 | 6.219 | 209.79 | 58 |
| <input type="checkbox"/> Rad54g | Rad54g | Rad54g | Human | AA024820 | 5.2% | 70 | actin | Human | U00000 | 5.219 | 209.38 | 59 |
| <input type="checkbox"/> Rad54h | Rad54h | Rad54h | Human | AA024821 | 4.2% | 71 | 119 | Human | U00000 | 4.219 | 208.33 | 68 |
| <input type="checkbox"/> Rad54i | Rad54i | Rad54i | Human | AA024822 | 3.2% | 72 | cyclin | Human | U00000 | 3.216 | 211.34 | 65 |
| <input type="checkbox"/> Rad54j | Rad54j | Rad54j | Human | AA024823 | 2.2% | 73 | KIAA0169 protein | Human | U00000 | 2.217 | 144.37 | 88 |
| <input type="checkbox"/> Rad54k | Rad54k | Rad54k | Human | AA024824 | 1.2% | 74 | Wt1 protein | Human | U00000 | 1.217 | 21.26 | 56 |
| <input type="checkbox"/> Rad54l | Rad54l | Rad54l | Human | AA024825 | 0.2% | 75 | KIAA037 protein | Human | U00000 | 0.219 | 137.179 | 34 |
| <input type="checkbox"/> Rad54m | Rad54m | Rad54m | Human | AA024826 | 0.1% | 76 | Ret gene | Human | U00000 | 0.121 | 174.97 | 23 |
| <input type="checkbox"/> Rad54n | Rad54n | Rad54n | Human | AA024827 | 0.1% | 77 | Ret gene | Human | U00000 | 0.122 | 203.98 | 27 |
| <input type="checkbox"/> Rad54o | Rad54o | Rad54o | Human | AA024828 | 0.1% | 78 | Ret gene | Human | U00000 | 0.123 | 162.154 | 32 |
| <input type="checkbox"/> Rad54p | Rad54p | Rad54p | Human | AA024829 | 0.1% | 79 | Ret gene | Human | U00000 | 0.127 | 203.46 | 31 |
| <input type="checkbox"/> Rad54q | Rad54q | Rad54q | Human | AA024830 | 0.1% | 80 | hypothetical protein | Human | U00000 | 0.128 | 20.21 | 29 |

FIG. 2B
 SUBSTITUTE SHEET (RULE 26)

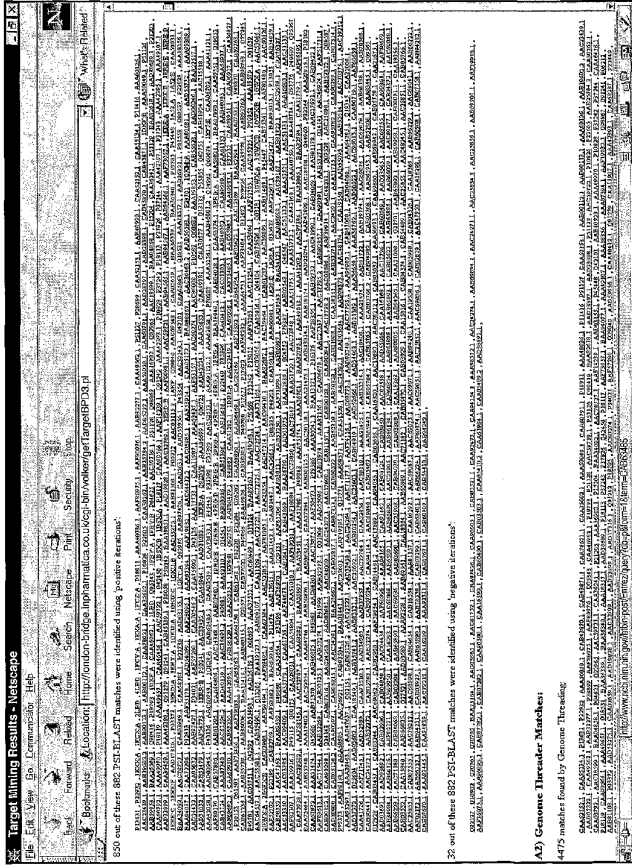


FIG. 20

SUBSTITUTE SHEET (RULE 26)

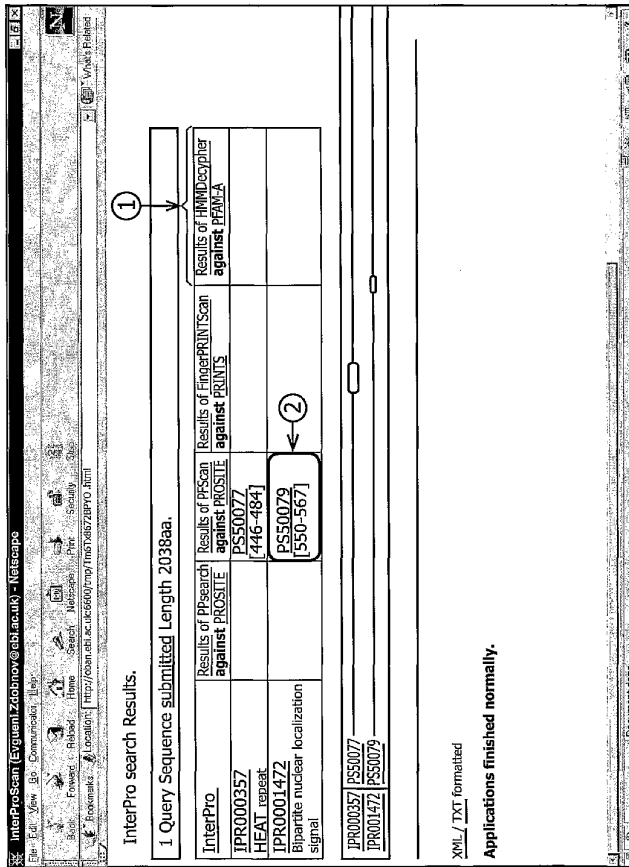


FIG. 4

Microsoft Word of SWISS-PROT.01408 - Microsoft Internet Explorer provided by @ipametrics Ltd

General information about this entry

Entry name: **CTCOC_HUMAN**

Primary accession number: **Q14608**

Secondary accession number(s): **P04608**

Created in SWISS-PROT in: **Release 36, Jul 1998**

Sequence will be re-annotated in: **Release 38, Jul 1998**

Version: **1**

Number of amino acids in the protein: **48**

Protein name: **CHELOG PROTEIN**

Systematic name: **COLONIC AND HEPATIC TUMOR OVER-EXPRESSED PROTEIN**

Gene name(s): **CTLA0197**

Form: **Polypeptide**

Accession: **Ensembl: Musculus, Chordata, Vertebrata, Euteleostomi, Mammalia, Rodentia, Primates, Eumastulata, Euteleostomi, Mus musculus**

NCBI TaxID: **10090**

SEQUENCE FROM A: **Medline**

1852867 [Nishizawa, T, et al., 1998, PubMed=9556852]

Medline=9556852 [Nishizawa, T, et al., 1998, PubMed=9556852]

Characterization of the cDNA and protein of expression of a new gene over-expressed in human hepatomas and colonic tumors *.

En, J., Nishizawa, T, et al., 1998, J. Hepatol., 29:460-471(1998).

SEQUENCE FROM A: **Medline**

1788527 [Kobayashi, T, et al., 1997, PubMed=91440120]

Medline=91440120 [Kobayashi, T, et al., 1997, PubMed=91440120]

Prediction of the coding sequence of a novel gene, CTLA0197, deduced by analysis of cDNA clones from human cell line FC-1 *.

En, J., Kobayashi, T, et al., 1997, J. Hepatol., 25:460-471(1997).

