

(19)日本国特許庁（J P）

(12) **公開特許公報** (A) (11)特許出願公開番号

特開2002－360268
(P2002－360268A)

(43)公開日 平成14年12月17日(2002.12.17)

(51)Int.Cl ⁷	識別記号	F I	テマコード [*] (参考)
C 1 2 N 15/09	ZNA	A 6 1 K 35/12	2 G 0 4 5
A 6 1 K 35/12		35/64	4 B 0 2 4
35/64		35/72	4 B 0 6 3
35/72		35/74	A 4 B 0 6 4
35/74		35/76	4 B 0 6 5
審査請求 未請求 請求項の数 27 O L (全 72数) 最終頁に続く			

(21)出願番号	特願2001－236788(P2001－236788)	(71)出願人	596175810 財団法人かずさディー・エヌ・エー研究所 千葉県木更津市かずさ鎌足2－6－7
(22)出願日	平成13年8月3日(2001.8.3)	(71)出願人	000002831 第一製薬株式会社 東京都中央区日本橋3丁目14番10号
(31)優先権主張番号	特願2000－236839(P2000－236839)	(72)発明者	小原 収 千葉県木更津市矢那1532番3号 財団法人か ずさディー・エヌ・エー研究所内
(32)優先日	平成12年8月4日(2000.8.4)	(74)代理人	100088904 弁理士 庄司 隆
(33)優先権主張国	日本(JP)		
(31)優先権主張番号	特願2001－108723(P2001－108723)		
(32)優先日	平成13年4月6日(2001.4.6)		
(33)優先権主張国	日本(JP)		
		最終頁に続く	

(54)【発明の名称】 家系診断マーカー

(57)【要約】

【課題】 新規なポリグルタミン病関連遺伝子を単離・同定し、該遺伝子がコードするポリペプチドを明らかにし、さらにこれらを利用したポリグルタミン病の検査方法、防止および／または治療剤を提供すること。

【解決手段】 ポリグルタミンをコードするCAGリピートを有する配列表の配列番号1、配列番号3、配列番号5、配列番号7、または配列番号9に記載の塩基配列からなるポリヌクレオチドを含む新規遺伝子並びに該遺伝子がコードするポリペプチドを見出し、該CAGリピートの数が異なることに由来する各遺伝子の多型を利用した家系診断手段、該診断手段に用いることの出来る該遺伝子の部分配列よりなるプローブ、並びに該遺伝子の発現および該遺伝子がコードするポリペプチドの機能を阻害することに基づくポリグルタミン病の防止および／または治療剤を提供する。

【特許請求の範囲】

【請求項1】 下記の群より選ばれるポリヌクレオチド；

- ①配列表の配列番号1、配列番号3、配列番号5、配列番号7、または配列番号9に記載の塩基配列からなるポリヌクレオチドまたはその相補鎖、
- ②配列表の配列番号1、配列番号3、配列番号5、配列番号7、または配列番号9に記載の塩基配列からなるポリヌクレオチドにコードされるポリペプチドと少なくとも70%以上のアミノ酸配列上の相同性を有し、且つポリグルタミン配列を有するポリペプチドをコードするポリヌクレオチドまたはその相補鎖、
- ③上記①のポリヌクレオチドまたはその相補鎖とストリンジェントな条件下でハイブリダイゼーションするポリヌクレオチド、
- ④上記①のポリヌクレオチドまたはその相補鎖の塩基配列において1ないし数個のアミノ酸の欠失、置換、付加、および／または挿入等の変異を有するポリヌクレオチド、および
- ⑤上記①から④のいずれか1のポリヌクレオチドを含有するポリヌクレオチドまたはその相補鎖。

【請求項2】 請求項1に記載のポリヌクレオチドをなす塩基配列中の、少なくとも約10個の連続する塩基配列を有するポリヌクレオチド。

【請求項3】 請求項1に記載のポリヌクレオチドの部分塩基配列からなるポリヌクレオチドであって、配列表の配列番号11から配列番号20のいずれか1に記載の塩基配列からなるポリヌクレオチド。

【請求項4】 下記の群より選ばれるポリペプチド；

- (A) 配列表の配列番号1、配列番号3、配列番号5、配列番号7、または配列番号9に記載の塩基配列からなるポリヌクレオチドにコードされるポリペプチドであって、配列表の配列番号2、配列番号4、配列番号6、配列番号8、または配列番号10に記載のアミノ酸配列からなるポリペプチド、
- (B) 上記(A)のポリペプチドと少なくとも70%以上のアミノ酸配列上の相同性を有し、且つポリグルタミン配列を有するポリペプチド、
- (C) 上記(A)のポリペプチドにおいて1ないし数個のアミノ酸の欠失、置換、付加、および／または挿入といった変異を有し、且つポリグルタミン配列を有するポリペプチド、および
- (D) 上記(A)、(B)、または(C)のポリペプチドを含有するポリペプチド。

【請求項5】 請求項1から3のいずれか1項に記載のポリヌクレオチドまたはその相補鎖を含有する組換えベクター。

【請求項6】 組換えベクターが発現組換えベクターである請求項5に記載の組換えベクター。

【請求項7】 請求項5または6に記載の組換えベクターにより形質転換された形質転換体。

【請求項8】 請求項6に記載の組換えベクターにより

形質転換された形質転換体を培養する工程を含む、請求項4に記載のポリペプチドの製造方法。

【請求項9】 請求項4に記載のポリペプチドを免疫学的に認識する抗体。

【請求項10】 請求項1に記載のポリヌクレオチドと相互作用してその発現を阻害する化合物、および／または請求項4に記載のポリペプチドと相互作用して該ポリペプチドのTAF130との結合を阻害する化合物の同定方法であって、請求項1に記載のポリヌクレオチド、請求項4に記載のポリペプチド、請求項5若しくは6に記載のベクター、請求項7に記載の形質転換体、請求項9に記載の抗体のうちの少なくともいずれか1つを用いることを特徴とする方法。

【請求項11】 請求項1に記載のポリヌクレオチドと相互作用してその発現を阻害する化合物の同定方法であって、化合物と該ポリヌクレオチドとの間の相互作用を可能にする条件下で、該ポリヌクレオチドとスクリーニングに供する化合物とを接触させて該化合物の相互作用を評価し、次いで、該化合物と該ポリヌクレオチドとの相互作用により生じるシグナルの存在または不存在または変化を検出することにより、該化合物が該ポリヌクレオチドと相互作用してその発現を阻害するかどうかを決定することを含む方法。

【請求項12】 請求項4に記載のポリペプチドと相互作用して該ポリペプチドとTAF130との結合を阻害する化合物の同定方法であって、化合物と該ポリペプチドとの間の相互作用を可能にする条件下で、該ポリペプチドとスクリーニングに供する化合物とを接触させて該化合物の相互作用を評価し、次いで、該化合物と該ポリペプチドとの相互作用により生じるシグナルの存在または不存在または変化を検出することにより、該化合物が該ポリペプチドと相互作用してTAF130との結合を阻害するかどうかを決定することを含む方法。

【請求項13】 請求項10から12のいずれか1項に記載の方法で同定された化合物。

【請求項14】 請求項1に記載のポリヌクレオチドと相互作用してその発現を阻害する化合物、または請求項4に記載のポリペプチドと相互作用して該ポリペプチドとTAF130との結合を阻害する化合物。

【請求項15】 請求項1から3のいずれか1項に記載のポリヌクレオチド、請求項4に記載のポリペプチド、請求項5または6に記載のベクター、請求項7に記載の形質転換体、請求項9に記載の抗体、または請求項13若しくは14に記載の化合物のうちの少なくともいずれか1つを含有することを特徴とする医薬組成物。

【請求項16】 請求項1に記載のポリヌクレオチドと相互作用してその発現を阻害する化合物、および／または請求項4に記載のポリペプチドと相互作用して該ポリペプチドとTAF130との結合を阻害する化合物を含有することを特徴とする医薬組成物。

【請求項17】 ポリグルタミン病の防止および／または治療剤である、請求項15および／または16に記載の医薬組成物。

【請求項18】 請求項1に記載のポリヌクレオチドまたは請求項4に記載のポリペプチドを定量的あるいは定性的に測定する手段。

【請求項19】 請求項18に記載の手段を使用することを特徴とする、請求項1に記載のポリヌクレオチドまたは請求項4に記載のポリペプチドが関連する疾病の診断手段。

【請求項20】 前記疾病がポリグルタミン病である請求項19に記載の診断手段。

【請求項21】 請求項1に記載のポリヌクレオチド中のポリグルタミンをコードする塩基配列長を解析することを特徴とするポリグルタミン病の診断手段。

【請求項22】 請求項21に記載のポリグルタミン病の診断手段において、請求項2および／または3に記載のポリヌクレオチドを使用することを特徴とするポリグルタミン病の診断手段。

【請求項23】 請求項1に記載のポリヌクレオチド中のポリグルタミンをコードする塩基配列長を多型マーカーとして解析することを特徴とする家系診断手段。

【請求項24】 配列表の配列番号13、配列番号14、配列番号15、配列番号16、配列番号17、および配列番号18に記載の塩基配列からなるポリヌクレオチドをプローブとして使用することを特徴とする請求項23に記載の家系診断手段。

【請求項25】 請求項24に記載の家系診断手段において、配列表の配列番号13と配列番号14に記載の塩基配列からなるポリヌクレオチドの組み合わせ、配列表の配列番号15と配列番号16に記載の塩基配列からなるポリヌクレオチドの組み合わせ、および配列表の配列番号17と配列番号18に記載の塩基配列からなるポリヌクレオチドの組み合わせを、それぞれ単独または複数組み合わせて使用することを特徴とする家系診断手段。

【請求項26】 配列表の配列番号11と配列番号12に記載の塩基配列からなるポリヌクレオチドの組み合わせ、および／または配列表の配列番号19と配列番号20に記載の塩基配列からなるポリヌクレオチドの組み合わせを陰性対照として使用することを特徴とする請求項24若しくは25に記載の家系診断手段。

【請求項27】 請求項18から26のいずれか1項に記載の手段に使用する試薬キットであって、請求項1から3のいずれか1項に記載のポリヌクレオチド、請求項4に記載のポリペプチド、または請求項9に記載の抗体を少なくとも1つ以上含んでなる試薬キット。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、新規な遺伝子および該

遺伝子がコードするポリペプチドに関するものであり、詳しくはポリグルタミン病関連遺伝子および該遺伝子がコードするポリペプチド並びにそれらの利用手段に関する。

【0002】

【従来の技術】 これまでの分子遺伝学的な解析によって、ハンチントン舞蹈病を代表とする8つの遺伝性神経変性疾患は、その原因遺伝子内のCAGリピートの異常な伸長によって引き起こされることが明らかになっている (Trend Genet 14:396-402, 1998)。これら疾患の原因遺伝子において、CAGリピートは翻訳領域に存在し、全てポリグルタミンに翻訳される。このCAGリピートの数は、脊髄小脳失調症6型を除いて、健康人では約40リピートを越えないが、それ以上のリピート数（長くても約130リピートまで）では遺伝性神経変性疾患をおこす。さらに、CAGリピートが長いほど発症が早くなり重篤化すること、世代を経るごとにCAGリピートの長さが伸張していくことが判明している。

【0003】 CAGリピート数に異常があるポリグルタミン病の発症を規定する要因としては、CAGリピートから翻訳されるポリグルタミン自身の直接的な関与が考えられる。ポリグルタミンを発現させた培養細胞ではアポトーシス性の細胞死が見られること、およびポリグルタミンを発現させたモデル動物において神経変性や小脳の萎縮が認められることが報告されている (Nat. Genet. 13, 196-202, 1996)。また、脳の領域特異的に個々の疾患原因遺伝子産物からポリグルタミンが切り出され、該特定の領域が疾患特異的に傷害されると考えるプロセッシングモデルが提唱されている。

【0004】 最近、伸長されたポリグルタミン鎖が神経細胞の核内に存在するTAF130と呼称される蛋白質 (TATA-binding protein-associated factor; TATA-結合蛋白質-関連因子) と結合してその機能を強く阻害することが報告された (Nat. Genet. 26, 29-36, 2000)。TAF130は、cAMP応答配列結合蛋白質依存性の転写活性化に関与する遺伝子発現活性化因子の1つであるため、神経細胞の生存や機能発現に重要な神経成長因子および神経細胞の可塑性において不可欠である。従って、TAF130の機能が阻害されれば、神経細胞に重大な障害がもたらされる。さらに上記報告において、ポリグルタミン病患者の剖検脳の神経細胞内に見い出された凝集体中にポリグルタミンに加えてTAF130が存在すること、また、ポリグルタミンの発現によりもたらされる遺伝子の転写・発現阻害および細胞死が、TAF130の導入により回復あるいは緩和されることが示された。従来、ポリグルタミン病の発症機序には、凝集体の形成による神経細胞のアポトーシス誘導

が関与すると考えられてきたが、核内に集積したポリグルタミンがTAF130と結合して遺伝子の転写・発現を阻害し、核の機能障害をもたらすことが関与している可能性が示唆された。

【0005】ポリグルタミン病の早期発見および／または事前予防は、遺伝子内のCAGリピート数が健常人のそれと比較して異常である遺伝子をマーカーとして利用すれば可能である。しかしながら、個体において遺伝子は多様であり、ポリグルタミン病の診断のためには、ポリグルタミン病関連遺伝子を多種見いだすことが必須である。また、ポリグルタミン病に対する有効な防止および／または治療剤はほとんど上市されておらず、その開発が待望されている。

【0006】

【発明が解決しようとする課題】本発明の課題は、新規なポリグルタミン病関連遺伝子を単離・同定して、該遺伝子がコードするポリペプチドを明らかにし、これらを利用して該ポリグルタミン病関連遺伝子が関与する疾患の防止および／または治療剤、並びに診断のための測定手段を提供することである。

【0007】

【解決のための手段】本発明者はCAGリピート数が多く存在するcDNAに注目し、ヒト脳cDNAライブラリーから得たCAGリピート数の多いcDNAが多型を示すことを見だし、これらを家系マーカーとして使用できること、また得られたcDNAにコードされるポリペプチドをポリグルタミン病の防止および／または治療剤のスクリーニングに利用できることを見だして、本発明を完成した。

【0008】すなわち、本発明は、

(1) 下記の群より選ばれるポリヌクレオチド；

①配列表の配列番号1、配列番号3、配列番号5、配列番号7、または配列番号9に記載の塩基配列からなるポリヌクレオチドまたはその相補鎖、

②配列表の配列番号1、配列番号3、配列番号5、配列番号7、または配列番号9に記載の塩基配列からなるポリヌクレオチドにコードされるポリペプチドと少なくとも70%以上のアミノ酸配列上の相同性を有し、且つポリグルタミン配列を有するポリペプチドをコードするポリヌクレオチドまたはその相補鎖、

③上記①のポリヌクレオチドまたはその相補鎖とストリジェントな条件下でハイブリダイゼーションするポリヌクレオチド、

④上記①のポリヌクレオチドまたはその相補鎖の塩基配列において1ないし数個のアミノ酸の欠失、置換、付加、および／または挿入等の変異を有するポリヌクレオチド、および

⑤上記①から④のいずれか1のポリヌクレオチドを含有するポリヌクレオチドまたはその相補鎖、

(2) 前記(1)のポリヌクレオチドをなす塩基配列中

の、少なくとも約10個の連続する塩基配列を有するポリヌクレオチド、

(3) 前記(1)のポリヌクレオチドの部分塩基配列からなるポリヌクレオチドであって、配列表の配列番号11から配列番号20のいずれか1に記載の塩基配列からなるポリヌクレオチド、

(4) 下記の群より選ばれるポリペプチド；

(A) 配列表の配列番号1、配列番号3、配列番号5、配列番号7、または配列番号9に記載の塩基配列からなるポリヌクレオチドにコードされるポリペプチドであって、配列表の配列番号2、配列番号4、配列番号6、配列番号8、または配列番号10に記載のアミノ酸配列からなるポリペプチド、(B) 上記(A)のポリペプチドと少なくとも70%以上のアミノ酸配列上の相同性を有し、且つポリグルタミン配列を有するポリペプチド、

(C) 上記(A)のポリペプチドにおいて1ないし数個のアミノ酸の欠失、置換、付加、および／または挿入といった変異を有し、且つポリグルタミン配列を有するポリペプチド、および(D) 上記(A)、(B)、または

(C)のポリペプチドを含有するポリペプチド、

(5) 前記(1)から(3)のいずれかのポリヌクレオチドまたはその相補鎖を含有する組換えベクター、

(6) 組換えベクターが発現組換えベクターである前記(5)の組換えベクター、

(7) 前記(5)または(6)の組換えベクターにより形質転換された形質転換体、

(8) 前記(6)の組換えベクターにより形質転換された形質転換体を培養する工程を含む、前記(4)のポリペプチドの製造方法、

(9) 前記(4)のポリペプチドを免疫学的に認識する抗体、

(10) 前記(1)のポリヌクレオチドと相互作用してその発現を阻害する化合物、および／または前記(4)のポリペプチドと相互作用して該ポリペプチドのTAF130との結合を阻害する化合物の同定方法であって、前記(1)のポリヌクレオチド、前記(4)のポリペプチド、前記(5)若しくは(6)のベクター、前記

(7)の形質転換体、前記(9)の抗体のうちの少なくともいずれか1つを用いることを特徴とする方法、

(11) 前記(1)のポリヌクレオチドと相互作用してその発現を阻害する化合物の同定方法であって、化合物と該ポリヌクレオチドとの間の相互作用を可能にする条件下で、該ポリヌクレオチドとスクリーニングに供する化合物とを接触させて該化合物の相互作用を評価し、次いで、該化合物と該ポリヌクレオチドとの相互作用により生じるシグナルの存在または不存在または変化を検出することにより、該化合物が該ポリヌクレオチドと相互作用してその発現を阻害するかどうかを決定することを含む方法、

(12) 前記(4)のポリペプチドと相互作用して該ポ

リペプチドとTAF130との結合を阻害する化合物の同定方法であって、化合物と該ポリペプチドとの間の相互作用を可能にする条件下で、該ポリペプチドとスクリーニングに供する化合物とを接触させて該化合物の相互作用を評価し、次いで、該化合物と該ポリペプチドとの相互作用により生じるシグナルの存在または不存在または変化を検出することにより、該化合物が該ポリペプチドと相互作用してTAF130との結合を阻害するかどうかを決定することを含む方法、

(13) 前記(10)から(12)のいずれかの方法で同定された化合物、

(14) 前記(1)のポリヌクレオチドと相互作用してその発現を阻害する化合物、または前記(4)のポリペプチドと相互作用して該ポリペプチドとTAF130との結合を阻害する化合物、

(15) 前記(1)から(3)のいずれかのポリヌクレオチド、前記(4)のポリペプチド、前記(5)または(6)のベクター、前記(7)の形質転換体、前記

(9)の抗体、または前記(13)若しくは(14)の化合物のうちの少なくともいずれか1つを含有すること

を特徴とする医薬組成物、
(16) 前記(1)のポリヌクレオチドと相互作用してその発現を阻害する化合物、および／または前記(4)のポリペプチドと相互作用して該ポリペプチドとTAF130との結合を阻害する化合物を含有することを特徴とする医薬組成物、

(17) ポリグルタミン病の防止および／または治療剤である前記(15)および／または(16)の医薬組成物、

(18) 前記(1)のポリヌクレオチドまたは前記(4)のポリペプチドを定量的あるいは定性的に測定する手段、

(19) 前記(18)の手段を使用することを特徴とする、前記(1)のポリヌクレオチドまたは前記(4)のポリペプチドに関連する疾病の診断手段、

(20) 前記疾病がポリグルタミン病である前記(19)の診断手段、

(21) 前記(1)のポリヌクレオチド中のポリグルタミンをコードする塩基配列長を解析することを特徴とするポリグルタミン病の診断手段、

(22) 前記(21)のポリグルタミン病の診断手段において、前記(2)および／または(3)のポリヌクレオチドを使用することを特徴とするポリグルタミン病の診断手段、

