

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4672866号
(P4672866)

(45) 発行日 平成23年4月20日(2011.4.20)

(24) 登録日 平成23年1月28日(2011.1.28)

(51) Int.Cl.

F I

C 1 2 N 15/09 (2006.01)

C 1 2 N 15/00 Z N A A

C O 7 K 14/705 (2006.01)

C O 7 K 14/705

C 1 2 Q 1/02 (2006.01)

C 1 2 Q 1/02

G O 1 N 33/15 (2006.01)

G O 1 N 33/15 Z

G O 1 N 33/50 (2006.01)

G O 1 N 33/50 Z

請求項の数 19 (全 281 頁) 最終頁に続く

(21) 出願番号 特願2000-576019 (P2000-576019)
 (86) (22) 出願日 平成11年10月12日(1999.10.12)
 (65) 公表番号 特表2003-531565 (P2003-531565A)
 (43) 公表日 平成15年10月28日(2003.10.28)
 (86) 国際出願番号 PCT/US1999/023938
 (87) 国際公開番号 W02000/022129
 (87) 国際公開日 平成12年4月20日(2000.4.20)
 審査請求日 平成18年7月21日(2006.7.21)
 (31) 優先権主張番号 09/170,496
 (32) 優先日 平成10年10月13日(1998.10.13)
 (33) 優先権主張国 米国(US)

微生物の受託番号 ATCC 203351

(73) 特許権者 500478097
 アリーナ ファーマシューティカルズ、
 インコーポレイテッド
 アメリカ合衆国カリフォルニア州9212
 1 サンディエゴ・ナンシーリッジドライブ
 6166
 (74) 代理人 100078282
 弁理士 山本 秀策
 (74) 代理人 100062409
 弁理士 安村 高明
 (74) 代理人 100113413
 弁理士 森下 夏樹

最終頁に続く

(54) 【発明の名称】 非内因性の構成的に活性化されたヒトGタンパク質共役受容体

(57) 【特許請求の範囲】

【請求項1】

内因性ヒトGタンパク質共役受容体(GPCR)の構成的に活性のある非内因性の変種を作製するための方法であって、該内因性GPCRは、膜貫通6領域及び細胞内ループ3領域を含んでなり、該方法は：

(a) 膜貫通6領域内のプロリン残基を含む内因性ヒトGPCRを選択すること；

(b) カルボキシ末端～アミノ末端の向きに、工程(a)のプロリン残基から16番目の内因性アミノ酸残基を同定すること；

(c) 工程(b)の同定されたアミノ酸残基を非内因性アミノ酸残基に改変して該内因性ヒトGPCRの非内因性の変種を作製すること；及び

(d) 該内因性GPCRによって誘導されるシグナルと比較した場合の該非内因性の変種に関して測定した細胞内シグナルにおける差を測定することにより、工程(c)の内因性ヒトGPCRの非内因性の変種が構成的に活性であるかどうかを決定すること、を含んでなる方法。

【請求項2】

非内因性の構成的に活性化されたヒトGタンパク質共役受容体に対する逆アゴニスト、アゴニスト及び部分アゴニストよりなる群から選択される化合物を直接同定する方法であって、該受容体が膜貫通6領域及び細胞内ループ3領域を含んでなり、該方法は、請求項1に記載の工程(a)～(d)を含み、さらに

(e) 工程(d)で同定された非内因性の構成的に活性のあるGPCRと候補化合物とを接触

させること；及び

(f) 該接触させた受容体での化合物効能の測定により、該化合物が該受容体の逆アゴニスト、アゴニストまたは部分アゴニストであるかどうかを決定すること、の工程を含んでなる方法。

【請求項 3】

膜貫通 6 領域中の該プロリン残基からカルボキシ末端～アミノ末端の向きに 2 残基であるアミノ酸残基がトリプトファンである請求項 1 の方法。

【請求項 4】

カルボキシ末端～アミノ末端の向きに、前記プロリン残基から 16 番目の内因性アミノ酸残基がリシン残基に改変されている請求項 1～3 のいずれか 1 項の方法。

10

【請求項 5】

カルボキシ末端～アミノ末端の向きに、前記プロリン残基から 16 番目の内因性アミノ酸残基がアラニン残基に改変されている請求項 1～3 のいずれか 1 項の方法。

【請求項 6】

カルボキシ末端～アミノ末端の向きに、前記プロリン残基から 16 番目の内因性アミノ酸残基がヒスチジン残基またはアルギニン残基に改変されている請求項 1～3 のいずれか 1 項の方法。

【請求項 7】

カルボキシ末端～アミノ末端の向きに、前記プロリン残基から 16 番目の内因性アミノ酸残基がアルギニン残基に改変されている請求項 1～3 のいずれか 1 項の方法。

20

【請求項 8】

カルボキシ末端～アミノ末端の向きに、前記プロリン残基から 16 番目の内因性アミノ酸残基がヒスチジン残基に改変されている請求項 1～3 のいずれか 1 項の方法。

【請求項 9】

内因性ヒト Gタンパク質共役受容体 (GPCR) の構成的に活性のある非内因性の変種を作製するための方法であって、該内因性 GPCR は、膜貫通 6 領域及び細胞内ループ 3 領域を含んでなり、該方法は：

(a) ポリヌクレオチドを提供することであって、該ポリヌクレオチドは、内因性ヒト GPCR をコードし、該内因性 GPCR は、膜貫通 6 領域および細胞内ループ 3 領域を含み、該膜貫通 6 領域はプロリン残基を含むこと；

30

(b) カルボキシ末端～アミノ末端の向きに、工程 (a) の GPCR の該プロリン残基から 16 番目の内因性アミノ酸残基に対応する、該ポリヌクレオチドのコドン进行を同定すること；

(c) 工程 (b) の同定されたコドン进行を非内因性アミノ酸残基をコードするように改変して非内因性ポリヌクレオチドを提供すること；

(d) 該非内因性ポリヌクレオチドを宿主細胞において発現し、それにより、該内因性ヒト GPCR の非内因性の変種を提供すること；および

(e) 該内因性 GPCR によって誘導されるシグナルと比較した場合の該非内因性の変種に関して測定した細胞内シグナルにおける差を測定することにより、工程 (d) の内因性ヒト GPCR の非内因性の変種が構成的に活性であるかどうかを決定すること、を含んでなる方法。

40

【請求項 10】

膜貫通 6 領域中の該プロリン残基からカルボキシ末端～アミノ末端の向きに 2 残基であるアミノ酸残基がトリプトファンである請求項 9 の方法。

【請求項 11】

工程 (b) の同定されたコドンがリシンをコードするコドンに改変されている請求項 9 または請求項 10 の方法。

【請求項 12】

工程 (b) の同定されたコドンがアラニンをコードするコドンに改変されている請求項 9 または請求項 10 の方法。

【請求項 13】

50

工程（b）の同定されたコドンがヒスチジンまたはアルギニンをコードするコドンに改変されている請求項 9 または 10 の方法。

【請求項 14】

工程（b）の同定されたコドンがアルギニンをコードするコドンに改変されている請求項 9 または請求項 10 の方法。

【請求項 15】

工程（b）の同定されたコドンがヒスチジンをコードするコドンに改変されている請求項 9 または請求項 10 の方法。

【請求項 16】

直接同定された化合物が逆アゴニストである請求項 2 の方法。

10

【請求項 17】

直接同定された化合物がアゴニストである請求項 2 の方法。

【請求項 18】

直接同定された化合物が部分アゴニストである請求項 2 の方法。

【請求項 19】

前記化合物を製薬学的組成物中に調合する工程（g）をさらに含んでなる請求項 2 の方法。

【発明の詳細な説明】

【0001】

1998年10月13日に出願された共通して所有する米国第09/170,496号、1997年4月14日に
出願された米国第08/839,449号（現在放棄された）、1998年4月14日に
出願された米国第09/060,188号；1998年6月26日に
出願された米国仮出願第60/090,783号；及び1998年8月7日に
出願された米国仮出願第60/095,677号の利益は本明細書により請求される。前述の出願の
各々は引用することによりそれらの全部が本明細書に組み込まれる。

20

【0002】

発明の分野

本特許書類において開示する本発明は膜貫通受容体、より詳細には、改変したヒトGタン
パク質共役受容体(GPCR)が構成的に活性化されるように改変されているGPCRに関する。最
も好ましくは、改変したヒトGPCRを治療化合物のスクリーニングに用いる。

【0003】

30

発明の背景

多数の受容体クラスがヒトにおいて存在するが、明らかに最も豊富で且つ治療的に関係の
あるのはGタンパク質共役受容体（GPCRまたはGPCRs）クラスにより代表される。ヒトゲノ
ム内にはおよそ100,000個の遺伝子があると概算され、これらのうち約2%、すなわち2,00
0個の遺伝子がGPCRをコードすると見積もられる。これらの中に、GPCRに結合する内因性
リガンドが同定されている約100個のGPCRがある。内因性GPCR及びその内因性リガンドの
発見の間に存在する著しい時間的ずれのために、残りの1,900個のGPCRは、これらの受容
体の内因性リガンドが同定されるずっと前に同定され、特性化されると推定することがで
きる。実際、ヒトゲノムプロジェクトが100,000個のヒト遺伝子をシークエンスしている
迅速さは、残りのヒトGPCRが次の数年以内に完全にシークエンスされることを示す。それ
にもかかわらず、そしてヒトゲノムをシークエンスする努力にもかかわらず、ヒト症状を
改善し、良くするために科学者等がどのようにこの情報を迅速に、効果的に且つ効率よく
利用することができるかに関してまだ非常に不明瞭である。本発明はこの重要な目的に適
合する。

40

【0004】

GPCRを包含する、内因性リガンドが同定されている受容体は「既知の」受容体と呼ばれ、
一方、内因性リガンドが同定されていない受容体は「オーファン」受容体と呼ばれる。こ
の区別は、特にGPCRの場合、単に語義ではない。GPCRは製薬学的製品の開発のための重要
な分野であり：100個の既知のGPCRのうち約20個から、全ての処方製薬の60%が開発され
ている。従って、オーファンGPCRは製薬産業にとって、19世紀末期に金がカリフォルニア

50

にとってそうであったこと-成長、拡張、拡大及び発展を導く機会である。しかしながら、新規な治療法の開発に関してオーファン受容体には重大な欠点が存在する。これは、製薬の発見及び開発への伝統的な方法が受容体及びその内因性リガンドの両方の入手を必要としているからである。従って、これまで、オーファンGPCRは製薬の発見のための興味をかき立てる未開発の供給源を当該技術分野に示している。

【0005】

可能性のある治療法の発見への伝統的な方法では、一般に受容体が最初に同定されると言える。薬剤発見努力を開始できる前に、典型的には、該受容体の内因性リガンドを同定し、単離し、作製するために複雑で、多くの時間を必要とし且つ費用のかかる手続きが行われ (put into place)、この手続きは受容体当たり約500万ドル (U.S.) の費用で受容体当たり3~10年を必要とする可能性がある。これらの時間及び財源は、薬剤発見への伝統的な方法が開始できる前に使い果たされるはずである。これは、伝統的な薬剤発見技術が、内在性リガンドが受容体に結合するのを妨げる化合物 (「アンタゴニスト」) または受容体に結合するリガンドの作用を高めるかもしくは模倣する化合物 (「アゴニスト」) のいずれかを発見しようと努力して推定の治療因子が受容体に対して「スクリーニングされる」、いわゆる「競合結合アッセイ」に依存するからである。全体的な目的は、リガンドが受容体に結合した場合に細胞活性化を妨げる化合物 (アンタゴニスト) またはリガンドが受容体と正しく結合している場合にそうでなければ生じる細胞活性を高めるかもしくは上げる化合物 (アゴニスト) を同定することである。オーファンGPCRの内因性リガンドは自明のこととして同定されていないので、伝統的な薬剤発見技術を用いてこれらの受容体に対する新規で且つ独特な治療法を発見できることはあり得ない。本発明は、以下にさらに詳細に記述するように、そのような伝統的な薬剤発見技術によりもたらされるこれら及び他の深刻な制約を克服する。

【0006】

GPCRは共通の構造モチーフを共有する。これらの受容体は全て、7個の α ヘリックスを形成する22~24個の間の疎水性アミノ酸の7配列を有し、それらの各々は膜にまたがる (各スパンは番号により同定される、すなわち、膜貫通-1 (TM-1)、膜貫通-2 (TM-2) 等)。これらの膜貫通ヘリックスは、細胞膜の外側、すなわち「細胞外」側で膜貫通-2と膜貫通-3、膜貫通-4と膜貫通-5、そして膜貫通-6と膜貫通-7の間でアミノ酸の鎖により連結されている (これらはそれぞれ「細胞外」領域1、2及び3 (EC-1、EC-2及びEC-3) と呼ばれる)。膜貫通ヘリックスはまた、細胞膜の内側、すなわち「細胞内」側で膜貫通-1と膜貫通-2、膜貫通-3と膜貫通-4、そして膜貫通-5と膜貫通-6の間でアミノ酸の鎖により連結されている (これらはそれぞれ「細胞内」領域1、2及び3 (IC-1、IC-2及びIC-3) と呼ばれる)。受容体の「カルボキシ」 (「C」) 末端は細胞内の細胞内空間にあり、そして受容体の「アミノ」 (「N」) 末端は細胞の外側の細胞外空間にある。Gタンパク質共役受容体の一般構造を図1に示す。

【0007】

一般に、内因性リガンドが受容体と結合すると (多くの場合、受容体の「活性化」と呼ばれる)、細胞内領域の構造の変化があり、これは細胞内領域と細胞内「Gタンパク質」間の共役を与える。他のGタンパク質も存在するが、現在、Gq、Gs、Gi及びGoが同定されているGタンパク質である。内因性リガンドにより活性化されたGタンパク質と共役するGPCRは、シグナリングカスケードプロセスを開始する (「シグナル伝達」と呼ばれる)。通常の条件下で、シグナル伝達は最終的に細胞活性化または細胞阻害をもたらす。受容体のIC-3ループ並びにカルボキシ末端がGタンパク質と相互作用すると考えられる。本発明の主要な焦点は、GPCRの膜貫通-6 (TM6) 領域及び細胞内-3 (IC3) 領域に向けられる。

【0008】

生理的条件下で、GPCRは2つの異なる構造: 「不活性」状態及び「活性」状態間で平衡を保って細胞膜中に存在する。図2に図式的に示すように、不活性状態の受容体は、細胞内シグナリング伝達経路に連結して生物学的応答を引き起こすことができない。受容体構造を活性状態に変えることにより (Gタンパク質を介して) 伝達経路に連結され、生物学的

応答を引き起こす。

【0009】

受容体は、内因性リガンドまたは薬剤のような化合物により活性状態に安定させることができる。受容体のアミノ酸配列に対する改変を包含するがこれだけに限定されるものではない最近の発見は、活性状態構造の受容体を増進し安定させるために内因性リガンドまたは薬剤以外の手段を提供する。これらの手段は、受容体に結合する内因性リガンドの作用をまねることにより活性状態の受容体を効果的に安定させる。そのようなリガンドに依存しない手段による安定化は、「構成的受容体活性化」と呼ばれる。

【0010】

上述のように、スクリーニング目的のためのオーファン受容体の使用はあり得なかった。これは、化合物のスクリーニングに関する伝統的な「原則」が、受容体のリガンドが既知であることを要求するからである。従って、自明のこととして、この方法はオーファン受容体に関して適用性がない。従って、治療法の発見へのこの独断的な方法に固執することにより、当該技術分野は、本質的に、受容体の内因性リガンドが発見されるまでオーファン受容体の使用をやめるように教示し且つ教示されてきた。概算して2,000個のGタンパク質共役受容体があり、これらの大多数がオーファン受容体であることを考えれば、そのような原則は、治療法の発見への独創的で独特な異なる方法を訂正する。

【0011】

様々なGPCRの核酸及び／またはアミノ酸配列に関する情報を以下の表Aに要約する。本明細書において開示する本発明の重要な焦点はオーファンGPCRに向けられるので、以下に記載する参考文献の多くはオーファンGPCRに関連する。しかしながら、この表は、本明細書において開示する本発明がオーファンGPCRまたは以下に記載する特定のGPCRにのみ適用できることを意味するものではなく、またこの表は、法的にせよそうでないにせよ、そのように解釈されるべきではない。さらに、単離されているある種の受容体は公開自体の主題ではなく；例えば、GPCRを載せている「ワールド-ワイドウェブ (world-wide web)」上のGタンパク質共役受容体データベース（指定された発明者も譲受人もこのサイトといかなる関係もない）を参照する。他のGPCRは本譲受人により所有される特許出願の主題であり、これらは以下に記載されていない（GPR3、GPR6及びGPR12を包含する；米国仮第60/094879号を参照）。