Comments:

- **EXPRESSION PATTERNS:** EXPRESSED IN HEPATOMAS AND COLONIC TUMORS. ALSO EXPRESSED IN SKELETAL MUSCLE, BRAIN, HEART, PLACENTA, LUNG.
- **TISSUE EXPRESSION:** LIVER, KIDNEY AND PANCREAS.
- **SIMILARITY:** TO C. ELEGANS F2B3.7 AND F2B3.8 AND SOME TO YEAST STU2.

Keywords:

The SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL database. There are no restrictions on its use by non-profit institutions as long as the contents are not republished without the permission of the copyright holder. For commercial entities a license agreement (GFI: <http://www.ebi.ac.uk/infocentre/copyright.htm>) or email request to swissprot@ebi.ac.uk is required.

FIG. 5B

SUBSTITUTE SHEET (RULE 26)



FIG. 6A

SUBSTITUTE SHEET (RULE 26)

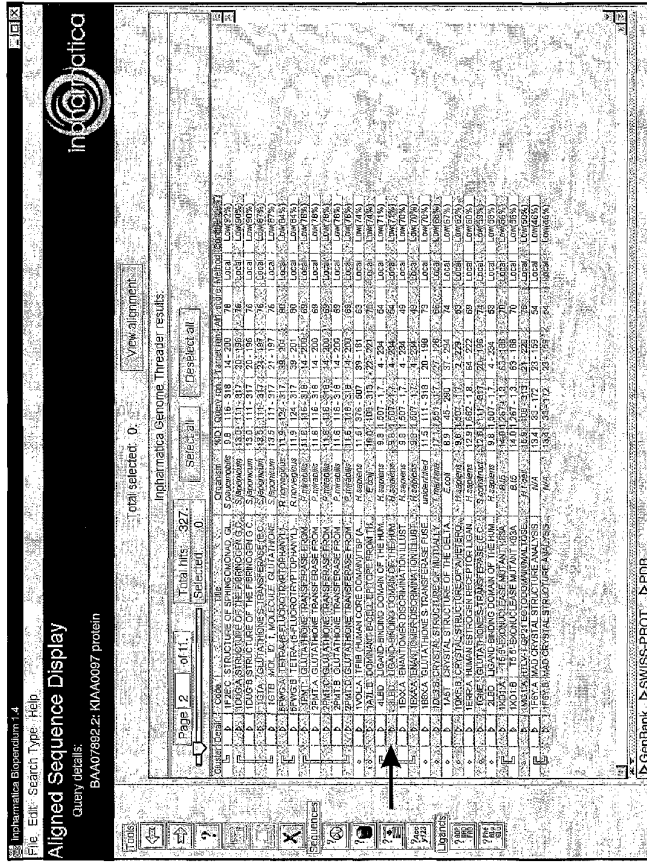


FIG. 6B

SUBSTITUTE SHEET (RULE 26)

AIEye output (February 20, 2001 1:05 PM)

FIG. 7-1

BA07862.2 G K P G S T M G D D S E W L K L P V D D K C E H K L W K A R L S S Y E A L K I F O K I K D E S P E W S K F L G L I K K F V T D S N A V V D L K G L E A A L
3.LBD 10 20 30 40 50 60 70

BA07862.2 V Y V E K A H V A G K T T G E V S G V S K V F N P K A K A K E L G I E I C L M Y I E I E K G E A Y O E E L L K G L D N K N P K I I V A C I E T L R K A L
3.LBD 80 90 100 110 120 130 140 150

BA07862.2 S E F G S K I I L K P I I K V L P K L F E S R E K A V R D E A K L I A V E I Y R W I R D A L R P P L Q N I N S V O L K E L E E E W V K L P T S A P R P T R F
3.LBD 160 170 180 190 200 210 220 230

BA07862.2 L R S O C E L E A K L E Q Q S A G G D A E G G D G D E V P Q I D A Y E L L E A V E I L S K L P K D F Y D K I E A K K W G E R K E A L E S V E V L I K N P
3.LBD 240 250 260 270 280 290 300 310

BA07862.2 K L E A G D Y A D L V K A L K Y V G K D T N V M L V A L A K C L T G L A V G R K K F G Q Y A G H V V P T I L E K F K E K K P Q V Y G A L G E A I D A I F
3.LBD 320 330 340 350 360 370 380 390

BA07862.2 L T T L L O N I S E D V L A Y M D N K N P T I K O O T S L F I A R S F R H C T A S T L P K S L L K P F C A A L L K H I N D S A P E V R D A F E A L G T A L K
3.LBD 400 410 420 430 440 450 460 470

BA07862.2 V Y G E K A V K P F L A D V D K L K L D K I K E C S E K V E L I H G K K A G L A A D K K E F K P L P G R T A A S G A A G D K D T K D I S A P K P G P L K K A P
3.LBD 480 490 500 510 520 530 540 550

BA07862.2 A A K A G G P P K K G K P A A P G G A G N T G T K N K G L E T K E I V E P E L S I E V G E E E K A S A V L P P T C I Q L L D S S N W K E R L A C M E E F O K A
3.LBD 560 570 580 590 600 610 620 630

BA07862.2 V E L M D R T E M P C C A L V R M L A K K P G W K E T N F Q V M Q M K L H I V A L I A R K N G F S K T S A Q V V D D G L V D K I G D V K C G R N A K E A M T A
3.LBD 640 650 660 670 680 690 700 710

AlEye output (February 20, 2001 1:05 PM)

FIG. 7-2

BAA07892.2 720 730 740 750 760 770 780 790
 3LED J A E A C M L P W T A E Q V V S M A F S Q K N P K K O S E T L N W L S N A I K E F G F S G L N V K A F I S N V K T A L A A T N P A V R T A A I T L L G V W Y L
 800 810 820 830 840 850 860
 SAA07892.2 Y V G P S L R M F F E D E K P A L L S Q I D A E F E K M Q Q S P P A P T R G I S K H S T S G T D E G E D E F D D S N D V D L L P R T E I S D K I T S
 3LED 870 880 890 900 910 920 930 940
 BAA07892.2 E L Y S K I G D K K W K I R K E G L D E V A G I I N D A K F I Q F N I G E L P T A L K G R L N D S N K I L V O O T L N I L Q O L A V A M G P N I K O H Y K N L
 3LED 850 860 870 880 890 900 910 920 930 940 950 960 970 980 990 1000 1010 1020
 BAA07892.2 G I F I I T V L G D S K N N V R A A L A T Y N A W A E Q T G M K E W L E G E D L S E E L K K E N P F L R G E L L G W L A E K L P T L R S T P T L I L C V P
 3LED 1030 1040 1050 1060 1070 1080 1090 1100
 BAA07892.2 H L Y S C L E D R N G D V R K K A Q D A L P F M M H L G Y E K M A K A T G K L K P S K Q Q V L A M L E K A K V N M P A K P A P T K A T S K P M G S S A P
 3LED 1110 1120 1130 1140 1150 1160 1170 1180
 BAA07892.2 A K F O P A S A P A E D C I S S T E P K P D P K K A K A P G L S S K A K S A O G K K M P S K T S L K E D E K S G P I F I V P P N G K E Q R M K D E K G L K
 3LED 1190 1200 1210 1220 1230 1240 1250 1260
 BAA07892.2 V L K W N F T T P R D E Y I E Q L K T O M S S C V A K W L Q D E M F H S D F Q H N K A L A V M V D H L S E K E G V I G C L D L I L K W L T L R F F D T N T
 3LED 1270 1280 1290 1300 1310 1320 1330 1340
 BAA07892.2 S Y L M K A L E Y L K L L F L L S E E Y H L T E N E A S S F I P Y L V V K V G E P K D V I R K D V R A I L N R M G L V Y P A S K M F P F I M E G T S K N
 3LED 1350 1360 1370 1380 1390 1400 1410 1420
 BAA07892.2 S K O R A E C L E E L G C L V E S Y G M N Y C O P T P G K A L K E I A V H I G D R D N A V R N A A L N T I V T V N V H G D G V F K L I G N L S E K D M S M L
 3LED