(23) 前記(1)のポリヌクレオチド中のポリグルタミンをコードする塩基配列長を多型マーカーとして解析することを特徴とする家系診断手段、

(24) 配列表の配列番号13、配列番号14、配列番号15、配列番号16、配列番号17、および配列番号18に記載の塩基配列からなるポリヌクレオチドをプロ

ーブとして使用することを特徴とする前記(23)の家系診断手段、

(25) 前記(24)の家系診断手段において、配列表の配列番号13と配列番号14に記載の塩基配列からなるポリヌクレオチドの組み合わせ、配列表の配列番号15と配列番号16に記載の塩基配列からなるポリヌクレオチドの組み合わせ、および配列表の配列番号17と配列番号18に記載の塩基配列からなるポリヌクレオチドの組み合わせを、それぞれ単独でまたは複数組み合わせで使用することを特徴とする家系診断手段、

(26) 配列表の配列番号11と配列番号12に記載の塩基配列からなるポリヌクレオチドの組み合わせ、および／または配列表の配列番号19と配列番号20に記載の塩基配列からなるポリヌクレオチドの組み合わせを陰性対照として使用することを特徴とする前記(24)若しくは(25)の家系診断手段、

(27) 前記(18)から(26)のいずれかの手段に使用する試薬キットであって、前記(1)から(3)のいずれかのポリヌクレオチド、前記(4)のポリペプチド、または前記(9)の抗体を少なくとも1つ以上含んでなる試薬キット、からなる。

【0009】

【発明の実施の形態】本発明に係るcDNAは、ヒト脳cDNAライブラリーから常法に従って得たcDNAのうち、長いCAGリピートを有するものに着目して選別した5個の遺伝子である。これらの遺伝子は、配列表の配列番号1、配列番号3、配列番号5、配列番号7、または配列番号9に記載の塩基配列をそれぞれそのオープンリーディングフレーム(ORF)として含む遺伝子であり、以下本明細書においてそれぞれHH00456、HG00622、HH01321、FH16716、FJ14474と呼ぶ。各遺伝子のポリグルタミンをコードしていると考えられる塩基配列ユニット中に、HH00456は、65リピート、HG00622は30リピート、HH01321は14リピート、FH16716は21リピート、およびFJ14474は27リピートのCAGリピートを有している。また、これら遺伝子の転写産物の組織分布を常法により調べたところ、HH00456については特に全脳、卵巣、および脊髄で、HG00622については特に全脳で、HH01321については特に全脳および卵巣で、FH16716については特に全脳、卵巣、および小脳で、FJ14474については特に全脳、小脳、および脊髄で、それぞれの転写産物の高い存在が認められた。

【0010】HG00622、HH01321、およびFH16716は、それぞれの遺伝子に存在するCAGリピートに隣接する5'側および3'側の塩基配列に基づいて設計し合成したポリヌクレオチドをプライマーとして用いたポリメラーゼ連鎖反応(PCR)増幅法によるCAGリピートの解析により、CAGリピートの数が

異なる遺伝子多型が存在することが判明した。健常人では、HG00622由来のPCR産物であるDNA断片は、そのサイズが153bp、156bp、168bpの3種が検出され、153bpのホモ、156bpのホモ、153bpと156bp、153bpと168bpを持つものに分類できた。また、HH01321由来のPCR産物であるDNA断片は、そのサイズが124bp、130bp、133bpの3種が検出され、124bpと130bp、130bpのホモ、124bpと133bp、130と133bp、133bpのホモを持つものに分類できた。FH16716由来のPCR産物であるDNA断片は、そのサイズが104bp、113bp、116bp、119bp、122bp、125bpの6種が検出され、104bpと113bp、113bpのホモ、113bpと116bp、113bpと119bp、113bpと122bp、113bpと125bp、122bpのホモに分類できた。ポリグルタミン病においては健常人と比較してCAGリピートの伸長が認められることから、上記PCR産物サイズより大きくなるものは、ポリグルタミン病の発症素因を有するとして診断可能と考えられる。さらに、上記遺伝子は、その多型を解析することにより、家系マーカーの1つとして使用することができる。また、HH00456およびFJ14474には遺伝子多型が存在しないが、ポリグルタミン病を診断するためのスクリーニングにおいて陰性対照として使用することができる。本発明に係る遺伝子の具体的な取得方法および解析方法は、実施例に詳述する。さらに、ポリグルタミンをコードしていると考えられるユニットにおけるCAGリピート数の増加や、該ユニット数の増加等の変化を指標としてポリグルタミン病診断が可能である。ここで、「ポリグルタミンをコードしているユニット」とは、1つまたは複数の伸長したCAGリピートを含む塩基配列で構成される塩基配列の単位を意味する。

【0011】本発明に係るポリヌクレオチドは、配列表の配列番号1、配列番号3、配列番号5、配列番号7、または配列番号9に記載の塩基配列からなるポリヌクレオチドまたはその相補鎖；配列表の配列番号1、配列番号3、配列番号5、配列番号7、または配列番号9に記載の塩基配列からなるポリヌクレオチドによりコードされるポリペプチドと少なくとも70%以上のアミノ酸上の相同性を有し、且つポリグルタミン配列を有するポリペプチドをコードするポリヌクレオチドまたはその相補鎖；上記ポリヌクレオチドまたはその相補鎖とストリンジェントな条件下でハイブリダイゼーションするポリヌクレオチド；上記ポリヌクレオチドまたはその相補鎖の塩基配列において1ないし数個のヌクレオチドの欠失、置換、付加、および／または挿入等の変異を有するポリヌクレオチド；上記ポリヌクレオチドまたはその相補鎖のいずれか1を含有するポリヌクレオチドまたはその相

補鎖である。なお、「DNA分子にストリンジェントな条件下でハイブリダイズするDNA分子」は、例えばMolecular Cloning: A Laboratory Manual (コールドスプリングハーバーラボラトリー, 1989)等に記載の方法によって得ることができる。ここで、「ストリンジェントな条件下でハイブリタイズする」とは、例えば、 $6\times\text{SSC}$ 、0.5% SDSおよび50%ホルムアミドの溶液中で42℃にて加温した後、 $0.1\times\text{SSC}$ 、0.5% SDSの溶液中で68℃にて洗浄する条件でも依然として陽性のハイブリタイズのシグナルが観察されることを意味する。

【0012】上記ポリヌクレオチドは、いずれも家系診断のための遺伝子またはポリグルタミン病関連遺伝子として有用な遺伝子情報を提供するものであり、家系診断またはポリグルタミン病診断のためのマーカーとして使用できる。また、上記ポリヌクレオチドは、該ポリヌクレオチドがコードするポリペプチドの製造において有用であり、あるいは核酸に関する試薬・標準品としても利用できる。

【0013】また、上記ポリヌクレオチドをなす塩基配列において指定された領域に対応する連続する10個以上、好ましくは15個以上、より好ましくは20個以上のヌクレオチドからなるポリヌクレオチドおよびそれらの相補鎖も本発明の範囲に包含される。これらは、本発明に係る遺伝子を検出するためのプローブ若しくはプライマー、または遺伝子発現を調節するためのアンチセンスオリゴヌクレオチド等として使用できる。その意味で、本発明に係るポリヌクレオチドは翻訳領域のみでなく、非翻訳領域に対応するものも包含する。アンチセンスオリゴヌクレオチドは、上記各遺伝子の発現を特異的に阻害する目的で使用するときは、上記各遺伝子に固有な領域の塩基配列からなるものであることが好ましい。プローブ若しくはプライマーは、例えば、上記ポリヌクレオチドの塩基配列において少なくともCAGリピート部位に隣接する5'側、あるいは3'側の塩基配列から設計し合成したポリヌクレオチドであり得る。該ポリヌクレオチドをPCR用のプライマーとして用い、PCR遺伝子増幅法によりCAGリピート部分のポリヌクレオチドをPCR産物として得、CAGリピートのサイズを分析することにより、ポリグルタミン病の診断および家系診断を行うことができる。プローブとして特に好ましくは、HH00456から設計された配列表の配列番号11若しくは配列番号12に記載の塩基配列からなるポリヌクレオチド、HG00622から設計された配列表の配列番号13若しくは配列番号14に記載の塩基配列からなるポリヌクレオチド、HH01321から設計された配列表の配列番号15若しくは配列番号16に記載の塩基配列からなるポリヌクレオチド、FH16716から設計された配列表の配列番号17若しくは配列番号18に記載の塩基配列からなるポリヌクレオチド、また

はF J 1 4 4 7 4から設計された配列表の配列番号1 9若しくは配列番号2 0に記載の塩基配列からなるポリヌクレオチドである。

【0014】本発明に係るポリヌクレオチドは、家系診断手段または該ポリヌクレオチドが関連する疾病、例えばポリグルタミン病の診断手段として有用である。ここにおいて手段とは、目的を達成するための方法または媒体を意味する。また、本発明における媒体とは、例えば、診断用キットやポリヌクレオチドを含む組成物等を意味する。

【0015】上記診断手段は、例えば上記ポリヌクレオチドとの相互作用や反応性を利用して、相応する核酸の存在量を決定および／または該核酸の変異を測定し、存在量の変化または変異の有無を検出することにより実施できる。すなわち、本発明に係る遺伝子を診断マーカーとして定性的にあるいは定量的に測定する。試料中の当該核酸の定量的または定性的な測定法は当業者に周知の方法を利用することができる。このような測定法には、例えば増幅、PCR、RT-PCR、RNアーゼ保護、ノーザンブロッティングおよびその他のハイブリダイゼーション法が挙げられる。

【0016】測定される試料として、個体由来の細胞、例えば血液、尿、唾液、髄液、組織生検または剖検材料等を挙げることができる。また、測定される核酸は、上記各試料から自体公知の核酸調製法により得られる。核酸は、ゲノムDNAを検出に直接使用してもよく、あるいは分析前にPCR若しくはその他の増幅法を用いることにより酵素的に増幅してもよい。RNAまたはcDNAを同様に用いてもよい。正常遺伝子型との比較において、増幅生成物のサイズ変化により欠失および挿入を検出することができる。増幅DNAを標識した上記ポリペプチドをコードするDNAにハイブリダイゼーションさせることにより点突然変異を同定することができる。

【0017】具体的には例えば、本発明に係る遺伝子のCAGリピート部位に隣接する5'側または3'側の塩基配列から設計し合成した上記本発明に係るポリヌクレオチドをPCR用のプライマーとして用い、PCR遺伝子増幅法によりCAGリピート部分のポリヌクレオチドをPCR産物として得て電気泳動を行い、CAGリピートのサイズを解析用ソフトで解析することにより、家系診断を行う。またPCR産物サイズが健常人のものより大きくなるものは、ポリグルタミン病の発症素因を有するとして診断可能と考えられる。本発明に係るポリヌクレオチドを用いるCAGリピート部分の検出法および解析法は、特に上記に限定されるものではなく、公知の手法を使用できる。

【0018】本発明に係るポリグルタミン病の家系診断手段には、好ましくは、HG00622から設計された配列表の配列番号13および配列番号14に記載の塩基配列からなるポリヌクレオチドを組み合わせ、またH

H01321から設計された配列表の配列番号15および配列番号16に記載の塩基配列からなるポリヌクレオチドを組み合わせ、またFH16716から設計された配列表の配列番号17および配列番号18に記載の塩基配列からなるポリヌクレオチドを組み合わせ、使用する。これらの組み合わせは、単独で用いてもよいし、複数を組み合わせで使用することもできる。また、HH00456から設計された配列表の配列番号11と配列番号12に記載の塩基配列からなるポリヌクレオチドの組み合わせ、および／またはF J 1 4 4 7 4から設計された配列表の配列番号19と配列番号20に記載の塩基配列からなるポリヌクレオチドの組み合わせを陰性対照として使用することが好ましい。

【0019】本発明は、家系診断等に使用する、上記本発明に係るポリヌクレオチドを含む診断用試薬または診断用試薬キットも提供するものである。該診断用試薬または診断用試薬キットは、少なくとも本発明に係るポリヌクレオチドを含んでなるものであり、さらに本発明の遺伝子の多型を検出するための測定に必要な緩衝液や核酸等を含んでいてもよい。

【0020】本発明に係るポリペプチドは、上記ポリヌクレオチドによりコードされるポリペプチドである。好ましくは上記HH00456、HG00622、HH01321、FH16716、またはF J 1 4 4 7 4にそれぞれコードされる配列表の配列番号2、配列番号4、配列番号6、配列番号8、または配列番号10のいずれか1に記載のアミノ酸配列からなるポリペプチドである。また、これらのポリペプチドと少なくとも70%以上のアミノ酸配列上の相同性を有し、且つポリグルタミン配列を有するポリペプチドも本発明の範囲に包含される。本明細書において、ペプチド結合または修飾されたペプチド結合により互いに結合している2個またはそれ以上のアミノ酸を含む任意のペプチドのうち長鎖ペプチドをポリペプチドという。例えば蛋白質も本明細書においてはポリペプチドに含まれる。上記本発明に係るポリペプチドに基づいて、少なくとも含まれるポリグルタミンの長さを指標とすることにより、1ないし数個のアミノ酸の欠失、置換、付加、および／または挿入等の変異を有するアミノ酸配列からなるポリペプチドも提供される。欠失、置換、付加、および／または挿入等の変異を導入する手段は自体公知であり、例えばウルマーの技術(Science 219:666, 1983)を利用することができる。このような変異の導入において、当該ポリペプチドの基本的な性質(物性、活性、または免疫学的活性等)を変化させないという観点から、例えば、同族アミノ酸(極性アミノ酸、非極性アミノ酸、疎水性アミノ酸、親水性アミノ酸、陽性荷電アミノ酸、陰性荷電アミノ酸、芳香族アミノ酸等)の間での相互置換は容易に想定される。さらに、これらポリペプチドは、その構成アミノ基若しくはカルボキシル基等を修飾する

等、機能の著しい変更を伴わない程度に改変が可能である。さらに、上記ポリペプチドを含有するポリペプチドも、本発明の範囲に包含される。

【0021】最近、伸長したポリグルタミン鎖が神経細胞における遺伝子発現に関与する蛋白質TAF130と結合して、TAF130が関与する転写活性化を阻害することが報告された(Nat. Genet. 26, 29-36, 2000)。上記CAGリピートを含む遺伝子がコードする上記ポリペプチドはポリグルタミンを含んでおり、TAF130と結合して遺伝子発現を抑制する
10 と考えられる。ポリグルタミン病では伸長したポリグルタミン鎖とTAF130とから形成された凝集体が見い出されており、この凝集体の毒性による細胞のアポトーシス誘導またはポリグルタミン鎖とTAF130との結合による遺伝子発現の阻害が、ポリグルタミン病の原因である可能性が指摘されている。従って、本発明に係るポリペプチドの機能を阻害する物質や上記結合を阻害する物質は、ポリグルタミン病の防止および／または治療剤に使用できる。本発明に係るポリペプチドは上記ポリペプチドの機能を阻害する物質のスクリーニングに有用
20 である。

【0022】本発明に係るポリペプチドは、自体公知の合成、抽出単離、あるいは遺伝子工学的手法により得ることができる。例えば、適当な発現ベクターDNAに上記各ポリヌクレオチドを導入し、次いで該ベクターDNAを対応する宿主に形質導入し、該形質導入された形質転換体を培養することにより得られる。上記ポリヌクレオチドが組み込まれた発現ベクターDNAを導入した形質転換体は、自体公知の各宿主の培養条件に最適な条件を選択して培養される。培養は、形質転換体により発現
30 される本発明に係るポリペプチドの作用、例えばTAF130との結合能、または宿主中若しくは宿主外に産生された該ポリペプチド量を指標にして行ってもよいが、培地中の形質転換体量を指標にして継代培養またはバッチ培養を行ってもよい。または、例えば、上記各ポリヌクレオチドから自体公知の方法でmRNAを合成し、該mRNAを用いて公知の無細胞蛋白質合成方法により得ることもできる。上記ポリペプチドを含む培養液や合成反応液からのポリペプチドの回収は、該ポリペプチドの機能、例えばTAF130との結合能を指標にして、分
40 子篩、イオンカラム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等を組み合わせるか、溶解度差に基づく硫酸、アルコール等の分画手段によっても精製回収できる。より好ましくは、これらのアミノ酸配列の情報に基づき、該アミノ酸配列に対する抗体を作製し、得られたポリクローナル抗体またはモノクローナル抗体によって、特異的に吸着回収する方法を用いる。

【0023】上記ポリヌクレオチドを含む組換えベクターは、宿主の種類および使用目的により適宜選択されたベクターDNAに、上記ポリヌクレオチドを導入するこ
50

とにより得られる。ベクターDNAは、天然に存在するものを抽出したもののほか、増殖に必要な部分以外のDNAの部分が一部欠落しているものでもよい。例えば、染色体、エピソームおよびウイルス由来のベクター、例えば細菌プラスミド由来、バクテリオファージ由来、トランスポゾン由来、酵母エピソーム由来、挿入エレメント由来、酵母染色体エレメント由来、例えばバキュロウイルス、パポウイルス、SV40、ワクシニアウイルス、アデノウイルス、鶏痘ウイルス、仮性狂犬病ウイルスおよびレトロウイルス等のウイルス由来のベクター、並びにそれらを組み合わせたベクター、例えばプラスミドおよびバクテリオファージの遺伝学的エレメント由来のベクター、例えばコスミドおよびファージミド等を挙げるができる。また、目的により発現ベクターやクローニングベクター等を用いることができる。

【0024】組換えベクターは、目的の遺伝子配列と複製そして制御に関する情報を担持した遺伝子配列、例えばプロモーター、リボソーム結合部位、ターミネーター、シグナル配列、エンハンサー等、とを構成要素とし、これらを自体公知の方法により組み合わせて作製される。上記ベクターDNAに本発明に係るポリヌクレオチドを導入する方法は、自体公知の方法を適用し得る。例えば、適当な制限酵素を選択、処理してDNAを特定部位で切断し、次いで同様に処理したベクターとして用いるDNAと混合し、リガーゼによって再結合する方法が用いられる。あるいは、目的ポリヌクレオチドに適当なリンカーをライゲーションし、これを目的に適したベクターのマルチクローニングサイトへ挿入することによっても、所望の組換えベクターを得ることができる。

【0025】上記ポリヌクレオチドが組み込まれたベクターDNAにより、自体公知の宿主、例えば大腸菌、酵母、枯草菌、昆虫細胞、または動物細胞等を自体公知の方法で形質転換することにより形質転換体を得られる。形質転換を行う場合、より好ましい系としては遺伝子の安定性を考慮するならば染色体内へのインテグレート法が挙げられるが、簡便には核外遺伝子を利用した自律複製系を用いることができる。ベクターDNAの宿主細胞への導入は、例えば、Molecular Cloning: A Laboratory Manual (Sambrookら編、コールド・スプリング・ハーバー・ラボラトリー・プレス、コールド・スプリング・ハーバー、ニューヨーク、1989年)等に記載されている標準的な方法により行うことができる。具体的には、リン酸カルシウムトランスフェクション、DEAE-デキストラン媒介トランスフェクション、マイクロインジェクション、陽イオン脂質媒介トランスフェクション、エレクトロポレーション、形質導入、スクレイプ負荷(scraper loading)、バリスティック導入(ballistic introduction)および感染等を挙げることができる。

【0026】上記ポリペプチドに対する抗体は、該ポリペプチドを抗原として用いて作製する。抗原は上記ポリペプチド自体でもまたはその断片でもよく、少なくとも5個、より好ましくは少なくとも8～10個のアミノ酸で構成される。上記ポリペプチドに特異的な抗体を作製するためには、該ポリペプチドに固有なアミノ酸配列からなる領域を用いることが好ましい。このアミノ酸配列は、必ずしも該ポリペプチドのアミノ酸配列と相同である必要はなく、該ポリペプチドの立体構造上の外部への露出部位が好ましく、露出部位のアミノ酸配列が一次構造上で不連続であっても、該露出部位について連続的なアミノ酸配列であればよい。抗体は、免疫学的に該ポリペプチドを結合または認識する限り特に限定されない。この結合または認識の有無は、公知の抗原抗体結合反応によって決定される。

【0027】抗体を産生するためには、自体公知の抗体作製法を利用できる。例えば、抗原をアジュバントの存在または非存在下で単独または担体に結合して動物に投与し、体液性応答および／または細胞性応答等の免疫誘導を行うことにより得られる。担体は、それ自体が宿主に対して有害作用をおこさなければ特に限定されず、例えばセルロース、重合アミノ酸、アルブミン等が例示される。免疫される動物は、マウス、ラット、ウサギ、ヤギ、ウマ等が好適に用いられる。