【0012】

【表1】

10

20

30

表A

受容体名	公開参考文献
GPR1	23 Genomics 609 (1994)
GPR4	14 DNA and Cell Biology 25 (1995)
GPR5	14 DNA and Cell Biology 25 (1995)
GPR7	28 Genomics 84 (1995)
GPR8	28 Genomics 84 (1995)
GPR9	184 J. Exp. Med. 963 (1996)
GPR10	29 Genomics 335 (1995)
GPR15	32 Genomics 462 (1996)
GPR17	70 J Neurochem. 1357 (1998)
GPR18	42 Genomics 462 (1997)
GPR20	187 Gene 75 (1997)
GPR21	187 Gene 75 (1997)
GPR22	187 Gene 75 (1997)
GPR24	398 FEBS Lett. 253 (1996)
GPR30	45 Genomics 607 (1997)
GPR31	42 Genomics 519 (1997)
GPR32	50 Genomics 281 (1997)
GPR40	239 Biochem. Biophys. Res. Commun. 543 (1997)
GPR41	239 Biochem. Biophys. Res. Commun. 543 (1997)
GPR43	239 Biochem. Biophys. Res. Commun. 543 (1997)
APJ	136 Gene 355 (1993)
BLR1	22 Eur. J. Immunol. 2759 (1992)
CEPR	231 Biochem. Biophys. Res. Commun. 651 (1997)
EBI1	23 Genomics 643 (1994)
EBI2	67 J. Virol. 2209 (1993)
ETBR-LP2	424 FEBS Lett. 193 (1998)
GPCR-CNS	54 Brain Res. Mol. Brain Res. 152 (1998); 45 Genomics 68 (1997)
GPR-NGA	394 FEBS Lett. 325 (1996)
H9	386 FEBS Lett 219 (1996)
HBA954	1261 Biochim. Biophys. Acta 121 (1995)
HG38	247 Biochem. Biophys. Res. Commun. 266 (1998)
HM74	5 Int. Immunol. 1239 (1993)
OGR1	35 Genomics 397 (1996)
V28	163 Gene 295 (1995)

【0013】

以下にさらに詳細に記述し、開示するように、ヒトGPCRの内因性配列を改変するための突然変異カセットの利用により、ヒトGPCRの構成的に活性化されたバージョンがもたらされる。ヒトGPCRのこれらの非内因性の構成的に活性化されたバージョンは、とりわけ、例えば治療関連性の化合物を直接同定するための候補化合物のスクリーニングに利用することができる。

【0014】

発明の要約

本明細書において開示するのは、GPCRの膜貫通-6 (TM6) 及び細胞内ループ-3 (IC3) 領域を横切る (a) 最も好ましいアミノ酸配列領域として (C末端～N末端の向き)、そして/または (b) 最も好ましい核酸配列領域として (3'～5'の向き) :

(a) P¹AA₁₅X

ここで:

10

20

30

40

50

(1) P^1 はGPCRのTM6領域内に位置するアミノ酸残基であり、ここで、 P^1 は(i)内因性GPCRのプロリン残基及び(ii)プロリン以外の非内因性アミノ酸残基よりなる群から選択される；

(2) AA_{15} は(a)内因性GPCRのアミノ酸、(b)非内因性アミノ酸残基並びに(c)内因性GPCRのアミノ酸及び非内因性アミノ酸の組み合わせよりなる群から選択される15個のアミノ酸であり、ただし、GPCRのTM6領域内に位置する15個の内因性アミノ酸残基のいずれもプロリンでない；そして

(3) Xは該GPCRのIC3領域内に位置する非内因性アミノ酸残基であり、好ましくはリシン、ヒスチジン及びアルギニンよりなる群から選択され、最も好ましくはリシンであり、ただし、位置Xの内因性アミノ酸がリシンである場合には、Xはリシン以外のアミノ酸、好ましくはアラニンである

及び/または

(b) $P^{codon} (AA-codon)_{15} X_{codon}$

ここで：

(1) P^{codon} はGPCRのTM6領域内の核酸配列であり、ここで、 P^{codon} は(i)内因性GPCRのプロリン残基及び(ii)プロリン以外の非内因性アミノ酸残基よりなる群から選択されるアミノ酸をコードする；

(2) $(AA-codon)_{15}$ は(a)内因性GPCRのアミノ酸、(b)非内因性アミノ酸残基並びに(c)内因性GPCRのアミノ酸及び非内因性アミノ酸の組み合わせよりなる群から選択される15個のアミノ酸をコードする15個のコドンであり、ただし、GPCRのTM6領域内の15個の内因性コドンのいずれもプロリンアミノ酸残基をコードしない；そして

(3) X_{codon} は該GPCRのIC3領域内に位置する領域残基をコードする核酸であり、ここで、 X_{codon} は非内因性アミノ酸、好ましくは、リシン、ヒスチジン及びアルギニンよりなる群から選択される非内因性アミノ酸、最も好ましくはリシンでをコードし、ただし、位置 X_{codon} の内因性コード領域がアミノ酸リシンをコードする場合には、 X_{codon} はリシン以外のアミノ酸、好ましくはアラニンをコードする

を含んでなる非内因性のヒトGタンパク質共役受容体である。

これらの配列カセットに関連する内因性及び非内因性という用語は、内因性GPCRに関してである。例えば、いったん内因性プロリン残基が特定のGPCRのTM6領域内に位置決定され、それから16番目のアミノ酸が受容体を構成的に活性化する突然変異のために同定されると、プロリン残基が突然変異されないことが最も好ましいが、この内因性のプロリン残基を突然変異させることもまた可能である（すなわち、いったんマーカが位置決定され、そして突然変異させる16番目のアミノ酸が同定されると、マーカ自体を突然変異させることができる）。同様に、 AA_{15} はそれらの内因性形態で保たれることが最も好ましいが、これらのアミノ酸もまた突然変異させることができる。ヒトGPCRの非内因性バージョンにおいて突然変異させなければならない唯一のアミノ酸はXであり、すなわち、本明細書においてさらに開示するように、 P^1 から16残基である内因性アミノ酸はその内因性形態で保つことができず、突然変異させなければならない。再び言明すると、ヒトGPCRの非内因性バージョンにおいて、 P^1 及び AA_{15} はそれらの内因性形態のままである（すなわち、それらの野性型形態と同一である）ことが好ましいが、いったんXが同定され、突然変異されると、 P^1 及び AA_{15} のいずれか及び/または全てを突然変異させることができる。これは核酸配列にも同様に当てはまる。位置Xの内因性アミノ酸がリシンである場合、そのようなGPCRの非内因性バージョンにおいて、Xはリシン以外のアミノ酸、好ましくはアラニンである。

【0015】

従って、仮定の例として、内因性GPCRが上記の位置で以下の内因性アミノ酸配列：P-AACCTTGGRRRDDDE-Qを有する場合、以下の代表的な仮定のカセットはいずれも本開示の範囲内に入る（非内因性アミノ酸をボールド体で記述する）：

P-AACCTTGGRRRDDDE-K

P-AACCTTHIGRRDDDE-K

P-ADEETTGRRRDDDE-A
P-LLKFMSTWZLVAAPQ-K
A-LLKFMSTWZLVAAPQ-K

AA₁₅内にアミノ酸残基を付加することもまた可能であるが、そのような方法は特に提示されない。実際、最も好ましい態様として、ヒトGPCRの内因性バージョンと比較した場合にそのGPCRの非内因性バージョンにおいて異なる唯一のアミノ酸は位置Xのアミノ酸であり；このアミノ酸自体の突然変異が受容体の構成的活性化を導く。

【0016】

従って、特に好ましい態様として、P¹及びP^{codon}はそれぞれ内因性プロリン及びプロリンをコードする内因性核酸コード領域であり；そしてX及びX_{codon}はそれぞれ非内因性のリシンまたはアラニン及びリシンまたはアラニンをコードする非内因性核酸コード領域であり、リシンが最も好ましい。これらの突然変異を含むヒトGPCRの非内因性バージョンは哺乳類細胞中に取り込まれ、候補化合物のスクリーニングに利用されることが最も好ましいので、突然変異を含む非内因性ヒトGPCRはそれ自体精製され、単離される必要はない（すなわち、これらは哺乳類細胞の細胞膜内に含まれる）が、そのような精製され、単離された非内因性ヒトGPCRは本開示の十分に範囲内である。非内因性ヒトGPCRを含む遺伝子ターゲティングした及びトランスジェニックの非ヒト哺乳類（好ましくはラット及びマウス）もまた本発明の範囲内であり；特に、遺伝子ターゲティングした哺乳類は、これらの動物が非ヒト哺乳類の内因性GPCRをコードする領域の代わりにヒトGPCRの非内因性バージョンを含む点で最も好ましい（非ヒト哺乳類のタンパク質コード領域をヒトコード領域で置き換えるためにそのような非ヒト哺乳類を作製する技術は周知である；例えば米国特許第5,777,194号を参照）。

【0017】

内因性ヒトGPCRに対するこれらの変異はGPCRを構成的に活性化し、その結果、本明細書においてさらに開示するように、ヒトGPCRの非内因性の構成的に活性化されたバージョンを、とりわけ、内因性リガンドの必要なしに候補化合物の直接スクリーニングに利用できることが発見された。従って、これらの材料を用いる方法及びこれらの方法により同定される生成物もまた以下の開示の範囲内である。

【0018】

詳細な記述

受容体に関して出ている科学文献は、受容体に対して様々な作用を有するリガンドを表すために多数の用語を採用している。明瞭化及び一貫性のために、以下の定義を本特許書類の全体にわたって用いる。これらの定義がこれらの用語の他の定義と矛盾するのであれば、以下の定義が支配するものとする：

アゴニストは、受容体に結合すると細胞内応答を活性化するかまたは膜へのGTP結合を高める化合物を意味するものとする。

【0019】

本明細書において用いるアミノ酸省略形を以下に示す。

【0020】

【表2】

アラニン	ALA	A
アルギニン	ARG	R
アスパラギン	ASN	N
アスパラギン酸	ASP	D
システイン	CYS	C
グルタミン酸	GLU	E
グルタミン	GLN	Q
グリシン	GLY	G
ヒスチジン	HIS	H
イソロイシン	ILE	I
ロイシン	LEU	L
リシン	LYS	K
メチオニン	MET	M
フェニルアラニン	PHE	F
プロリン	PRO	P
セリン	SER	S
トレオニン	THR	T
トリプトファン	TRP	W
チロシン	TYR	Y
バリン	VAL	V

10

20

【0021】

部分アゴニストは、受容体に結合すると細胞内応答をアゴニストより少ない度合い/程度に活性化するかまたは膜へのGTP結合をアゴニストより少ない度合い/程度に高める化合物を意味するものとする。

30

【0022】

アゴニストは、アゴニストと同じ部位で受容体に競合的に結合するが、受容体の活性形態により開始される細胞内応答を活性化せず、それによりアゴニストまたは部分アゴニストによる細胞内応答を阻害することができる化合物を意味するものとする。アンタゴニストは、アゴニストまたは部分アゴニストの非存在下での基準細胞内応答を減少させない。

【0023】

候補化合物は、スクリーニング技術が容易にできる分子を意味するものとする（例えば、限定ではなく、化学化合物）。好ましくは、「候補化合物」という語句には、間接的同定方法により以前に決定されたような、受容体に対する逆アゴニスト、アゴニストまたはアンタゴニストよりなる群から選択される化合物であることが公的に既知であった化合物（「間接的に同定された化合物」）は包含されず；より好ましくは、少なくとも1つの哺乳類において治療効能を有することが以前に決定されている間接的に同定された化合物は包含されず；そして最も好ましくは、ヒトにおいて治療有用性を有することが以前に決定されている間接的に同定された化合物は包含されない。

40

【0024】

コドンは、一般にリン酸基に結合したヌクレオシド（アデノシン(A)、グアノシン(G)、シチジン(C)、ウリジン(U)及びチミジン(T))を含んでなり、翻訳された場合にアミノ酸をコードする3個のヌクレオチド（またはヌクレオチドの同等物）のグループを意味するものとする。

【0025】

50

化合物効能は、受容体結合親和性とは対照的に、受容体機能性を阻害するかまたは刺激する化合物の能力の測定を意味するものとする。化合物効能を検出する好ましい手段は、本特許書類の実施例の節においてさらに開示するように、例えば $[^{35}\text{S}]$ GTP γ S結合の測定による。

【0026】

構成的に活性化された受容体は、構成的受容体活性化を受ける受容体を意味するものとする。本明細書において開示する本発明では、非内因性の構成的に活性化されたヒトGタンパク質共役受容体は、以下にさらに詳細に記述するように、アミノ酸カセットP¹AA₁₅Xを含むように突然変異されているものである。

【0027】

構成的受容体活性化は、受容体の内因性リガンドまたはその化学的同等物と受容体との結合以外の手段による活性状態の受容体の安定化を意味するものとする。好ましくは、本明細書において開示する本発明の構成的受容体活性化を受けたGタンパク質共役受容体は、そのGPCRの内因性形態と比較した場合に構成的活性化に関して測定したシグナルに対する応答の少なくとも10%の差（場合によって、増加または減少）、より好ましくは、そのような比較応答の約25%の差、最も好ましくはそのような比較応答の約50%の差を示す。候補化合物を直接同定する目的のために用いる場合、選択した候補化合物間を区別するために十分な差が内因性シグナルと非内因性シグナル間であるようにシグナルの差が少なくとも約50%であることが最も好ましい。大部分の場合において、「差」はシグナルの増加であるが；しかしながら、Gs共役GPCRに関しては、以下にさらに詳細に記述するように、測定される「差」は好ましくは減少である。

【0028】

接触または接触するは、インビトロ系にせよインビボ系にせよ、少なくとも2つの成分を一緒にすることを意味するものとする。

【0029】

「候補化合物」という語句に関連して、直接同定するまたは直接同定されたは、構成的に活性化されたGタンパク質共役受容体に対する候補化合物のスクリーニング及びそのような化合物の化合物効能を評価することを意味するものとする。この語句は、いかなる状況下でも、「間接的に同定する」または「間接的に同定された」という語句により包含されるかまたはそれを包含すると解釈されるかまたは理解されるべきではない。

【0030】

内因性のは、種のゲノムにより生来生産される物質を意味するものとする。例えば、限定ではなく、GPCRに関して内因性のは、ヒト、昆虫、植物、細菌またはウイルスにより生来生産されるものを意味するものとする。それに反して、これに関連して非内因性のもという用語は、種のゲノムにより生来生産されないものを意味するものとする。例えば、限定ではなく、最も好ましくは、内因性形態では構成的に活性でないが、本明細書において開示するカセットを用いることにより突然変異させると、その後、構成的に活性になる受容体が、本明細書において「非内因性の構成的に活性化された受容体」と呼ばれる。両方の用語を「インビボ」及び「インビトロ」系の両方を表すために利用することができる。例えば、限定ではなく、スクリーニング方法において、内因性または非内因性受容体は、受容体が哺乳類細胞の細胞表面上に発現されるインビトロスクリーニング系に関してであることができる。さらなる例として、限定ではなく、哺乳類のゲノムが非内因性の構成的に活性化された受容体を含むように操作されている場合、インビボ系の手段による候補化合物のスクリーニングが実行可能である。

【0031】

宿主細胞は、その中にプラスミド及び／またはベクターを含むことができる細胞を意味するものとする。原核生物宿主細胞の場合、プラスミドは典型的に宿主細胞が複製する場合に自律的分子として複製され（一般に、プラスミドは真核生物宿主細胞中への導入のためにその後単離される）；真核生物宿主細胞の場合、プラスミドは、真核生物宿主細胞が複製する場合にプラスミドが複製するように宿主細胞の細胞DNA中に組込まれる。好ましく

は、本明細書において開示する本発明の目的のために、宿主細胞は真核生物の、より好ましくは哺乳類のであり、最も好ましくは293、293T及びCOS-7細胞よりなる群から選択される。

【0032】

間接的に同定するまたは間接的に同定されたは、内因性受容体に特異的な内因性リガンドの同定、リガンド-受容体相互作用を妨げるもの及び／またはそれと競合するものを決定するための受容体に対する候補化合物のスクリーニング並びに活性化された受容体と関連する少なくとも1つの第二メッセンジャー経路に影響を及ぼすことに関して化合物の効能を評価することを含む薬剤発見プロセスへの伝統的な方法を意味する。

【0033】

「応答」という用語に関連して阻害するまたは阻害しているは、化合物の非存在下とは対照的に化合物の存在下で応答が減少されるかまたは妨げられることを意味するものとする。

【0034】

逆アゴニストは、受容体の内因性形態または受容体の構成的に活性化された形態のいずれかに結合し、受容体の活性形態により開始される基準細胞内応答をアゴニストもしくは部分アゴニストの非存在下で認められる活性の通常基礎レベル未満に阻害するかまたは膜へのGTP結合を減少させる化合物を意味するものとする。好ましくは、基準細胞内応答は、逆アゴニストの非存在下での基準応答と比較した場合に、逆アゴニストの存在下で少なくとも30%、より好ましくは少なくとも50%、最も好ましくは少なくとも75%阻害される。