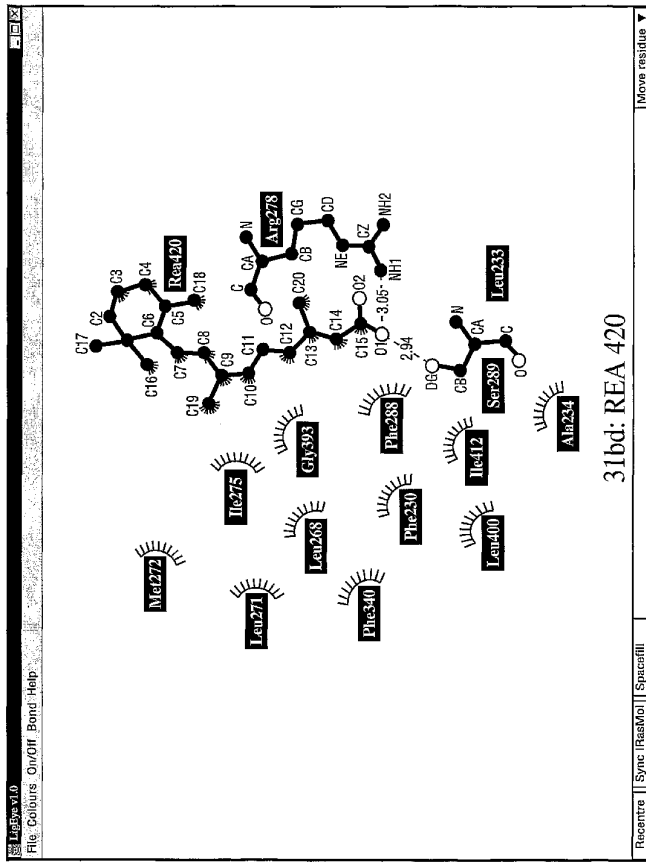


FIG. 8A

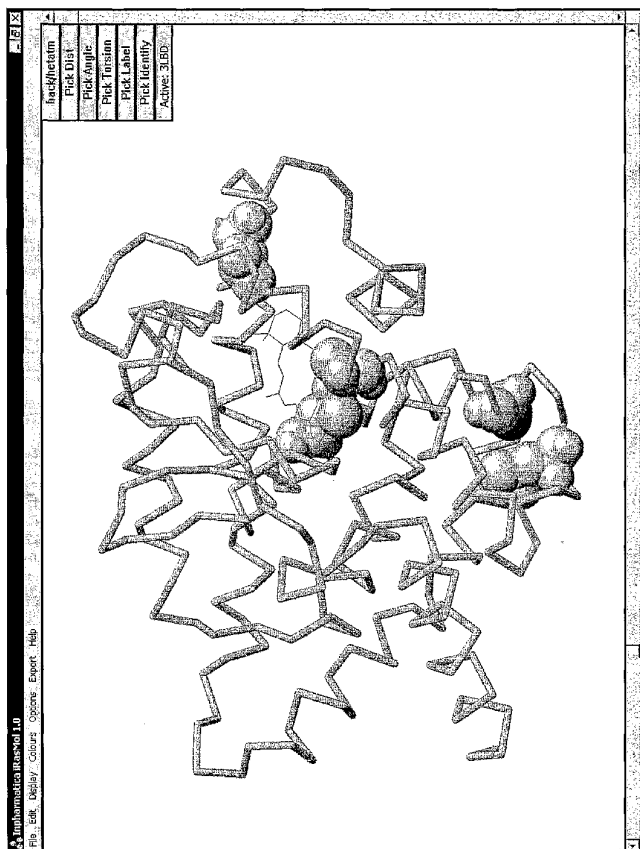


FIG. 8B

SUBSTITUTE SHEET (RULE 26)

ALEye output (February 21, 2001 10:52 AM)

FIG. 9-1

80 90 100 110 120 130 140 150
 BAA07892.2 GKPGSTMGDDSEWLKLPVDDQKCEHKLWAKRLSGVEEAKLFOKIKDEKSPWSKFLGLIKKRVTDNSNAVQLKGLAAL
 CAB61894.1 - - - - -MGDDSEWKLPIIDQKCEHKLWAKRLSGVEEAKLFOKIKDEKSPWSKFLGLIKKRVTDNSNAVQLKGLAAL
 CAB55772.1 - - - - -MAEDTEYKLPVEERCVHLKWKARVGGYEAAKIFRELDDEKSPWSKFLGLIKKRVTDNSNAVQLKGLAAL

 160 170 180 190 200 210 220 230
 BAA07892.2 VYVENAHVAGKTTGEVSVGVSKVFNQPKAKKELGIEICLMYIEIEKGEAVQEEELKGLDNKNPKIIVACIETLRKAL
 CAB61894.1 VYVENAHVAGKTTGEVSVGVSKVFNQPKAKKELGIEICLMYIEIEKGEAVQEEELKGLDNKNPKIIVACIETLRKAL
 CAB55772.1 IFVENSGLAGRTVGDVMTGIYOKCIAPKTKTKELSYQVALMYEIEKGEAVQEEELKGLDNKNPKIIVACIETLRKAL

 240 250 260 270 280 290 300 310
 BAA07892.2 SEFQSKIIILKPIIKVLPKLFESREKAVRDEAKLIAVEIYRWIRDALRPPLOININSVOLKLEEEWVKLPTSAPRPTRF
 CAB61894.1 SEFQSKIIILKPIIKVLPKLFESREKAVRDEAKLIAVEIYRWIRDALRPPLOININSVOLKLEEEWVKLPTSAPRPTRF
 CAB55772.1 REFGHKVIGVKPLIKKLAPLMSDRDKTVRDEKQKLAVEIYRWIRDALRPPLOININSVOLKLEEEWVKLPTSAPRPTRF

 320 330 340 350 360 370 380 390
 BAA07892.2 LRSQGELEAKLEQQSAGGD AEGGGDDGDEVPQIDAYELLEAVEILSKLPKDFYDKIEAKKWQERKEALEVEVLI-KN
 CAB61894.1 LRSQDDLKAKFEQQQAAGDDG-GDDGEEIYPOVDAYELLEAVEILSKLPKDFYDKIEAKKWQERKEALEVEVLI-KN
 CAB55772.1 LKSOQEKQAKI-ADAAATEDAYNEDDGEAGVEEIDPMDLDPVDILSKMPKDFYDKIEAKKWQERKEALEVEVLI-KN

 400 410 420 430 440 450 460 470
 BAA07892.2 PKLEAGDYADLVKALKKVVGGKDTNVMVLVALAAKCLTGLAVGLKRFKGGYAGHVYPTILEKFKKPKQVVOALQEAIDAI
 CAB61894.1 PKIEAGDFADLVKALKKVVGGKDTNVMVLVALAAKCLTGLAVGLKRFKGGYAGHVYPTILEKFKKPKQVVOALQEAIDAI
 CAB55772.1 PKLENGEYGVALVSALKKVIITKDSNVVLYVAMAGKCOLALLAKGLAKRFSNYSACVPSLLEKFKKPKQVVOALQEAIDAI

 480 490 500 510 520 530 540 550
 BAA07892.2 FLTTTLQNI SEDVILAVMDNKNPTIKQQTSLFIARSFRHCTASTLPKSLKPKFCRALLKHI NDSAPEVRDAAFEALGTAL
 CAB61894.1 FLTTTLQNI SEDVILAVMDNKNPAIKQQTSLFIARSFRHCTASTLPKSLKPKFCRALLKHI NDSAPEVRDAAFEALGTAL
 CAB55772.1 YASTSLEAQQESI VESLANKNPSYKSETALFIARALTRTQPTALNKLLKLLTTSVLTALNEPPTVRDSSAEALGTLI