【0028】ポリクローナル抗体は、上記免疫手段が施された動物の血清から自体公知の抗体回収法によって取得される。好ましい手段として免疫アフィニティクロマトグラフィー法が挙げられる。モノクローナル抗体を生産するためには、上記の免疫手段が施された動物から抗体産生細胞（例えば、脾臓またはリンパ節由来のリンパ球）を回収し、自体公知の永久増殖性細胞（例えば、P3/X63-Ag8株等のミエローマ株）への形質転換手段を導入することによ利行われる。例えば、抗体産生細胞と永久増殖性細胞とを自体公知の方法で融合させてハイブリドーマを作成し、これをクローン化した後、上記ポリペプチドを特異的に認識する抗体を産生するハイブリドーマを選別し、該ハイブリドーマの培養液から抗体を回収する。かくして得られた、上記ポリペプチドを認識し結合し得るポリクローナル抗体またはモノクローナル抗体は、該ポリペプチドの精製用抗体、試薬、または標識マーカー等として利用できる。

【0029】上記ポリヌクレオチドおよびその相補鎖、上記ポリペプチド、これらの塩基配列およびアミノ酸配列の情報に基づき形質転換させた細胞、またはこれらを免疫学的に認識する抗体は、単独または複数を組み合わせることによって、該ポリペプチドの機能、例えばTAF130との結合、を阻害または促進する物質、ポリヌクレオチドの発現を阻害または促進する物質のスクリーニングに有効な手段を提供する。スクリーニング方法は、自体公知の医薬品スクリーニングシステムを利用して構

築可能である。上記ポリヌクレオチドと相互作用してその発現を阻害する化合物を同定するには、化合物と該ポリヌクレオチドとの間の相互作用を可能にする条件下で、該ポリヌクレオチドとスクリーニングに供する化合物とを接触させて該化合物と該ポリヌクレオチドの相互作用を評価し、次いで、該化合物と該ポリヌクレオチドとの相互作用により生じるシグナルの存在または不存在または変化を検出することにより、該化合物が該ポリヌクレオチドと相互作用してその発現を阻害するかどうかを決定する方法を用いることができる。また、上記ポリペプチドと相互作用して該ポリペプチドとTAF130との結合を阻害する化合物の同定するには、化合物と該ポリペプチドとの間の相互作用を可能にする条件下で、該ポリペプチドとスクリーニングに供する化合物とを接触させて該化合物と該ポリペプチドの相互作用を評価し、次いで、該化合物と該ポリペプチドとの相互作用により生じるシグナルの存在または不存在または変化を検出することにより、該化合物が該ポリペプチドと相互作用してTAF130との結合を阻害するかどうかを決定する方法を用いることができる。具体的には例えば、自体公知のtwo-hybrid systemを利用し、上記ポリヌクレオチドと、TAF130およびTAF130が関与する遺伝子発現に関連する因子群と、該遺伝子発現を検出するシグナルとを、宿主、例えば酵母に共遺伝子導入してこれらを発現させる実験系(Nat. Genet. 26, 29-36, 2000)を構築し、この実験系に被検物質を加えて該シグナルの変化を検出することにより、上記ポリヌクレオチドがTAF130に結合することによる該遺伝子発現の阻害を防止する物質を選別できる。この実験系はスクリーニング方法の1つを説明するものであり、本発明に係るスクリーニング方法はこれに限定されるものではない。

【0030】本発明は、上記スクリーニング方法によって得られた化合物も対象とする。該化合物は、上記ポリペプチドと相互作用して該ポリペプチドとTAF130との結合を阻害する化合物、または上記ポリヌクレオチドと相互作用してその発現を阻害する化合物等であり得る。かくして選別された化合物は、生物学的有用性と毒性のバランスを考慮して選別することによって、医薬組成物として調製可能である。

【0031】また、本発明に係るポリペプチド、本発明に係るポリヌクレオチドおよびその相補鎖、これらのアミノ酸配列および塩基配列の情報に基づき作製した組換えベクター、該組換えベクターにより形質転換させた細胞、または該ポリペプチドを免疫学的に認識する抗体、該ポリペプチドの機能、例えばTAF130との結合、を阻害する化合物、または該ポリヌクレオチドと相互作用してその発現を阻害する化合物を、単独または複数組み合わせて利用することによって、これらのうち少なくとも1つを含有する医薬組成物を提供できる。該医薬組成

物のうち例えば、上記ポリペプチドのTAF130への結合を阻害する化合物、および／または上記ポリヌクレオチドの発現を阻害する化合物を含む医薬組成物は、TAF130と該ポリペプチドとの凝集体形成および／またはTAF130が関与する遺伝子発現に対する該ポリペプチドによる阻害を防止できるため、ポリグルタミン病の防止および／または治療剤として有用である。なお、医薬組成物の製剤化にあたっては、自体公知の、蛋白質、ポリヌクレオチド、抗体、または化合物等の各対象に応じた製剤化手段を導入すればよい。その投与量は、投与経路、症状、処方、性質、対象の症状の性質等に応じて適宜選択できるが、一般的には活性本体として0.01mg～100mg/日/成人ヒト、好ましくは0.1mg～10mg/日/成人ヒトである。当該分野においてよく知られた最適化のための一般的な常套の実験を用いてこれらの用量の変更を行うことができる。

【0032】また、本発明に係るポリペプチドは、診断マーカーとして定性的にあるいは定量的に測定することにより、該ポリペプチドの発現または機能に関連した疾患、例えばポリグルタミン病の診断手段に使用できる。例えば当該ポリペプチドについて個体中の生体内分布を決定および／または個体由来試料中の当該ポリペプチドの存在量を決定し、正常なものと比較することにより、診断を行うことができる。試料中の当該ポリペプチドの定量的または定性的な測定法は当業者に周知の方法を利用することができる。このような測定法には、ラジオイムノアッセイ、競争結合アッセイ、ウェスタンブロット分析およびELISAアッセイ等がある。測定される試料としては、上記核酸の測定に使用される試料と同様のものが使用可能である。

【0033】本発明は、該ポリペプチドの発現または機能に関連した疾患、例えばポリグルタミン病等に使用する診断用試薬または診断用試薬キットも提供するものである。該診断用試薬または診断用試薬キットは、本発明に係るポリヌクレオチドまたは本発明に係る抗体を少なくとも含んでなるものであり、さらに測定に必要な緩衝液等を含んでいてもよい。

【0034】

【実施例】以下、本発明を実施例および実験例に基づき具体的に説明するが、本発明は下記の実施例および実験例に限定されない。

【0035】

【実施例1】(cDNAライブラリーの構築) 二重鎖cDNA断片を、逆転写酵素(Superscript II, Invitrogen, USA)、5'部分にNot I部位を有する(dT)₁₅プライマー(5'-pGACTAGTTCTAGATCGCGAGCGGC CGCC(T)₁₅-3')、およびヒト全脳のポリ(A)⁺RNA(Clontech, USA)を用いて、製品説明書に従って作製した。得られたcDNAに

Sal IアダプターをライゲーションしてNot Iで消化した後、1%アガロースゲル(low-melting agarose gel)で電気泳動し、3kb以下のcDNA断片を除いた。次いで、3kb以上のcDNA断片をSal IおよびNot Iで消化したpBluescript II SK+ベクターにライゲーションし、エレクトロポレーション法により大腸菌に導入した(ElectroMax DH10B細胞, Invitrogen, USA)。大腸菌を寒天培地で培養して得られたアンピシリン耐性コロニー8×10⁶個から、アルカリ/硫化ドデシルナトリウム塩法でプラスミドを抽出した。挿入cDNAサイズを3kb以上に大きくするため、該プラスミドを、更にアガロースゲル電気泳動により10画分に分離した。各画分のプラスミドはそれぞれ大腸菌に再導入し、該大腸菌を37℃で3時間、アンピシリン100μg/mlを含むLB培地中で振盪培養して、得られたコロニー(各画分毎に10⁶以上)から再度プラスミドを回収した。上記のサイズ選別工程を2回以上繰り返して行うことにより約8kbまでのサイズのcDNAを得た。

【0036】(DNAシーケンシング) プラスミドDNAについて、色素プライマーサイクルシーケンシング反応(dye-primer cycle sequencing reaction)をABI PRISMTMサイクルシーケンシングキット(Perkin-Elmer Cetus, USA)とロボット反応システム(robotic reaction system)(CATALYST Turbo, Perkin-Elmer Cetus, USA)を用いて行った。反応産物を、ABI 373Aまたは377 DNAシーケンサーにより、ABI 配列分析システムINHERITを使用して分析した。色素標識したM13フォワードプライマーおよびリバースプライマーは、末端からの一方向シーケンシング(single-pass sequencing)に用い、またショットガンストラテジーによる挿入cDNAの全長配列決定にも使用した。全長配列の推定のために、挿入cDNAをベクターから切り出し、アガロースゲル電気泳動により精製した。回収した挿入cDNAはセルフ・ライゲーションさせ、ウシ小腸アルカリホスファターゼにより脱リン酸化した後、超音波処理して断片化した。得られた700bp～1000bpの断片は平滑末端であり、アガロースゲル電気泳動を行い、次いで脱リン酸化したSma I-digested M13mp18ベクターとライゲーションした。1kbのcDNA毎に16のランダムに分離したショットガンクローンを分析した。残存するギャップは、ギャップを埋めるために設計されたPCR産物の適切なプライマーを使用して、色素を結合したターミネータまたは色素を結合したプライマー・シーケンシングにより埋めた。cDNAクローンの全長配列は両鎖について

のデータから構築した。得られた配列についての相同性検索等の分析はGCGソフトウェアパッケージを用いて行った。

【0037】(CAGリピートを有する新規遺伝子)得られたcDNAクローンの塩基配列においてCAGリピートの多いものを選別し、HH00456、HG00622、HH01321、FH16716、およびFJ14474を得た。各遺伝子のポリグルタミンをコードしていると考えられる塩基配列ユニット中に、HH00456は、65リピート、HG00622は30リピート、HH01321は14リピート、FH16716は21リピート、FJ14474は27リピートのCAGリピートを有している。各クローンのORF部分の塩基配列を配列表に示した。配列表の配列番号1はHH00456、配列番号3はHG00622、配列番号5はH*

*H01321、配列番号7はFH16716、および配列番号9はFJ14474に対応する。これらのクローンは、相同検索の結果、いずれも未だ報告されていない新規な遺伝子であった。またこれら遺伝子がコードするアミノ酸配列を推定し、それぞれ配列表の配列番号2、配列番号4、配列番号6、配列番号8、および配列番号10に記載した。

【0038】(プライマーの設計)これらのCAGリピートを含む塩基配列に基づき、CAGリピート部分に隣接する5'側(U)および3'側(R)の塩基配列からプライマーとして使用するポリヌクレオチドを設計し、自体公知の方法で合成した。各プライマーの配列を表1に示す。

【0039】

【表1】

HH00456-U:5'-GCC TTC TCT CCT ACA CTA C
AC CC-3' (配列番号11)
HH00456-R:5'-GCT GGC TTG GTA GAG ATT G
GG C-3' (配列番号12)
HG00622-U:5'-GCA CAT TTC CAG CTT CTC A
GG C-3' (配列番号13)
HG00622-R:5'-GCC CTG GAT CTG TGG AAC T
TG C-3' (配列番号14)
HH01321-U:5'-CTT CAT GCC TAC CAG TCG G
GC C-3' (配列番号15)
HH01321-R:5'-GGC GAG GTT TTG GTA ATG G

【0040】 AG GG-3' (配列番号16)

【実験例1】(遺伝子の増幅と解析)PCR反応に用いたDNA Aの80ngを鋳型DNAとして、(配列番号13および14)のプライマーを用いて、FHR6C1R反応を行った。PCR反応液は、総量20μl、10×PCR Buffer (配列番号18) 2μl、(Takara社製FJ14474-5:51MOMGAGGC₂CG各AAC 0.4mMのdNTPの混合物、0(配列番号19)プライマー、0.2μMのF4474-R:5'-TGGATGATGATC-TGT および0.05U/μl Taq (配列番号20) Taq (Takara社製)からなる。PCR反応は、94℃で5分間、94℃で1分間-65℃で30秒-72℃で3分間を20サイクル、72℃で7分間の条件で実施した。

【0041】得られたPCR産物を、次にALF II DNAシーケンサー(Amersham Pharmacia)で、6%シーケンス・ゲルを用いて電気泳動した。PCR産物の1μlに反応停止用溶液5μlを加えて95℃で5分間反応させたものを泳動用サンプルとした。泳動結果を解析用ソフトフラグメントマネージャー(Fragment Manager)(Amersham Pharmacia)を使用して解析し、PCR産物のサイズを解析した。サイズマーカーとしては、Sizer50-500(Amersham Ph

armacia)を使用した。

【0042】HG00622由来のプライマー(配列表の配列番号13および14)を用いて、105例の健康人検体について遺伝子解析を行った。増幅されたDNA断片は、サイズが153bp、156bp、168bpの3種であり、この遺伝子に多型が存在することが明らかとなった。このDNA断片のサイズの差は、CAGリピート数の差を反映している。各検体において検出されたDNA断片のサイズの組み合わせにより、検体は4種類、すなわち、153bpのホモ、156bpのホモ、153bpと156bp、または153bpと168bpに分類できた。各組み合わせのDNA断片を有するヒトの割合は、それぞれ34.3%、15.2%、47.6%、または2.90%であった。

【0043】HG00622を多型マーカーとして、CAGリピートを含むDNA断片の解析を、親子DNAサンプルを用いて行った。HG00622由来のプライマー(配列表の配列番号13および14)を用いた分析結果を図1に示す。父親(A1)の対立遺伝子上のこの遺伝子に由来するCAGリピートを含むDNA断片は150bpと156bpであることが判明した。母親(A2)のこの遺伝子に由来するCAGリピートを含むDNA断片は153bpと156bpであった。一方、子供

たちは両親のそれぞれの染色体を反映する形でCAGリピートを含むDNA断片が決定されている。例えば、本遺伝子の場合、第一子（A3）、第二子（A4）とも父親由来の150bpと母親由来の153bpを獲得している。この分析により親から子への伴性的な遺伝に関与する断片が明らかになった。

【0044】

【実験例2】HH01321由来のプライマー（配列表の配列番号15および16）を用いて、実験例1と同様に104例の健常人検体について遺伝子解析を行った。その結果、サイズが124bp、130bp、133bpの3種のDNA断片が検出され、この遺伝子に多型が存在することが明らかとなった。各検体において検出されたDNA断片のサイズの組み合わせにより、検体は5種類、すなわち、124bpと130bp、130bpのホモ、124bpと133bp、130bpと133bp、または133bpのホモに分類できた。各組み合わせのDNA断片を有するヒトの割合は、それぞれ16.3%、64.4%、1.00%、17.3%、または1.00%であった。

【0045】HH01321を多型マーカーとして、CAGリピートを含むDNA断片の解析を、親子DNAサンプルを用いて行った。HH01321由来のプライマー（配列表の配列番号15および16）を用いた分析結果を図2に示す。父親（A1）の対立遺伝子上のこの遺伝子に由来するCAGリピートを含むDNA断片は130bpと133bpであることが判明した。母親（A2）のこの遺伝子に由来するCAGリピートを含むDNA断片は124bpと130bpであった。一方、子供たちは両親のそれぞれの染色体を反映する形でCAGリピートを含むDNA断片が決定されている。例えば、本遺伝子の場合、第一子（A3）では母親由来の124bpと父親由来の133bpが獲得されており、第二子（A4）では母親由来の130bpと父親由来の133bpが獲得されている。この分析により親から子への伴性的な遺伝に関与する断片が明らかになった。

【0046】

【実験例3】FH16716由来のプライマー（配列表の配列番号17および18）を用いて、実験例1と同様に34例の健常人検体について遺伝子解析を行った。その結果、サイズが104bp、113bp、116bp、119bp、122bp、125bpの6種のDNA断片が検出され、この遺伝子に多型が存在することが明らかとなった。各検体において検出されたDNA断片のサイズの組み合わせにより、検体は7種類、すなわち、104bpと113bp、113bpのホモ、113bpと116bp、113bpと119bp、113bpと122bp、113bpと125bp、または122bpのホモに分類できた。各組み合わせのDNA断片を有するヒトの割合は、それぞれ2.94%、58. 50

8%、2.94%、2.94%、26.5%、2.94%、または2.94%であった。

【0047】FH16716を多型マーカーとして、CAGリピートを含むDNA断片の解析を、親子DNAサンプルを用いて行った。FH16716由来のプライマー（配列表の配列番号17および18）を用いた分析結果を図3に示す。父親（A1）の対立遺伝子上のこの遺伝子に由来するCAGリピートを含むDNA断片は122bpのホモであることが判明した。母親（A2）のこの遺伝子に由来するCAGリピートを含むDNA断片は113bpのホモであった。一方、子供たちは両親のそれぞれの染色体を反映する形でCAGリピートを含むDNA断片が決定されている。例えば、本遺伝子の場合、第一子（A3）、第二子（A4）とも父親由来の122bpと母親由来の113bpを獲得している。この分析により親から子への伴性的な遺伝に関与する断片が明らかになった。

【0048】

【実験例4】HH00456由来のプライマー（配列表の配列番号11および12）を用いて、実験例1と同様に107例のヒト検体について遺伝子解析を行った。その結果、サイズが289bpまたは292bpの2種が取得されたのみであった。各検体において検出されたDNA断片のサイズの組み合わせは、289bpのホモが99.0%、そして289bpと292bpの組み合わせが1.00%であった。

【0049】

【実施例5】FJ14474由来のプライマー（配列表の配列番号19および20）を用いて、実験例1と同様に、親子DNAサンプルについてCAGリピートを含むDNA断片の解析を行った。その結果を図4に示す。父親（A1）、母親（A2）、第1子（A3）、および第2子（A4）の対立遺伝子上のこの遺伝子に由来するCAGリピートを含むDNA断片はいずれも141bpを示しており、この遺伝子に多型は認められなかった。

【0050】

【発明の効果】以上説示したように本発明は、CAGリピートが数多く存在し且つ多型に富む新規遺伝子および該遺伝子がコードするポリペプチドを見出し、これらを利用することにより達成されたものである。本発明によれば、本発明に係るポリヌクレオチドをプローブとして本発明に係る遺伝子の多型を解析することにより、家系診断およびCAGリピート数に由来する疾患、例えばポリグルタミン病の遺伝子診断が可能となる。またさらに、本発明は、上記遺伝子および該遺伝子がコードするポリペプチドを利用したポリグルタミン病の防止および／または治療剤のスクリーニング方法およびポリグルタミン病の防止および／または治療剤を提供することができる。すなわち、本発明は、ポリグルタミン病の早期発見に寄与するものであり、またポリグルタミン病の解明

並びにポリグルタミンの防止および／または治療にも有用なものである。

【0051】

【配列表フリーテキスト】配列番号11；配列表の配列番号1に記載の塩基配列に基づいて設計されたポリヌクレオチド。

配列番号12；配列表の配列番号1に記載の塩基配列に基づいて設計されたポリヌクレオチド。

配列番号13；配列表の配列番号3に記載の塩基配列に基づいて設計されたポリヌクレオチド。

配列番号14；配列表の配列番号3に記載の塩基配列に基づいて設計されたポリヌクレオチド。

配列番号15；配列表の配列番号5に記載の塩基配列に

基づいて設計されたポリヌクレオチド。

配列番号16；配列表の配列番号5に記載の塩基配列に基づいて設計されたポリヌクレオチド。

配列番号17；配列表の配列番号7に記載の塩基配列に基づいて設計されたポリヌクレオチド。

配列番号18；配列表の配列番号7に記載の塩基配列に基づいて設計されたポリヌクレオチド。

配列番号19；配列表の配列番号9に記載の塩基配列に基づいて設計されたポリヌクレオチド。

10 配列番号20；配列表の配列番号9に記載の塩基配列に基づいて設計されたポリヌクレオチド。

【0052】

【配列表】

SEQUENCE LISTING

<110> KAZUSA DNA RESEARCH INSTITUTE

DAIICHI PHARMACEUTICAL CO., LTD

<120> Marker for genealogical diagnosis

<130> NP01-1082

<140>

<141>

<150> JP P2000-236839

<151> 2000-08-04

<150> JP P2001-108723

<151> 2001-04-06

<160> 20

<170> PatentIn Ver. 2.1

<210> 1

<211> 3459

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(3459)