【0035】

既知の受容体は、その受容体に特異的な内因性リガンドが同定されている内因性受容体を意味するものとする。

【0036】

リガンドは、内因性の天然に存在する受容体に特異的な内因性の天然に存在する分子を意味するものとする。

【0037】

内因性受容体の核酸及び／またはアミノ酸配列に関する突然変異体または突然変異は、内因性の構成的に活性化されていない受容体の突然変異形態が受容体の構成的活性化を示すような、そのような内因性配列に対する特定の1つの変異または複数の変異を意味するものとする。特定の配列の同等物に関して、(a) 受容体の次に突然変異させた形態の構成的活性化のレベルが受容体の第一の突然変異により示されるものと実質的に同じである場合；そして(b) 受容体の次に突然変異させた形態と受容体の第一の突然変異の間の配列(アミノ酸及び／または核酸) 相同性%が少なくとも約80%、より好ましくは少なくとも約90%、最も好ましくは少なくとも95%である場合、ヒト受容体の次に突然変異させた形態はヒト受容体の第一の突然変異と同等であると考えられる。理想的には、そして構成的活性化を行うために本明細書において開示する最も好ましいカセットはGPCRの内因性及び非内因性形態間で単一のアミノ酸及び／またはコドン変異(すなわち、XまたはX_{codon})を含むということのために、配列相同性%は少なくとも98%であるはずである。

【0038】

オーファン受容体は、その受容体に特異的な内因性リガンドが同定されていないかまたは既知でない内因性受容体を意味するものとする。

【0039】

製薬学的組成物は、少なくとも1つの有効成分を含んでなる組成物を意味するものとし、それにより組成物は哺乳類(例えば、限定ではなくヒト)における特定の有効な結果に関して容易に調べることができる。当業者は、当業者の必要性に基づいて有効成分が所望する有効な結果を有するかどうかを決定するために適当な技術を理解し、認識する。

【0040】

プラスミドは、ベクターとcDNAの組み合わせを意味するものとする。一般に、プラスミドは、複製及び／またはタンパク質としてのcDNAの発現の目的のために宿主細胞中に導入さ

10

20

30

40

50

れる。

【0041】

「応答」という用語に関連して、刺激するまたは刺激しているは、化合物の非存在下とは対照的に化合物の存在下で応答が増加されることを意味するものとする。

【0042】

特定の核酸配列または特定のアミノ酸配列のいずれかに関して、横切るまたは横切っているは、該配列が少なくとも2つの異なる特定の領域内に位置することを意味するものとする。例えば、10個の成分のうち3個がGPCRのTM6領域中であり、そして残りの7個の成分がGPCRのIC3領域中である、10アミノ酸成分の長さであるアミノ酸配列において、この10アミノ酸成分はGPCRのTM6及びIC3領域を横切ると記述することができる。

10

【0043】

cDNAに関連するベクターは、少なくとも1つのcDNAを含むことができ且つ宿主細胞中に取り込むことができる環状DNAを意味するものとする。

【0044】

以下の節の順序は表象的効率のために記述され、以下の開示または請求に対する限定として意図されず、また解釈されるべきではない。

A. 概論

受容体の伝統的研究は、いつも、受容体に作用することができるアンタゴニスト及び他の分子を発見により次に見出すことができるより前にまず内因性リガンドが同定されなければならないという（歴史的に基づく）前提から生じている。アンタゴニストが最初に既知である可能性がある場合でさえ、探索はすぐに内因性リガンドを探すことに及ぶ。この考え方は、構成的に活性化された受容体の発見後でさえ受容体研究において存続している。これまで認識されていないことは、受容体のアゴニスト、部分アゴニスト及び逆アゴニストを見出すために最も有用であるのは受容体の活性状態であるということである。過度に活性である受容体または不十分に活性である受容体に起因する疾病に対して、治療薬において所望されるものは、それぞれ、受容体の活性状態を減らすかまたは受容体の活性を高めるように作用する化合物であり、必ずしも内因性リガンドに対するアンタゴニストである薬剤ではない。これは、活性のある受容体状態の活性を減らすまたは高める化合物は内因性リガンドと同じ部位で結合する必要がないからである。従って、本発明の方法により教示されるように、治療化合物のあらゆる探索は、リガンドに依存しない活性状態に対して化合物をスクリーニングすることにより出発するはずである。

20

30

【0045】

非内因性の構成的に活性化されたGPCRに対して候補化合物をスクリーニングすることにより、受容体の内因性リガンドのいかなる事前の知識または使用も必要とせず、これらの細胞表面受容体で作用する候補化合物を直接同定することができる。そのようなGPCRの内因性バージョンが発現される及び／または過剰に発現される体内の領域を決定することにより、これらの受容体の発現及び／または過剰発現と関係する関連疾病／疾患状態を決定することが可能であり；そのような方法を本特許書類において開示する。

B. 疾病／疾患の同定及び／または選択

最も好ましくは、本発明の材料を用いて非内因性の構成的に活性化されたGPCRに対する逆アゴニストを同定することができる。そのような逆アゴニストは、これらの受容体に関連する疾病を処置するための薬剤発見プログラムにおける先導化合物（lead compound）として理想的な候補である。これらの受容体に対する逆アゴニスト、部分アゴニストまたはアゴニストを直接同定することができ、それにより製薬学的組成物を開発することができるので、これらの受容体と関連する疾病及び疾患の検索が可能である。例えば、これらの受容体の存在に関して罹病及び正常組織サンプルの両方を詳しく調べることは、今や学術的試行またはオープン受容体の場合には内因性リガンドを同定する道筋に沿って実施されるかもしれないもの以上になる。組織精査は、広範囲の健康な組織及び罹患組織にわたって実施することができる。そのような組織精査は、疾病及び／または疾患と特定の受容体とを結び付けることにおける好ましい第一段階を与える。

40

50

【0046】

好ましくは、組織サンプルにおけるGPCRの発現の放射性標識したcDNAまたはRT-PCR同定のいずれかのプローブを作製するために内因性GPCRのDNA配列を用いる。好ましくは、罹病組織における受容体の存在または正常組織と比較して罹病組織における増大したもしくは減少した濃度での受容体の存在を、その疾病との相関関係を同定するために利用することができる。この技術により受容体を器官の全域に同等に適切に位置決定することができる。受容体が位置決定される特定の組織の既知の機能に基づき、受容体の一般に考えられる機能的役割を推定することができる。

C. 「ヒトGPCRプロリンマーカー」アルゴリズム及び非内因性の構成的に活性のあるヒトGPCRの作製

10

バイオテクノロジー技術に立ち向かう多くの難題の中に、ある種から遺伝情報を収集し、そしてその情報を別の種に相関させることに関して予測できないことがあり、この問題は、この技術のいかなる場所にも核酸及びタンパク質をコードする遺伝子配列におけるほど厄介な悪化を示さない。従って、一貫性のため、そしてこの技術の非常に予測できない性質のために、以下の発明は、哺乳類に関してはヒトGPCRに限定され、他の哺乳類種への本発明の適用性は、潜在的な可能性はあるが、単なる機械的適用を超えると考えられる。

【0047】

一般に、ある関連したタンパク質配列から別のものにまたはある種から別のものに共通の「規則」を適用しようとする場合、当該技術は典型的に配列整列に頼っており、すなわち、配列を線状にし、次に2つまたはそれ以上の配列間の共通の性質の領域を見い出そうと試みる。有用ではあるが、この方法はいつも意味のある情報をもたらすとは限らない。GPCRの場合、一般的な構造モチーフは全てのGPCRに対して同一であるが、TM、EC及びICの長さの変動は、あるGPCRから別のものにそのような整列方法をせいぜい困難にする。従って、例えば構成的活性化への一貫した方法があるGPCRから別のものに適用することが望ましい可能性があるが、あるGPCRから次のものへの配列の長さ、正確さ等の大きな違いのために、一般的に適用でき容易にうまくいく突然変異整列方法は本質的にあり得ない。類似として、そのような方法は、縮尺または距離マーカーが地図のどこにもなく、旅行者に地点Bへの多数の異なる地図を与えることにより地点Aで旅行者に旅行を開始させ、次にそれらの地図を用いることによってのみ目的地Bへの最短で且つ最も効率のよい経路を見出すように旅行者に求めることと同じである。そのような状況において、この作業は、(a) 各地図上の共通の「場所マーカー」及び(b) その場所マーカーから目的地Bへの距離を測定する能力を有することにより容易に簡単にすることができ、従って、これにより旅行者は出発地点Aから目的地Bへの最も効率のよいものを選択することができる。

20

30

【0048】

本質的に、本発明の特徴は、ヒトGPCRの構成的に活性のある形態を容易に作製することができるヒトGPCR内のそのような位置を提供することである。

【0049】

当業者が認識するように、細胞の膜貫通領域は非常に疎水性であり；従って、標準的な疎水性プロット技術を用いて、当業者はGPCRのTM領域、特にTM6を容易に決定することができる（この同じ方法は、GPCRのEC及びIC領域を決定することにも適用できる）。ヒトGPCRのTM6領域内で、（一般にTM6の真ん中近くの）共通のプロリン残基が構成的活性化「マーカー」の役割を果たすことが発見された。このプロリンマーカーから15アミノ酸を数えることにより、（IC3ループ中に位置する）16番目のアミノ酸が、その内因性形態から非内因性形態に突然変異させると、受容体の構成的活性化を導く。便宜上、これを「ヒトGPCRプロリンマーカー」アルゴリズムと称する。この位置の非内因性アミノ酸はアミノ酸のいずれであってもよいが、最も好ましくは、非内因性アミノ酸はリシンである。いかなる理論によっても拘束されることを望まないが、我々は、この位置自体が独特であり、そしてこの位置での突然変異は構成的活性化を与えるように受容体に影響を与えると考える。

40

【0050】

例えば、16番目の位置の内因性アミノ酸がすでにリシンである場合（GPR4及びGPR32がそ

50

うであるが)、Xが非内因性アミノ酸であるためには、それはリシン以外でなければならず;従って、内因性GPCRが16番目の位置で内因性のリシン残基を有する状況では、そのGPCRの非内因性バージョンは、好ましくは、この位置でリシン以外のアミノ酸、好ましくはアラニン、ヒスチジン及びアルギニンを含むことを特筆する。さらに重要なことに (of further note)、GPR4はGsに共役し、その内因性形態で活性であるようであると決定された(データは示さない)。

【0051】

(天然に存在しないアミノ酸の使用もまた実行可能であるが)20個の天然に存在するアミノ酸しかないので、この16番目の位置での置換のために特定の非内因性アミノ酸の選択は実行可能であり、研究者の要求に適合する非内因性アミノ酸を効率よく選択することができる。しかしながら、記載したように、16番目の位置のより好ましい非内因性アミノ酸はリシン、ヒスチジン、アルギニン及びアラニンであり、リシンが最も好ましい。当業者は、所望する突然変異を行うためにコドンの配列を変える熟練した方法を容易に決定することができると考えられる。

【0052】

いつもとは限らないが、時折、TM6中でプロリン残基マーカーの前にW2が位置すること(すなわち、W2P¹AA₁₅X)もまた見出され、ここで、Wはトリプトファンであり、2はあらゆるアミノ酸残基である。

【0053】

我々の発見は、とりわけ、当該技術分野により一般的に用いられる予測できず且つ面倒な配列整列法の必要性を否定する。実際、我々の発見の長所は、アルゴリズムが事実上あるが(while an algorithm in nature)、独特で且つ非常に有用な最終生成物、すなわち、ヒトGPCRの構成的に活性化されたバージョンを得るために、当業者が熟練した扱いやすさで、ヒトGPCRにそれを容易に適用できることである。ヒトゲノムプロジェクトが明らかにしているヒトGPCRの内因性リガンドを決定するためには多年及びかなりの額の金銭が必要とされるので、開示した本発明は、この配列情報を実践的に利用するために必要な時間を減らすだけでなく、著しく費用を節減する。この方法は、例えば疾病におけるGPCRの役割を理解するためだけではなく、ヒト症状を改善する機会を与えるためにも遺伝子情報の利用を可能にするので、ヒトゲノムプロジェクトの重要性を適切に実証する。

D. 候補化合物のスクリーニング

1. 一般的なGPCRスクリーニングアッセイ技術

Gタンパク質受容体が構成的に活性になると、それはGタンパク質(例えばGq、Gs、Gi、Go)に共役し、GDPの遊離及び続いて起こるGタンパク質へのGTPの結合を刺激する。Gタンパク質は次にGTPアーゼとして働き、GTPをGDPにゆっくり加水分解し、それにより受容体は通常の条件下で非活性化されるようになる。しかしながら、本発明の非内因性の構成的に活性のあるヒトGPCRを包含する構成的に活性化された受容体は、GDPをGTPに交換し続ける。GTPの加水分解できない類似体 [³⁵S] GTP γ Sは、構成的に活性化された受容体を発現する膜上に存在するGタンパク質への増大した結合をモニターするために用いることができる。 [³⁵S] GTP γ Sは、リガンドの非存在下及び存在下で膜へのGタンパク質結合をモニターするために使用できることが報告されている。このモニタリングの例は、当業者が周知で且つ利用できる他の例の中、1995年にTraynor及びNahorskiにより報告された。このアッセイ系は、受容体の細胞内ドメインと相互作用する特定のGタンパク質にかかわらず全てのGタンパク質共役受容体に一般的に適用できるので、この系の好ましい使用は候補化合物の初期スクリーニングに対してである。

【0054】

B 2. 特定のGPCRスクリーニングアッセイ技術

C 「一般的な」Gタンパク質共役受容体アッセイ(すなわち、アゴニスト、部分アゴニストまたは逆アゴニストである化合物を選択するためのアッセイ)を用いていったん候補化合物が同定されると、それらの化合物が受容体部位で相互作用していることを確かめるためのさらなるスクリーニングが好ましい。例えば、「一般的な」アッセイにより同定され

た化合物は受容体に結合することができないが、その代わりに細胞内ドメインからGタンパク質を単に「離す」ことができる。

【0055】

a. Gs及びGi.

Gsは酵素アデニルシクラーゼを刺激する。一方、Gi（及びGo）はこの酵素を阻害する。アデニルシクラーゼはcAMPへのATPの転化を触媒し；従って、Gsタンパク質と共役する構成的に活性化されたGPCRはcAMPの増加した細胞レベルと関係する。一方、Gi（またはGo）タンパク質と共役する構成的に活性化されたGPCRはcAMPの減少した細胞レベルと関係する。一般的に、「Indirect Mechanisms of Synaptic Transmission」, 8章, From Neuron To Brain (第3版) Nichols, J. G. et al編集, Sinauer Associates, Inc. (1992)を参照。従って、候補化合物が受容体に対する例えば逆アゴニストであるかどうかを決定するためにcAMPを検出するアッセイを利用することができる（すなわち、そのような化合物はcAMPのレベルを下げる）。cAMPを測定するために当該技術分野において既知の様々な方法を利用することができ；最も好ましい方法は、ELISAに基づく形態の抗-cAMP抗体の使用による。利用することができる別のタイプのアッセイは、全細胞第二メッセンジャーレポーター系アッセイである。遺伝子上のプロモーターは、特定の遺伝子がコードするタンパク質の発現を導く。サイクリックAMPは、cAMP応答DNA結合タンパク質または転写因子（CREB）の結合を促進することにより遺伝子発現を導き、それは次にcAMP応答配列と呼ばれる特定の位置でプロモーターに結合し、遺伝子の発現を導く。レポーター遺伝子、例えばβ-ガラクトシダーゼまたはルシフェラーゼの前に多数のcAMP応答配列を含有するプロモーターを有するレポーター系を構築することができる。従って、構成的に活性化されたGs共役受容体（Gs-linked receptor）はcAMPの蓄積をもたらし、これは次にレポータータンパク質の遺伝子及び発現を活性化する。次に、標準的な生化学アッセイを用いてβ-ガラクトシダーゼまたはルシフェラーゼのようなレポータータンパク質を検出することができる（Chen et al. 1995）。Gi（またはGo）に共役し、従ってcAMPのレベルを下げるGPCRに関して、Gsに共役する（従って、cAMPのレベルを上げる）受容体の利用に基づく例えば逆アゴニストのスクリーニングの方法をGPR17及びGPR30について実施例の節において開示する。

【0056】

b. Go及びGq

Go及びGqは酵素ホスホリパーゼCの活性化と関係し、これは次にリン脂質PIP₂を加水分解し、2つの細胞内メッセンジャー：ジアシルグリセロール（DAG）及びイノシトール1,4,5-三リン酸（IP₃）を遊離する。IP₃の増加した蓄積は、Gq-及びGo-共役受容体（Go-associated receptors）の活性化と関係する。一般的に、「Indirect Mechanisms of Synaptic Transmission」, 8章, From Neuron To Brain (第3版) Nichols, J. G. et al編集, Sinauer Associates, Inc. (1992)を参照。候補化合物がGq-またはGo-共役受容体に対する例えば逆アゴニストであるかどうかを決定するためにIP₃蓄積を検出するアッセイを利用することができる（すなわち、そのような化合物はIP₃のレベルを下げる）。Gq依存性ホスホリパーゼCはAP1要素を含有する遺伝子の活性化をもたらす；従って、活性化されたGq-共役受容体はそのような遺伝子の発現の増加を示し、それにより、その逆アゴニストはそのような発現の減少を示し、そしてアゴニストはそのような発現の増加を示すので、AP1レポーターアッセイを用いてGq-共役受容体を調べることもできる。そのような検出のために市販されているアッセイが利用できる。

E. 医薬品化学

必ずしもそうとは限らないが、一般に、候補化合物の直接的同定は、好ましくは、組み合わせ化学技術により作製された化合物と共に実施され、それにより何千もの化合物がそのような分析のために無作為に調製される。一般に、そのようなスクリーニングの結果は、独特なコア構造を有する化合物であり；その後、これらの化合物は、好ましくは、その医薬品特性をさらに高めるために好ましいコア構造（1つまたは複数）の辺りのさらなる化学的改変に供される。そのような技術は当業者に既知であり、本特許書類において詳細に