 BAA07892.2 KVVGEKAVKPFPLADVDKLLDKIKCEKSEKVELIHGKAGLAAADKKEFKPLPGRTAAS-...GAAGDKDTKDISAPKPG
 CAB61894.1 KVVGEKAVNPFPLAEVDKLLDRIKCEADKELANGKGGGAAGKKEFKPLPGRTAAS-...GAAGDKDTKDISAPKPG
 CAB55772.1 KLMGDKAVTFLADVDPLKMAKIKCEQEKAE-...IKIKYAGPKKTRPASPAPTAKAAAAPAKTYAGSVDPKPYTRPATT

AIEye output (February 21, 2001 10:52 AM)

FIG. 9-2

560 570 580 590 600 610 620 630
 BAA07892.2 P L K K P A A K A S G P P K K G K P A A P G S A G N T G T K N K K G L E T K E I V E P E I S I E V C E E K A S A V L P P T C I Q L L D S S N W K E R L A C M
 CAB61894.1 A P K K A P A A K P G G V Y K K A K A P A S S G A T . . . A K G K A V E N K I I E C E I S P E A C E E R A A Y P A S C M Q Q L D S S N W K E R L A S M
 CAB55721.1 G A R K V L K K P A T V S G G A T S A P T A A L K A S G K P L A T E R E I T P E E L G K S E I L P A E I N S L Y S N W K N R L A A Y 740 750 760 770 780

540 550 560 570 580 590 600 610 620 630
 BAA07892.2 E E F O K A V E L M D R T E . . M P C Q A L V R M L A . . K F P G W K E T N F Q V M Q M K L H I V A L I A Q K G N F S K T S A Q V V L D G L V D K I G D V K C G N
 CAB61894.1 E E F O K T V E S M E R N D . . I P C Q A L V R M L A . . K F P G F K E T N F Q V M Q M K L H I V A L I A Q K G N F S K T S A C A V L D G L V D K V G D V K C G G
 CAB55721.1 E O L L G E I S G F D A K Q A G I S Q I L I R T I S G R K F G L K E M N F Q V L K F K L D I J R S V A E N Y P L T T T T V D L V I N E I I E K L A D A R K N G A 740 750 760 770 780

720 730 740 750 760 770 780 790
 BAA07892.2 N A K E A M T A I A E A C M L P W T A E Q V Y S M A F S O K N P K N Q S E T I N W L S N A I K E F G F S Q L N V K A F I S N V K T A L A A T N P A V R T A A C
 CAB61894.1 N A K E A L S G I A E A C T L P M T A E Q V Y S L A F A O K N P K N Q S E T I N W L S N A I K E F G F T G I N K K A F I S N V K T A L A A T N P A I R T S A I
 CAB55721.1 A A A D V L S A F A E A T K L E Y V G K V L S F A F E G K S P A Y Q S E A F N W V W R S I I E F G F . . O L O P K T L I E D V R K G V G S T N P T V R A S A I 740 750 760 770 780 790

800 810 820 830 840 850 860 870 880 890 900 910 920 930 940
 BAA07892.2 T L L G V M Y L Y G P S L R M F F E D E K P A L L S Q I D A E F E K M G Q S P P A P T R G I S K H S T S G T D E G E D E P D . D G S N D V
 CAB61894.1 T L L G V M Y L Y M G A P L R M F F E E K P A L L S Q I D A E F E K M G Q T P P Y S I R G . . . S K H S G D E G E G E E D E D A P A P D V
 CAB55721.1 Q M V G T M S M Y M G K A L M F F D S E K P A L K S Q I Q V E F D K N V G E K P P P V R G V Q . R S S G G T A G N S P D N E D D D G G A A G E E E P I N M 840 850 860 870 880 890 900 910 920 930 940

950 960 970 980 990 1000 1010 1020
 BAA07892.2 Y D L L P R T E I S D K I T S E L V S K I G D K N W K I R K E G L D E V A G I N D A K F I O P N I G E L P T A L K R L R L D S N K I L Y O O T L N I L Q Q L
 CAB61894.1 Y D L L P R T D I S D K I S S D L V S K I E D K N W K I R K E G L D E V T A I N E A K F I P S T I G E L P S A L K R L R L D S N K I L Y O O T L T I L Q Q L
 CAB55721.1 A D L L P R V D I A P Q I T E A L L K E M S D K W K T R N E G L T K L Q A I I S E A R L I K P S I G D L A P A L A H R L R L V D S N A K I A Q T T L A I G E Q L 950 960 970 980 990 1000 1010 1020

1030 1040 1050 1060 1070 1080 1090 1100
 BAA07892.2 A V A M G P N I K O H V K N L G I I I T V L G D S K N N V R A A A L A T V N A W A E Q T G M K E W L E G E D L S E E L K K E N P F L R Q E L L G W L A E K L
 CAB61894.1 S T A M G H N I K O H V K N L M P I I I T V L G D S K A N V R A A A L G T L K S W D Q T G M K W L E G E D L S E E L K K E N P F L R Q E L L G W L A E K L
 CAB55721.1 A T A M G A G C R N H V R N L F P G F L H A L G D N K S F V R A A A L N C I N S F G E K G S Y K E F F E S E M I A D A L K G G S P A L K T E L W A W L A D K L 1030 1040 1050 1060 1070 1080 1090 1100

1110 1120 1130 1140 1150 1160 1170 1180 1190 1200
 BAA07892.2 P T I R S T P T D L I L O V P H L Y S C L E D R N G D V R K K A Q D A L P F F M M H L G Y E K M K A T G K L R P T S K D Q V L A M L E R A K K V M P A
 CAB61894.1 P S M A T Y P S D L Q L C V P L Y N C L E D R N G D V R K K A Q E A L P F F M M H I G F E K M S A T S K L R P A S K D Q V V A L L E R A K A S M P A
 CAB55721.1 P G L P P K S V S N E D I H S M V P H L Y A H I C D R N A D V R K N A E A V L G I M I H L G F D A M N R A L D Q K R P A S K K D I L A A L E K A R P N L P Y 1110 1120 1130 1140 1150 1160 1170 1180 1190 1200

AlEye output (February 21, 2001 10:52 AM)