<400> 1

atg ggg gac aca gcg ccc ccg cag gcc ccc

gca gga ggg cta ggg ggg 48

Met Gly Asp Thr Ala Pro Pro Gln Ala Pro

Ala Gly Gly Leu Gly Gly

gcc tct ggg gcg ggg ctc ctt gga ggg ggc
 tca gtc acc ccg aga gtg 96
 Ala Ser Gly Ala Gly Leu Leu Gly Gly Gly
 Ser Val Thr Pro Arg Val
 20 25 30

cac agt gct atc gtg gag cgc ctc cgg gct
 cgg atc gct gtc tgc cgc 144
 His Ser Ala Ile Val Glu Arg Leu Arg Ala
 Arg Ile Ala Val Cys Arg
 35 40 45

caa cac cac ctg agc tgt gaa gga cga tat
 gaa cga ggt agg gcc gag 192
 Gln His His Leu Ser Cys Glu Gly Arg Tyr
 Glu Arg Gly Arg Ala Glu
 50 55 60

agc tca gac cgg gaa aga gaa agc acc ttg
 cag ctc ctg agc ctt gta 240
 Ser Ser Asp Arg Glu Arg Glu Ser Thr Leu
 Gln Leu Leu Ser Leu Val
 65 70 75 80

cag cat ggc cag ggg gca agg aaa gct ggc
 aaa cac acc aag gcc acc 288
 Gln His Gly Gln Gly Ala Arg Lys Ala Gly
 Lys His Thr Lys Ala Thr
 85 90 95

gcc act gct gcc acc act aca gcc cct cca
 ccg ccc cct gct gcc cct 336
 Ala Thr Ala Ala Thr Thr Thr Ala Pro Pro
 Pro Pro Pro Ala Ala Pro
 100 105 110

cct gcg gcc tcc caa gca gca gca aca gca
 gcc cca ccg ccc cca cca 384
 Pro Ala Ala Ser Gln Ala Ala Ala Thr Ala
 Ala Pro Pro Pro Pro Pro
 115 120 125

gac tat cac cat cac cac cag cag cac ctg
 ctg aac agt agc aat aat 432
 Asp Tyr His His His His Gln Gln His Leu
 Leu Asn Ser Ser Asn Asn
 130 135 140

ggt ggc agt ggt ggg ata aac gga gag cag

210	215	220	
caa agt cag att atg tca ggg acc ttg cct atg agc caa gca ccc ctg 720 Gln Ser Gln Ile Met Ser Gly Thr Leu Pro Met Ser Gln Ala Pro Leu 225 230 235 2			
40			
cga aag act aac act ctg cca tcc cat aca cat tct cct ggc aat ggc 768 Arg Lys Thr Asn Thr Leu Pro Ser His Thr His Ser Pro Gly Asn Gly 245 250 255			
ctg ttt aac atg ggc tta aag gag gta aag aag gag cca gga gag act 816 Leu Phe Asn Met Gly Leu Lys Glu Val Lys Lys Glu Pro Gly Glu Thr 260 265 270			
ctg tct tgc agt aag cac atg gat ggc caa atg acc caa gag aat att 864 Leu Ser Cys Ser Lys His Met Asp Gly Gln Met Thr Gln Glu Asn Ile 275 280 285			
ttt cct aat agg tac gga gac gac cct gga gaa caa ctg atg gat cct 912 Phe Pro Asn Arg Tyr Gly Asp Asp Pro Gly Glu Gln Leu Met Asp Pro 290 295 300			
gag ctg cag gaa ctg ttc aat gaa ctg acc aac ata tct gtg cct ccc 960 Glu Leu Gln Glu Leu Phe Asn Glu Leu Thr Asn Ile Ser Val Pro Pro 305 310 315 3			
20			
atg agt gac ctt gaa ctg gag aac atg atc aat gcc acc ata aag cag 1008 Met Ser Asp Leu Glu Leu Glu Asn Met Ile Asn Ala Thr Ile Lys Gln 325 330 335			
gat gac cca ttt aac att gac ttg ggt cag			

ggc tcc gga gca agc cgg gcc ttg cca agc
 tgg cag gaa gta tcc cat 1296
 Gly Ser Gly Ala Ser Arg Ala Leu Pro Ser
 Trp Gln Glu Val Ser His

420

425

430

gcc cag cag ctc aaa cag ata gct gct aat
 cgt cag cag cat gcc cgg 1344
 Ala Gln Gln Leu Lys Gln Ile Ala Ala Asn
 Arg Gln Gln His Ala Arg

435

440

445

atg cag cag cac cag cag cag cac cag cct
 acc aac tgg tca gcc ttg 1392
 Met Gln Gln His Gln Gln Gln His Gln Pro
 Thr Asn Trp Ser Ala Leu

450

455

460

ccc tcc tct gct gga cca tca cca ggt cca
 ttt ggg cag gag aaa atc 1440
 Pro Ser Ser Ala Gly Pro Ser Pro Gly Pro
 Phe Gly Gln Glu Lys Ile

465

470

475

4

80

ccc agc cct tct ttt ggt cag cag aca ttc
 agc cca cag agc tcc ccc 1488
 Pro Ser Pro Ser Phe Gly Gln Gln Thr Phe
 Ser Pro Gln Ser Ser Pro

485

490

495

atg cct ggg gta gct ggc ggc agc ggc cag
 tcg aaa gta atg gct aac 1536
 Met Pro Gly Val Ala Gly Gly Ser Gly Gln
 Ser Lys Val Met Ala Asn

500

505

510

tac atg tac aag gcc ggc ccc tca gcc cag
 ggt ggg cac cta gat gtc 1584
 Tyr Met Tyr Lys Ala Gly Pro Ser Ala Gln
 Gly Gly His Leu Asp Val

515

520

525

ctc atg cag caa aag cct cag gat ctc agt
 cga agt ttt att aac aac 1632
 Leu Met Gln Gln Lys Pro Gln Asp Leu Ser
 Arg Ser Phe Ile Asn Asn

610	615	620	
caa cag caa cag cag cag agc tca att tca			
gcc caa cag cag cag cag 1920			
Gln Gln Gln Gln Gln Gln Ser Ser Ile Ser			
Ala Gln Gln Gln Gln Gln			
625	630	635	6
40			
cag caa caa cag cag cag cag cag caa caa			
caa cag caa caa cag cag 1968			
Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln			
Gln Gln Gln Gln Gln Gln			
	645	650	655
cag cag cag cag caa caa cca tct tct cag			
cct gcc caa tct cta cca 2016			
Gln Gln Gln Gln Gln Gln Pro Ser Ser Gln			
Pro Ala Gln Ser Leu Pro			
	660	665	670
agc cag cct ttg cta agg tca cct ttg cca			
ctt cag caa aag ctc cta 2064			
Ser Gln Pro Leu Leu Arg Ser Pro Leu Pro			
Leu Gln Gln Lys Leu Leu			
	675	680	685
ctt cag caa atg cag aat cag ccc att gca			
gga atg gga tac caa gtc 2112			
Leu Gln Gln Met Gln Asn Gln Pro Ile Ala			
Gly Met Gly Tyr Gln Val			
690	695	700	
tcc caa caa cag aga cag gat caa cac tct			
gtg gta ggc cag aac aca 2160			
Ser Gln Gln Gln Arg Gln Asp Gln His Ser			
Val Val Gly Gln Asn Thr			
705	710	715	7
20			
ggc ccc agt cca agt cct aac ccc tgc tca			
aat cca aac act gga agt 2208			
Gly Pro Ser Pro Ser Pro Asn Pro Cys Ser			
Asn Pro Asn Thr Gly Ser			
	725	730	735
ggt tac atg aac tcc cag caa tca ctg ttg			

ggc tca tcc aca ata agc tta aac tct aac
 cag gct ttg gca aac cca 2496
 Gly Ser Ser Thr Ile Ser Leu Asn Ser Asn
 Gln Ala Leu Ala Asn Pro

820

825

830

gtt tca aca cac acc att tta act ccc aat
 tcc agc ctc ctg tct act 2544
 Val Ser Thr His Thr Ile Leu Thr Pro Asn
 Ser Ser Leu Leu Ser Thr

835

840

845

tct cac ggg aca aga atg cca tca tta tct
 aca gca gtt cag aat atg 2592
 Ser His Gly Thr Arg Met Pro Ser Leu Ser
 Thr Ala Val Gln Asn Met

850

855

860

ggg atg tat gga aat ctg cct tgt aat caa
 cct aac aca tac agt gtc 2640
 Gly Met Tyr Gly Asn Leu Pro Cys Asn Gln
 Pro Asn Thr Tyr Ser Val

865

870

875

8

80

act tca gga atg aat caa ttg acc caa cag
 aga aac cca aag caa ttg 2688
 Thr Ser Gly Met Asn Gln Leu Thr Gln Gln
 Arg Asn Pro Lys Gln Leu

885

890

895

tta gca aat caa aac aac cct atg atg cca
 cgg cca cct acc tta ggg 2736
 Leu Ala Asn Gln Asn Asn Pro Met Met Pro
 Arg Pro Pro Thr Leu Gly

900

905

910

cca agt aat aat aac aat gta gcc act ttt
 gga gct gga tct gtt ggt 2784
 Pro Ser Asn Asn Asn Asn Val Ala Thr Phe
 Gly Ala Gly Ser Val Gly

915

920

925

aat tca caa caa ttg aga cca aat tta acc
 cat agt atg gca agc atg 2832
 Asn Ser Gln Gln Leu Arg Pro Asn Leu Thr
 His Ser Met Ala Ser Met

1010	1015	1020
ccc atg aac caa atg agc caa aca cta aat		
ggg caa acc atg ggt ccc 3120		
Pro Met Asn Gln Met Ser Gln Thr Leu Asn		
Gly Gln Thr Met Gly Pro		
1025	1030	1035
1040		
ctc agg ggt ctg aat ctc aga ccc aat cag		
cta agc aca cag att ttg 3168		
Leu Arg Gly Leu Asn Leu Arg Pro Asn Gln		
Leu Ser Thr Gln Ile Leu		
	1045	1050 1055
cct aat ttg aat cag tca gga aca ggg ttg		
aat cag tcg agg acg ggc 3216		
Pro Asn Leu Asn Gln Ser Gly Thr Gly Leu		
Asn Gln Ser Arg Thr Gly		
	1060	1065 1070
atc aac cag cca cca tcc ctg acg ccc agc		
aat ttt cct tca ccc aac 3264		
Ile Asn Gln Pro Pro Ser Leu Thr Pro Ser		
Asn Phe Pro Ser Pro Asn		
	1075	1080 1085
caa agt tcc agg gct ttt caa gga act gac		
cac agc agt gac tta gct 3312		
Gln Ser Ser Arg Ala Phe Gln Gly Thr Asp		
His Ser Ser Asp Leu Ala		
1090	1095	1100
ttt gac ttc ctc agc caa caa aat gat aac		
atg ggc cct gcc cta aac 3360		
Phe Asp Phe Leu Ser Gln Gln Asn Asp Asn		
Met Gly Pro Ala Leu Asn		
1105	1110	1115
1120		
agt gat gct gat ttc att gat tct tta ttg		
aag aca gag cct ggt aat 3408		
Ser Asp Ala Asp Phe Ile Asp Ser Leu Leu		
Lys Thr Glu Pro Gly Asn		
	1125	1130 1135
gat gac tgg atg aaa gac atc aat ctt gat		

His Ser Ala Ile Val Glu Arg Leu Arg Ala
Arg Ile Ala Val Cys Arg

35

40

45

Gln His His Leu Ser Cys Glu Gly Arg Tyr
Glu Arg Gly Arg Ala Glu

50

55

60

Ser Ser Asp Arg Glu Arg Glu Ser Thr Leu
Gln Leu Leu Ser Leu Val

65

70

75

80

Gln His Gly Gln Gly Ala Arg Lys Ala Gly
Lys His Thr Lys Ala Thr

85

90

95

Ala Thr Ala Ala Thr Thr Thr Ala Pro Pro
Pro Pro Pro Ala Ala Pro

100

105

110

Pro Ala Ala Ser Gln Ala Ala Ala Thr Ala
Ala Pro Pro Pro Pro Pro

115

120

125

Asp Tyr His His His His Gln Gln His Leu
Leu Asn Ser Ser Asn Asn

130

135

140

Gly Gly Ser Gly Gly Ile Asn Gly Glu Gln
Gln Pro Pro Ala Ser Thr

145

150

155

1

60

Pro Gly Asp Gln Arg Asn Ser Ala Leu Ile
Ala Leu Gln Gly Ser Leu

165

170

175

Lys Arg Lys Gln Val Val Asn Leu Ser Pro
Ala Asn Ser Lys Arg Pro

180

185

190

Asn Gly Phe Val Asp Asn Ser Phe Leu Asp
Ile Lys Arg Ile Arg Val

195

200

205

Glu Leu Gln Glu Leu Phe Asn Glu Leu Thr
 Asn Ile Ser Val Pro Pro
 305 310 315 3
 20

Met Ser Asp Leu Glu Leu Glu Asn Met Ile
 Asn Ala Thr Ile Lys Gln
 325 330 335

Asp Asp Pro Phe Asn Ile Asp Leu Gly Gln
 Gln Ser Gln Arg Ser Thr
 340 345 350

Pro Arg Pro Ser Leu Pro Met Glu Lys Ile
 Val Ile Lys Ser Glu Tyr
 355 360 365

Ser Pro Gly Leu Thr Gln Gly Pro Ser Gly
 Ser Pro Gln Leu Arg Pro
 370 375 380

Pro Ser Ala Gly Pro Ala Phe Ser Met Ala
 Asn Ser Ala Leu Ser Thr
 385 390 395 4
 00

Ser Ser Pro Ile Pro Ser Val Pro Gln Ser
 Gln Ala Gln Pro Gln Thr
 405 410 415

Gly Ser Gly Ala Ser Arg Ala Leu Pro Ser
 Trp Gln Glu Val Ser His
 420 425 430

Ala Gln Gln Leu Lys Gln Ile Ala Ala Asn
 Arg Gln Gln His Ala Arg
 435 440 445

Met Gln Gln His Gln Gln Gln His Gln Pro
 Thr Asn Trp Ser Ala Leu
 450 455 460

Pro Ser Ser Ala Gly Pro Ser Pro Gly Pro
 Phe Gly Gln Glu Lys Ile

	565	570	575
Ser Gln Asn Lys Pro Ser Leu Leu His Tyr			
Thr Gln Gln Gln Gln Gln			
	580	585	590
Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln			
Gln Gln Gln Gln Gln Gln			
	595	600	605
Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln			
Ser Ser Ile Ser Ala Gln			
	610	615	620
Gln Gln Gln Gln Gln Gln Ser Ser Ile Ser			
Ala Gln Gln Gln Gln Gln			
625	630	635	6
40			
Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln			
Gln Gln Gln Gln Gln Gln			
	645	650	655
Gln Gln Gln Gln Gln Gln Pro Ser Ser Gln			
Pro Ala Gln Ser Leu Pro			
	660	665	670
Ser Gln Pro Leu Leu Arg Ser Pro Leu Pro			
Leu Gln Gln Lys Leu Leu			
	675	680	685
Leu Gln Gln Met Gln Asn Gln Pro Ile Ala			
Gly Met Gly Tyr Gln Val			
	690	695	700
Ser Gln Gln Gln Arg Gln Asp Gln His Ser			
Val Val Gly Gln Asn Thr			
705	710	715	7
20			
Gly Pro Ser Pro Ser Pro Asn Pro Cys Ser			
Asn Pro Asn Thr Gly Ser			
	725	730	735

Val Ser Thr His Thr Ile Leu Thr Pro Asn
Ser Ser Leu Leu Ser Thr

835 840 845

Ser His Gly Thr Arg Met Pro Ser Leu Ser
Thr Ala Val Gln Asn Met

850 855 860

Gly Met Tyr Gly Asn Leu Pro Cys Asn Gln
Pro Asn Thr Tyr Ser Val

865 870 875 8
80

Thr Ser Gly Met Asn Gln Leu Thr Gln Gln
Arg Asn Pro Lys Gln Leu

885 890 895

Leu Ala Asn Gln Asn Asn Pro Met Met Pro
Arg Pro Pro Thr Leu Gly

900 905 910

Pro Ser Asn Asn Asn Asn Val Ala Thr Phe
Gly Ala Gly Ser Val Gly

915 920 925

Asn Ser Gln Gln Leu Arg Pro Asn Leu Thr
His Ser Met Ala Ser Met

930 935 940

Pro Pro Gln Arg Thr Ser Asn Val Met Ile
Thr Ser Asn Thr Thr Ala

945 950 955 9
60

Pro Asn Trp Ala Ser Gln Glu Gly Thr Ser
Lys Gln Gln Glu Ala Leu

965 970 975

Thr Ser Ala Gly Val Arg Phe Pro Thr Gly
Thr Pro Ala Ala Tyr Thr

980 985 990

Pro Asn Gln Ser Leu Gln Gln Ala Val Gly
Ser Gln Gln Phe Ser Gln

995 1000 1005

Phe Asp Phe Leu Ser Gln Gln Asn Asp Asn
 Met Gly Pro Ala Leu Asn
 1105 1110 1115
 1120

Ser Asp Ala Asp Phe Ile Asp Ser Leu Leu
 Lys Thr Glu Pro Gly Asn
 1125 1130 1135

Asp Asp Trp Met Lys Asp Ile Asn Leu Asp
 Glu Ile Leu Gly Asn Asn
 1140 1145 1150

Ser

<210> 3
 <211> 3327
 <212> DNA
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (1)..(3327)

<400> 3
 atg gct ttc tta act cag cga acc atc cag
 gag ctg ttt gaa gtt tat 48
 Met Ala Phe Leu Thr Gln Arg Thr Ile Gln
 Glu Leu Phe Glu Val Tyr
 1 5 10 15

tct ccc atg gat gat gct ggc ttc ccg gtc
 aaa gct gag gag ttt gtg 96
 Ser Pro Met Asp Asp Ala Gly Phe Pro Val
 Lys Ala Glu Glu Phe Val
 20 25 30

gtg ctt tct cag gaa cct tct gtc acg gaa
 acc att gca ccc aaa att 144
 Val Leu Ser Gln Glu Pro Ser Val Thr Glu
 Thr Ile Ala Pro Lys Ile
 35 40 45

gca aga cct ttc ata gag gcc ctc aag agt
 att gag tat ctg gag gag 192
 Ala Arg Pro Phe Ile Glu Ala Leu Lys Ser
 Ile Glu Tyr Leu Glu Glu

Gln Leu Glu Glu Leu Ala Asp Phe Met Glu
Gln Leu Thr Pro Ile Glu

100

105

110

aaa tat gct tta aat tac ctg gaa tta ttc
cat act tct att gag caa 384
Lys Tyr Ala Leu Asn Tyr Leu Glu Leu Phe
His Thr Ser Ile Glu Gln

115

120

125

gaa aag gag aga aac agt gag gac gca gtg
atg act gca gtg agg gca 432
Glu Lys Glu Arg Asn Ser Glu Asp Ala Val
Met Thr Ala Val Arg Ala

130

135

140

tgg gag ttc tgg aac ctg aag acc ctg cag
gag agg gag gcc cgg ctg 480
Tyr Gln Ser Met Asn Tyr Lys Thr Glu Asp
Glu Asp Gly Ala Thr Glu

145

180

150

185

155

190

1

60

gag atg gag cag gag gag gcg gag atc ctg
acg tag gag gga ggg gat 528
Arg Met Thr Glu Thr Thr Ala Thr Leu
Pro Gln Asp Asp Ser Asp

195

165

200

170

205

175

gac tac agc gag gag gat gtg tgc gaa gag
gat gaa ggg aag aca gaa 578
Ile Tyr Leu Asp Ser Val Met Cys Leu Met
Tyr Glu Ala Thr Pro Ile