10

20

30

40

50

は検討されない。

F. 製薬学的組成物

さらなる開発のために選択された候補化合物は、当業者に周知の技術を用いて製薬学的組成物中に調合することができる。当業者は適当な製薬学的に許容しうる担体を利用できる；例えば、Remington's Pharmaceutical Sciences, 第16版, 1980, Mack Publishing Co., (Oslo et al.編集)を参照。

G. 他の有用性

開示したヒトGPCRの非内因性バージョンの好ましい用途は、逆アゴニスト、アゴニストまたは部分アゴニストとして候補化合物を直接同定するため（好ましくは、製薬としての使用のため）であるが、これらの受容体を研究環境において利用することもできる。例えば、これらの受容体を含むインビボ及びインビトロ系は、正常及び罹病の両方のヒト状態におけるこれらの受容体の役割をさらに明らかにし理解するため並びにそれがシグナリングカスケードを理解することに適用される場合に構成的活性化の役割を理解することを利用することができる。これらの非内因性受容体の価値は、それらの独特な特徴のために、開示した受容体をその内因性リガンドが同定される前に人体における特定の受容体の役割を理解するために用いることができるので研究手段としてのそれらの有用性が高められることである。開示した受容体の他の用途は、とりわけ本特許書類の説明に基づいて当業者に明らかになる。

【0057】

実施例

以下の実施例は、本発明を明らかにする目的のために示し、限定ではない。本明細書において特定の核酸及びアミノ酸配列が開示されるが、ヒトGPCRを構成的に活性化するためにTM6中のプロリン残基に対する位置に基づき受容体のIC3ループにおいて突然変異カセットを利用することができるという本特許書類の教示に従って、当業者は以下に報告するものと同じまたは実質的に同様な結果を得ながらこれらの配列に対してわずかな改変を行うことができると考えられる。配列突然変異の特定の方法は、当業者の特定の必要性に基づき当業者の認識範囲内である。

【0058】

実施例1

内因性ヒトGPCRの調製

様々なGPCRを以下の実施例において利用した。いくつかの内因性ヒトGPCRは（以下に謝辞を示すように）発現ベクター中で親切にも提供され、そして他の内因性ヒトGPCRは公的に利用できる配列情報を用いて新規に合成した。

【0059】

1. GPR1（ジーンバンク受託番号：U13666）

GPR1のヒトcDNA配列は、Brian O'Dowd (University of Toronto) によりpRcCMV中で提供された。GPR1 cDNA(1.4kBフラグメント)をpRcCMVベクターからNdeI-XbaIフラグメントとして切り出し、pCMVベクター（図3参照）のNdeI-XbaI部位中にサブクローン化した。その後、ヒトGPR1の核酸（配列番号:1）及びアミノ酸（配列番号:2）配列を決定し、確認した。

【0060】

2. GPR4（ジーンバンク受託番号：L36148、U35399、U21051）

GPR4のヒトcDNA配列は、Brian O'Dowd (University of Toronto) によりpRcCMV中で提供された。GPR1 cDNA(1.4kBフラグメント)をpRcCMVベクターからApaI（平滑にした）-XbaIフラグメントとして切り出し、pCMVベクターのHindIII（平滑にした）-XbaI部位中に（5'非翻訳領域の大部分を除いて）サブクローン化した。その後、ヒトGPR4の核酸（配列番号:3）及びアミノ酸（配列番号:4）配列を決定し、確認した。

【0061】

3. GPR5（ジーンバンク受託番号：L36149）

ヒトGPR5のcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鋳型とし

てゲノムDNA及びrTthポリメラーゼ (Perkin Elmer) を製造業者により提供されるバッファー系、0.25 μ Mの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94°Cで1分間；64°Cで1分間；そして72°Cで1.5分間を30サイクルであった。5' PCRプライマーは配列：

5'-TATGAATTCAGATGCTCTAAACGTCCCTGC-3' (配列番号:5)

でEcoRI部位を含有し、そして3'プライマーは配列：

5'-TCCGGATCCACCTGCACCTGCGCCTGCACC-3' (配列番号:6)

でBamHI部位を含有した。1.1kbのPCRフラグメントをEcoRI及びBamHIで消化し、pCMV発現ベクターのEcoRI-BamHI部位中にクローン化した。その後、ヒトGPR5の核酸 (配列番号:7) 及びアミノ酸 (配列番号:8) 配列を決定し、確認した。

10

【0062】

4. GPR 7 (ジーンバンク受託番号: U22491)

ヒトGPR 7のcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：PCR条件-鋳型としてゲノムDNA及びrTthポリメラーゼ (Perkin Elmer) を製造業者により提供されるバッファー系、0.25 μ Mの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94°Cで1分間；62°Cで1分間；そして72°Cで1分20秒間を30サイクルであった。5' PCRプライマーは配列：

5'-GCAAGCTTGGGGGACGCCAGGTCGCCGGCT-3' (配列番号:9)

でHindIII部位を含有し、そして3'プライマーは配列：

5'-GCGGATCCGGACGCTGGGGGAGTCAGGCTGC-3' (配列番号:10)

20

でBamHI部位を含有した。1.1kbのPCRフラグメントをHindIII及びBamHIで消化し、pCMV発現ベクターのHindIII-BamHI部位中にクローン化した。その後、ヒトGPR7の核酸 (配列番号:11) 及びアミノ酸 (配列番号:12) 配列を決定し、確認した。

【0063】

5. GPR8 (ジーンバンク受託番号: U22492)

ヒトGPR8のcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鋳型としてゲノムDNA及びrTthポリメラーゼ (Perkin Elmer) を製造業者により提供されるバッファー系、0.25 μ Mの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94°Cで1分間；62°Cで1分間；そして72°Cで1分20秒間を30サイクルであった。5' PCRプライマーは配列：

30

5'-CGGAATTCGTCAACGGTCCCAGCTACAATG-3' (配列番号:13)

でEcoRI部位を含有し、そして3'プライマーは配列：

5'-ATGGATCCCAGGCCCTTCAGCACCGCAATAT-3' (配列番号:14)

でBamHI部位を含有した。1.1kbのPCRフラグメントをEcoRI及びBamHIで消化し、pCMV発現ベクターのEcoRI-BamHI部位中にクローン化した。シークエンスした4個全てのcDNAクローンは、ArgからGlnへのアミノ酸206の変異を伴う、可能性がある多型を含有した。この違いを除いて、その後、ヒトGPR8の核酸 (配列番号:15) 及びアミノ酸 (配列番号:16) 配列を決定し、確認した。

【0064】

6. GPR 9 (ジーンバンク受託番号: X95876)

40

ヒトGPR 9のcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鋳型として (Brian O'Dowdにより提供された) クローン及びpfuポリメラーゼ (Stratagene) を10% DMSOを補足した製造業者により提供されるバッファー系、0.25 μ Mの各プライマー及び0.5mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94°Cで1分間；56°Cで1分間；そして72°Cで2.5分間を25サイクルであった。5' PCRプライマーは配列：

5'-ACGAATTCAGCCATGGTCCTTGAGGTGAGTGACCACCAAGTGCTAAAT-3' (配列番号:17) でEcoRI部位を含有し、そして3'プライマーは配列：

5'-GAGGATCCTGGAATGCGGGGAAGTCAG-3' (配列番号:18)

でBamHI部位を含有した。1.2kbのPCRフラグメントをEcoRIで消化し、pCMV発現ベクターのEcoRI-SmaI部位中にクローン化した。その後、ヒトGPR9の核酸 (配列番号:19) 及びアミ

50

ノ酸（配列番号:20）配列を決定し、確認した。

【0065】

7. GPR9-6（ジーンバンク受託番号：U45982）

ヒトGPR9-6のcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鋳型としてゲノムDNA及びrTthポリメラーゼ（Perkin Elmer）を製造業者により提供されるバッファー系、0.25 μ Mの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94°Cで1分間；62°Cで1分間；そして72°Cで1分20秒間を30サイクルであった。5' PCRプライマーは配列：

5'-TTAAGCTTGACCTAATGCCATCTTGTGTCC-3'（配列番号:21）

でキナーゼ処理し（kinased）、そして3'プライマーは配列：

5'-TTGGATCCAAAAGAACCATGCACCTCAGAG-3'（配列番号:22）

でBamHI部位を含有した。1.2kbのPCRフラグメントをBamHIで消化し、pCMV発現ベクターのEcoRV-BamHI部位中にクローン化した。その後、ヒトGPR9-6の核酸（配列番号:23）及びアミノ酸（配列番号:24）配列を決定し、確認した。

【0066】

8. GPR10（ジーンバンク受託番号：U32672）

GPR10のヒトcDNA配列は、Brian O'Dowd（University of Toronto）によりpRcCMV中で提供された。GPR10 cDNA(1.3kbフラグメント)をpRcCMVベクターからEcoRI-XbaIフラグメントとして切り出し、pCMVベクターのEcoRI-XbaI部位中にサブクローン化した。その後、ヒトGPR10の核酸（配列番号:25）及びアミノ酸（配列番号:26）配列を決定し、確認した。

【0067】

9. GPR15（ジーンバンク受託番号：U34806）

GPR15のヒトcDNA配列は、Brian O'Dowd（University of Toronto）によりpCDNA3中で提供された。GPR15 cDNA(1.5kbフラグメント)をpCDNA3ベクターからHindIII-Bamフラグメントとして切り出し、pCMVベクターのHindIII-Bam部位中にサブクローン化した。その後、ヒトGPR15の核酸（配列番号:27）及びアミノ酸（配列番号:28）配列を決定し、確認した。

【0068】

10. GPR17（ジーンバンク受託番号：Z94154）

ヒトGPR17のcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鋳型としてゲノムDNA及びrTthポリメラーゼ（Perkin Elmer）を製造業者により提供されるバッファー系、0.25 μ Mの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94°Cで1分間；56°Cで1分間；そして72°Cで1分20秒間を30サイクルであった。5' PCRプライマーは配列：

5'-CTAGAATTCTGACTCCAGCCAAAGCATGAAT-3'（配列番号:29）

でEcoRI部位を含有し、そして3'プライマーは配列：

5'-GCTGGATCCTAAACAGTCTGCGCTCGGCCT-3'（配列番号:30）

でBamHI部位を含有した。1.1kbのPCRフラグメントをEcoRI及びBamHIで消化し、pCMV発現ベクターのEcoRI-BamHI部位中にクローン化した。その後、ヒトGPR17の核酸（配列番号:31）及びアミノ酸（配列番号:32）配列を決定し、確認した。

【0069】

11. GPR18（ジーンバンク受託番号：L42324）

ヒトGPR18のcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鋳型としてゲノムDNA及びrTthポリメラーゼ（Perkin Elmer）を製造業者により提供されるバッファー系、0.25 μ Mの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94°Cで1分間；54°Cで1分間；そして72°Cで1分20秒間を30サイクルであった。5' PCRプライマーは配列：

5'-ATAAGATGATCACCTGAACAATCAAGAT-3'（配列番号:33）

でキナーゼ処理し、そして3'プライマーは配列：

5'-TCCGAATTCATAACATTTCACTGTTTATATTGC-3'（配列番号:34）

でEcoRI部位を含有した。1.0kbのPCRフラグメントをEcoRIで消化し、pCMV発現ベクターの

10

20

30

40

50

平滑-EcoRI部位中にクローン化した。シークエンスした8個全てのcDNAクローンは、ThrからProへのアミノ酸12、AlaからGluへのアミノ酸86、IleからLeuへのアミノ酸97及びLeuからMetへのアミノ酸310の変異を伴う、可能性がある4つの多型を含有した。これらの変異を除いて、その後、ヒトGPR18の核酸（配列番号:35）及びアミノ酸（配列番号:36）配列を決定し、確認した。

【0070】

12. GPR20（ジーンバンク受託番号：U66579）

ヒトGPR20のcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鑄型としてゲノムDNA及びrTthポリメラーゼ（Perkin Elmer）を製造業者により提供されるバッファー系、0.25 μ Mの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94℃で1分間；62℃で1分間；そして72℃で1分20秒間を30サイクルであった。5' PCRプライマーは配列：

5'-CCAAGCTTCCAGGCCTGGGGTGTGCTGG-3'（配列番号:37）

でキナーゼ処理し、そして3'プライマーは配列：

5'-ATGGATCTGACCTTCGGCCCTGGCAGA-3'（配列番号:38）

でBamHI部位を含有した。1.2kbのPCRフラグメントをBamHIで消化し、pCMV発現ベクターのEcoRV-BamHI部位中にクローン化した。その後、ヒトGPR20の核酸（配列番号:39）及びアミノ酸（配列番号:40）配列を決定し、確認した。

【0071】

13. GPR21（ジーンバンク受託番号：U66580）

ヒトGPR21のcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鑄型としてゲノムDNA及びrTthポリメラーゼ（Perkin Elmer）を製造業者により提供されるバッファー系、0.25 μ Mの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94℃で1分間；62℃で1分間；そして72℃で1分20秒間を30サイクルであった。5' PCRプライマーは配列：

5'-GAGAATTCACCTCCTGAGCTCAAGATGAACT-3'（配列番号:41）

でキナーゼ処理し、そして3'プライマーは配列：

5'-CGGGATCCCCGTAAGTGAAGCCACTTCAGAT-3'（配列番号:42）

でBamHI部位を含有した。1.1kbのPCRフラグメントをBamHIで消化し、pCMV発現ベクターのEcoRV-BamHI部位中にクローン化した。その後、ヒトGPR21の核酸（配列番号:43）及びアミノ酸（配列番号:44）配列を決定し、確認した。

【0072】

14. GPR22（ジーンバンク受託番号：U66581）

ヒトGPR22のcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鑄型としてゲノムDNA及びrTthポリメラーゼ（Perkin Elmer）を製造業者により提供されるバッファー系、0.25 μ Mの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94℃で1分間；50℃で1分間；そして72℃で1.5分間を30サイクルであった。5' PCRプライマーは配列：

5'-TCCCCCGGGAAAAAACCAACTGCTCCAAA-3'（配列番号:45）

でキナーゼ処理し、そして3'プライマーは配列：

5'-TAGGATCCATTTGAATGTGGATTTGGTGAAA-3'（配列番号:46）

でBamHI部位を含有した。1.38kbのPCRフラグメントをBamHIで消化し、pCMV発現ベクターのEcoRV-BamHI部位中にクローン化した。その後、ヒトGPR22の核酸（配列番号:47）及びアミノ酸（配列番号:48）配列を決定し、確認した。

【0073】

15. GPR24（ジーンバンク受託番号：U71092）

ヒトGPR24のcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鑄型としてゲノムDNA及びrTthポリメラーゼ（Perkin Elmer）を製造業者により提供されるバッファー系、0.25 μ Mの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94℃で1分間；56℃で1分間；そして72℃で1分20秒間を30サイクル

であった。5' PCRプライマーは配列：

5'-GTGAAGCTTGCTCTGGTGCCTGCAGGAGG-3' (配列番号:49)

でHindIII部位を含有し、そして3'プライマーは配列：

5'-GCAGAATTCCTGGTGGCGTGTGTGGTGGCC-3' (配列番号:50)

でEcoRI部位を含有した。1.3kbのPCRフラグメントをHindIII及びEcoRIで消化し、pCMV発現ベクターのHindIII-EcoRI部位中にクローン化した。その後、ヒトGPR24の核酸 (配列番号:51) 及びアミノ酸 (配列番号:52) 配列を決定し、確認した。

【0074】

16. GPR30 (ジーンバンク受託番号: U63917)

ヒトGPR30のcDNAを以下のように作製し、クローン化した：GPR30のコーディング配列 (11 28bpの長さ) をプライマー：

5'-GGCGGATCCATGGATGTGACTTCCCAA-3' (配列番号:53) 及び

5'-GGCGGATCCCTACACGGCACTGCTGAA-3' (配列番号:54)

を用いてゲノムDNAから増幅した。次に、増幅産物を市販されているベクターpCR2.1 (Invitrogen) 中に製造業者の説明書に従って「TOPO-TAクローニングキット」 (Invitrogen、#K4500-01) を用いてクローン化した。全長GPR30インサートをBamHIでの消化により遊離し、アガロースゲル電気泳動によりベクターから分離し、Sephaglas Bandprep™キット (Pharmacia、#27-9285-01) を製造業者の説明書に従って用いて精製した。その後、ヒトGPR30の核酸 (配列番号:55) 及びアミノ酸 (配列番号:56) 配列を決定し、確認した。

【0075】

17. GPR31 (ジーンバンク受託番号: U65402)

ヒトGPR31のcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鋳型としてゲノムDNA及びrTthポリメラーゼ (Perkin Elmer) を製造業者により提供されるバッファー系、0.25 µMの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94℃で1分間；58℃で1分間；そして72℃で2分間を30サイクルであった。5' PCRプライマーは配列：

5'-AAGGAATTCACGGCCGGGTGATGCCATTCCC-3' (配列番号:57)

でEcoRI部位を含有し、そして3'プライマーは配列：

5'-GGTGGATCCATAAACACGGGCGTTGAGGAC-3' (配列番号:58)