FIG. 9-3

```

1110      1120      1130      1140      1150      1160      1170      1180
BA07892.2  KPA-PPTKATSKPIGGGSAKAPAFQAPASARAEDCISSTSTPKDPKAKKAPGSSKAK-----
CAB51894.1  KPAGPPGRASSKQPPAYAGASAPPPASSSGGSSSTSDYKDPKTKTGTGAKKAKTSYSEGNSTLSPNNTSLTPSK
CAB55772.1  KP-----LPKGRHQAPIPEEPKLTIVRGGAGG-----APGQKASATARVAGG-----

1190      1200      1210      1220      1230      1240      1250      1260
BA07892.2  -----SAGGKKMPSKTSLKEDKSGPIFIVPNGKQRKDEKGLKVLKWNFTTTPHDEYIEQLKTMSS-
CAB51894.1  ANTSLSKAKPAKQTLPGKKAPSKPNAKDEKSGPIYIIVPNGKQRKDEKGLKVLKWNFTTTPHDEYIEQLKTMSS-
CAB55772.1  -----QDKQVPAR-----KKDEDIDTSPLLCANSAKNORLDEQKMKVVKWTFVTPPEFTELLRDOMMTA

1270      1280      1290      1300      1310      1320      1330      1340
BA07892.2  CVAKWLODEMFHSDFOHKNKALAVMVDHLESEKGYIGCLDILIKWLTLEFDNTSVLMKALEYKLLFTLLSEEVH
CAB51894.1  CIARWLODELFHADFQGIKGLAVMTEHLESEKGYISCLDILYKWFLEFDNTSVLMKALEYKLLFTLLSEEVH
CAB55772.1  MYNKALIANMFHDDFRHLYLKVIEQLSEDLAGNSKALVGNLDLILKWLTRFDYFNPSVLIKGLYLVGVVQVVLDEEYI

1350      1360      1370      1380      1390      1400      1410      1420
BA07892.2  LTEMASFFIPLYVVKVGEPKDVIKRDYRAILNRMCLVYPASKMFPIFIMEGTSKNSKQRAECLLEQLVESYGMNYC
CAB51894.1  LTEMGTSFLPYLMLKVGEPKDVIKRDYRAILTKMCQVYPASKMFVMEGTSKNSKQRAECLLEQLVESYGMNYC
CAB55772.1  LAENEGSSFYPHLLLK-ANPKDAVRNGVRRVLRQVILVFPFVKVFGYVMEGLKSNARQRTECLDELFLIESYGMNIC

1430      1440      1450      1460      1470      1480      1490      1500
BA07892.2  OPTPGKALKEIAVHIGDDNDNVRNAALNTIVTVYVNHGDDEKLIIGNLSEKDMMLERIKRSAKR-PSAAPIKQVEEK
CAB51894.1  OPTPAKALKEIAHIGDDDTTVRNAALNTIVTVYVNHGEVFKLIIGNLSEKDMMLERIKRSAKR-PSAAPIKQVEEK
CAB55772.1  ---POSAMREIARQISDRDNSVRAALNCIVQVFFLSEKTYKMIHGLNEKDLMLDERIKRAKTKKPTPPSPSYDVFPA

1510      1520      1530      1540      1550      1560      1570      1580
BA07892.2  PORAQNISSNANML-----RKGPAE-----DMS-----SKLNOARMSHGHPEAAQMVRRFQLDLDEIENDNGTV
CAB51894.1  PORVQ--SANASIL-----RKAPPE-----DMS-----SKLNOARMMGGHTERSHSVPRFQLDLDEIENDNGTV
CAB55772.1  PORHDSIEIEDAEVGNCGDELPPDEEDGTFDQAPSSQLLLQQLQQLQQAQQKPSGPFGLDSQVISELEKQWVRVD

1590      1600      1610      1620      1630      1640      1650
BA07892.2  RCMEPELVQHKLDDIIFEPV-----
CAB51894.1  RCMEPALVQHKLDEIIFEPV-----
CAB55772.1  QMECKPILLNVDISSLDEPIKVRPTFRAGIHYPQEKFDRLISRQHYMQQTLTTSPTAGMTSGVSPYRSPMLRQHQDQDQ

```


AIEye output (February 21, 2001 4:29 PM)

FIG. 10-1

10 20 30 40 50 60 70
 BA078622 GKFGSTMGGDDSEWLKLPVDDQKCEHKLWKARLISGYEEALKIFOKIKDEKSPWSKFLGLIKKFFVDSNAVVQLKGLLEAAL
 CA4532121MGDDSEWLKLPVDDQKCEHKLWKARLISGYEEALKIFOKIKDEKSPWSKFLGLIKKFFVDSNAVVQLKGLLEAAL
 80 90 100 110 120 130 140 150
 BA078622 VYVENAHVAGTTTGEVVSQVSKVFNQPKAKAKELGIEICIMYIEIEGSEAVQVEELLKGLDNKNPKIIVACIETLRKAL
 CA4532121 VYVENAHVAGTTTGEVVSQVSKVFNQPKAKAKELGIEICIMYIEIEGSEAVQVEELLKGLDNKNPKIIVACIETLRKAL
 160 170 180 190 200 210 220 230
 BA078622 SEFGSKILLKPIIKVLPKLFESREKAVRDEAKLI AVEIYRWIRDALPPLONINSVOLKELEEEWVKLPTSAPRPTRF
 CA4532121 SEFGSKILLKPIIKVLPKLFESREKAVRDEAKLI AVEIYRWIRDALPPLONINSVOLKELEEEWVKLPTSAPRPTRF
 240 250 260 270 280 290 300 310
 BA078622 LRSQOELEAKLEQQQSAGGDAEGGDDGDEVPQIDAYELLEAVEILLKPKDFYDKIEAKKQWERKEALESVEVLKKNP
 CA4532121 LRSQOELEAKLEQQQSAGGDAEGGDDGDEVPQIDAYELLEAVEILLKPKDFYDKIEAKKQWERKEALESVEVLKKNP
 320 330 340 350 360 370 380 390
 BA078622 KLEAGDYADLYKALKKVVYKDTNMYMLVALAAKCLTGLAVGRKFGQYAGHVPTIILEKFKKPPQVYQALQEAIDAIF
 CA4532121 KLEAGDYADLYKALKKVVYKDTNMYMLVALAAKCLTGLAVGRKFGQYAGHVPTIILEKFKKPPQVYQALQEAIDAIF
 400 410 420 430 440 450 460 470
 BA078622 LTTLLONISEDVLAVMDNKNPTIKQOTSIFARSFRHCTASTLPSLLKPECAALLKHINDSAPVSDAAFEALGTALK
 CA4532121 LTTLLONISEDVLAVMDNKNPTIKQOTSIFARSFRHCTASTLPSLLKPECAALLKHINDSAPVSDAAFEALGTALK
 480 490 500 510 520 530 540 550
 BA078622 VVGEKAVKPFLLADVDKLLDKIKECSEKVELIHGKKAGLAADKKEFKPLPGRRTAASGAAGDKTDIISAPKPGPLKKAP
 CA4532121 VVGEKAVKPFLLADVDKLLDKIKECSEKVELIHGKKAGLAADKKEFKPLPGRRTAASGAAGDKTDIISAPKPGPLKKAP
 560 570 580 590 600 610 620 630
 BA078622 AAKAGPPKKGKPAAPGAGNTGTKNKKGLTKEIPEPELSIEVCEEKASAVLPPPTC IQLLDSNNWKERLACMEEFFOKA
 CA4532121 AAKAGPPKKGKPAAPGAGNTGTKNKKGLTKEIPEPELSIEVCEEKASAVLPPPTC IQLLDSNNWKERLACMEEFFOKA
 640 650 660 670 680 690 700 710
 BA078622 VELWDRTEMPCOALVRMLAKKPGWKETNFOVMOMKHLIVALIADKGNFESKTSQVVLDGLVDKIGDYKCGNNAKEAMTA
 CA4532121 VELWDRTEMPCOALVRMLAKKPGWKETNFOVMOMKHLIVALIADKGNFESKTSQVVLDGLVDKIGDYKCGNNAKEAMTA

SUBSTITUTE SHEET (RULE 26)

AlEye output (February 21, 2001 4:29 PM)

FIG. 10-2

BA07892.2 I A E A C M L P M T A E Q V Y S M A F S Q K N P K N Q S E T L N W L S N A I K E F G F S G L N V K A F I S N V K T A L A A T M P A V R T A A I T L L G V M Y L 700 710 720 730 740 750 760 770 780 790
 CA0632121 I A E A C M L P M T A E Q V Y S M A F S Q K N P K N Q S E T L N W L S N A I K E F G F S G L N V K A F I S N V K T A L A A T M P A V R T A A I T L L G V M Y L 800 810 820 830 840 850 860 870 880 890 900

BA07892.2 Y Y G P S L R M F F E D E K F A L L S Q I D A E F E K M Q G S P P A P T R G I S K H S T S G T D E G E D D E P D D S S N D V D L L P R T E I S D K I T S 870 880 890 900 910 920 930 940
 CA0632121 Y Y G P S L R M F F E D E K F A L L S Q I D A E F E K M Q G S P P A P T R G I S K H S T S G T D E G E D D E P D D S S N D V D L L P R T E I S D K I T S 950 960 970 980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080 1090 1100

BA07892.2 E L V S K I G D K N W K I R K E G L D E V A G I I N D A K F I Q P N I G E L P T A L K G R L N D S N K I L Y Q Q T L N I L Q O L A V A M G P N I K O H V K N L 1100 1110 1120 1130 1140 1150 1160 1170 1180 1190 1200
 CA0632121 E L V S K I G D K N W K I R K E G L D E V A G I I N D A K F I Q P N I G E L P T A L K G R L N D S N K I L Y Q Q T L N I L Q O L A V A M G P N I K O H V K N L 1210 1220 1230 1240 1250 1260 1270 1280 1290 1300 1310 1320 1330 1340 1350