210

215

220

cca gag gct aag ctg ccc cct gtg tac gtg
agg aag gag cgg aag cga 720
Pro Glu Ala Lys Leu Pro Pro Val Tyr Val
Arg Lys Glu Arg Lys Arg

225

230

235

2

40

cac aaa aca gac ccc tca gct gca ggc agg
aag aag aag cag cgt cac 768
His Lys Thr Asp Pro Ser Ala Ala Gly Arg
Lys Lys Lys Gln Arg His

245

250

255

290	295	300	
aaa ccc aca gct gag cct ggt caa gac aac			
ccc gag tgg ctc atc agt 960			
Lys Pro Thr Ala Glu Pro Gly Gln Asp Asn			
Pro Glu Trp Leu Ile Ser			
305	310	315	3
20			
gag gac tgg gcg ctg ctg cag gct gta aag			
cag tta ctg gag ctg cct 1008			
Glu Asp Trp Ala Leu Leu Gln Ala Val Lys			
Gln Leu Leu Glu Leu Pro			
	325	330	335
ttg aac ctc aca atc gtg tca cct gct cac			
aca cct aat tgg gat ctt 1056			
Leu Asn Leu Thr Ile Val Ser Pro Ala His			
Thr Pro Asn Trp Asp Leu			
	340	345	350
gtc agt gac gtt gtt aac tcc tgt agc cga			
atc tac cgc tct tcc aaa 1104			
Val Ser Asp Val Val Asn Ser Cys Ser Arg			
Ile Tyr Arg Ser Ser Lys			
	355	360	365
cag tgc cgg aat cgc tac gag aat gtc atc			
att cca cga gag gag ggg 1152			
Gln Cys Arg Asn Arg Tyr Glu Asn Val Ile			
Ile Pro Arg Glu Glu Gly			
370	375	380	
aag agt aaa aac aac cgt cct ctc cgt acg			
agc cag atc tat gcc cag 1200			
Lys Ser Lys Asn Asn Arg Pro Leu Arg Thr			
Ser Gln Ile Tyr Ala Gln			
385	390	395	4
00			
gat gag aat gcc aca cac acc cag ctg tac			
acg agc cac ttt gac tta 1248			
Asp Glu Asn Ala Thr His Thr Gln Leu Tyr			
Thr Ser His Phe Asp Leu			
	405	410	415
atg aaa atg act gct ggc aag agg agt ccc			

ccg cag ccc cca cca ccc ccg cag cag cca
 ccg cca ccg ctg cca caa 1536
 Pro Gln Pro Pro Pro Pro Pro Gln Gln Pro
 Pro Pro Pro Leu Pro Gln

500

505

510

cca cag gca gcg ggc agc cag ccg cca gca
 ggg cca cca gct gtc cag 1584
 Pro Gln Ala Ala Gly Ser Gln Pro Pro Ala
 Gly Pro Pro Ala Val Gln

515

520

525

ccc caa ccc cag cca cag ccc cag acc cag
 cca cag cct gtg cag gcc 1632
 Pro Gln Pro Gln Pro Gln Pro Gln Thr Gln
 Pro Gln Pro Val Gln Ala

530

535

540

cca gcg aag gcg cag ccc gca atc acg acg
 ggg ggc agt gca gcc gta 1680
 Pro Ala Lys Ala Gln Pro Ala Ile Thr Thr
 Gly Gly Ser Ala Ala Val

545

550

555

5

60

ctg gca gga acc att aaa aca tca gtt act
 ggg acg agc atg ccc act 1728
 Leu Ala Gly Thr Ile Lys Thr Ser Val Thr
 Gly Thr Ser Met Pro Thr

565

570

575

ggt gcc gtg agt gga aat gtg atc gtg aac
 acc atc gca ggg gtc cca 1776
 Gly Ala Val Ser Gly Asn Val Ile Val Asn
 Thr Ile Ala Gly Val Pro

580

585

590

gct gcc acc ttc cag tcc atc aac aag cgc
 ctg gcg tcg cca gtg gct 1824
 Ala Ala Thr Phe Gln Ser Ile Asn Lys Arg
 Leu Ala Ser Pro Val Ala

595

600

605

cct ggg gcc ttg act acg ccg gga ggc tct
 gct ccc gcc cag gtg gtg 1872
 Pro Gly Ala Leu Thr Thr Pro Gly Gly Ser
 Ala Pro Ala Gln Val Val

690

695

700

cag cag cag cag cag cag caa caa cag cag
 cag cag cag cag cag cag 2160
 Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln
 Gln Gln Gln Gln Gln Gln
 705 710 715 7
 20

cag cag cag cag cag cag cag caa cag cag
 cag cag caa cag acg acg 2208
 Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln
 Gln Gln Gln Gln Thr Thr
 725 730 735

acg acc tct cag gtg caa gtt cca cag atc
 cag ggc cag gcc cag tcc 2256
 Thr Thr Ser Gln Val Gln Val Pro Gln Ile
 Gln Gly Gln Ala Gln Ser
 740 745 750

cca gca cag atc aaa gct gtg ggc aag ctg
 acg ccg gaa cac ctc atc 2304
 Pro Ala Gln Ile Lys Ala Val Gly Lys Leu
 Thr Pro Glu His Leu Ile
 755 760 765

aaa atg cag aag cag aaa ctg cag atg ccc
 ccg cag ccc cca ccg cca 2352
 Lys Met Gln Lys Gln Lys Leu Gln Met Pro
 Pro Gln Pro Pro Pro Pro
 770 775 780

cag gcc cag tct gcg ccc ccg cag cca aca
 gcc caa gtg caa gtg cag 2400
 Gln Ala Gln Ser Ala Pro Pro Gln Pro Thr
 Ala Gln Val Gln Val Gln
 785 790 795 8
 00

acc tcg cag ccg ccg cag cag cag agc ccc
 cag ctc acg acg gtc acg 2448
 Thr Ser Gln Pro Pro Gln Gln Gln Ser Pro
 Gln Leu Thr Thr Val Thr
 805 810 815

gcc cca agg cct ggt gcc ctg ctg acg ggc

cag cca cag aag gca gca gga cag acc gtg
 gtg gcc cag ccc gtg cac 2736
 Gln Pro Gln Lys Ala Ala Gly Gln Thr Val
 Val Ala Gln Pro Val His

900

905

910

atg cag cag ctg ctg aag ctg aag cag cag
 gcc gtc cag cag cag aag 2784
 Met Gln Gln Leu Leu Lys Leu Lys Gln Gln
 Ala Val Gln Gln Gln Lys

915

920

925

gcc atc cag ccc cag gct gca cag ggc ccg
 gca gcc gtc cag cag aag 2832
 Ala Ile Gln Pro Gln Ala Ala Gln Gly Pro
 Ala Ala Val Gln Gln Lys

930

935

940

atc acc gca cag cag atc acc acc cct ggc
 gcg cag cag aag gtt gcc 2880
 Ile Thr Ala Gln Gln Ile Thr Thr Pro Gly
 Ala Gln Gln Lys Val Ala

945

950

955

9

60

tac gcc gcg cag ccg gcc ctt aag acc cag
 ttt ctt acc aca ccc atc 2928
 Tyr Ala Ala Gln Pro Ala Leu Lys Thr Gln
 Phe Leu Thr Thr Pro Ile

965

970

975

tcc cag gcc cag aaa ctg gcc ggg gcc cag
 caa gtg cag acc cag atc 2976
 Ser Gln Ala Gln Lys Leu Ala Gly Ala Gln
 Gln Val Gln Thr Gln Ile

980

985

990

cag gtt gca aaa ctt cct caa gtt gtt caa
 cag caa aca ccc gtg gcc 3024
 Gln Val Ala Lys Leu Pro Gln Val Val Gln
 Gln Gln Thr Pro Val Ala

995

1000

1005

agc atc cag caa gtt gcc tct gct tcc cag
 cag gct tct cca cag act 3072
 Ser Ile Gln Gln Val Ala Ser Ala Ser Gln
 Gln Ala Ser Pro Gln Thr

1090	1095	1100	
aag cct ccg tgc cag			
3327			
Lys Pro Pro Cys Gln			
1105			
<210> 4			
<211> 1109			
<212> PRT			
<213> Homo sapiens			
<400> 4			
Met Ala Phe Leu Thr Gln Arg Thr Ile Gln			
Glu Leu Phe Glu Val Tyr			
1	5	10	15
Ser Pro Met Asp Asp Ala Gly Phe Pro Val			
Lys Ala Glu Glu Phe Val			
	20	25	30
Val Leu Ser Gln Glu Pro Ser Val Thr Glu			
Thr Ile Ala Pro Lys Ile			
	35	40	45
Ala Arg Pro Phe Ile Glu Ala Leu Lys Ser			
Ile Glu Tyr Leu Glu Glu			
	50	55	60
Asp Ala Gln Lys Ser Ala Gln Glu Gly Val			
Leu Gly Pro His Thr Asp			
65	70	75	80
Ala Leu Ser Ser Asp Ser Glu Asn Met Pro			
Cys Asp Glu Glu Pro Ser			
	85	90	95
Gln Leu Glu Glu Leu Ala Asp Phe Met Glu			
Gln Leu Thr Pro Ile Glu			
	100	105	110
Lys Tyr Ala Leu Asn Tyr Leu Glu Leu Phe			
His Thr Ser Ile Glu Gln			
	115	120	125

195	200	205	
Ile Tyr Leu Asp Ser Val Met Cys Leu Met			
Tyr Glu Ala Thr Pro Ile			
210	215	220	
Pro Glu Ala Lys Leu Pro Pro Val Tyr Val			
Arg Lys Glu Arg Lys Arg			
225	230	235	2
40			
His Lys Thr Asp Pro Ser Ala Ala Gly Arg			
Lys Lys Lys Gln Arg His			
245	250	255	
Gly Glu Ala Val Val Pro Pro Arg Ser Leu			
Phe Asp Arg Ala Thr Pro			
260	265	270	
Gly Leu Leu Lys Ile Arg Arg Glu Gly Lys			
Glu Gln Lys Lys Asn Ile			
275	280	285	
Leu Leu Lys Gln Gln Val Pro Phe Ala Lys			
Pro Leu Pro Thr Phe Ala			
290	295	300	
Lys Pro Thr Ala Glu Pro Gly Gln Asp Asn			
Pro Glu Trp Leu Ile Ser			
305	310	315	3
20			
Glu Asp Trp Ala Leu Leu Gln Ala Val Lys			
Gln Leu Leu Glu Leu Pro			
325	330	335	
Leu Asn Leu Thr Ile Val Ser Pro Ala His			
Thr Pro Asn Trp Asp Leu			
340	345	350	
Val Ser Asp Val Val Asn Ser Cys Ser Arg			
Ile Tyr Arg Ser Ser Lys			
355	360	365	

Ser Leu Arg Ala Glu Arg Ile Ala Lys Glu
 Lys Lys Ala Leu Ala Asp
 465 470 475 4
 80

Gln Gln Lys Ala Gln Gln Pro Ala Val Ala
 Gln Pro Pro Pro Pro Gln
 485 490 495

Pro Gln Pro Pro Pro Pro Pro Gln Gln Pro
 Pro Pro Pro Leu Pro Gln
 500 505 510

Pro Gln Ala Ala Gly Ser Gln Pro Pro Ala
 Gly Pro Pro Ala Val Gln
 515 520 525

Pro Gln Pro Gln Pro Gln Pro Gln Thr Gln
 Pro Gln Pro Val Gln Ala
 530 535 540

Pro Ala Lys Ala Gln Pro Ala Ile Thr Thr
 Gly Gly Ser Ala Ala Val
 545 550 555 5
 60

Leu Ala Gly Thr Ile Lys Thr Ser Val Thr
 Gly Thr Ser Met Pro Thr
 565 570 575

Gly Ala Val Ser Gly Asn Val Ile Val Asn
 Thr Ile Ala Gly Val Pro
 580 585 590

Ala Ala Thr Phe Gln Ser Ile Asn Lys Arg
 Leu Ala Ser Pro Val Ala
 595 600 605

Pro Gly Ala Leu Thr Thr Pro Gly Gly Ser
 Ala Pro Ala Gln Val Val
 610 615 620

His Thr Gln Pro Pro Pro Arg Ala Val Gly
 Ser Pro Ala Thr Ala Thr
 625 630 635 6

Thr Thr Ser Gln Val Gln Val Pro Gln Ile
Gln Gly Gln Ala Gln Ser

740

745

750

Pro Ala Gln Ile Lys Ala Val Gly Lys Leu
Thr Pro Glu His Leu Ile

755

760

765

Lys Met Gln Lys Gln Lys Leu Gln Met Pro
Pro Gln Pro Pro Pro Pro

770

775

780

Gln Ala Gln Ser Ala Pro Pro Gln Pro Thr
Ala Gln Val Gln Val Gln

785

790

795

8

00

Thr Ser Gln Pro Pro Gln Gln Gln Ser Pro
Gln Leu Thr Thr Val Thr

805

810

815

Ala Pro Arg Pro Gly Ala Leu Leu Thr Gly
Thr Thr Val Ala Asn Leu

820

825

830

Gln Val Ala Arg Leu Thr Arg Val Pro Thr
Ser Gln Leu Gln Ala Gln

835

840

845

Gly Gln Met Gln Thr Gln Ala Pro Gln Pro
Ala Gln Val Ala Leu Ala

850

855

860

Lys Pro Pro Val Val Ser Val Pro Ala Ala
Val Val Ser Ser Pro Gly

865

870

875

8

80

Val Thr Thr Leu Pro Met Asn Val Ala Gly
Ile Ser Val Ala Ile Gly

885

890

895

Gln Pro Gln Lys Ala Ala Gly Gln Thr Val
Val Ala Gln Pro Val His

995	1000	1005
Ser Ile Gln Gln Val Ala Ser Ala Ser Gln		
Gln Ala Ser Pro Gln Thr		
1010	1015	1020
Val Ala Leu Thr Gln Ala Thr Ala Ala Gly		
Gln Gln Val Gln Met Ile		
1025	1030	1035
1040		
Pro Ala Val Thr Ala Thr Ala Gln Val Val		
Gln Gln Lys Leu Ile Gln		
	1045	1050
		1055
Gln Gln Val Val Thr Thr Ala Ser Ala Pro		
Leu Gln Thr Pro Gly Ala		
1060	1065	1070
Pro Asn Pro Ala Gln Val Pro Ala Ser Ser		
Asp Ser Pro Ser Gln Gln		
1075	1080	1085
Pro Lys Leu Gln Met Arg Val Pro Ala Val		
Arg Leu Lys Thr Pro Thr		
1090	1095	1100

Lys Pro Pro Cys Gln

1105

<210> 5

<211> 4920

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(4920)

<400> 5

atg cag tct ttt cga gaa agg tgt ggt ttc

cat ggc aaa caa cag aac 48

Met Gln Ser Phe Arg Glu Arg Cys Gly Phe

His Gly Lys Gln Gln Asn

acg ccc tct ggc act gca gcc gcg gtg gcc
gcc gac aag tac cac cga 240
Thr Pro Ser Gly Thr Ala Ala Ala Val Ala
Ala Asp Lys Tyr His Arg
65 70 75 80

ggc agc aag gcc ctg ccc aca cag caa ggc
ctg cag ggg agg ccg gct 288
Gly Ser Lys Ala Leu Pro Thr Gln Gln Gly
Leu Gln Gly Arg Pro Ala
85 90 95
26

t	c	g	t	c	a	g	a		
A	C	G	S	L	G	A	T		
R	P	G	T	E	V	A	S	S	
Pro	Tyr	Phe	Gly	Arg	Tyr		120		125
			100					105	110

cca cag ccg cag cca cta cct gca ggg gtg
gct ggt tag gat ggg aat caa gct tgg ggg
pro gln arg gln pro leu pr8841a Gly Val
Ala Lys Tyr Asp Glu Asn
130 135 140

```

ttg atg aaa aag aca gca gtg ccc ccc agc
agg cag tat gca gag cag 480
Leu Met Lys Lys Thr Ala Val Pro Pro Ser
Arg Gln Tyr Ala Glu Gln
145 150 155 160

```

ggc gcc cag gtg ccc ttt cgg act cac tcc
ctg cac gtc cag cag cca 528
Gly Ala Gln Val Pro Phe Arg Thr His Ser
Leu His Val Gln Gln Pro
165 170 175

cgc cgc ccc cag cag ccc ctg gca tac ccc
aag ctc caa agg cag aag 576
Pro Pro Pro Gln Gln Pro Leu Ala Tyr Pro
Lys Leu Gln Arg Gln Lys
180 185 190

ctg cag aac gac att gcc tcc cct ctg ccc
 ttc ccc cag ggt acc cac 624
 Leu Gln Asn Asp Ile Ala Ser Pro Leu Pro
 Phe Pro Gln Gly Thr His
 195 200 205

ctg gcc ccg ggg cag cgg gtc cag aat ctt
 cat gcc tac cag tcg ggc 816
 Leu Ala Pro Gly Gln Arg Val Gln Asn Leu
 His Ala Tyr Gln Ser Gly

260

265

270

cgc ctc agc tat gac cag cag cag cag cag
 cag cag cag cag cag cag 864
 Arg Leu Ser Tyr Asp Gln Gln Gln Gln Gln
 Gln Gln Gln Gln Gln Gln

275

280

285

cag cag caa gcc ctt cag agc cgg cac cat
 gcc cag gaa acc ctc cat 912
 Gln Gln Gln Ala Leu Gln Ser Arg His His
 Ala Gln Glu Thr Leu His

290

295

300

tac caa aac ctc gcc aag tat cag cac tac
 ggg cag caa ggc cag ggc 960
 Tyr Gln Asn Leu Ala Lys Tyr Gln His Tyr
 Gly Gln Gln Gly Gln Gly

305

310

315

3

20

tac tgc cag ccg gac gca gcc gtc cgg acc
 cca gag cag tac tac cag 1008
 Tyr Cys Gln Pro Asp Ala Ala Val Arg Thr
 Pro Glu Gln Tyr Tyr Gln

325

330

335

acc ttc agc ccc agc tcc agc cac tca ccc
 gcc cgc tcc gtg ggc cgc 1056
 Thr Phe Ser Pro Ser Ser Ser His Ser Pro
 Ala Arg Ser Val Gly Arg

340

345

350

tca cct tcc tac agt tcc aca ccg tcg ccg
 ctg atg cca aac ctg gag 1104
 Ser Pro Ser Tyr Ser Ser Thr Pro Ser Pro
 Leu Met Pro Asn Leu Glu

355

360

365

aac ttt ccc tac agc cag cag ccg ctc agc
 acc ggg gcc ttc ccc gca 1152
 Asn Phe Pro Tyr Ser Gln Gln Pro Leu Ser
 Thr Gly Ala Phe Pro Ala

450

455

460

gtc aag aac ctc gtg tcc agg acc cca gag
cag cat aaa agc cag cac 1440

Val Lys Asn Leu Val Ser Arg Thr Pro Glu
Gln His Lys Ser Gln His

465

470

475

4

80

tgc agc ccc gaa ggg agc ggc tac tca gcc
gag ccc gca ggc aca ccg 1488

Cys Ser Pro Glu Gly Ser Gly Tyr Ser Ala
Glu Pro Ala Gly Thr Pro

485

490

495

ctg tca gag ccg ccg agc agc acg cca cag
tcc acg cat gcg gag ccg 1536

Leu Ser Glu Pro Pro Ser Ser Thr Pro Gln
Ser Thr His Ala Glu Pro

500

505

510

cag gag gcc gac tac ctg agc ggc tcc gag
gac cca ctg gag cgc agc 1584

Gln Glu Ala Asp Tyr Leu Ser Gly Ser Glu
Asp Pro Leu Glu Arg Ser

515

520

525

ttc ctc tac tgc aac cag gcc cgt ggc agc
cct gcc agg gtc aac agc 1632

Phe Leu Tyr Cys Asn Gln Ala Arg Gly Ser
Pro Ala Arg Val Asn Ser

530

535

540

aac tcg aag gcc aag ccc gag tcc gtg tcc
acc tgt tct gtg acc tct 1680

Asn Ser Lys Ala Lys Pro Glu Ser Val Ser
Thr Cys Ser Val Thr Ser

545

550

555

5

60

cct gac gac atg tcc acc aaa tct gac gac
tcc ttc cag agc cta cac 1728

Pro Asp Asp Met Ser Thr Lys Ser Asp Asp
Ser Phe Gln Ser Leu His

565

570

575

ggc agt ctg ccg ctc gac agc ttc tcc aag

gcg tgg ccc cgg cct ggg gag cca gag gcc
 ctg ccc gac tcc ttg cag 2016
 Ala Trp Pro Arg Pro Gly Glu Pro Glu Ala
 Leu Pro Asp Ser Leu Gln