でBamHI部位を含有した。1.0kbのPCRフラグメントをEcoRI及びBamHIで消化し、pCMV発現ベクターのEcoRI-BamHI部位中にクローン化した。その後、ヒトGPR31の核酸 (配列番号:59) 及びアミノ酸 (配列番号:60) 配列を決定し、確認した。

【0076】

18. GPR32 (ジーンバンク受託番号: AF045764)

ヒトGPR32のcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鋳型としてゲノムDNA及びrTthポリメラーゼ (Perkin Elmer) を製造業者により提供されるバッファー系、0.25 µMの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94℃で1分間；56℃で1分間；そして72℃で1分20秒間を30サイクルであった。5' PCRプライマーは配列：

5'-TAAGAATTCATATAAAATTATGGAATGG-3' (配列番号:243)

でEcoRI部位を含有し、そして3'プライマーは配列：

5'-CCAGGATCCAGCTGAAGTCTTCCATCATTC-3' (配列番号:244)

でBamHI部位を含有した。1.1kbのPCRフラグメントをEcoRI及びBamHIで消化し、pCMV発現ベクターのEcoRI-BamHI部位中にクローン化した。その後、ヒトGPR32の核酸 (配列番号:245) 及びアミノ酸 (配列番号:246) 配列を決定し、確認した。

【0077】

19. GPR40 (ジーンバンク受託番号: AF024687)

ヒトGPR40のcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鋳型としてゲノムDNA及びrTthポリメラーゼ (Perkin Elmer) を製造業者により提供されるバッファー系、0.25 µMの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施

した。サイクル条件は94℃で1分間；65℃で1分間；そして72℃で1分10秒間を30サイクルであった。5' PCRプライマーは配列：

5'-GCAGAAATTCGGCGGCCCCATGGACCTGCCCC-3' (配列番号:247)

でEcoRI部位を含有し、そして3'プライマーは配列：

5'-GCTGGATCCCCGAGCAGTGGCGTTACTTC-3' (配列番号:248)

でBamHI部位を含有した。1kbのPCRフラグメントをEcoRI及びBamHIで消化し、pCMV発現ベクターのEcoRI-BamHI部位中にクローン化した。その後、ヒトGPR40の核酸 (配列番号:249) 及びアミノ酸 (配列番号:250) 配列を決定し、確認した。

【0078】

20. GPR41 (ジーンバンク受託番号: AF024688)

10

ヒトGPR41のcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鋳型としてゲノムDNA及びrTthポリメラーゼ (Perkin Elmer) を製造業者により提供されるバッファー系、0.25 µMの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94℃で1分間；65℃で1分間；そして72℃で1分10秒間を30サイクルであった。5' PCRプライマーは配列：

5'-CTCAAGCTTACTCTCTCTCACCAGTGGCCAC-3' (配列番号:251)

でHindIII部位を含有し、そして3'プライマーは配列：

5'-CCCTCTCCCCCGAGGACCTAGC-3' (配列番号:252)

でキナーゼ処理した。1kbのPCRフラグメントをHindIIIで消化し、pCMV発現ベクターのHindIII-平滑部位中にクローン化した。その後、ヒトGPR41の核酸 (配列番号:253) 及びアミノ酸 (配列番号:254) 配列を決定し、確認した。

20

【0079】

21. GPR43 (ジーンバンク受託番号: AF024690)

ヒトGPR43のcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鋳型としてゲノムDNA及びrTthポリメラーゼ (Perkin Elmer) を製造業者により提供されるバッファー系、0.25 µMの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94℃で1分間；65℃で1分間；そして72℃で1分10秒間を30サイクルであった。5' PCRプライマーは配列：

5'-TTTAAGCTTCCCCTCCAGGATGCTGCCGGAC-3' (配列番号:255)

でHindIII部位を含有し、そして3'プライマーは配列：

5'-GGCGAATTCTGAAGGTCCAGGAACTGCTA-3' (配列番号:256)

でEcoRI部位を含有した。1kbのPCRフラグメントをHindIII及びEcoRIで消化し、pCMV発現ベクターのHindIII-EcoRI部位中にクローン化した。その後、ヒトGPR43の核酸 (配列番号:257) 及びアミノ酸 (配列番号:258) 配列を決定し、確認した。

30

【0080】

22. APJ (ジーンバンク受託番号: U03642)

ヒトAPJ cDNA (pRcCMVベクター中) は、Brian O'Dowd (University of Toronto) により提供された。ヒトAPJ cDNAをpRcCMVベクターからEcoRI-XbaI (平滑にした) フラグメントとして切り出し、pCMVベクターのEcoRI-SmaI部位中にサブクローン化した。その後、ヒトAPJの核酸 (配列番号:61) 及びアミノ酸 (配列番号:62) 配列を決定し、確認した。

40

【0081】

23. BLR1 (ジーンバンク受託番号: X68149)

ヒトBLR1のcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鋳型として胸腺cDNA及びrTthポリメラーゼ (Perkin Elmer) を製造業者により提供されるバッファー系、0.25 µMの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94℃で1分間；62℃で1分間；そして72℃で1分20秒間を30サイクルであった。5' PCRプライマーは配列：

5'-TGAGAATTCTGGTGACTCACAGCCGGCACAG-3' (配列番号:63)

でEcoRI部位を含有し、そして3'プライマーは配列：

5'-GCCGGATCCAAGGAAAAGCAGCAATAAAAGG-3' (配列番号:64)

50

でBamHI部位を含有した。1.2kbのPCRフラグメントをEcoRI及びBamHIで消化し、pCMV発現ベクターのEcoRI-BamHI部位中にクローン化した。その後、ヒトBLR1の核酸（配列番号:65）及びアミノ酸（配列番号:66）配列を決定し、確認した。

【0082】

24. CEPR（ジーンバンク受託番号：U77827）

ヒトCEPRのcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鋳型としてゲノムDNA及びrTthポリメラーゼ（Perkin Elmer）を製造業者により提供されるバッファー系、0.25 μ Mの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94℃で1分間；65℃で1分間；そして72℃で1分20秒間を30サイクルであった。5' PCRプライマーは配列：

5'-CAAAGCTTGAAAGCTGCACGGTGCAGAGAC-3'（配列番号:67）

でキナーゼ処理し、そして3'プライマーは配列：

5'-GCGGATCCCGAGTCACACCCTGGCTGGGCC-3'（配列番号:68）

でBamHI部位を含有した。1.2kbのPCRフラグメントをBamHIで消化し、pCMV発現ベクターのEcoRV-BamHI部位中にクローン化した。その後、ヒトCEPRの核酸（配列番号:69）及びアミノ酸（配列番号:70）配列を決定し、確認した。

【0083】

25. EBI1（ジーンバンク受託番号：L31581）

ヒトEBI1のcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鋳型として胸腺cDNA及びrTthポリメラーゼ（Perkin Elmer）を製造業者により提供されるバッファー系、0.25 μ Mの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94℃で1分間；62℃で1分間；そして72℃で1分20秒間を30サイクルであった。5' PCRプライマーは配列：

5'-ACAGAATTCCTGTGTGGTTTTACCGCCAG-3'（配列番号:71）

でEcoRI部位を含有し、そして3'プライマーは配列：

5'-CTCGGATCCAGGCAGAAGAGTCGCCTATGG-3'（配列番号:72）

でBamHI部位を含有した。1.2kbのPCRフラグメントをEcoRI及びBamHIで消化し、pCMV発現ベクターのEcoRI-BamHI部位中にクローン化した。その後、ヒトEBI1の核酸（配列番号:73）及びアミノ酸（配列番号:74）配列を決定し、確認した。

【0084】

26. EBI2（ジーンバンク受託番号：L08177）

ヒトEBI2のcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鋳型としてcDNAクローン（Kevin Lynch、University of Virginia Health Sciences Centerにより親切に提供された；利用したベクターは供給者により同定されなかった）及びpfuポリメラーゼ（Stratagene）を10% DMSOを補足した製造業者により提供されるバッファー系、0.25 μ Mの各プライマー及び0.5mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94℃で1分間；60℃で1分間；そして72℃で1分20秒間を30サイクルであった。5' PCRプライマーは配列：

5'-CTGGAATTCACCTGGACCACCAATGGATA-3'（配列番号:75）

でEcoRI部位を含有し、そして3'プライマーは配列：

5'-CTCGGATCCTGCAAAGTTTGTATACAGTT-3'（配列番号:76）

でBamHI部位を含有した。1.2kbのPCRフラグメントをEcoRI及びBamHIで消化し、pCMV発現ベクターのEcoRI-BamHI部位中にクローン化した。その後、ヒトEBI2の核酸（配列番号:77）及びアミノ酸（配列番号:78）配列を決定し、確認した。

【0085】

27. ETBR-LP2（ジーンバンク受託番号：D38449）

ヒトETBR-LP2のcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鋳型として脳cDNA及びrTthポリメラーゼ（Perkin Elmer）を製造業者により提供されるバッファー系、0.25 μ Mの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94℃で1分間；65℃で1分間；そして72℃で1.5分間を30サイクルであ

10

20

30

40

50

った。5' PCRプライマーは配列：

5'-CTGGAATTCTCTGCTCATCCAGCCATGCGG-3' (配列番号:79)

でEcoRI部位を含有し、そして3'プライマーは配列：

5'-CCTGGATCCCCACCCCTACTGGGGCCTCAG-3' (配列番号:80)

でBamHI部位を含有した。1.5kbのPCRフラグメントをEcoRI及びBamHIで消化し、pCMV発現ベクターのEcoRI-BamHI部位中にクローン化した。その後、ヒトETBR-LP2の核酸 (配列番号:81) 及びアミノ酸 (配列番号:82) 配列を決定し、確認した。

【0086】

28. GHSR (ジーンバンク受託番号:U60179)

ヒトGHSRのcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鋳型として海馬cDNA及びTaqPlus Precisionポリメラーゼ (Stratagene) を製造業者により提供されるバッファー系、0.25 µMの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94℃で1分間；68℃で1分間；そして72℃で1分10秒間を30サイクルであった。1回目のPCRのために、5' PCRプライマー配列は：

5'-ATGTGGAACGCGACGCCCAGCG-3' (配列番号:83)

であり、そして3'プライマー配列は：

5'-TCATGTATTAATACTAGATTCT-3' (配列番号:84)

であった。1回目のPCRのうち2 µlを2回目のPCRに鋳型として用い、その場合、5'プライマーは配列：

5'-TACCATGTGGAACGCGACGCCCAGCGAAGAGCCGGGGT-3' (配列番号:85)

でキナーゼ処理し、そして3'プライマーは配列：

5'-CGGAATTCATGTATTAATACTAGATTCTGTCCAGGCCCG-3' (配列番号:86)

でEcoRI部位を含有した。1.1kbのPCRフラグメントをEcoRIで消化し、pCMV発現ベクターの平滑-EcoRI部位中にクローン化した。その後、ヒトGHSRの核酸 (配列番号:87) 及びアミノ酸 (配列番号:88) 配列を決定し、確認した。

【0087】

29. GPCR-CNS (ジーンバンク受託番号:AF017262)

ヒトGPCR-CNSのcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鋳型として脳cDNA及びrTthポリメラーゼ (Perkin Elmer) を製造業者により提供されるバッファー系、0.25 µMの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94℃で1分間；65℃で1分間；そして72℃で2分間を30サイクルであった。5' PCRプライマーは配列：

5'-GCAAGCTTGTGCCCTCACCAAGCCATGCGAGCC-3' (配列番号:89)

でHindIII部位を含有し、そして3'プライマーは配列：

5'-CGGAATTCAGCAATGAGTTCCGACAGAAGC-3' (配列番号:90)

でEcoRI部位を含有した。1.9kbのPCRフラグメントをHindIII及びEcoRIで消化し、pCMV発現ベクターのHindIII-EcoRI部位中にクローン化した。シークエンスした9個全てのクローンは、S284C変異を伴う、可能性がある多型を含有した。この違いを除いて、その後、ヒトGPCR-CNSの核酸 (配列番号:91) 及びアミノ酸 (配列番号:92) 配列を決定し、確認した。

【0088】

30. GPR-NGA (ジーンバンク受託番号:U55312)

ヒトGPR-NGAのcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鋳型としてゲノムcDNA及びrTthポリメラーゼ (Perkin Elmer) を製造業者により提供されるバッファー系、0.25 µMの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94℃で1分間；56℃で1分間；そして72℃で1.5分間を30サイクルであった。5' PCRプライマーは配列：

5'-CAGAATTCAGAGAAAAAAGTGAATATGGTTTTT-3' (配列番号:93)

でEcoRI部位を含有し、そして3'プライマーは配列：

5'-TTGGATCCCTGGTGCATAACAATTGAAAGAAT-3' (配列番号:94)

10

20

30

40

50

でBamHI部位を含有した。1.3kbのPCRフラグメントをEcoRI及びBamHIで消化し、pCMV発現ベクターのEcoRI-BamHI部位中にクローン化した。その後、ヒトGPR-NGAの核酸（配列番号:95）及びアミノ酸（配列番号:96）配列を決定し、確認した。

【0089】

31. H9（ジーンバンク受託番号：U52219）

ヒトHB954のcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鋳型として下垂体cDNA及びrTthポリメラーゼ（Perkin Elmer）を製造業者により提供されるバッファー系、0.25 μ Mの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94℃で1分間；62℃で1分間；そして72℃で2分間を30サイクルであった。5'プライマーは配列：

5'-GGAAAGCTTAACGATCCCCAGGAGCAACAT-3'（配列番号:97）

でHindIII部位を含有し、そして3'プライマーは配列：

5'-CTGGGATCCTACGAGAGCATTTTTCACACAG-3'（配列番号:98）

でBamHI部位を含有した。1.9kbのPCRフラグメントをHindIII及びBamHIで消化し、pCMV発現ベクターのHindIII-BamHI部位中にクローン化した。公開配列と比較した場合に、細胞質尾部中に12bpのインフレーム挿入を有する異なるアイソフォームも同定され、「H9b」と称した。両方のアイソフォームは、アミノ酸P320S及びアミノ酸G448Aの変異を伴う2つの可能性がある多型を含有した。アイソフォームH9aはアミノ酸S493Nの別の可能性がある多型を含有し、一方、アイソフォームH9bはアミノ酸I502T及びアミノ酸A532T（アイソフォームH9aのアミノ酸528に対応する）の変異を伴う2つのさらなる可能性がある多型を含有した。その後、ヒトH9の核酸（配列番号:99）及びアミノ酸（配列番号:100）配列を決定し、確認した（以下の節において、両方のアイソフォームをヒトGPCRプロリンマーカールゴリズムに従って突然変異させた）。

【0090】

32. HB954（ジーンバンク受託番号：D38449）

ヒトHB954のcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鋳型として脳cDNA及びrTthポリメラーゼ（Perkin Elmer）を製造業者により提供されるバッファー系、0.25 μ Mの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94℃で1分間；58℃で1分間；そして72℃で2分間を30サイクルであった。5' PCRプライマーは配列：

5'-TCCAAGCTTCGCCATGGGACATAACGGGAGCT-3'（配列番号:101）

でHindIII部位を含有し、そして3'プライマーは配列：

5'-CGTGAATTCCAAGAATTTACAATCCTTGCT-3'（配列番号:102）

でEcoRI部位を含有した。1.6kbのPCRフラグメントをHindIII及びEcoRIで消化し、pCMV発現ベクターのHindIII-EcoRI部位中にクローン化した。その後、ヒトHB954の核酸（配列番号:103）及びアミノ酸（配列番号:104）配列を決定し、確認した。

【0091】

33. HG38（ジーンバンク受託番号：AF062006）

ヒトHG38のcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鋳型として脳cDNA及びrTthポリメラーゼ（Perkin Elmer）を製造業者により提供されるバッファー系、0.25 μ Mの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94℃で1分間；56℃で1分間；そして72℃で1分30秒間を30サイクルであった。5'及び3'フラグメントを別個に得るために2つのPCR反応を実施した。5'フラグメントのために、5' PCRプライマーは配列：

5'-CCCAAGCTTCGGGCACCATGGACACCTCCC-3'（配列番号:259）

でHindIII部位を含有し、そして3'プライマーは配列：

5'-ACAGGATCCAAATGCACAGCACTGGTAAGC-3'（配列番号:260）

でBamHI部位を含有した。この1.5 kbの5' PCRフラグメントをHindIII及びBamHIで消化し、pCMVのHindIII-BamHI部位中にクローン化した。3'フラグメントのために、5' PCRプライマーは配列：

ヒトOGR1のcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鋳型としてゲノムcDNA及びrTthポリメラーゼ (Perkin Elmer) を製造業者により提供されるバッフ

50

ァー系、0.25 μ Mの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94°Cで1分間；65°Cで1分間；そして72°Cで1分20秒間を30サイクルであった。5' PCRプライマーは配列：

5'-GGAAGCTTCAGGCCCCAAAGATGGGGAACAT-3' (配列番号:115)

でキナーゼ処理し、そして3'プライマーは配列：

5'-GTGGATCCACCCGCGGAGGCCAGGCTAG-3' (配列番号:116)

でBamHI部位を含有した。1.1kbのPCRフラグメントをBamHIで消化し、pCMV発現ベクターのEcoRV-BamHI部位中にクローン化した。その後、ヒトOGR1の核酸 (配列番号:117) 及びアミノ酸 (配列番号:118) 配列を決定し、確認した。