BA07892.2 G I P I I T V L G D S K N N V R A A A L A T V N A W A E O T G M K E W L E G E D L S E E L K K E N P F L R G E L L G W L A E K L P T L R S T P T D L I L C V P 1350 1360 1370 1380 1390 1400 1410 1420
 CA0632121 G I P I I T V L G D S K N N V R A A A L A T V N A W A E O T G M K E W L E G E D L S E E L K K E N P F L R G E L L G W L A E K L P T L R S T P T D L I L C V P 1430 1440 1450 1460 1470 1480 1490 1500

BA07892.2 H L Y S C L E D R N G D V R K K A Q D A L P F F M H H L G Y E K M A K A T G K L P T S K D O V L A M L E K A K V N M P A K P A P T K A T S K P M G G S A P 1510 1520 1530 1540 1550 1560 1570 1580 1590 1600
 CA0632121 H L Y S C L E D R N G D V R K K A Q D A L P F F M H H L G Y E K M A K A T G K L P T S K D O V L A M L E K A K V N M P A K P A P T K A T S K P M G G S A P 1610 1620 1630 1640 1650 1660 1670 1680 1690 1700

BA07892.2 A K F Q P A S A P A E D C I S S S T E P R P D P K K A K A P G L S S A K A S A Q G K K M P S T S L K E D E D K S G P I F I V Y P N G K E G R M A D E K G L K 1710 1720 1730 1740 1750 1760 1770 1780 1790 1800
 CA0632121 A K F Q P A S A P A E D C I S S S T E P R P D P K K A K A P G L S S A K A S A Q G K K M P S T S L K E D E D K S G P I F I V Y P N G K E G R M A D E K G L K 1810 1820 1830 1840 1850 1860 1870 1880 1890 1900

BA07892.2 V L K W N F T T P R D E Y I E Q L K T Q M S S C V A K W L O D E M F H S D F Q H H N K A L A V M V D H L S E K E G V I G C L D L I L K W L T L R F F D T N T 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000
 CA0632121 V L K W N F T T P R D E Y I E Q L K T Q M S S C V A K W L O D E M F H S D F Q H H N K A L A V M V D H L S E K E G V I G C L D L I L K W L T L R F F D T N T 2010 2020 2030 2040 2050 2060 2070 2080 2090 2100

BA07892.2 S V L M K A L E Y L K L L F T L L S E E E Y H L T E N E A S S F I P Y L V V K V G E P K D V I R K D V R A I L N R M C L V Y P A S K M F P F I M E G T S K N 2110 2120 2130 2140 2150 2160 2170 2180 2190 2200
 CA0632121 S V L M K A L E Y L K L L F T L L S E E E Y H L T E N E A S S F I P Y L V V K V G E P K D V I R K D V R A I L N R M C L V Y P A S K M F P F I M E G T S K N 2210 2220 2230 2240 2250 2260 2270 2280 2290 2300

BA07892.2 S K O R A E C L E E I L G C L V E S Y G M V G C P T P G K A L K E I A V H I G D R D N A V R A A L T I V T V Y N V H G D V F K L I G N L S E K D M S M L 2310 2320 2330 2340 2350 2360 2370 2380 2390 2400
 CA0632121 S K O R A E C L E E I L G C L V E S Y G M V G C P T P G K A L K E I A V H I G D R D N A V R A A L T I V T V Y N V H G D V F K L I G N L S E K D M S M L 2410 2420 2430 2440 2450 2460 2470 2480 2490 2500

FIG. 10-3

AIEye output (February 21, 2001 4:29 PM)

1430 1440 1450 1460 1470 1480 1490 1500
 BA07892.2 EERI KRSARRPSAAPIKQVVEERFORAQNISSANMLRGGPREDSSKLNQARSMSSGHPEAQMVRREFQLDLDLEIENDN
 CA463212.1 EERI KRSARRPSAAPIKQVVEERFORAQNISSANMLRGGPREDSSKLNQARSMSSGHPEAQMVRREFQLDLDLEIENDN
 1510 1520 1530 1540 1550 1560 1570 1580
 BA07892.2 GTVRCMPPELVQHKLDDIFEPVLIPEPKIRAVSPHFDDMHNTASTINFISSQSSGDINTSIOALTOIDVILFQEDKKA
 CA463212.1 GTVRCMPPELVQHKLDDIFEPVLIPEPKIRAVSPHFDDMHNTASTINFISSQSSGDINTSIOALTOIDVILFQEDKKA
 1590 1600 1610 1620 1630 1640 1650
 BA07892.2 EEMVSGHIDGFOCATQALRLIYNTHMADEKLEKDEIKLYSCIIIGNMISLFOIESLAREASTGVKOLMHLITLMLDS
 CA463212.1 EEMVSGHIDGFOCATQALRLIYNTHMADEKLEKDEIKLYSCIIIGNMISLFOIESLAREASTGVKOLMHLITLMLDS
 1660 1670 1680 1690 1700 1710 1720 1730
 BA07892.2 RLEDLEEGQVIRSVNLLVVKYLEKSDOTNILLSALLVLDLSSLLATASSPFSELYMVKCLWRMVRQVDPDTINSINLDR
 CA463212.1 RLEDLEEGQVIRSVNLLVVKYLEKSDOTNILLSALLVLDLSSLLATASSPFSELYMVKCLWRMVRQVDPDTINSINLDR
 1740 1750 1760 1770 1780 1790 1800 1810
 BA07892.2 QLDIHIFRKYVFKKELKCKSEFFPIRTLLHTLCKLGGPKILDHMTIDNKNKESELEAHLGRMMKHSMDQIGSKSDK
 CA463212.1 QLDIHIFRKYVFKKELKCKSEFFPIRTLLHTLCKLGGPKILDHMTIDNKNKESELEAHLGRMMKHSMDQIGSKSDK
 1820 1830 1840 1850 1860 1870 1880 1890
 BA07892.2 ETEKASRIDEKSSKAKVNDFLAEIFKIGSKENTKEGLAELYKYYKKYSDADIEPFLKNSSOFFQSYVERGLRVIEME
 CA463212.1 ETEKASRIDEKSSKAKVNDFLAEIFKIGSKENTKEGLAELYKYYKKYSDADIEPFLKNSSOFFQSYVERGLRVIEME
 1900 1910 1920 1930 1940 1950 1960 1970
 BA07892.2 REGKGRISTSTGIFSPOMEVTCVPTPTSTVSSIGNTNGEVEGSPSYLERLKLROCGLDNTKQDDRPPLTSLLSKPAPV
 CA463212.1 REGKGRISTSTGIFSPOMEVTCVPTPTSTVSSIGNTNGEVEGSPSYLERLKLROCGLDNTKQDDRPPLTSLLSKPAPV
 1980 1990 2000 2010 2020 2030 2040
 BA07892.2 TVASSTDMHLKLSQLRESRECHGHSDLDLSNQTSSGTVTSSTANIDDLKRLERIKSSRK
 CA463212.1 TVASSTDMHLKLSQLRESRECHGHSDLDLSNQTSSGTVTSSTANIDDLKRLERIKSSRK

The screenshot shows a web browser window displaying a PubMed search result. The browser's address bar shows the URL: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=8555249&from_uid=8555249. The page header includes the NCBI logo and navigation links such as 'Home', 'About Entrez', 'Help/FAQ', 'New! News/Notify', 'PubMed Services', 'Journal Browser', 'MESH Browser', 'Simple ClinTrials', 'Basic ClinTrials/Medline', 'Clinical queries', 'Chubby', 'Related Resources', 'Order Documents', 'PubMed Alerts', 'Clinical Alerts', 'ClinicalTrials.gov', and 'Privacy Policy'. The main content area displays the following information:

Characterization of the cDNA and pattern of expression of a new gene over-expressed in human hepatomas and colonic tumors.
 Charasse S, Mizel M, Tavitian S, Berra P, Chou T, Laruega C
 Ush: INSERM 129, Montpellier, France.
 A cDNA clone of 6.449 kb, ch-TCO2 (for colorectal and hepatic tumor over-expressed gene) initially selected from various human libraries and completed by 5' rapid amplification of cDNA ends (RACE) PCR is described. The original cDNA clone was retracted from an expression library constructed from a human tumoral liver. The library was screened with an antibody raised against the epitope 21-50aa that was shown to be over-expressed in chronically induced mouse hepatic tumors. Using this cDNA as a probe, a full-length cDNA was characterized. Its nucleotide sequence shows no significant similarity with any of the gene sequences collected in the various DNA data bases. In addition, other larger open reading frame areas to a putative protein of 1972 amino acids (non-coding mass = 218153 Da). Hybridization analysis on Southern blot and on Northern blot of the tumor tissues showed that the ch-TCO2 gene is over-expressed in the same tissues. The expression of ch-TCO2 transcripts are present in several human tissues. Over-expression of ch-TCO2 in neoplastic liver and colon, compared with the corresponding normal tissues, was demonstrated. The level of the expression of ch-TCO2 transcripts was also studied in the various differentiation stages of the human colonic adenocarcinomas cell line Colo-2.
 PMID: 8324682

At the bottom of the page, there is a small box containing the following text:
 SOURCE: NIH, NIH/NCBI
 NLM/BIOMED
 DEPARTMENT OF HEALTH & HUMAN SERVICES
 NATIONAL INSTITUTES OF HEALTH

FIG. 12

SUBSTITUTE SHEET (RULE 26)

【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 02/070558 A3

- (51) International Patent Classification: C07K 14/705
- (21) International Application Number: PCT/GB02/00941
- (22) International Filing Date: 5 March 2002 (05.03.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
0105401.4 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 [US/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB); PHELPS, Christopher, Benjamin [GB/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB).
- (74) Agents: MERCER, Christopher, Paul et al.; Carpmaels & Ransford, 43 Bloomsbury Square, London WC1 2RA (GB).
- (81) Designated States (national): AF, 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, GI, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, 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, 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, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CI, 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:
19 December 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 02/070558 A3

(54) Title: NUCLEAR HORMONE RECEPTOR LIGAND BINDING DOMAIN

(57) Abstract: The invention relates to a novel protein, termed BAA07892.2, 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 02/00941 |
|--|---|--|
| A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C07K14/705 | | |
| 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) SEQUENCE SEARCH, EPO-Internal, MPI Data, PAJ, BIOSIS | | |
| C. DOCUMENTS CONSIDERED TO BE RELEVANT | | |
| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
| X | NAGASE T ET AL: "PREDICTION OF THE CODING SEQUENCES OF UNIDENTIFIED HUMAN GENES. III - THE CODING SEQUENCES OF 40 NEW GENES (KIAA0081-KIAA0120) DEDUCED BY ANALYSIS OF CDNA CLONES FROM HUMAN CELL LINE K6-1" DNA RESEARCH, UNIVERSAL ACADEMY PRESS, JP, vol. 2, no. 1, 1995, pages 37-43, XP000874164 ISSN: 1340-2838 cited in the application | 1-13, 15, 40-44 |
| Y | table 3 | 19-28, 31-39, 45 |
| X | -& DATABASE NCBI GENE BANK 'Online! 6 January 2000 (2000-01-06) NAGASE ET AL.: "KIAA0097 protein" Database accession no. D43948 XP002214318 cited in the application | 1-13, 15, 40-44 |
| Y | abstract | 19-28, |
| -/- | | |
| <input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input type="checkbox"/> Patent family members are listed in annex. | | |
| * Special categories of cited documents: | | |
| ** document defining the general state of the art which is not considered to be of particular relevance | | ** later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention |
| *E* earlier document but published on or after the international filing date | | ** document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone |
| *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) | | ** 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. |
| *O* document relating to an oral disclosure, use, exhibition or other means | | ** document member of the same patent family |
| ** document published prior to the international filing date but later than the priority date claimed | | |
| Date of the actual completion of the international search 23 September 2002 | | Date of mailing of the international search report 07/10/2002 |
| Name and mailing address of the ISA European Patent Office, P.B. 5616 Patentleaz 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 apo nl, Fax: (+31-70) 340-3016 | | Authorized officer Sprinks, M |

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 02/00941

| C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT | | |
|--|--|------------------------------|
| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
| | | 31-39, 45 |
| X | CHARRASSE ET AL.: "Characterization of the cDNA and pattern of expression of a new gene over-expressed in human hepatomas and colonic tumors" EUROPEAN JOURNAL OF BIOCHEMISTRY, vol. 234, 1995, pages 406-413, XP002214317 cited in the application abstract; figure 1 | 1-13, 15, 19-28, 31-45 |
| Y | | 19-28, 31-39, 45 |
| X | page 413, column 2, line 10 - line 15 -& DATABASE EMBL 'Online! 30 October 1995 (1995-10-30) CHARRASSE ET AL.: "H.sapiens mRNA for ch-TO6 protein" Database accession no. X92474 XP002214319 cited in the application abstract | 1-13, 15, 19-28, 31-45 |
| Y | | 19-28, 31-39, 45 |

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

INTERNATIONAL SEARCH REPORT

International application No.
PCT/GB 02/00941**Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)**

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
see FURTHER INFORMATION sheet PCT/ISA/210
2. Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210
3. Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this International application, as follows:

1. As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- The additional search fees were accompanied by the applicant's protest.
- No protest accompanied the payment of additional search fees.

International Application No. PCT/JP 02/00941

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

Although claims 20, 22-28 and 38 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 35-37 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

Present claims 14 and 16-18 relate to compounds defined merely by reference to a desirable characteristic or property, namely the ability to bind/inhibit/agonise/antagonise etc. the polypeptide of the invention.

The claims cover all compounds having this characteristic or property, whereas the application provides support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for only a very limited number of such compounds i.e. antibodies. 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 compounds by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, the search has been carried out for those parts of said claims which appear to be clear, supported and disclosed, namely those parts relating to antibodies.

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.

フロントページの続き

| (51)Int.Cl. ⁷ | F I | テーマコード(参考) |
|--------------------------|----------------|------------|
| A 6 1 P 1/04 | A 6 1 P 3/04 | 4 H 0 4 5 |
| A 6 1 P 3/04 | A 6 1 P 3/08 | |
| A 6 1 P 3/08 | 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 5/16 | |
| A 6 1 P 5/16 | 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 9/12 | |
| A 6 1 P 9/12 | A 6 1 P 11/00 | |
| A 6 1 P 11/00 | A 6 1 P 11/06 | |
| A 6 1 P 11/06 | A 6 1 P 13/12 | |
| A 6 1 P 13/12 | A 6 1 P 17/00 | |
| A 6 1 P 17/00 | A 6 1 P 17/02 | |
| A 6 1 P 17/02 | 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 19/10 | |
| A 6 1 P 19/10 | A 6 1 P 25/16 | |
| A 6 1 P 25/16 | 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 31/00 | |
| A 6 1 P 31/00 | A 6 1 P 35/00 | |
| A 6 1 P 35/00 | A 6 1 P 37/04 | |
| A 6 1 P 37/04 | A 6 1 P 43/00 | 1 1 1 |
| A 6 1 P 43/00 | C 0 7 K 14/72 | |
| C 0 7 K 14/72 | C 0 7 K 16/28 | |
| C 0 7 K 16/28 | 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 Q 1/02 | |
| C 1 2 N 5/10 | C 1 2 Q 1/68 | A |
| C 1 2 Q 1/02 | G 0 1 N 33/15 | Z |
| C 1 2 Q 1/68 | G 0 1 N 33/50 | Z |
| G 0 1 N 33/15 | G 0 1 N 33/53 | D |
| G 0 1 N 33/50 | G 0 1 N 33/53 | M |
| G 0 1 N 33/53 | G 0 1 N 33/566 | |
| G 0 1 N 33/566 | C 1 2 N 5/00 | A |
| | A 6 1 K 37/02 | |