660

665

670

ctg gac aag ggc ggc aat gcc aag gac ttc
 agc cca ggg ctg ttt gaa 2064
 Leu Asp Lys Gly Gly Asn Ala Lys Asp Phe
 Ser Pro Gly Leu Phe Glu

675

680

685

gac cct tcc gtg gcc ttc gct acg cct gac
 ccc aaa aag aca act ggt 2112
 Asp Pro Ser Val Ala Phe Ala Thr Pro Asp
 Pro Lys Lys Thr Thr Gly

690

695

700

cct ctc tcc ttt ggt acc aag ccc acc ctt
 ggg gtt cct gct cca gac 2160
 Pro Leu Ser Phe Gly Thr Lys Pro Thr Leu
 Gly Val Pro Ala Pro Asp

705

710

715

7

20

ccc act aca gca gct ttt gac tgt ttc ccg
 gac aca acc gct gcc agc 2208
 Pro Thr Thr Ala Ala Phe Asp Cys Phe Pro
 Asp Thr Thr Ala Ala Ser

725

730

735

tca gcg gac agc gcc aac ccc ttt gcc tgg
 cca gag gaa aac ctg ggg 2256
 Ser Ala Asp Ser Ala Asn Pro Phe Ala Trp
 Pro Glu Glu Asn Leu Gly

740

745

750

gat gct tgt ccc agg tgg gga ttg cac cct
 ggc gag ctt acc aag ggc 2304
 Asp Ala Cys Pro Arg Trp Gly Leu His Pro
 Gly Glu Leu Thr Lys Gly

755

760

765

ctg gag cag ggt ggg aag gcc tca gat ggc
 atc agc aaa ggg gac acc 2352
 Leu Glu Gln Gly Gly Lys Ala Ser Asp Gly
 Ile Ser Lys Gly Asp Thr

850

855

860

ctg gaa gct gag gag gag tac tcc tcc cta

tgt gag ctc ctg ggc agc 2640

Leu Glu Ala Glu Glu Glu Tyr Ser Ser Leu

Cys Glu Leu Leu Gly Ser

865

870

875

8

80

ccc gag cag agg cct ggc atg cag gac ccg

ctg tca ccc aag gcc cca 2688

Pro Glu Gln Arg Pro Gly Met Gln Asp Pro

Leu Ser Pro Lys Ala Pro

885

890

895

ctc atc tgc acc aag gag gag gtg gag gag

gtg ctg gac tcc aag gcc 2736

Leu Ile Cys Thr Lys Glu Glu Val Glu Glu

Val Leu Asp Ser Lys Ala

900

905

910

ggc tgg ggc tct ccg tgc cac ctc tca ggg

gag tcc gtc atc ctg ctg 2784

Gly Trp Gly Ser Pro Cys His Leu Ser Gly

Glu Ser Val Ile Leu Leu

915

920

925

ggc cct aca gtg ggc acc gag tca aag gtc

cag agc tgg ttt gag tcc 2832

Gly Pro Thr Val Gly Thr Glu Ser Lys Val

Gln Ser Trp Phe Glu Ser

930

935

940

tct ctg tca cac atg aag cca ggt gaa gag

ggg cct gat ggg gag cga 2880

Ser Leu Ser His Met Lys Pro Gly Glu Glu

Gly Pro Asp Gly Glu Arg

945

950

955

9

60

gct cca ggg gat tcc acc acc tcg gac gcc

tct ctg gcc cag aag ccc 2928

Ala Pro Gly Asp Ser Thr Thr Ser Asp Ala

Ser Leu Ala Gln Lys Pro

965

970

975

aac aag cct gct gtg ccc gag gcg ccc atc

act cgt tct ctc acg gcc ctg agt gag ccc
 cgc acg ccc gga ccc cca 3216
 Thr Arg Ser Leu Thr Ala Leu Ser Glu Pro
 Arg Thr Pro Gly Pro Pro

1060

1065

1070

ggc ctg acc acc acc cct gca ccc cca gac
 aaa ctg ggg ggc aag cag 3264
 Gly Leu Thr Thr Thr Pro Ala Pro Pro Asp
 Lys Leu Gly Gly Lys Gln

1075

1080

1085

cga gcc gcc ttc aag tcg ggc aag cgg gtg
 ggg aag ccc tca ccc aag 3312
 Arg Ala Ala Phe Lys Ser Gly Lys Arg Val
 Gly Lys Pro Ser Pro Lys

1090

1095

1100

gct gcc tcc agc ccc agc aac ccg gcc gcc
 ctg cct gtg gcc tcc gac 3360
 Ala Ala Ser Ser Pro Ser Asn Pro Ala Ala
 Leu Pro Val Ala Ser Asp

1105

1110

1115

1120

agc agc ccg atg ggc tcc aag acc aag gag
 aca gac tca ccc agc acg 3408
 Ser Ser Pro Met Gly Ser Lys Thr Lys Glu
 Thr Asp Ser Pro Ser Thr

1125

1130

1135

cct ggc aag gac cag cgc tcc atg atc ctt
 cgg tca cgc acc aaa acc 3456
 Pro Gly Lys Asp Gln Arg Ser Met Ile Leu
 Arg Ser Arg Thr Lys Thr

1140

1145

1150

cag gag atc ttc cac tcc aag cgg cgg agg
 ccc tct gag ggc cgg ctc 3504
 Gln Glu Ile Phe His Ser Lys Arg Arg Arg
 Pro Ser Glu Gly Arg Leu

1155

1160

1165

ccc aac tgc cgt gcc acc aag aag ctc ctc
 gac aac agc cac ttg ccc 3552
 Pro Asn Cys Arg Ala Thr Lys Lys Leu Leu
 Asp Asn Ser His Leu Pro

1250	1255	1260
ggt tcc ccc acc ctc ttc aag agg atg tct tct ccc aag aaa gcc aag 3840 Gly Ser Pro Thr Leu Phe Lys Arg Met Ser Ser Pro Lys Lys Ala Lys 1265 1270 1275 1280		
ccc acc aag ggc aat ggc gag cct gcc aca aag ctc cca ccc ccg gag 3888 Pro Thr Lys Gly Asn Gly Glu Pro Ala Thr Lys Leu Pro Pro Pro Glu 1285 1290 1295		
acc ccc gat gcc tgc ctc aag ctc gcc tct cgg gca gcc ttc cag ggg 3936 Thr Pro Asp Ala Cys Leu Lys Leu Ala Ser Arg Ala Ala Phe Gln Gly 1300 1305 1310		
gcc atg aag acc aag gtg ctg cca ccc cgg aag ggc cgg ggc ctg aag 3984 Ala Met Lys Thr Lys Val Leu Pro Pro Arg Lys Gly Arg Gly Leu Lys 1315 1320 1325		
ctg gaa gcc atc gtg cag aag atc acc tcg ccc agc ctc aag aag ttc 4032 Leu Glu Ala Ile Val Gln Lys Ile Thr Ser Pro Ser Leu Lys Lys Phe 1330 1335 1340		
gca tgt aaa gcg cca ggg gcc tct cct ggt aat cct ctg agc cca tcc 4080 Ala Cys Lys Ala Pro Gly Ala Ser Pro Gly Asn Pro Leu Ser Pro Ser 1345 1350 1355 1360		
ctt tcc gac aaa gac cgt ggg ctc aag ggt gct ggg ggc agc cca gtg 4128 Leu Ser Asp Lys Asp Arg Gly Leu Lys Gly Ala Gly Gly Ser Pro Val 1365 1370 1375		
ggg gtg gaa gaa ggc ctg gta aat gtg ggc		

gga ctc tcc aaa ggc ccg ctg gag aag cgg
 ccc tat ctt ggc ccg gct 4416
 Gly Leu Ser Lys Gly Pro Leu Glu Lys Arg
 Pro Tyr Leu Gly Pro Ala

1460

1465

1470

ctg ctc ctg act ccc cga gac agg gcc agt
 ggc aca caa ggg gcc agt 4464
 Leu Leu Leu Thr Pro Arg Asp Arg Ala Ser
 Gly Thr Gln Gly Ala Ser

1475

1480

1485

gag gac aac tct ggt gga gga ggc aag aag
 cca aag atg gag gag ctg 4512
 Glu Asp Asn Ser Gly Gly Gly Lys Lys
 Pro Lys Met Glu Glu Leu

1490

1495

1500

ggc ctg gcc tcc cag ccc ccg gag ggc agg
 ccc tgc cag ccc cag aca 4560
 Gly Leu Ala Ser Gln Pro Pro Glu Gly Arg
 Pro Cys Gln Pro Gln Thr

1505

1510

1515

1520

agg gca cag aaa cag cca ggc cac acc aac
 tac agc agc tat tcc aag 4608
 Arg Ala Gln Lys Gln Pro Gly His Thr Asn
 Tyr Ser Ser Tyr Ser Lys

1525

1530

1535

cgg aag cgc ctc act cgg ggc cgg gcc aag
 aac acc acc tct tca ccc 4656
 Arg Lys Arg Leu Thr Arg Gly Arg Ala Lys
 Asn Thr Thr Ser Ser Pro

1540

1545

1550

tgt aag ggg cgt gcc aag cga cga cga cag
 cag cag gtg ctg ccc ctg 4704
 Cys Lys Gly Arg Ala Lys Arg Arg Arg Gln
 Gln Gln Val Leu Pro Leu

1555

1560

1565

gat ccc gca gag cct gaa atc cgc ctc aag
 tac att tcc tct tgc aag 4752
 Asp Pro Ala Glu Pro Glu Ile Arg Leu Lys
 Tyr Ile Ser Ser Cys Lys

<211> 1640

<212> PRT

<213> Homo sapiens

<400> 6

Met Gln Ser Phe Arg Glu Arg Cys Gly Phe

His Gly Lys Gln Gln Asn

1

5

10

15

Tyr Gln Gln Thr Ser Gln Glu Thr Ser Arg

Leu Glu Asn Tyr Arg Gln

20

25

30

Pro Ser Gln Ala Gly Leu Ser Cys Asp Arg

Gln Arg Leu Leu Ala Lys

35

40

45

Asp Tyr Tyr Asn Pro Gln Pro Tyr Pro Ser

Tyr Glu Gly Gly Ala Gly

50

55

60

Thr Pro Ser Gly Thr Ala Ala Ala Val Ala

Ala Asp Lys Tyr His Arg

65

70

75

80

Gly Ser Lys Ala Leu Pro Thr Gln Gln Gly

Leu Gln Gly Arg Pro Ala

85

90

95

Phe Pro Gly Tyr Gly Val Gln Asp Ser Ser

Pro Tyr Pro Gly Arg Tyr

100

105

110

Ala Gly Glu Glu Ser Leu Gln Ala Trp Gly

Ala Pro Gln Pro Pro Pro

115

120

125

Pro Gln Pro Gln Pro Leu Pro Ala Gly Val

Ala Lys Tyr Asp Glu Asn

130

135

140

Leu Met Lys Lys Thr Ala Val Pro Pro Ser

Arg Gln Tyr Ala Glu Gln

145

150

155

1

Ala Pro Thr Ala Gln Pro His Asp Arg Pro
Leu Thr Ala Ser Ser Ser

245 250 255

Leu Ala Pro Gly Gln Arg Val Gln Asn Leu
His Ala Tyr Gln Ser Gly

260 265 270

Arg Leu Ser Tyr Asp Gln Gln Gln Gln
Gln Gln Gln Gln Gln

275 280 285

Gln Gln Gln Ala Leu Gln Ser Arg His His
Ala Gln Glu Thr Leu His

290 295 300

Tyr Gln Asn Leu Ala Lys Tyr Gln His Tyr
Gly Gln Gln Gly Gln Gly

305 310 315 3
20

Tyr Cys Gln Pro Asp Ala Ala Val Arg Thr
Pro Glu Gln Tyr Tyr Gln

325 330 335

Thr Phe Ser Pro Ser Ser Ser His Ser Pro
Ala Arg Ser Val Gly Arg

340 345 350

Ser Pro Ser Tyr Ser Ser Thr Pro Ser Pro
Leu Met Pro Asn Leu Glu

355 360 365

Asn Phe Pro Tyr Ser Gln Gln Pro Leu Ser
Thr Gly Ala Phe Pro Ala

370 375 380

Gly Ile Thr Asp His Ser His Phe Met Pro
Leu Leu Asn Pro Ser Pro

385 390 395 4
00

Thr Asp Ala Thr Ser Ser Val Asp Thr Gln
Ala Gly Asn Cys Lys Pro

405 410 415

Gln Glu Ala Asp Tyr Leu Ser Gly Ser Glu
Asp Pro Leu Glu Arg Ser

515 520 525

Phe Leu Tyr Cys Asn Gln Ala Arg Gly Ser
Pro Ala Arg Val Asn Ser

530 535 540

Asn Ser Lys Ala Lys Pro Glu Ser Val Ser
Thr Cys Ser Val Thr Ser

545 550 555 560

Pro Asp Asp Met Ser Thr Lys Ser Asp Asp
Ser Phe Gln Ser Leu His

565 570 575

Gly Ser Leu Pro Leu Asp Ser Phe Ser Lys
Phe Val Ala Gly Glu Arg

580 585 590

Asp Cys Pro Arg Leu Leu Leu Ser Ala Leu
Ala Gln Glu Asp Leu Ala

595 600 605

Ser Glu Ile Leu Gly Leu Gln Glu Ala Ile
Gly Glu Lys Ala Asp Lys

610 615 620

Ala Trp Ala Glu Ala Pro Ser Leu Val Lys
Asp Ser Ser Lys Pro Pro

625 630 635 640

Phe Ser Leu Glu Asn His Ser Ala Cys Leu
Asp Ser Val Ala Lys Ser

645 650 655

Ala Trp Pro Arg Pro Gly Glu Pro Glu Ala
Leu Pro Asp Ser Leu Gln

660 665 670

Leu Asp Lys Gly Gly Asn Ala Lys Asp Phe
Ser Pro Gly Leu Phe Glu

770	775	780	
His Glu Ala Ser Ala Cys Leu Gly Phe Gln			
Glu Glu Asp Pro Pro Gly			
785	790	795	8
00			
Glu Lys Val Ala Ser Leu Pro Gly Asp Phe			
Lys Gln Glu Glu Val Gly			
	805	810	815
Gly Val Lys Glu Glu Ala Gly Gly Leu Leu			
Gln Cys Pro Glu Val Ala			
	820	825	830
Lys Ala Asp Arg Trp Leu Glu Asp Ser Arg			
His Cys Cys Ser Thr Ala			
	835	840	845
Asp Phe Gly Asp Leu Pro Leu Leu Pro Pro			
Thr Ser Arg Lys Glu Asp			
	850	855	860
Leu Glu Ala Glu Glu Glu Tyr Ser Ser Leu			
Cys Glu Leu Leu Gly Ser			
865	870	875	8
80			
Pro Glu Gln Arg Pro Gly Met Gln Asp Pro			
Leu Ser Pro Lys Ala Pro			
	885	890	895
Leu Ile Cys Thr Lys Glu Glu Val Glu Glu			
Val Leu Asp Ser Lys Ala			
	900	905	910
Gly Trp Gly Ser Pro Cys His Leu Ser Gly			
Glu Ser Val Ile Leu Leu			
	915	920	925
Gly Pro Thr Val Gly Thr Glu Ser Lys Val			
Gln Ser Trp Phe Glu Ser			
	930	935	940

Ala Gly Ala Pro Gly Arg Gly Ala Ser Glu			
Gly Leu Pro Arg Met Cys			
	1045	1050	1055

Thr Arg Ser Leu Thr Ala Leu Ser Glu Pro			
Arg Thr Pro Gly Pro Pro			
	1060	1065	1070

Gly Leu Thr Thr Thr Pro Ala Pro Pro Asp			
Lys Leu Gly Gly Lys Gln			
	1075	1080	1085

Arg Ala Ala Phe Lys Ser Gly Lys Arg Val			
Gly Lys Pro Ser Pro Lys			
	1090	1095	1100

Ala Ala Ser Ser Pro Ser Asn Pro Ala Ala			
Leu Pro Val Ala Ser Asp			
	1105	1110	1115
	1120		

Ser Ser Pro Met Gly Ser Lys Thr Lys Glu			
Thr Asp Ser Pro Ser Thr			
	1125	1130	1135

Pro Gly Lys Asp Gln Arg Ser Met Ile Leu			
Arg Ser Arg Thr Lys Thr			
	1140	1145	1150

Gln Glu Ile Phe His Ser Lys Arg Arg Arg			
Pro Ser Glu Gly Arg Leu			
	1155	1160	1165

Pro Asn Cys Arg Ala Thr Lys Lys Leu Leu			
Asp Asn Ser His Leu Pro			
	1170	1175	1180

Ala Thr Phe Lys Val Ser Ser Ser Pro Gln			
Lys Glu Gly Arg Val Ser			
	1185	1190	1195
	1200		

Gln Arg Ala Arg Val Pro Lys Pro Gly Ala			
Gly Ser Lys Leu Ser Asp			
	1205	1210	1215

Ala Met Lys Thr Lys Val Leu Pro Pro Arg
Lys Gly Arg Gly Leu Lys

1315

1320

1325

Leu Glu Ala Ile Val Gln Lys Ile Thr Ser
Pro Ser Leu Lys Lys Phe

1330

1335

1340

Ala Cys Lys Ala Pro Gly Ala Ser Pro Gly
Asn Pro Leu Ser Pro Ser

1345

1350

1355

1360

Leu Ser Asp Lys Asp Arg Gly Leu Lys Gly
Ala Gly Gly Ser Pro Val

1365

1370

1375

Gly Val Glu Glu Gly Leu Val Asn Val Gly
Thr Gly Gln Lys Leu Pro

1380

1385

1390

Thr Ser Gly Ala Asp Pro Leu Cys Arg Asn
Pro Thr Asn Arg Ser Leu

1395

1400

1405

Lys Gly Lys Leu Met Asn Ser Lys Lys Leu
Ser Ser Thr Asp Cys Phe

1410

1415

1420

Lys Thr Glu Ala Phe Thr Ser Pro Glu Ala
Leu Gln Pro Gly Gly Thr

1425

1430

1435

1440

Ala Leu Ala Pro Lys Lys Arg Ser Arg Lys
Gly Arg Ala Gly Ala His

1445

1450

1455

Gly Leu Ser Lys Gly Pro Leu Glu Lys Arg
Pro Tyr Leu Gly Pro Ala

1460

1465

1470

Leu Leu Leu Thr Pro Arg Asp Arg Ala Ser
Gly Thr Gln Gly Ala Ser

1570	1575	1580	
Arg Leu Arg Ser Asp Ser Arg Thr Pro Ala			
Phe Ser Pro Leu Glu Pro			
1585	1590	1595	
1600			
Ser Leu Gly Ala Gln Asn Pro Arg Ser Gly			
Gln Asn Ala Pro Pro Ala			
	1605	1610	1615
Pro Ala Asp Ala Arg Pro Leu Cys Thr Thr			
Arg Asp Arg Arg Ala Tyr			
	1620	1625	1630
Ser Ala Arg Glu Gln Gly Gln Arg			
	1635	1640	
<210> 7			
<211> 3399			
<212> DNA			
<213> Homo sapiens			
<220>			
<221> CDS			
<222> (1)..(3399)			
<400> 7			
atg ggg gat ttc gca gcc ccc gct gct gcc			
gcg aat ggc agt agt att 48			
Met Gly Asp Phe Ala Ala Pro Ala Ala Ala			
Ala Asn Gly Ser Ser Ile			
1	5	10	15
tgc atc aac agt agc ctg aac agc agc ctc			
ggc ggg gcc ggg atc ggt 96			
Cys Ile Asn Ser Ser Leu Asn Ser Ser Leu			
Gly Gly Ala Gly Ile Gly			
	20	25	30
gtg aat aat act ccc aat agt act ccc gct			
gct ccg agt agc aat cac 144			
Val Asn Asn Thr Pro Asn Ser Thr Pro Ala			
Ala Pro Ser Ser Asn His			
	35	40	45

cag gct cag gtg gag cag ctg gag ctg gag
 cgc cgg gac acc gtg agc 336
 Gln Ala Gln Val Glu Gln Leu Glu Leu Glu
 Arg Arg Asp Thr Val Ser