【0095】

10

37. セロトニン5HT_{2A}

内因性ヒト5HT_{2A}受容体をコードするcDNAをヒト脳ポリ-A⁺ RNA；XhoI制限部位を有する5'非翻訳領域からの5'プライマー：

5'-GACCTCGAGTCCTTCTACACCTCATC-3' (配列番号:119)

及びXbaI部位を含有する3'非翻訳領域からの3'プライマー：

5'-TGCTCTAGATTCCAGATAGGTGAAAACCTTG-3' (配列番号:120)

を用いてRT-PCRにより得た。TaqPllusTM Precisionポリメラーゼ (Stratagene) またはrTthTMポリメラーゼ (Perkin Elmer) のいずれかを製造業者により提供されるバッファー系、0.25 μ Mの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94°Cで1分間；57°Cで1分間；そして72°Cで2分間を30サイクルであった。1.5kbのPCRフラグメントをXbaIで消化し、pBluescriptのEcoRV-XbaI部位中にサブクローン化した。得られたcDNAクローンを完全にシークエンスし、公開配列から2個のアミノ酸変異をコードすることが分かった。第一のものはN末端細胞外ドメイン中のT25N突然変異であり；第二のものはH452Y突然変異である。2つの異なる市販の製造業者からのTaqポリメラーゼ (StratageneからのTaqPllusTM及びPerkin ElmerからのrTthTM) を用いて2つの独立したPCR反応から得られたcDNAクローンが同じ2つの突然変異を含有したので、これらの突然変異はPCRの誤りよりむしろ配列多型であると思われる。これらを除いて、その後、ヒト5HT_{2A}の核酸 (配列番号:121) 及びアミノ酸 (配列番号:122) 配列を決定し、確認した。

20

【0096】

30

38. セロトニン5HT_{2C}

内因性ヒト5HT_{2C}受容体をコードするcDNAをヒト脳ポリ-A⁺ RNAからRT-PCRにより得た。5'及び3'プライマーは5'及び3'非翻訳領域から得られ、以下の配列：

5'-GACCTCGAGGTTGCTTAAGACTGAAGC-3' (配列番号:123)

5'-ATTTCTAGACATATGTAGCTTGTACCG-3' (配列番号:124)

を含有した。その後、ヒト5HT_{2C}の核酸 (配列番号:125) 及びアミノ酸 (配列番号:126) 配列を決定し、確認した。

【0097】

39. V28 (ジーンバンク受託番号：U20350)

ヒトV28のcDNAを以下のように作製し、pCMV発現ベクター中にクローン化した：鋳型として脳cDNA及びrTthポリメラーゼ (Perkin Elmer) を製造業者により提供されるバッファー系、0.25 μ Mの各プライマー及び0.2mMの各4ヌクレオチドと共に用いてPCRを実施した。サイクル条件は94°Cで1分間；65°Cで1分間；そして72°Cで1分20秒間を30サイクルであった。5' PCRプライマーは配列：

40

5'-GGTAAGCTTGGCAGTCCACGCCAGGCCTTC-3' (配列番号:127)

でHindIII部位を含有し、そして3'プライマーは配列：

5'-TCCGAATTCTCTGTAGACACAAGGCTTTGG-3' (配列番号:128)

でEcoRI部位を含有した。1.1kbのPCRフラグメントをHindIII及びEcoRIで消化し、pCMV発現ベクターのHindIII-EcoRI部位中にクローン化した。その後、ヒトV28の核酸 (配列番号:129) 及びアミノ酸 (配列番号:130) 配列を決定し、確認した。

50

【0098】

実施例2

非内因性ヒトGPCRの調製

1. 部位特異的突然変異誘発

本明細書において開示するヒトGPCRプロリンマーカ法に基づく突然変異誘発をTransformer部位特異的突然変異誘発キット (Clontech) を製造業者の説明書に従って用いて前述の内因性ヒトGPCRに対して実施した。この突然変異誘発法のために、突然変異プローブ及び選択マーカプローブ（他に示さないかぎり、配列番号:132のプローブは終始同じであった）を利用し、特定の配列に対するこれらの配列を以下の表Bに記載する（挿入番号は配列番号である）。便宜上、ヒトGPCR中に含まれるコドン突然変異もまた標準的な形態で記載する：

【0099】

【表3】

表B

受容体識別名 (コドン突然変異)	突然変異プローブ配列 (5'-3') (配列番号)	選択マーカプローブ配列 (5'-3') (配列番号)
20 GPR1 (F245K)	GATCTCCAGTAGGCATA <u>AAGT</u> GGACAAATTCCTGG (131)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAG (132)
GPR4 (K2223A)	AGAAAGCCAAGATC <u>GCGCGG</u> CTGGCCCTCA (133)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
25 GPR5 (V224K)	CGGCGCCACCGCACG <u>AAAAA</u> GCTCATCTTC	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT

【0100】

【表4】

	(134)	
GPR7 (T250K)	GCCAAGAAGCGG <u>G</u> TGA <u>AAGTT</u> CCTGGTGGTGGCA (135)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
GPR8 (T259K)	CAGGCGGAAGGTGA <u>AAAGTCC</u> TGGTCCTCGT (136)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
30 GPR9 (M254K)	CGGCGCCTGCGGGCC <u>AAGCG</u> GCTGGTGGTGGTG (137)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
GPR9-6 (L241K)	CCAAGCACAAGGCC <u>AAGAAA</u> GTGACCATCAC (138)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
35 GPR10 (F276K)	GCGCCGGCGCACC <u>AAATGCT</u> TGCTGGTGGT (139)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
GPR15 (I240K)	CAAAAAGCTGAAGAAATCTA <u>AGAAGATCATCTTTATTGTCG</u> (140)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
GPR17 (V234K)	CAAGACCAAGGCA <u>AAACGCA</u> TGATCGCCAT (141)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
40 GPR18 (I231K)	GTCAAGGAGAAGTCC <u>AAAAG</u> GATCATCATC (142)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
GPR20 (M240K)	CGCCGCGTGCGGGCC <u>AAGCA</u> GCTCCTGCTC (143)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
45 GPR21 (A251K)	CCTGATAAGCGCTATA <u>AAAAT</u> GGTCCTGTTTCGA (144)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
GPR22 (F312K)	GAAAGACAAAAGAGAGTCA <u>AGAGGATGTCTTTATTG</u> (145)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
GPR24 (T304K)	CGGAGAAAAGAGGGTG <u>AAAC</u> GCACAGCCATCGCC (146)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
50 GPR30 (L258K)	別の方法;以下参照	別の方法;以下参照
GPR31 (Q221K)	AAGCTTCAGCGGGCC <u>AAGGC</u> ACTGGTCACC (147)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
55 GPR32 (K255A)	CATGCCAACC GGCCC <u>GCGAG</u> GCTGCTGCTGGT (279)	ACCAGCAGCAGCCTCGCGGG CCGGTTGGCATG (280)
GPR40 (A223K)	CGGAAGCTGCGGGCC <u>AAATG</u> GGTGGCCGGC (265)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
GPR41	CAGAGGAGGGTG <u>AAGGGGCT</u> GTTGGCG	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT

【 0 1 0 1 】

【 表 5 】

(A223K)	(266)	
GPR43 (V221K)	GGCGGCGCCGAGCC <u>AAGGGG</u> CTGGCTGTGG (267)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
APJ (L247K)	別の方法:以下参照	別の方法:以下参照
BLR1 (V258K)	CAGCGGCAGAAAGGCA <u>AAAA</u> GGGTGGCCATC (148)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
CEPR (L258K)	CGGCAGAAAGGCG <u>AAGCGCAT</u> GATCCTCGCG (149)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
EBI1 (L262K)	GAGCGCAACAAGGCC <u>AAAA</u> AGGTGATCATC (150)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
EBI2 (L243K)	GGTGTAACAAAAAGGCT <u>AA</u> <u>AAACACAATTATTC</u> TATT (151)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
ETBR-LP2 (N358K)	GAGAGCCAGCTC <u>AAGAGCAC</u> CGTGGTG (152)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
GHSR (V262K)	CCACAAGCAAACC <u>AAGAAAA</u> TGCTGGCTGT (153)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
GPCR-CNS (N491K)	CTAGAGAGTCAGATG <u>AAGTG</u> TACAGTAGTGGCAC (155)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
GPR-NGA (L275K)	CGGACAAAAGTGAAAACT <u>AA</u> <u>AAAGATGTTCCTCATT</u> (156)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
H9a and H9b (F236K)	GCTGAGGTTCGCAAT <u>AAACT</u> AACCATGTTTGTG (157)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
HB954 (H265K)	GGGAGGCCGAGCTG <u>AAAGCC</u> ACCCTGCTC (158)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
HG38 (V765K)	GGGACTGCTCTATG <u>AAAAAA</u> CACATTGCCCTG (268)	CATCAAGTGATCATGTGCC AAGTACGCCC (154)
HM74 (L230K)	CAAGATCAAGAGAGCC <u>AAAA</u> CCTTCATCATG (159)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
MIG (T273K)	CCGGAGACAAGTG <u>AAGAAG</u> ATGCTGTTTGTG (160)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
OGR1 (Q227K)	GCAAGGACCAGATC <u>AAGCGG</u> CTGGTGCTCA (161)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
セロニン5HT _{2A} (C322K)	別の方法:以下参照	別の方法:以下参照
セロニン5HT _{2C} (S310K)	別の方法:以下参照	別の方法:以下参照
V28 (L230K)	CAAGAAAGCCAAAGCC <u>AAG</u> AAACTGATCCTTCTG (162)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT

【0102】

次に非内因性ヒトGPCRをシーケンスし、得られた、確認された核酸及びアミノ酸配列を以下の表Cに要約するように本特許書類に添付の「配列表」付表中に記載する：

【0103】

【表6】

表C

	突然変異させたGPCR	核酸配列表	アミノ酸配列表
	GPR1 (F245K)	配列番号 : 163	配列番号 : 164
10	GPR4 (K223A)	配列番号 : 165	配列番号 : 166
	GPR5 (V224K)	配列番号 : 167	配列番号 : 168
	GPR7 (T250K)	配列番号 : 169	配列番号 : 170
15	GPR8 (T259K)	配列番号 : 171	配列番号 : 172
	GPR9 (M254K)	配列番号 : 173	配列番号 : 174
20	GPR9-6 (L241K)	配列番号 : 175	配列番号 : 176
	GPR10 (F276K)	配列番号 : 177	配列番号 : 178
	GPR15 (I240K)	配列番号 : 179	配列番号 : 180
25	GPR17 (V234K)	配列番号 : 181	配列番号 : 182
	GPR18 (I231K)	配列番号 : 183	配列番号 : 184
30	GPR20 (M240K)	配列番号 : 185	配列番号 : 186
	GPR21 (A251K)	配列番号 : 187	配列番号 : 188
	GPR22 (F312K)	配列番号 : 189	配列番号 : 190
35	GPR24 (T304K))	配列番号 : 191	配列番号 : 192
	GPR30	配列番号 : 193	配列番号 : 194

【 0 1 0 4 】

【表 7】

	(L258K)		
	GPR31 (Q221K)	配列番号 : 195	配列番号 : 196
5	GPR32 (K255A)	配列番号 : 269	配列番号 : 270
	GPR40 (A223K)	配列番号 : 271	配列番号 : 272
	GPR41 (A223K)	配列番号 : 273	配列番号 : 274
10	GPR43 (V221K)	配列番号 : 275	配列番号 : 276
	APJ (L247K)	配列番号 : 197	配列番号 : 198
15	BLR1 (V258K)	配列番号 : 199	配列番号 : 200
	CEPR (L258K)	配列番号 : 201	配列番号 : 202
	EBI1 (I262K)	配列番号 : 203	配列番号 : 204
20	EBI2 (L243K)	配列番号 : 205	配列番号 : 206
	ETBR-LP2 (N358K)	配列番号 : 207	配列番号 : 208
25	GHSR (V262K)	配列番号 : 209	配列番号 : 210
	GPCR-CNS (N491K)	配列番号 : 211	配列番号 : 212
	GPR-NGA (I275K)	配列番号 : 213	配列番号 : 214
30	H9a (F236K)	配列番号 : 215	配列番号 : 216
	H9b (F236K)	配列番号 : 217	配列番号 : 218
35	HB954 (H265K)	配列番号 : 219	配列番号 : 220
	HG38 (V765K)	配列番号 : 277	配列番号 : 278
	HM74 (I230K)	配列番号 : 221	配列番号 : 222
40	MIG (T273K)	配列番号 : 223	配列番号 : 224
	OGR1 (Q227K)	配列番号 : 225	配列番号 : 226
45	セロトニン5HT _{2A} (C322K)	配列番号 : 227	配列番号 : 228
	セロトニン5HT _{2C} (S310K)	配列番号 : 229	配列番号 : 230
	V28 (I230K)	配列番号 : 231	配列番号 : 232

【0105】

2. プロリンマーカールゴリズムの利用のための別の突然変異方法：APJ；セロトニン5HT_{2A}；セロトニン5HT_{2C}；及びGPR30

上記の部位特異的突然変異誘発法が特に好ましいが、そのような突然変異を作製するために他の方法を利用することができ；当業者は、当業者の特定の必要性に適合するGPCRを突然変異させる方法を容易に選択することができると考えられる。

【0106】

a. APJ

L247Kに突然変異させることにより非内因性ヒトAPJ受容体の調製を実施した。この突然変異を含有する2個のオリゴヌクレオチドを合成した：

5'-GGCTTAAGAGCATCATCGTGGTGCTGGTG-3' (配列番号:233)

5'-GTCACCACCAGCACCACGATGATGCTCTTAAGCC-3' (配列番号:234)

これら2個のオリゴヌクレオチドをアニーリングさせ、ヒト内因性APJのNaeI-BstEIIフラグメントを置き換えるために用いてヒトAPJの非内因性バージョンを作製した。

【0107】

b. セロトニン5HT_{2A}

アミノ酸322を包含する制限酵素部位SphIを利用することにより点突然変異C322Kを含有するcDNAを構築した。C322K突然変異を含有するプライマー：

5'-CAAAGAAAGTACTGGGCATCGTCTTCTTCCT-3' (配列番号:235)

を受容体の3'非翻訳領域からのプライマー：

5'-TGCTCTAGATTCCAGATAGGTGAAAACTTG-3' (配列番号:236)

と一緒に用いて (上記の条件下で) PCRを実施した。次に、T4ポリメラーゼで平滑にしたSphI部位を介して、得られたPCRフラグメントを用いて内因性5HT_{2A} cDNAの3'末端を置き換えた。

【0108】

c. セロトニン5HT_{2C}

アミノ酸310を含有するStyI制限フラグメントを所望する突然変異をコードする合成の二本鎖オリゴヌクレオチドで置き換えることによりS310K突然変異を含有するcDNAを構築した。利用したセンス鎖配列は以下の配列：

5'-CTAGGGGCACCATGCAGGCTATCAACAATGAAAGAAAAGCTAAGAAAGTC-3' (配列番号:237)

を有し、そして利用したアンチセンス鎖配列は以下の配列：

5'-CAAGGACTTTCTTAGCTTTTCTTTCATTGTTGATAGCCTGCATGGTGCCC-3' (配列番号:238)

を有した。

【0109】

d. GPR30

非内因性GPR30を作製する前に、PCRにより生じる突然変異を含まないクローンを同定するためにいくつかの独立したpCR2.1/GPR30単離体をそれらの全体をシークエンスした。突然変異を含まないクローンをEcoRIで消化し、そしてpCI-NeoをEcoRIで消化し、pCR2.1/GPR30からEcoRIで遊離したGPR30フラグメントをサブクローン化することにより内因性GPR30 cDNAフラグメントをCMVにより制御される発現プラスミドpCI-neo (Promega) 中に移し、pCI/GPR30を作製した。その後、製造業者の説明書に従ってQuick-ChangeTM 部位特異的突然変異誘発キット (Stratagene、#200518) 及び以下のプライマー：

5'-CGGCGGCAGAAGGCGAAACGCATGATCCTCGCGGT-3' (配列番号:239) 及び

5'-ACCGCGAGGATCATGCGTTTCGCCTTCTGCCGCCG-3' (配列番号:240)

を用いてコドン258のロイシンをリシンに突然変異させた。

【0110】

実施例 3

(内因性及び突然変異させた) 受容体発現

タンパク質の発現のために様々な細胞を当該技術分野は利用できるが、哺乳類細胞を利用することが最も好ましい。この主な理由は実用性に基づき、すなわち、GPCRの発現のための例えば酵母細胞の利用は、可能ではあるが、哺乳類系に対して進化した受容体共役、遺伝子機構及び分泌経路を含有しない可能性がある (実際、酵母の場合には含有しない) 非哺乳類細胞をプロトコル中に導入し、従って、非哺乳類細胞において得られた結果は、有用である可能性はあるが、哺乳類細胞から得られたものほど好ましくない。哺乳類細胞のうち、COS-7、293及び293T細胞が特に好ましいが、利用する特定の哺乳類細胞は当業者の特定の必要性に基づくことができる。