(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
- F ターム(参考) 2G045 AA25 AA40 BA11 CB01 CB17 DA12 DA13 DA14 DA36 DA77
FB02 FB03
4B024 AA01 AA11 BA63 BA80 CA02 CA09 CA12 CA20 DA01 DA02
DA05 DA11 DA12 GA11 HA11 HA12 HA13 HA17
4B063 QA01 QA05 QA13 QA17 QA19 QQ02 QQ08 QQ21 QQ41 QQ43
QQ53 QQ61 QQ79 QQ89 QQ91 QR08 QR32 QR35 QR40 QR42
QR48 QR56 QR62 QR72 QR77 QR80 QR84 QS16 QS25 QS33
QS34 QS36 QX01 QX02
4B065 AA01X AA58X AA72X AA87X AA93Y AB01 AC14 BA02 CA43 CA44
CA46
4C084 AA02 AA13 AA17 NA14 ZA022 ZA112 ZA122 ZA152 ZA162 ZA362
ZA422 ZA592 ZA682 ZA702 ZA892 ZA942 ZA972 ZB082 ZB092 ZB112
ZB262 ZC062 ZC212 ZC332 ZC352 ZC422 ZC542
4H045 AA10 AA11 AA20 AA30 BA10 CA40 DA50 DA75 EA20 EA50
FA71 FA74

| | | | |
|----------------|---|---------|------------|
| 专利名称(译) | 核激素受体配体结合域 | | |
| 公开(公告)号 | JP2005500011A | 公开(公告)日 | 2005-01-06 |
| 申请号 | JP2002569877 | 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 | | |
| 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 | | |
| IPC分类号 | A01K67/027 A61K31/7088 A61K35/76 A61K38/00 A61K39/00 A61K45/00 A61K48/00 A61P1/00 A61P1/04 A61P1/18 A61P3/00 A61P3/04 A61P3/06 A61P3/08 A61P3/10 A61P3/14 A61P5/00 A61P5/14 A61P5/16 A61P7/00 A61P7/02 A61P7/04 A61P9/00 A61P9/06 A61P9/10 A61P9/12 A61P11/00 A61P11/06 A61P13/12 A61P15/00 A61P17/00 A61P17/02 A61P17/06 A61P17/10 A61P19/02 A61P19/10 A61P25/00 A61P25/04 A61P25/16 A61P25/22 A61P25/24 A61P25/28 A61P29/00 A61P29/02 A61P31/00 A61P31/04 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P35/00 A61P35/02 A61P37/00 A61P37/02 A61P37/04 A61P37/06 A61P37/08 A61P43/00 C07K14/47 C07K14/705 C07K14/72 C07K16/28 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N15/09 C12N15/12 C12Q1/02 C12Q1/68 G01N33/15 G01N33/50 G01N33/53 G01N33/566 G01N37/00 | | |
| CPC分类号 | A61K38/00 A61P1/00 A61P1/04 A61P1/18 A61P3/00 A61P3/14 A61P11/00 A61P11/06 A61P13/12 A61P15/00 A61P17/00 A61P17/02 A61P17/06 A61P17/10 A61P19/02 A61P19/10 A61P25/00 A61P25/04 A61P25/16 A61P25/22 A61P25/24 A61P25/28 A61P29/00 A61P29/02 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 A61K45/00 A61K48/00 A61P1/04 A61P3/04 A61P3/08 A61P3/10 A61P3/14 A61P5/16 A61P9/00 A61P9/06 A61P9/12 A61P11/00 A61P11/06 A61P13/12 A61P17/00 A61P17/02 A61P17/06 A61P17/10 A61P19/10 A61P25/16 A61P25/22 A61P25/24 A61P25/28 A61P31/00 A61P31/04 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P35/00 A61P35/02 C07K14/72 C07K16/28 C12N1/15 C12N1/19 C12N1/21 C12Q1/02 C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.D G01N33/53.M G01N33/566 C12N5/00.A A61K37/02 | | |
| F-TERM分类号 | 2G045/AA25 2G045/AA40 2G045/BA11 2G045/CB01 2G045/CB17 2G045/DA12 2G045/DA13 2G045/DA14 2G045/DA36 2G045/DA77 2G045/FB02 2G045/FB03 4B024/AA01 4B024/AA11 4B024/BA63 4B024/BA80 4B024/CA02 4B024/CA09 4B024/CA12 4B024/CA20 4B024/DA01 4B024/DA02 4B024/DA05 4B024/DA11 4B024/DA12 4B024/GA11 4B024/HA11 4B024/HA12 4B024/HA13 4B024/HA17 4B063/QA01 4B063/QA05 4B063/QA13 4B063/QA17 4B063/QA19 4B063/QQ02 4B063/QQ08 4B063/QQ21 4B063/QQ41 4B063/QQ43 4B063/QQ53 4B063/QQ61 4B063/QQ79 4B063/QQ89 4B063/QQ91 4B063/QR08 4B063/QR32 4B063/QR35 4B063/QR40 4B063/QR42 4B063/QR48 4B063/QR56 4B063/QR62 4B063/QR72 4B063/QR77 4B063/QR80 4B063/QR84 4B063/QS16 4B063/QS25 4B063/QS33 4B063/QS34 4B063/QS36 4B063/QX01 4B063/QX02 4B065/AA01X 4B065/AA58X 4B065/AA72X 4B065/AA87X 4B065/AA93Y 4B065/AB01 4B065/AC14 4B065/BA02 4B065/CA43 4B065/CA44 4B065/CA46 4C084/AA02 4C084/AA13 4C084/AA17 4C084/NA14 4C084/ZA022 4C084/ZA112 4C084/ZA122 4C084/ZA152 4C084/ZA162 4C084/ZA362 4C084/ZA422 4C084/ZA592 4C084/ZA682 4C084/ZA702 4C084/ZA892 4C084/ZA942 4C084/ZA972 4C084/ZB082 4C084/ZB092 4C084/ZB112 4C084/ZB262 4C084/ZC062 4C084/ZC212 4C084/ZC332 4C084/ZC352 4C084/ZC422 4C084/ZC542 4H045/AA10 4H045/AA11 4H045/AA20 4H045/AA30 4H045/BA10 4H045/CA40 4H045/DA50 4H045/DA75 4H045/EA20 4H045/EA50 4H045/FA71 4H045/FA74 | | |
| 代理人(译) | 中村稔 | | |

小川伸男
西岛隆义

優先権 2001005401 2001-03-05 GB

其他公开文献 JP2005500011A5

外部链接 Espacenet

摘要(译)

本发明涉及新颖的蛋白质，称为BAA91543.1，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。

(P2005-00011A)
(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) |
|----------------------------|-----------------|-------------|
| C12N 15/09 | C12N 15/00 ZNAA | 2G045 |
| A01K 67/027 | A01K 67/027 | 4B024 |
| A61K 38/00 | A61K 45/00 | 4B063 |
| A61K 45/00 | A61K 48/00 | 4B065 |
| A61K 48/00 | A61P 1/04 | 4C084 |

審査請求 未請求 予備審査請求 未請求 (全146頁) 最終頁に続く

| | |
|--|---------------------|
| (21) 出願番号 特願2002-569877 (P2002-569877) | (71) 出願人 502334043 |
| (86) (22) 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ディー |
| (86) 国際出願番号 PCT/GB2002/000941 | 2 エヌユー チャーロット ストリート |
| (87) 国際公開番号 W02002/070558 | 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 0105401.4 | 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 喜一 |
| | (74) 代理人 100074228 |
| | 弁理士 今城 俊夫 |

最終頁に続く