100

105

110

ctc tac cag cgg acc ctg gag cag agg gcc
 aag aaa tcg ggc gcc ggc 384
 Leu Tyr Gln Arg Thr Leu Glu Gln Arg Ala
 Lys Lys Ser Gly Ala Gly

115

120

125

acc ggc aaa cag cag cac ccg agc aaa ccc
 cag caa gat gcg gag gct 432
 Thr Gly Lys Gln Gln His Pro Ser Lys Pro
 Gln Gln Asp Ala Glu Ala

130

135

140

gcc tcg gcg gag cag agg aac cac acg ctg
 atc atg cta caa gag act 480
 Ala Ser Ala Glu Gln Arg Asn His Thr Leu
 Ile Met Leu Gln Glu Thr

145

150

155

1

60

gtg aaa agg aag ttg gaa gga gct cga tca
 cca ctt aat gga gac cag 528
 Val Lys Arg Lys Leu Glu Gly Ala Arg Ser
 Pro Leu Asn Gly Asp Gln

165

170

175

cag aat ggt gct tgt gat ggg aat ttt tct
 ccg act agc aaa cga att 576
 Gln Asn Gly Ala Cys Asp Gly Asn Phe Ser
 Pro Thr Ser Lys Arg Ile

180

185

190

cga aag gac att tct gcg ggg atg gaa gcc
 atc aac aat ttg ccc agt 624
 Arg Lys Asp Ile Ser Ala Gly Met Glu Ala
 Ile Asn Asn Leu Pro Ser

195

200

205

aac atg cca ctg cct tca gct tct cct ctt
 cac caa ctt gac ctg aaa 672
 Asn Met Pro Leu Pro Ser Ala Ser Pro Leu
 His Gln Leu Asp Leu Lys

290	295	300	
aat gat cag gag tgg caa gaa tta ata gat			
gaa ttg gcc aac acg gtt 960			
Asn Asp Gln Glu Trp Gln Glu Leu Ile Asp			
Glu Leu Ala Asn Thr Val			
305	310	315	3
20			
cct gag gat gac ata cag gac ctg ttc aac			
gaa gac ttt gaa gag aag 1008			
Pro Glu Asp Asp Ile Gln Asp Leu Phe Asn			
Glu Asp Phe Glu Glu Lys			
	325	330	335
aag gag cca gaa ttc tcg cag cca gca act			
gag acc cct ctc tcc cag 1056			
Lys Glu Pro Glu Phe Ser Gln Pro Ala Thr			
Glu Thr Pro Leu Ser Gln			
	340	345	350
gag agt gcg agc gtg aag agc gac ccc tct			
cac tct ccc ttc gca cat 1104			
Glu Ser Ala Ser Val Lys Ser Asp Pro Ser			
His Ser Pro Phe Ala His			
	355	360	365
gtc tcc atg gga tct ccc cag gcg agg cct			
tct tct tct ggt cct ccc 1152			
Val Ser Met Gly Ser Pro Gln Ala Arg Pro			
Ser Ser Ser Gly Pro Pro			
370	375	380	
ttt tct act gtc tcc acg gcc act agt tta			
cct tct gtt gcc agc act 1200			
Phe Ser Thr Val Ser Thr Ala Thr Ser Leu			
Pro Ser Val Ala Ser Thr			
385	390	395	4
00			
ccc gca gct cca aac cct gca agc tca cca			
gca aac tgt gct gtc cag 1248			
Pro Ala Ala Pro Asn Pro Ala Ser Ser Pro			
Ala Asn Cys Ala Val Gln			
	405	410	415
tcc cct caa act cca aac caa gcc cac act			

cag caa cag cag cag cag cag caa cag
 cac tca aat cag act tca 1536
 Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln
 His Ser Asn Gln Thr Ser

500

505

510

aat tgg tct ccc tta gga cct ccc tct agt
 cca tat gga gca gct ttt 1584
 Asn Trp Ser Pro Leu Gly Pro Pro Ser Ser
 Pro Tyr Gly Ala Ala Phe

515

520

525

act gca gaa aaa cca aat agc cca atg atg
 tac ccc caa gcc ttt aac 1632
 Thr Ala Glu Lys Pro Asn Ser Pro Met Met
 Tyr Pro Gln Ala Phe Asn

530

535

540

aac caa aac cct ata gtg cct cca atg gca
 aac aac ctg cag aag aca 1680
 Asn Gln Asn Pro Ile Val Pro Pro Met Ala
 Asn Asn Leu Gln Lys Thr

545

550

555

5

60

aca atg aat aac tac ctc cct cag aat cac
 atg aat atg atc aat cag 1728
 Thr Met Asn Asn Tyr Leu Pro Gln Asn His
 Met Asn Met Ile Asn Gln

565

570

575

cag cca aat aac ttg ggt aca aac tcc tta
 aac aaa cag cac aat att 1776
 Gln Pro Asn Asn Leu Gly Thr Asn Ser Leu
 Asn Lys Gln His Asn Ile

580

585

590

ctg act tat ggc aac act aaa ccc ctg acc
 cac ttc aat gca gac ctg 1824
 Leu Thr Tyr Gly Asn Thr Lys Pro Leu Thr
 His Phe Asn Ala Asp Leu

595

600

605

agt cag agg atg aca cca cca gtg gcc aac
 ccc aac aaa aac ccc ttg 1872
 Ser Gln Arg Met Thr Pro Pro Val Ala Asn
 Pro Asn Lys Asn Pro Leu

690

695

700

tcc tcc gtg ccc cca ggc tca ggt ggc atg
 gtc tca gga gcc agt ccc 2160
 Ser Ser Val Pro Pro Gly Ser Gly Gly Met
 Val Ser Gly Ala Ser Pro
 705 710 715 7
 20

gca ggc ccc ggc ttc ctg ggc agc cag ccc
 caa gca gcc atc atg aag 2208
 Ala Gly Pro Gly Phe Leu Gly Ser Gln Pro
 Gln Ala Ala Ile Met Lys
 725 730 735

cag atg ctc att gat cag cgg gcc cag ttg
 ata gag cag cag aag caa 2256
 Gln Met Leu Ile Asp Gln Arg Ala Gln Leu
 Ile Glu Gln Gln Lys Gln
 740 745 750

cag ttc ctg cgg gag caa agg cag cag cag
 cag cag cag cag cag att 2304
 Gln Phe Leu Arg Glu Gln Arg Gln Gln Gln
 Gln Gln Gln Gln Gln Ile
 755 760 765

ttg gcg gaa cag cag ttg cag caa tca cat
 cta ccc cgg cag cac ctc 2352
 Leu Ala Glu Gln Gln Leu Gln Gln Ser His
 Leu Pro Arg Gln His Leu
 770 775 780

cag cca cag cgg aat cca tac cca gtg cag
 cag gtc aat cag ttt caa 2400
 Gln Pro Gln Arg Asn Pro Tyr Pro Val Gln
 Gln Val Asn Gln Phe Gln
 785 790 795 8
 00

ggt tct ccc cag gat ata gca gcc gta aga
 agc caa gca gcc ctc cag 2448
 Gly Ser Pro Gln Asp Ile Ala Ala Val Arg
 Ser Gln Ala Ala Leu Gln
 805 810 815

agc atg cga acg tca cgg ctg atg gca cag

caa cct gcc tct ggg cag ggg gtt gga atg
 gtg agt ggc ttt ggt cag 2736
 Gln Pro Ala Ser Gly Gln Gly Val Gly Met
 Val Ser Gly Phe Gly Gln

900

905

910

agc atg ctg gtg aac tca gcc att acc cag
 caa cat cca cag atg aaa 2784
 Ser Met Leu Val Asn Ser Ala Ile Thr Gln
 Gln His Pro Gln Met Lys

915

920

925

ggg cca gta ggc cag gcc ttg cct agg ccc
 caa gcc cct cca agg ctg 2832
 Gly Pro Val Gly Gln Ala Leu Pro Arg Pro
 Gln Ala Pro Pro Arg Leu

930

935

940

cag agc ctt atg gga aca gtc cag caa gga
 gca caa agc tgg caa cag 2880
 Gln Ser Leu Met Gly Thr Val Gln Gln Gly
 Ala Gln Ser Trp Gln Gln

945

950

955

9

60

agg agc ttg cag ggc atg cct ggg agg act
 agt gga gaa ttg gga cca 2928
 Arg Ser Leu Gln Gly Met Pro Gly Arg Thr
 Ser Gly Glu Leu Gly Pro

965

970

975

ttc aac aat ggc gcc agc tac cct ctt caa
 gct ggg cag ccg aga ctg 2976
 Phe Asn Asn Gly Ala Ser Tyr Pro Leu Gln
 Ala Gly Gln Pro Arg Leu

980

985

990

acc aag cag cac ttc cca cag gga ctg agc
 cag tca gtc gtg gat gct 3024
 Thr Lys Gln His Phe Pro Gln Gly Leu Ser
 Gln Ser Val Val Asp Ala

995

1000

1005

aac acg ggc aca gtg agg acc ctc aac cca
 gct gcc atg ggt cgg cag 3072
 Asn Thr Gly Thr Val Arg Thr Leu Asn Pro
 Ala Ala Met Gly Arg Gln

1090	1095	1100	
gac ggt gca gac ctt gtg gac tcc atc atc			
aaa ggc ggg cca ggg gac 3360			
Asp Gly Ala Asp Leu Val Asp Ser Ile Ile			
Lys Gly Gly Pro Gly Asp			
1105	1110	1115	
1120			
gag tgg atg cag gag ctt gat gaa ttg ttt			
ggt aac ccc 3399			
Glu Trp Met Gln Glu Leu Asp Glu Leu Phe			
Gly Asn Pro			
	1125	1130	
<210> 8			
<211> 1133			
<212> PRT			
<213> Homo sapiens			
<400> 8			
Met Gly Asp Phe Ala Ala Pro Ala Ala Ala			
Ala Asn Gly Ser Ser Ile			
1	5	10	15
Cys Ile Asn Ser Ser Leu Asn Ser Ser Leu			
Gly Gly Ala Gly Ile Gly			
	20	25	30
Val Asn Asn Thr Pro Asn Ser Thr Pro Ala			
Ala Pro Ser Ser Asn His			
	35	40	45
Pro Ala Ala Gly Gly Cys Gly Gly Ser Gly			
Gly Pro Gly Gly Gly Ser			
50	55	60	
Ala Ala Val Pro Lys His Ser Thr Val Val			
Glu Arg Leu Arg Gln Arg			
65	70	75	80
Ile Glu Gly Cys Arg Arg His His Val Asn			
Cys Glu Asn Arg Tyr Gln			
	85	90	95

Gln Asn Gly Ala Cys Asp Gly Asn Phe Ser
Pro Thr Ser Lys Arg Ile

180

185

190

Arg Lys Asp Ile Ser Ala Gly Met Glu Ala
Ile Asn Asn Leu Pro Ser

195

200

205

Asn Met Pro Leu Pro Ser Ala Ser Pro Leu
His Gln Leu Asp Leu Lys

210

215

220

Pro Ser Leu Pro Leu Gln Asn Ser Gly Thr
His Thr Pro Gly Leu Leu

225

230

235

2

40

Glu Asp Leu Ser Lys Asn Gly Arg Leu Pro
Glu Ile Lys Leu Pro Val

245

250

255

Asn Gly Cys Ser Asp Leu Glu Asp Ser Phe
Thr Ile Leu Gln Ser Lys

260

265

270

Asp Leu Lys Gln Glu Pro Leu Asp Asp Pro
Thr Cys Ile Asp Thr Ser

275

280

285

Glu Thr Ser Leu Ser Asn Gln Asn Lys Leu
Phe Ser Asp Ile Asn Leu

290

295

300

Asn Asp Gln Glu Trp Gln Glu Leu Ile Asp
Glu Leu Ala Asn Thr Val

305

310

315

3

20

Pro Glu Asp Asp Ile Gln Asp Leu Phe Asn
Glu Asp Phe Glu Glu Lys

325

330

335

Lys Glu Pro Glu Phe Ser Gln Pro Ala Thr
Glu Thr Pro Leu Ser Gln

340

345

350

Gly Ser Ala Ser Gly Pro Val Ala Val Pro
Ser Ser Asp Met Ser Pro

450 455 460

Ala Glu Gln Leu Lys Gln Met Ala Ala Gln
Gln Gln Gln Arg Ala Lys

465 470 475 480

Leu Met Gln Gln Lys Gln Gln Gln Gln Gln
Gln Gln Gln Gln Gln Gln

485 490 495

Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln
His Ser Asn Gln Thr Ser

500 505 510

Asn Trp Ser Pro Leu Gly Pro Pro Ser Ser
Pro Tyr Gly Ala Ala Phe

515 520 525

Thr Ala Glu Lys Pro Asn Ser Pro Met Met
Tyr Pro Gln Ala Phe Asn

530 535 540

Asn Gln Asn Pro Ile Val Pro Pro Met Ala
Asn Asn Leu Gln Lys Thr

545 550 555 560

Thr Met Asn Asn Tyr Leu Pro Gln Asn His
Met Asn Met Ile Asn Gln

565 570 575

Gln Pro Asn Asn Leu Gly Thr Asn Ser Leu
Asn Lys Gln His Asn Ile

580 585 590

Leu Thr Tyr Gly Asn Thr Lys Pro Leu Thr
His Phe Asn Ala Asp Leu

595 600 605

Ser Gln Arg Met Thr Pro Pro Val Ala Asn
Pro Asn Lys Asn Pro Leu

705 710 715 7
20

Ala Gly Pro Gly Phe Leu Gly Ser Gln Pro
Gln Ala Ala Ile Met Lys

725 730 735

Gln Met Leu Ile Asp Gln Arg Ala Gln Leu
Ile Glu Gln Gln Lys Gln

740 745 750

Gln Phe Leu Arg Glu Gln Arg Gln Gln Gln
Gln Gln Gln Gln Gln Ile

755 760 765

Leu Ala Glu Gln Gln Leu Gln Gln Ser His
Leu Pro Arg Gln His Leu

770 775 780

Gln Pro Gln Arg Asn Pro Tyr Pro Val Gln
Gln Val Asn Gln Phe Gln

785 790 795 8
00

Gly Ser Pro Gln Asp Ile Ala Ala Val Arg
Ser Gln Ala Ala Leu Gln

805 810 815

Ser Met Arg Thr Ser Arg Leu Met Ala Gln
Asn Ala Gly Met Met Gly

820 825 830

Ile Gly Pro Ser Gln Asn Pro Gly Thr Met
Ala Thr Ala Ala Ala Gln

835 840 845

Ser Glu Met Gly Leu Ala Pro Tyr Ser Thr
Thr Pro Thr Ser Gln Pro

850 855 860

Gly Met Tyr Asn Met Ser Thr Gly Met Thr
Gln Met Leu Gln His Pro

865 870 875 8
80

Phe Asn Asn Gly Ala Ser Tyr Pro Leu Gln
Ala Gly Gln Pro Arg Leu

980

985

990

Thr Lys Gln His Phe Pro Gln Gly Leu Ser
Gln Ser Val Val Asp Ala

995

1000

1005

Asn Thr Gly Thr Val Arg Thr Leu Asn Pro
Ala Ala Met Gly Arg Gln

1010

1015

1020

Met Met Pro Ser Leu Pro Gly Gln Gln Gly
Thr Ser Gln Ala Arg Pro

1025

1030

1035

1040

Met Val Met Ser Gly Leu Ser Gln Gly Val
Pro Gly Met Pro Ala Phe

1045

1050

1055

Ser Gln His Pro Ala Gln Gln Gln Ile Pro
Ser Gly Ser Phe Ala Pro

1060

1065

1070

Ser Ser Gln Ser Gln Ala Tyr Glu Arg Asn
Ala Pro Gln Asp Val Ser

1075

1080

1085

Tyr Asn Tyr Ser Gly Asp Gly Ala Gly Gly
Ser Phe Pro Gly Leu Pro

1090

1095

1100

Asp Gly Ala Asp Leu Val Asp Ser Ile Ile
Lys Gly Gly Pro Gly Asp

1105

1110

1115

1120

Glu Trp Met Gln Glu Leu Asp Glu Leu Phe
Gly Asn Pro

1125

1130

atc ctc gac cca aag agc agt gtg acc cct
 gcc atc atc tcg gcc gcc 144
 Ile Leu Asp Pro Lys Ser Ser Val Thr Pro
 Ala Ile Ile Ser Ala Ala
 35 40 45

cta cag caa gtg gtt cac aat aag agt cta
 gtc act gct ggt ggg gct 192
 Leu Gln Gln Val Val His Asn Lys Ser Leu
 Val Thr Ala Gly Gly Ala
 50 55 60

ttg ggg aac ccc ccc agc agg ggt gag aga
 agg ctg gag gcc agc atg 240
 Leu Gly Asn Pro Pro Ser Arg Gly Glu Arg
 Arg Leu Glu Ala Ser Met
 65 70 75 80

ggg agg cca gag gtt agc atg atg agc agc
 agt gcc agt aag aat ctg 288
 Gly Arg Pro Glu Val Ser Met Met Ser Ser
 Ser Ala Ser Lys Asn Leu
 85 90 95

aag ttc aaa att agc ccc agt gct cca gag
 acc tca tgg aat tct caa 336
 Lys Phe Lys Ile Ser Pro Ser Ala Pro Glu
 Thr Ser Trp Asn Ser Gln
 100 105 110

cat cag ctg ggt gca gag gtc tct tcc agc
 ccc aga gca ccc aca ggc 384
 His Gln Leu Gly Ala Glu Val Ser Ser Ser
 Pro Arg Ala Pro Thr Gly
 115 120 125

agc cgg gct gac agc ctg cac ctc tcc caa
 caa gag gac agt ctg cct 432
 Ser Arg Ala Asp Ser Leu His Leu Ser Gln
 Gln Glu Asp Ser Leu Pro
 130 135 140

gtt caa aat ttc cct ccc aaa agc tat ctt
 ttg cga aca agc cga gag 480
 Val Gln Asn Phe Pro Pro Lys Ser Tyr Leu
 Leu Arg Thr Ser Arg Glu
 145 150 155 1

60

ggc acc gag tac ctg caa cct cca gca cct

agc agc ccc atc aat gtc cag cgc att cgt
 tct acc agc ctg gag tcc 1344
 Ser Ser Pro Ile Asn Val Gln Arg Ile Arg
 Ser Thr Ser Leu Glu Ser
 435 440 445

cga gag tgc cga tcg gac cct gag agt ggt
 gtt tcg tgc ctg acc acg 1392
 Arg Glu Cys Arg Ser Asp Pro Glu Ser Gly
 Val Ser Cys Leu Thr Thr
 450 455 460

tgt gcc tcg ggg ggc gag tgt ctg gga gct
 ccc aat tac agg aaa ctg 1440
 Cys Ala Ser Gly Gly Glu Cys Leu Gly Ala
 Pro Asn Tyr Arg Lys Leu
 465 470 475 4
 80

atg cgc cgc tac agt atc agt gag ctg gac
 cag ggt gac agg gcc tcg 1488
 Met Arg Arg Tyr Ser Ile Ser Glu Leu Asp
 Gln Gly Asp Arg Ala Ser
 485 490 495

ctg acc tcg gat gtc tac cca cat cct ccc
 ctg ggc atg ctg ccc agg 1536
 Leu Thr Ser Asp Val Tyr Pro His Pro Pro
 Leu Gly Met Leu Pro Arg
 500 505 510

gag gcc aag gag gta gag gca agc ctc ccc
 ata gcc ttg ggt ccc aaa 1584
 Glu Ala Lys Glu Val Glu Ala Ser Leu Pro
 Ile Ala Leu Gly Pro Lys
 515 520 525

agc agg tct ctg gag tca ccg acg ctg gga
 gac ccc tcc tac gtc cag 1632
 Ser Arg Ser Leu Glu Ser Pro Thr Leu Gly
 Asp Pro Ser Tyr Val Gln
 530 535 540

gtt gcc cca gag acc aaa ggc ccc aga cag
 atg gcc gtg ttc tca ctg 1680
 Val Ala Pro Glu Thr Lys Gly Pro Arg Gln
 Met Ala Val Phe Ser Leu