【0111】

本明細書において他に記載しないかぎり、以下のプロトコルを内因性及び非内因性ヒトGPCRの発現のために利用した。表Dは、GPCR発現のための哺乳類細胞及び (150mmプレート当たり) 利用した数を記載する。

10

20

30

40

50

【0112】

【表8】

表D

受容体名 (内因性または 非内因性)	哺乳類細胞 (利用した数)
GPR17	293 (2 x 10 ⁴)
GPR30	293 (4 x 10 ⁴)
APJ	COS-7 (5X10 ⁶)
ETBR-LP2	293 (1 x 10 ⁷)
	293T (1 x 10 ⁷)
GHSR	293 (1 x 10 ⁷)
	293T (1 x 10 ⁷)
MIG	293 (1 x 10 ⁷)
セロニン5HT _{2A}	293T (1 x 10 ⁷)
セロニン5HT _{2c}	293T (1 x 10 ⁷)

10

20

【0113】

1日目に、哺乳類細胞を平板培養した。2日目に、2本の反応チューブを調製し（各チューブの以下の割合はプレート当たりである）：チューブAは、20 μgのDNA（例えば、pCMVベクター；内因性受容体cDNAを有するpCMVベクター及び非内因性受容体cDNAを有するpCMVベクター）を1.2mlの無血清DMEM（Irvine Scientific, Irvine, CA）中に混合することにより調製し；チューブBは、120 μlのリポフェクトアミン（lipofectamine）（Gibco BRL）を1.2mlの無血清DMEM中に混合することにより調製した。次に、チューブA及びBを反転（数回）により混合し、続いて室温で30-45分間インキュベーションした。この混合物を「トランスフェクション混合物」と称する。平板培養した細胞を1X PBSで洗浄し、続いて10 mlの無血清DMEMを加えた。次に、2.4mlのトランスフェクション混合物をこれらの細胞に加え、続いて37℃/5% CO₂で4時間インキュベーションした。次に、トランスフェクション混合物を吸引により除き、続いて25mlのDMEM/10%ウシ胎仔血清を加えた。次に、細胞を37℃/5% CO₂でインキュベーションした。72時間のインキュベーション後、次に細胞を集め、分析のために利用した。

30

1. Gi共役受容体：Gs共役受容体との共トランスフェクション

GPR30の場合、この受容体はGタンパク質Giに共役することが決定されている。Giは、cAMPへのATPの転化を触媒するために必要である酵素アデニルシクラーゼを阻害することが知られている。従って、GPR30の非内因性の構成的に活性化された形態は、cAMPの減少したレベルと関連すると予想される。cAMPの減少するレベルの測定により直接GPR30の非内因性の構成的に活性化された形態をアッセイ確認することは、実行可能であるが、好ましくは、Gs共役受容体の共同使用により測定することができる。例えば、Gsに共役する受容体はアデニルシクラーゼを刺激し、従って、cAMPの増加と関連する。本願の譲受人は、オーファン受容体GPR6が内因性の構成的に活性化されたGPCRであることを発見している。GPR6はGsタンパク質に共役する。従って、共トランスフェクションすると、推定されるGPR30-突然変異がその構成的活性化を導くことを容易に確かめることができ、すなわち、内因性の構成的に活性化されたGPR6/内因性の構成的に活性化されていないGPR30細胞は、内因性の構成的に活性のあるGPR6/非内因性の構成的に活性化されたGPR30と比較した場合にcAMPの増大したレベルを示す（後者は比較的より低いレベルのcAMPを示す）。候補化合物

40

50

がGs共役受容体に対する例えば逆アゴニスト（すなわち、そのような化合物はcAMPのレベルを下げると思われる）またはGi共役受容体（もしくはGo共役受容体）に対する例えば逆アゴニスト（すなわち、そのような候補化合物はcAMPのレベルを上げるとみられる）であるかどうかを決定するためにcAMPを検出するアッセイを利用することができる。cAMPを測定するための当該技術分野において既知の様々な方法を利用することができ；好ましい方法は抗-cAMP抗体の使用による。別の方法で、最も好ましいものは、全細胞第二メッセンジャーレポーター系アッセイを利用する。遺伝子上のプロモーターは、特定の遺伝子がコードするタンパク質の発現を導く。サイクリックAMPは、cAMP応答DNA結合タンパク質または転写因子（CREB）の結合を促進することにより遺伝子発現を導き、それは次にcAMP応答配列と呼ばれる特定の位置でプロモーターに結合し、遺伝子の発現を導く。レポーター遺伝子、例えば β -ガラクトシダーゼまたはルシフェラーゼの前に多数のcAMP応答配列を含む有するプロモーターを有するレポーター系を構築することができる。従って、GPR6のような活性化された受容体は、cAMPの蓄積をもたらし、これは次にレポータータンパク質の遺伝子及び発現を活性化する。最も好ましくは、293細胞を好ましくは1:1の比率の、最も好ましくは1:4の比率のGPR6（または別のGs共役受容体）及びGPR30（または別のGi共役受容体）プラスミドで共トランスフェクションする。GPR6はcAMPの生成を刺激する内因性の構成的に活性のある受容体であるので、GPR6はレポーター遺伝子及びその発現を強力に活性化する。次に、標準的な生化学アッセイを用いて β -ガラクトシダーゼまたはルシフェラーゼのようなレポータータンパク質を検出することができる（Chen et al. 1995）。内因性の構成的に活性でないGPR30と内因性の構成的に活性のあるGPR6との共トランスフェクションは、ルシフェラーゼレポータータンパク質の増加を示す。逆に、非内因性の構成的に活性のあるGPR30と内因性の構成的に活性のあるGPR6との共トランスフェクションは、ルシフェラーゼの発現の劇的な減少を示す。いくつかのレポータープラスミドが既知であり、第二メッセンジャーアッセイを測定するために当該技術分野において利用できる。主として当業者の特定の必要性に基づいて特定の遺伝子発現のために適当なレポータープラスミドを決定することは当業者の十分に範囲内であると考えられる。様々な細胞が発現のために利用できるが、哺乳類細胞が最も好ましく、そしてこれらのタイプのうち、293細胞が最も好ましい。哺乳類トランスフェクションTMキット（Stratagene、#200285）CaPO₄沈殿プロトコルを製造業者の説明書に従って用いて293細胞をレポータープラスミドpCRE-Luc/GPR6及び非内因性の構成的に活性化されたGPR30でトランスフェクションした（公開された内因性GPR6配列については28 Genomics 347(1995)を参照）。この沈殿物は、400ngの受容体、80ngのCMV-発現プラスミド（1:4のGPR6:内因性GPR30または非内因性GPR30比を有する）及び20ngのCMV-SEAP（分泌されるアルカリホスファターゼをコードするトランスフェクションコントロールプラスミド）を含有した。沈殿物の50%を（4 X 10⁴細胞/ウェルを含有する）96穴組織培養皿の3ウェル中に分け；残りの50%を捨てた。翌朝、培地を交換した。トランスフェクションの開始後48時間で細胞を溶解し、LuciteTMキット（Packard、カタログ#6016911）並びにTrilux 1450 MicrobetaTM液体シンチレーション及び発光カウンター（Wallac）を販売業者の説明書どおりに用いてルシフェラーゼ活性に関して調べた。GraphPad Prism 2.0a（GraphPad Software Inc.）を用いてデータを分析した。

【0114】

Giに共役することが同様に決定されているGPR17に関しては、とりわけ、別のGs共役内因性受容体GPR3（23 Genomics 609(1994)及び24 Genomics 391(1994)を参照）の使用に基づき、前述の方法の改変を利用した。最も好ましくは293細胞を利用する。これらの細胞を96穴プレート上にウェル当たり2 X 10⁴細胞の密度で平板培養し、翌日リポフェクトアミン試薬（BRL）を製造業者の説明書に従って用いてトランスフェクションした。DNA/脂質混合物を各6ウェルトランスフェクションに対して以下のように調製した：100 μ lのDMEM中260ngのプラスミドDNAを100 μ lのDMEM中2 μ lの脂質と穏やかに混合した（260ngのプラスミドDNAは、200ngの8xCRE-Lucレポータープラスミド（以下参照）、50ngの内因性受容体もしくは非内因性受容体を含んでなるpCMVまたはpCMVのみ及び10ngのGPRS発現プラスミド（

pcDNA3 (Invitrogen) 中のGPRS) からなった)。8xCRE-Lucレポータープラスミドは以下のように調製した：p β gal-Basicベクター(Clontech)中のBglV-HindIII部位でラットソマトスタチンプロモーター (-71/+51) をクローン化することによりベクター-SRIF- β -galを得た。8コピーのcAMP応答配列をアデノウイルス鋳型AdpCF126CCRE8(7 Human Gene Therapy 1883(1996)を参照)からPCRにより得、Kpn-BglV部位でSRIF- β -galベクター中にクローン化し、8xCRE- β -galレポーターベクターをもたらしした。8xCRE-Lucレポータープラスミドは、8xCRE- β -galレポーターベクター中の β -ガラクトシダーゼ遺伝子をHindIII-BamHI部位でpGL3-basicベクター (Promega) から得たルシフェラーゼ遺伝子で置き換えることにより作製した。室温で30分のインキュベーション後、DNA/脂質混合物を400 μ lのDMEMで希釈し、希釈した混合物のうち100 μ lを各ウェルに加えた。細胞培養インキュベーター中で4時間のインキュベーション後に10%のFCSを含む100 μ lのDMEMを各ウェルに加えた。翌朝、トランスフェクションした細胞を10%のFCSを含む200 μ l/ウェルのDMEMで交換した。8時間後に、PBSで1回洗浄した後、フェノールレッドを含まない100 μ l/ウェルのDMEMにウェルを交換した。翌日、LucLiteTMレポーター遺伝子アッセイキット (Packard) を製造業者の説明書に従って用いてルシフェラーゼ活性を測定し、1450 MicroBetaTMシンチレーション及び発光カウンター (Wallac) で読み取った。

【0115】

図4は、構成的に活性のあるGPR30が293細胞においてGPR6によりもたらされるCRE-Lucレポーターの活性化を阻害することを示す。ルシフェラーゼは、発現ベクターpCMVにおいて約4.1相対光単位で測定された。内因性GPR30はルシフェラーゼを約8.5相対光単位で発現し、一方、非内因性の構成的に活性のあるGPR30 (L258K) はルシフェラーゼをそれぞれ約3.8及び3.1相対光単位で発現した。1:4の比率で、内因性GPR6と内因性GPR30との共トランスフェクションは、ルシフェラーゼ発現を約104.1相対光単位まで劇的に増加した。同じ比率で、内因性GPR6と非内因性GPR30 (L258K) との共トランスフェクションは、発現を劇的に減少し、これはそれぞれ約18.2及び29.5相対光単位で明白である。図5に記載するように、GPR3との共トランスフェクションに関してGPR17について同様の結果が認められた。

【0116】

実施例3

非内因性GPCRの構成的活性を測定するためのアッセイ

A. 膜結合アッセイ

1. [³⁵S] GTP γ Sアッセイ

Gタンパク質共役受容体がある場合、リガンド結合または構成的活性化のいずれかの結果として、該受容体はGタンパク質に共役し、GDPの遊離及びそれに続くGタンパク質へのGTPの結合を刺激する。Gタンパク質-受容体複合体の α サブユニットはGTPアーゼとして働き、GTPをGDPにゆっくり加水分解し、この時点で受容体は通常非活性化される。構成的に活性化された受容体は、GDPをGTPに交換し続ける。加水分解できないGTP類似体 [³⁵S] GTP γ Sを利用して、構成的に活性化された受容体を発現する膜への [³⁵S] GTP γ Sの増大した結合を示すことができる。構成的活性化を測定するために [³⁵S] GTP γ S結合を用いる利点は：(a) それが全てのGタンパク質共役受容体に一般的に適用できること；(b) それが膜表面で近接し、細胞内カスケードに影響を及ぼす分子を捕らえる可能性を低くすることである。

【0117】

該アッセイは、問題とされる受容体を発現する膜への [³⁵S] GTP γ S結合を刺激するGタンパク質共役受容体の能力を利用する。従って、該アッセイは、既知の、オーファン及び構成的に活性化されたGタンパク質共役受容体に対して候補化合物をスクリーニングする直接同定法において用いることができる。該アッセイは一般的であり、全てのGタンパク質共役受容体での薬剤発見に適応性を有する。

【0118】

[³⁵S] GTP γ Sアッセイを約0.3~約1.2nMの間の [³⁵S] GTP γ S (1.2が好ましいが、結果の

最適化のためにこの量を調整することができる)及び12.5~75 µgの膜タンパク質(例えば、受容体を発現するCOS-7細胞;75 µgが好ましいが、最適化のためにこの量を調整することができる)及び1 µMのGDP(最適化のためにこの量を変えることができる)と共に20mM HEPES及び1~約20mMの間のMgCl₂(20mMが好ましいが、結果の最適化のためにこの量を調整することができる)pH7.4、結合バッファー中で1時間インキュベーションした。次に、コムギ胚芽凝集素ビーズ(25 µl; Amersham)を加え、混合物を室温でさらに30分間インキュベーションした。次に、チューブを室温で1500 x gで5分間遠心分離し、次にシンチレーションカウンターで計数した。

【0119】

大規模スクリーニングの要求を同様に満たす、かかる費用がより少ないが同等に適用できる代案が同定されている。Flash platesTM及びWallacTMシンチストリップ(scintistrip)を高処理量 [³⁵S] GTP γ S結合アッセイを構成するために利用することができる。さらに、この技術を用いて、アッセイを [³⁵S] GTP γ S結合による効能をモニターすると同時に受容体へのトリチウム化したりガンド結合を同時にモニターするために既知のGPCRに対して利用することができる。Wallac β カウンターは、トリチウムで標識したプローブ及び³⁵Sで標識したプローブの両方を調べるためにエネルギーウィンドウを切り換えることができるのでこれが可能である。このアッセイはまた、受容体活性化をもたらす他のタイプの膜活性化事象を検出するために用いることもできる。例えば、様々な受容体(Gタンパク質共役受容体及びチロシンキナーゼ受容体の両方)の³²Pリン酸化をモニターするために該アッセイを用いることができる。膜がウェルの底に遠心分離されると、結合した [³⁵S] GTP γ Sまたは³²Pでリン酸化された受容体は、ウェルに被覆されているシンチラント(scintillant)を活性化する。Scinti^Rストリップ(Wallac)は、この原理を実証するために用いられている。さらに、該アッセイはまた、放射性標識したりガンドを用いて受容体へのリガンド結合を測定するためにも有用性を有する。同様に、放射性標識した結合したりガンドがウェルの底に遠心分離されると、シンチストリップ標識は放射性標識したりガンドと近接するようになり、活性化及び検出をもたらす。

【0120】

前述のプロトコルに基づいて、コントロール(pCMV)、内因性APJ及び非内因性APJを比較するグラフの代表的な結果を図6に記載する。

【0121】

2. アデニルシクラーゼ

細胞に基づくアッセイのために設計されたFlash PlateTMアデニルシクラーゼキット(New England Nuclear;カタログ番号SMP004A)を未精製の細胞膜での使用のために改変した。Flash Plateウェルは、cAMPを認識する特定の抗体も含有するシンチラント被覆を含有した。ウェル中に生成したcAMPをcAMP抗体への放射性cAMPトレーサーの結合に対する直接競合により定量した。以下のものは、受容体を発現する膜におけるcAMPレベルの変化の測定のための簡潔なプロトコルとして使える。

【0122】

トランスフェクション後約3日目にトランスフェクションした細胞を集めた。20mM HEPES、pH 7.4及び10mM MgCl₂を含有するバッファー中での懸濁した細胞の均質化により膜を調製した。Brinkman PolytronTMを約10秒間用いて均質化を氷上で実施した。得られたホモジネートを4℃で49,000 X gで15分間遠心分離した。得られたペレットを次に20mM HEPES、pH 7.4及び0.1mM EDTAを含有するバッファー中に再懸濁し、10秒間均質化し、続いて4℃で49,000 X gで15分間遠心分離した。得られたペレットを利用するまで-80℃で保存することができる。測定の日、膜ペレットを室温でゆっくり解凍し、20mM HEPES、pH 7.4及び10mM MgCl₂(本明細書に記載する値が好ましいが、これらの量を最適化することができる)を含有するバッファー中に再懸濁し、0.60mg/mlの最終タンパク質濃度をもたらした(再懸濁した膜は使用するまで氷上に置いた)。

【0123】

cAMP基準及び(11mlの検出バッファーに対して2 µCiのトレーサー [¹²⁵I] cAMP (100 µl

10

20

30

40

50

）を含んでなる）検出バッファーを調製し、製造業者の説明書に従って維持した。アッセイバッファーをスクリーニングのために新しく調製し、それは20mM HEPES、pH 7.4、10mM MgCl_2 、20mM (Sigma)、0.1ユニット/mlクレアチンホスホキナーゼ (Sigma)、50 μM GT P (Sigma) 及び0.2mM ATP (Sigma) を含有し；アッセイバッファーは、利用するまで氷上で保存することができる。NEN Flashプレートに50 μl のアッセイバッファーを加え、続いて50 μl の膜懸濁液を加えることによりアッセイを開始した。得られたアッセイ混合物を室温で60分間インキュベーションし、続いて100 μl の検出バッファーを加える。次に、プレートをさらに2-4時間インキュベーションし、続いてWallac MicroBetaシンチレーションカウンターで計数する。各アッセイプレート内に含まれる基準cAMP曲線からcAMP/ウェルの値を外挿する。前述のアッセイをMIGの分析に関して利用した。