625 630 635 6
40

agc agc agc cct gag gcc tcc cgc act cag
gag att gac ctc cgt gtg 1968
Ser Ser Ser Pro Glu Ala Ser Arg Thr Gln
Glu Ile Asp Leu Arg Val

645 650 655

tcc acc ttc gag ggg agc ctg gcc aag atc
aat gcc ctg cgg gcc cat 2016
Ser Thr Phe Glu Gly Ser Leu Ala Lys Ile
Asn Ala Leu Arg Ala His

660 665 670

gcc tat ggc ctc cct gat ggc ttc ctg gct
gcc cgg ctg gac acc aac 2064
Ala Tyr Gly Leu Pro Asp Gly Phe Leu Ala
Ala Arg Leu Asp Thr Asn

675 680 685

gag ctg ctg aca gtc ctg cgg cag tgt gtg
gcc agc ccc gag gcc cgt 2112
Glu Leu Leu Thr Val Leu Arg Gln Cys Val
Ala Ser Pro Glu Ala Arg

690 695 700

gcc ccc aag ccc tat gtg tct cag atc tcc
gag tat aag ctt gag cta 2160
Ala Pro Lys Pro Tyr Val Ser Gln Ile Ser
Glu Tyr Lys Leu Glu Leu

705 710 715 7
20

gct ctc aag ttc aag gag ctc cgg gcc tcc
tgc cgc cgt gtg gcc aat 2208
Ala Leu Lys Phe Lys Glu Leu Arg Ala Ser
Cys Arg Arg Val Ala Asn

725 730 735

gtg gac aag agc cca act cac atg ctg gca
gcc atc acg ggc agc ttc 2256
Val Asp Lys Ser Pro Thr His Met Leu Ala
Ala Ile Thr Gly Ser Phe

740 745 750

cag gtg ctg agc agc ctc att gag acc ttc

gtg gca gca gct gca ggg gca gcc aca gag
 cat cca cca ggc tcc cca 2544
 Val Ala Ala Ala Ala Gly Ala Ala Thr Glu
 His Pro Pro Gly Ser Pro
 835 840 845

act tcg gcg act gtt atg agc aca ttc acc
 cac tcc tta aaa acc ctt 2592
 Thr Ser Ala Thr Val Met Ser Thr Phe Thr
 His Ser Leu Lys Thr Leu
 850 855 860

att aag
 2598
 Ile Lys
 865

<210> 10
 <211> 866
 <212> PRT
 <213> Homo sapiens

<400> 10
 Met Ser Leu Arg Ala Ala Thr Ser Ser Leu
 Ser Glu Glu Gln Val Ser
 1 5 10 15

Glu Leu Arg Asp Asn Leu Pro Lys Glu Val
 Arg Leu Ser Pro Lys Leu
 20 25 30

Ile Leu Asp Pro Lys Ser Ser Val Thr Pro
 Ala Ile Ile Ser Ala Ala
 35 40 45

Leu Gln Gln Val Val His Asn Lys Ser Leu
 Val Thr Ala Gly Gly Ala
 50 55 60

Leu Gly Asn Pro Pro Ser Arg Gly Glu Arg
 Arg Leu Glu Ala Ser Met
 65 70 75 80

Gly Arg Pro Glu Val Ser Met Met Ser Ser
 Ser Ala Ser Lys Asn Leu

	165	170	175
Val Gly Gly Lys Pro Thr Leu Gln Lys Gln Gly Thr Ile Ser Ser Gln			
	180	185	190
Gly Glu Lys Ala Gln Leu Glu Ser Thr Pro Lys Arg Ser Lys Leu Glu			
	195	200	205
Glu Thr Ser Leu Val Pro Arg Ala Thr Tyr Pro Met Ala Leu Gln Ser			
	210	215	220
Pro Ser Cys Gln Ser Arg Ser His Ser Pro Ser Cys Gln Pro His Gly			
225	230	235	2
40			
His Ser Pro Ser Ser Gln Ser Arg Gly Gln Ser Pro Ser Cys Gln Pro			
	245	250	255
Arg Gly Gln Ser Pro Leu Arg Ser Gln Ala Ala Ser Arg Gln Val Ser			
	260	265	270
Thr Met Pro Ser Arg Lys Leu Glu Thr Thr Leu Asn Gly Ala His Ser			
	275	280	285
Thr Ser Glu Gly Pro Ala Lys Pro Lys Ser Ser Arg Gly Pro Phe Arg			
	290	295	300
Leu Arg Asn Leu Phe Ser Ala Thr Phe Pro Thr Arg Gln Lys Lys Glu			
305	310	315	3
20			
Thr Asp Glu Arg Gln Ala Gln Leu Gln Lys Val Lys Gln Tyr Glu Leu			
	325	330	335

Ser Ser Pro Ile Asn Val Gln Arg Ile Arg
Ser Thr Ser Leu Glu Ser

435

440

445

Arg Glu Cys Arg Ser Asp Pro Glu Ser Gly
Val Ser Cys Leu Thr Thr

450

455

460

Cys Ala Ser Gly Gly Glu Cys Leu Gly Ala
Pro Asn Tyr Arg Lys Leu

465

470

475

4

80

Met Arg Arg Tyr Ser Ile Ser Glu Leu Asp
Gln Gly Asp Arg Ala Ser

485

490

495

Leu Thr Ser Asp Val Tyr Pro His Pro Pro
Leu Gly Met Leu Pro Arg

500

505

510

Glu Ala Lys Glu Val Glu Ala Ser Leu Pro
Ile Ala Leu Gly Pro Lys

515

520

525

Ser Arg Ser Leu Glu Ser Pro Thr Leu Gly
Asp Pro Ser Tyr Val Gln

530

535

540

Val Ala Pro Glu Thr Lys Gly Pro Arg Gln
Met Ala Val Phe Ser Leu

545

550

555

5

60

Pro Glu Glu Val Tyr Arg Lys Pro Ala Glu
Leu Asp Glu Asp Ser Glu

565

570

575

Ser Ser Lys Cys Cys Ser Ile Arg Tyr Cys
Phe Tyr Tyr Arg Lys Cys

580

585

590

Asp Met Ala Asp Asp Ala Ser Asp Gly Lys
Asp Glu Leu Ser Tyr Ser

595

600

605

Ala Pro Lys Pro Tyr Val Ser Gln Ile Ser
 Glu Tyr Lys Leu Glu Leu
 705 710 715 7
 20

Ala Leu Lys Phe Lys Glu Leu Arg Ala Ser
 Cys Arg Arg Val Ala Asn
 725 730 735

Val Asp Lys Ser Pro Thr His Met Leu Ala
 Ala Ile Thr Gly Ser Phe
 740 745 750

Gln Val Leu Ser Ser Leu Ile Glu Thr Phe
 Val Arg Leu Val Phe Ile
 755 760 765

Val Arg Ser Glu Ala Gln Arg Gln Glu Leu
 Leu Ala Lys Val Glu Glu
 770 775 780

Val Val Arg Asn Tyr Thr Phe Leu Leu Arg
 Ala Ala Glu Glu Ser Thr
 785 790 795 8
 00

Ala Arg Asn Leu Asn Gln Gln Gln Gln
 Gln Gln Gln Gln Gln Gln
 805 810 815

Gln Gln Gln Gln Gln Gln Gln Gln Gln
 Gln Gln Gln Gln Gln Gln
 820 825 830

Val Ala Ala Ala Ala Gly Ala Ala Thr Glu
 His Pro Pro Gly Ser Pro
 835 840 845

Thr Ser Ala Thr Val Met Ser Thr Phe Thr
 His Ser Leu Lys Thr Leu
 850 855 860

Ile Lys

<210> 12
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial
Sequence:Designed
polynucleotide based on the base se
quence of the
sequence No.1 in the present sequen
ce listing

<400> 12
gctggcttgg tagagattgg gc
22

<210> 13
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial
Sequence:Designed
polynucleotide based on the base se
quence of the
sequence No.3 in the present sequen
ce listing

<400> 13
gcacatttcc agcttctcag gc
22

<210> 14
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial
Sequence:Designed
polynucleotide based on the base se
quence of the
sequence No.3 in the present sequen
ce listing

<400> 14
gccctggatc tgtggaactt gc
22

sequence NO.5 in the present sequence listing

<400> 15
cttcatgcct accagtcggg cc
22

<210> 16
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial
Sequence:Designed
polypeptide based on the aminoacid
sequence of the
sequence NO.6 in the present sequence listing

<400> 16
ggcgaggttt tggtaatgga ggg
23

<210> 17
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial
Sequence:Designed
polynucleotide based on the base sequence of the
sequence NO.7 in the present sequence listing

<400> 17
ggccaaactc atgcagcaga aac
23

<210> 18
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial
Sequence:Designed
polynucleotide based on the base se

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial
Sequence:Designed
polynucleotide based on the base se
quence of the
sequence No.9 of the present sequen
ce listing

<400> 19
tccacagccc gtaaccttaa cc
22

<210> 20
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial
Sequence:Designed

28

【図面の簡単な説明】 polynucleotide based on the base sequence of the present sequence listing

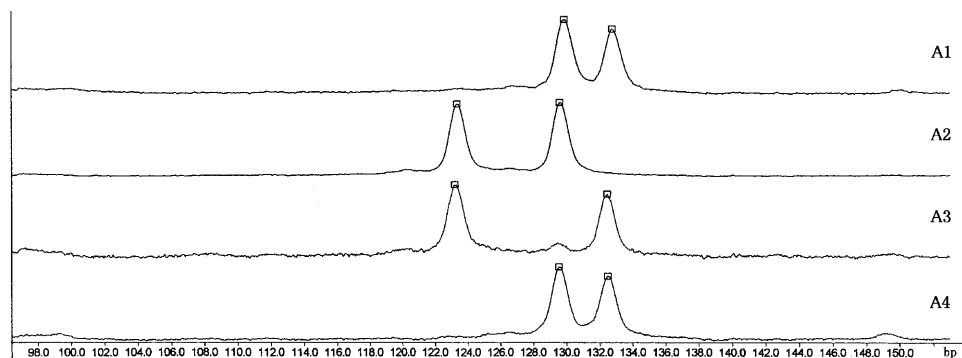
【図1】 HG006201 遺伝子多型の家系分析を、CAGリピートを含むDNA断片の解析により行った結果を示す図である。図中、A1は父親、A2は母親、A3およびA4は子供の解析結果を示す。

【図2】 HH013401 遺伝子多型の家系分析を、CAGリピートを含むDNA断片の解析により行った結果を示す図である。図中、A1は父親、A2は母親、A3およびA4は子供の解析結果を示す。

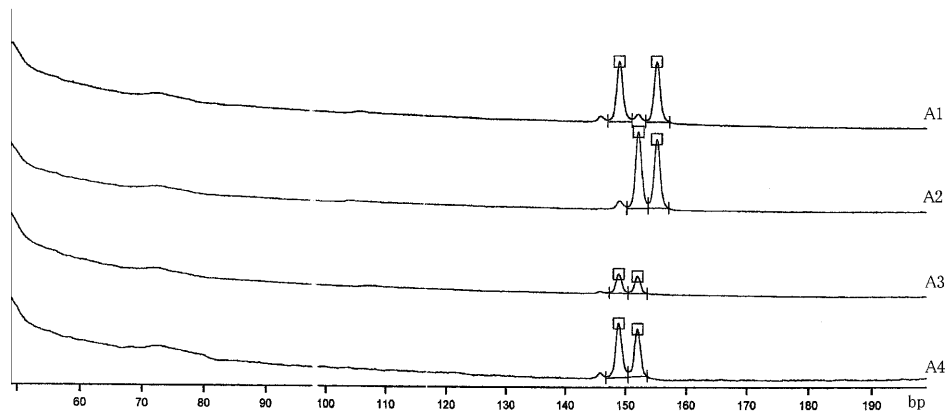
【図3】 FH16716 遺伝子多型の家系分析を、CAGリピートを含むDNA断片の解析により行った結果を示す図である。図中、A1は父親、A2は母親、A3およびA4は子供の解析結果を示す。

【図4】 FJ14474 遺伝子についての家系分析を、CAGリピートを含むDNA断片の解析により行った結果を示す図である。図中、A1は父親、A2は母親、A3およびA4は子供の解析結果を示す。

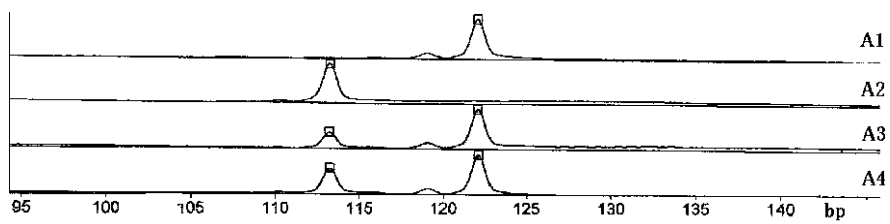
【図2】



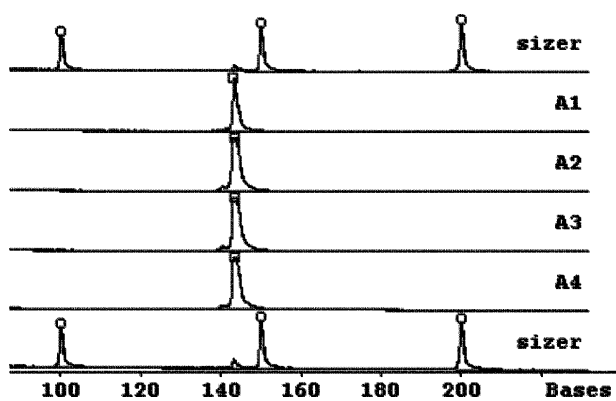
【図1】



【図3】



【図4】



フロントページの続き

(51)Int.Cl.⁷

A 6 1 K 35/76
38/00
39/395

識別記号

F I

A 6 1 K 39/395
45/00
48/00

テーマコード^{*} (参考)

D 4 C 0 8 4
N 4 C 0 8 5
4 C 0 8 7
4 H 0 4 5

45/00
48/00
A 6 1 P 25/28
C 0 7 K 14/47
16/18
C 1 2 N 1/15
1/19
1/21
5/10
C 1 2 P 21/02
C 1 2 Q 1/02
1/68
G 0 1 N 33/15
33/50
33/53
33/566

A 6 1 P 25/28
C 0 7 K 14/47
16/18
C 1 2 N 1/15
1/19
1/21
C 1 2 P 21/02 C
C 1 2 Q 1/02
1/68 A
G 0 1 N 33/15 Z
33/50 Z
33/53 D
M
33/566
C 1 2 N 15/00 Z N A A
5/00 A
A 6 1 K 37/02

- (72)発明者 長瀬 隆弘
千葉県木更津市矢那1532番3号 財団法人
かずさディー・エヌ・エー研究所内
- (72)発明者 大石 道夫
千葉県木更津市矢那1532番3号 財団法人
かずさディー・エヌ・エー研究所内
- (72)発明者 横田 博
東京都江戸川区北葛西1丁目16番13号 第
一製薬株式会社東京研究開発センター内
- (72)発明者 磯野 美智
東京都江戸川区北葛西1丁目16番13号 第
一製薬株式会社東京研究開発センター内

Fターム(参考) 2G045 AA40 DA12 DA13 DA36 FB02
FB03
4B024 AA01 AA11 BA44 BA80 CA04
GA11 HA12 HA15
4B063 QA01 QA12 QA19 QQ42 QQ79
QR55 QR59 QR80 QS24 QS25
QS34
4B064 AG01 AG31 CA19 CC24 DA01
DA13
4B065 AA93Y AB01 AC14 CA24
CA44 CA46
4C084 AA02 AA06 AA13 AA17 CA18
CA28 NA14 ZA022
4C085 AA13 AA14 CC05 CC22 DD33
DD86
4C087 BB33 BB63 BC11 BC34 BC65
BC83 NA14 ZA02
4H045 AA10 AA11 AA20 AA30 BA10
CA45 DA76 DA86 EA21 EA50
FA74

专利名称(译)	家庭经济标记		
公开(公告)号	JP2002360268A	公开(公告)日	2002-12-17
申请号	JP2001236788	申请日	2001-08-03
[标]申请(专利权)人(译)	上总DNA KENKYUSHO		
申请(专利权)人(译)	上总DNA研究所 第一制药有限公司		
[标]发明人	小原收 長瀬隆弘 大石道夫 横田博 磯野美智		
发明人	小原 收 長瀬 隆弘 大石 道夫 横田 博 磯野 美智		
IPC分类号	G01N33/50 A61K35/12 A61K35/64 A61K35/74 A61K35/76 A61K35/761 A61K36/06 A61K38/00 A61K39/395 A61K45/00 A61K48/00 A61P25/28 C07K14/47 C07K16/18 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N15/09 C12P21/02 C12Q1/02 C12Q1/68 G01N33/15 G01N33/53 G01N33/566 A61K35/72		
FI分类号	A61K35/12 A61K35/64 A61K35/72 A61K35/74.A A61K35/76 A61K39/395.D A61K39/395.N A61K45/00 A61K48/00 A61P25/28 C07K14/47 C07K16/18 C12N1/15 C12N1/19 C12N1/21 C12P21/02.C C12Q1/02 C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.D G01N33/53.M G01N33/566 C12N15/00.ZNA.A C12N5/00.A A61K37/02 A61K35/761 A61K36/06.Z A61K38/00 A61K38/17 C12N15/00.A C12N15/00.AZN.A C12N5/00.101 C12N5/10		
F-TERM分类号	2G045/AA40 2G045/DA12 2G045/DA13 2G045/DA36 2G045/FB02 2G045/FB03 4B024/AA01 4B024/AA11 4B024/BA44 4B024/BA80 4B024/CA04 4B024/GA11 4B024/HA12 4B024/HA15 4B063/QA01 4B063/QA12 4B063/QA19 4B063/QQ42 4B063/QQ79 4B063/QR55 4B063/QR59 4B063/QR80 4B063/QS24 4B063/QS25 4B063/QS34 4B064/AG01 4B064/AG31 4B064/CA19 4B064/CC24 4B064/DA01 4B064/DA13 4B065/AA93Y 4B065/AB01 4B065/AC14 4B065/CA24 4B065/CA44 4B065/CA46 4C084/AA02 4C084/AA06 4C084/AA13 4C084/AA17 4C084/CA18 4C084/CA28 4C084/NA14 4C084/ZA022 4C085/AA13 4C085/AA14 4C085/CC05 4C085/CC22 4C085/DD33 4C085/DD86 4C087/BB33 4C087/BB63 4C087/BC11 4C087/BC34 4C087/BC65 4C087/BC83 4C087/NA14 4C087/ZA02 4H045/AA10 4H045/AA11 4H045/AA20 4H045/AA30 4H045/BA10 4H045/CA45 4H045/DA76 4H045/DA86 4H045/EA21 4H045/EA50 4H045/FA74		
代理人(译)	庄司隆		
优先权	2000236839 2000-08-04 JP 2001108723 2001-04-06 JP		
外部链接	Espacenet		

摘要(译)

解决的问题：分离和鉴定新的与聚谷氨酰胺疾病有关的基因，阐明该基因编码的多肽，并提供检验方法，使用这些基因的聚谷氨酰胺疾病的预防和/或治疗剂。 那个 SOLUTION：一种新颖的基因，包含一个多核苷酸，该多核苷酸由序列表的SEQ ID NO：1，SEQ ID NO：3，SEQ ID NO：5，SEQ ID NO：7或SEQ ID NO：9所示的核苷酸序列组成，该序列表具有编码聚谷氨酰胺的CAG

重复序列，以及 为了找到由该基因编码的多肽，家庭诊断手段是利用源自CAG重复次数差异的每个基因的多态性，一种由该基因的部分序列组成的探针，可用于诊断手段， 以及针对谷氨酰胺疾病的预防

和/或治疗剂，其基于抑制基因的表达和由该基因编码的多肽的功能。