10

【0124】

B. レポーターに基づくアッセイ

1. CREBレポーターアッセイ (Gs共役受容体)

Gs刺激を検出する方法は、cAMPに依存して活性化される転写因子CREBの既知の特性による。293または293T細胞においてGs共役活性をアッセイするためにPathDetect CREB trans-Reporting System (Stratagene、カタログ#219010) を利用した。哺乳類トランスフェクションキット (Stratagene、カタログ#200285) を製造業者の説明書に従って用いて細胞をこの上記の系のプラスミド成分及び内因性または突然変異体受容体をコードする示した発現プラスミドでトランスフェクションした。簡潔に言えば、400ngのpFR-Luc (Gal4認識配列を含有するルシフェラーゼレポータープラスミド)、40ngのpFA2-CREB (Gal4 DNA結合ドメインを含有するGal4-CREB融合タンパク質)、80ngの(受容体を含んでなる) CMV-受容体発現プラスミド及び20ngのCMV-SEAP (分泌されるアルカリホスファターゼ発現プラスミド；アルカリホスファターゼ活性は、サンプル間のトランスフェクション効率の変動を制御するためにトランスフェクションした細胞の培地において測定される) をリン酸カルシウム沈殿物中にキットの説明書どおりに合わせた。沈殿物の半分を96穴プレート中の3ウェル上に等しく分配し、細胞上で一晩保ち、翌朝新しい培地で置き換えた。トランスフェクションの開始後48時間で、上記GPR30系に関して記述したように細胞を処置し、ルシフェラーゼ活性についてアッセイした。このアッセイをGHSRについて用いた。

20

【0125】

2. AP1レポーターアッセイ (Gq共役受容体)

Gq刺激を検出する方法は、プロモーター中にAP1要素を含有する遺伝子の活性化をもたらすGq依存性ホスホリパーゼCの既知の特性による。リン酸カルシウム沈殿物の成分が410ngのpAP1-Luc、80ngの受容体発現プラスミド及び20ngのCMV-SEAPであったことを除いて、CREBレポーターアッセイに関して上に記述したプロトコルに従ってPathdetect AP-1 cis-Reporting System (Stratagene、カタログ#219073) を利用した。このアッセイをETBR-LP2について用いた。

30

【0126】

C. 細胞内IP3蓄積アッセイ

1日目に、(内因性及び突然変異させた)セロトニン受容体を含んでなる細胞を24穴プレート上に通常 1×10^5 細胞/ウェル平板培養した。2日目に、まず第一に50 μl の無血清DMEM/ウェル中0.25 μg のDNAと50 μl の無血清DMEM/ウェル中2 μl のリポフェクトアミンを混合することにより細胞をトランスフェクションした。これらの溶液を穏やかに混合し、室温で15-30分間インキュベーションした。細胞を0.5mlのPBSで洗浄し、400 μl の無血清培地をトランスフェクション培地と混合し、細胞に加えた。次に、細胞を37°C/5% CO_2 で3-4時間インキュベーションし、次にトランスフェクション培地を除き、1ml/ウェルの通常の増殖培地で置き換えた。3日目に細胞を ^3H -ミオ-イノシトールで標識した。簡潔に言えば、培地を除き、細胞を0.5mlのPBSで洗浄した。次に、0.5mlのイノシトールを含まない/無血清培地 (GIBCO BRL) /ウェルを0.25 μCi の ^3H -ミオ-イノシトール/ウェルと共に加え、細胞を37°C/5% CO_2 で16-18時間インキュベーションした。4日目に細胞を0.5mlのPBSで洗浄し、イノシトールを含まない/無血清培地、10 μM のパルジリン、10mMの塩化リチウム

40

50

を含有する0.45mlのアッセイ培地または0.4mlのアッセイ培地と50 μ lの10xケタンセリン (ket) を10 μ Mの最終濃度に加えた。次に、細胞を37°Cで30分間インキュベーションした。次に、細胞を0.5mlのPBSで洗浄し、200 μ lの新しい/よく冷えた停止溶液 (1M KOH; 18mM ホウ酸Na; 3.8mM EDTA) /ウェルを加えた。この溶液を5-10分間または細胞が溶解されるまで氷上に保ち、次に200 μ lの新しい/よく冷えた中和溶液 (7.5% HCL) により中和した。次に、溶解産物を1.5mlエッペンドルフチューブ中に移し、1mlのクロロホルム/メタノール (1:2) /ウェルを加えた。この溶液を15秒間ボルテックスし、上相をBiorad AG1-X8陰イオン交換樹脂 (100-200メッシュ) に添加した。まず第一に、樹脂を1:1.25 W/Vで水で洗浄し、0.9mlの上相をカラム上に載せた。カラムを10mlの5mMミオ-イノシトール及び10mlの5mMホウ酸Na/60mMギ酸Naで洗浄した。イノシトールトリスリン酸塩を2mlの0.1Mギ酸/1Mギ酸アンモニウムで10mlのシンチレーションカクテルを含有するシンチレーションバイアル中に溶出した。カラムを10mlの0.1Mギ酸/3Mギ酸アンモニウムで洗浄することにより再生し、dd H₂Oで2回すすぎ、水中4°Cで保存した。

10

【0127】

図7は、C322K突然変異を含むヒト5-HT_{2A}受容体からのIP3生成の実例を示す。これらの結果は、プロリン突然変異アルゴリズム法がこの受容体を構成的に活性化することを示すが、可能性がある治療法の同定のためのスクリーニングにそのような受容体を用いる目的のためには、さらに確固たる差が好ましい。しかしながら、活性化された受容体は、構成的活性化の役割を理解し解明するため及びさらに調べることができる化合物の同定のために利用することができるので、この差はそれ自体ヒト5HT_{2A}受容体の内因性及び非内因性バージョンを区別することにおいて有用であると考ええる。

20

【0128】**D. 結果の要約**

試験したGPCRの結果を表Eに記載し、ここで、増加%は、内因性GPCRに比較した場合に非内因性GPCRに対して認められた結果の%差を示し; これらの値の後には利用したアッセイのタイプに関する挿入表示が続く。さらに、利用したアッセイ系を挿入して記載する (そして異なる宿主細胞を用いた場合には両方を記載する)。これらの結果が示すように、ヒトGPCRの非内因性バージョンの構成的活性を測定するために様々なアッセイを利用することができる。当業者は、前述の事項に基づきそして当該技術分野が利用できる情報を参照して、研究者の特定の必要性に適合する特定のアッセイ方法を選択すること及び/または最大限に活用することができると考えられる。

30

【0129】**【表9】**

表E

受容体識別名 (コドン突然変異)	%差
GPR17 (V234K)	74.5 (CRE-Luc)
GPR30 (L258K)	71.6 (CREB)
APJ (L247K)	49.0 (GTP γ S)
ETBR-LP2 (N358K)	48.4(AP1-Luc - 293) 61.1(AP1-Luc - 293T)

GHSR (V262K)	58.9(CREB - 293) 35.6(CREB - 293T)
-----------------	---------------------------------------

MIG (I230K)	39 (cAMP)
セロトニン5HT _{2A} (C322K)	33.2 (IP ₃)
セロトニン5HT _{2C} (S310K)	39.1(IP ₃)

【0130】

実施例6

内因性オーファンGPCRの組織分布

市販されているヒト組織ドットプロット形態を用いて、内因性オーファンGPCRをそのような受容体が局在する領域を決定するためにプローブで調べた。以下に示す場合を除いて、（放射性標識した）全受容体cDNAをプローブとして用い：Prime-It II™ランダムプライマーラベリングキット（Stratagene、#300385）を製造業者の説明書に従って使用して（ベクターから切り出した）完全な受容体cDNAを用いて放射性標識したプローブを作製した。ヒトRNA Master Blot™（Clontech、#7770-1）にGPCR放射性標識プローブをハイブリダイズさせ、製造業者の説明書に従ってストリンジентな条件下で洗浄した。プロットをKodak BioMaXオートラジオグラフィフィルムに-80℃で一晩感光させた。

【0131】

代表的なドットプロット形態結果を図8にGPR1（8A）、GPR30（8B）及びAPJ（8C）について示し、全ての受容体について結果を表Fに要約する。

【0132】

【表10】

表F

GPCR	組織分布 (ドットプロットにおいて他の組織に対して 最も高いレベル)
GPR1	胎盤, 卵巣, 副腎
GPR4	広範, 心臓, 肺, 副腎, 甲状腺, 脊髄において, 最も高い
GPR5	胎盤, 胸腺, 胎児胸腺, 脾臓, 胎児脾臓において, より少ないレベル
GPR7	肝臓, 脾臓, 脊髄, 胎盤
GPR8	いかなる発現も検出されない
GPR9-6	胸腺, 胎児胸腺, 小腸において, より少ないレベル
GPR18	脾臓, リンパ節, 胎児脾臓, 精巣
GPR20	広範
GPR21	広範; 非常に少量
GPR22	心臓, 胎児心臓, 脳において, より少ないレベル
GPR30	胃
GPR31	広範
BLR1	脾臓
CEPR	胃, 肝臓, 甲状腺, 被殻
EBI1	脾臓 リンパ組織において, より少ないレベル
EBI2	リンパ組織, 大動脈, 肺, 脊髄
ETBR-LP2	広範; 脳組織
GPCR-CNS	脳 精巣, 胎盤において, より少ないレベル
GPR-NGA	下垂体 脳において, より少ないレベル
H9	下垂体
HB954	大動脈, 小脳 大部分の他の組織において, より少ないレベル
HM74	脾臓, 白血球, 骨髄, 乳腺, 肺, 気管
MIG	腎臓, 肝臓, 脾臓, 肺, 脾臓において低レベル
ORG1	下垂体, 胃, 胎盤
V28	脳, 脾臓, 末梢白血球

10

20

30

【0133】

40

前述の情報に基づき、ヒトGPCRを罹病組織における分布に関しても評価することができ；次に「正常」及び罹病組織間の比較評価を利用して罹病状態における特定の受容体の過剰発現または不十分な発現の可能性を決定することができることを記す。可能性がある治療関連性の候補化合物を直接同定するスクリーニングの目的のためにヒトGPCRの非内因性バージョンを利用することが望ましい状況において、特定のヒトGPCRが過剰発現される疾病及び疾患の処置では逆アゴニストが有用であり、一方、特定のヒトGPCRが不十分に発現される疾病及び疾患の処置ではアゴニストまたは部分アゴニストが有用であることを記す。

【0134】

所望に応じて、目的の受容体が発現されるこれらの組織内の特定の細胞を同定するために、当業者に周知の技術（例えばインサイチュハイブリダイゼーション）を用いて、受容

50

体のより詳細な細胞位置決定を利用することができる。

【0135】

本特許書類中に挙げる特許、出願及び発行された (printed) 公開の各々は、これらを全部引用することにより本明細書に組み込まれるものとする。

【0136】

当業者が認識するように、本発明の精神からそれずに本発明の好ましい態様に対して多数の変更及び改変を行うことができる。全てのそのような変形は本発明の範囲内に入るものとする。

【0137】

内因性及び非内因性ヒトGPCRの両方に利用する目的のために、様々な発現ベクターを当業者は利用できるが、利用するベクターがpCMVであることが最も好ましい。このベクターは、特許手続き上の微生物の寄託の国際的承認に関するブダペスト条約の条項により1998年10月13日にAmerican Type Culture Collection(ATCC)(10801 University Blvd., Manassas, VA 20110-2209 USA)に寄託されている。このベクターは1998年にATCCにより試験され、1998年に生育できると確認された。ATCCはpCMVに以下の寄託番号： を付与している。

10

【図面の簡単な説明】

【図1】 膜貫通ヘリックス、細胞内ループ及び細胞外ループに番号を付与したGタンパク質共役受容体の一般化構造を示す。

【図2】 典型的なGタンパク質共役受容体の活性及び不活性の2つの状態並びに第二メッセンジャー伝達経路への活性状態の連結を図式的に示す。

20

【図3】 制限酵素部位の位置を含む、好ましいベクターpCMVの配列表である。

【図4】 pCMV、CRE-LucレポーターのGPR6によりもたらされる活性化の非内因性の構成的に活性のあるGPR30阻害をCRE-LucレポーターのGPR6によりもたらされる活性化の内因性GPR30阻害と比較する測定されたシグナルの図示である。

【図5】 pCMV、CRE-LucレポーターのGPR3によりもたらされる活性化の非内因性の構成的に活性化されたGPR17阻害をCRE-LucレポーターのGPR3によりもたらされる活性化の内因性GPR17阻害と比較する測定されたシグナルの図示である。

【図6】 コントロールのpCMV、内因性APJ及び非内因性APJを比較する測定されたシグナルの図式結果を示す。

30

【図7】 非内因性ヒト5-HT_{2A}受容体からのIP₃生成をこの受容体の内因性バージョンと比較した場合の実例を示す。

【図8】 GPR1(8A)、GPR30(8B)及びAPJ(8C)のドットプロット形態の結果である。

【配列表】

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Behan, Dominic P.
Chalmers, Derek T.
Liaw, Chen W.
- (ii) TITLE OF INVENTION: Non-Endogenous, Constitutively
Activated Human G Protein-Coupled
Orphan Receptors
- (iii) NUMBER OF SEQUENCES: 280
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Arena Pharmaceuticals, Inc.
(B) STREET: 6166 Nancy Ridge Drive
(C) CITY: San Diego
(D) STATE: CA
(E) COUNTRY: USA
(F) ZIP: 92122
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US
(B) FILING DATE:
(C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Burgoon, Richard P.
(B) REGISTRATION NUMBER: 34,787
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (619)453-7200
(B) TELEFAX: (619)453-7210

10

20

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1068 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

30

ATGGAAGATT TGGAGGAAAC ATTATTTGAA GAATTTGAAA ACTATTCCTA TGACCTAGAC 60

TATTACTCTC TGGAGTCTGA TTTGGAGGAG AAAGTCCAGC TGGGAGTTGT TCACTGGGTC 120
 TCCCTGGTGT TATATTGTTT GGCTTTTGT CTGGGAATTC CAGGAAATGC CATCGTCATT 180
 TGGTTACGGG GGCTCAAGTG GAAGAAGACA GTCACCACTC TGTGGTTCCT CAATCTAGCC 240
 ATTGCGGATT TCATTTTTCT TCTCTTCTG CCCCTGTACA TCTCCTATGT GGCCATGAAT 300
 TTCCACTGGC CCTTTGGCAT CTGGCTGTGC AAAGCCAATT CCTTCACTGC CCAGTTGAAC 360
 ATGTTTGCCA GTGTTTTTTT CCTGACAGTG ATCAGCCTGG ACCACTATAT CCACTTGATC 420
 CATCCTGTCT TATCTCATCG GCATCGAACC CTCAAGAACT CTCTGATTGT CATTATATTC 480
 ATCTGGCTTT TGGCTTCTCT AATTGGCGGT CCTGCCCTGT ACTTCCGGGA CACTGTGGAG 540
 TTCAATAATC ATACTCTTTG CTATAACAAT TTTCAGAAGC ATGATCCTGA CCTCACTTTG 600
 ATCAGGCACC ATGTTCTGAC TTGGGTGAAA TTTATCATTG GCTATCTCTT CCCTTTGCTA 660
 ACAATGAGTA TTTGCTACTT GTGTCTCATC TTCAAGGTGA AGAAGCGAAC AGTCCTGATC 720
 TCCAGTAGGC ATTTCTGGAC AATTCTGGTT GTGGTTGTGG CCTTTGTGGT TTGCTGGACT 780
 CCTTATCACC TGTTTAGCAT TTGGGAGCTC ACCATTCACT ACAATAGCTA TTCCCACCAT 840
 GTGATGCAGG CTGGAATCCC CCTCTCCACT GGTTTGGCAT TCCTCAATAG TTGCTTGAAC 900
 CCCATCCTTT ATGTCCTAAT TAGTAAGAAG TTCCAAGCTC GCTTCCGGTC CTCAGTTGCT 960
 GAGATACTCA AGTACACACT GTGGGAAGTC AGCTGTTCTG GCACAGTGAG TGAACAGCTC 1020
 AGGAACTCAG AAACCAAGAA TCTGTGTCTC CTGGAAACAG CTCAATAA 1068

10

(3) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 355 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Glu Asp Leu Glu Glu Thr Leu Phe Glu Glu Phe Glu Asn Tyr Ser
 1 5 10 15
 Tyr Asp Leu Asp Tyr Tyr Ser Leu Glu Ser Asp Leu Glu Glu Lys Val
 20 25 30
 Gln Leu Gly Val Val His Trp Val Ser Leu Val Leu Tyr Cys Leu Ala
 35 40 45

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

10

20

30

340

345

350

Thr Ala Gln
355

(4) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1089 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATGGGCAACC ACACGTGGGA GGGCTGCCAC GTGGACTCGC GCGTGGACCA CCTCTTTCCG 60
CCATCCCTCT ACATCTTTGT CATCGGCGTG GGGCTGCCCA CCAACTGCCT GGCTCTGTGG 120
GCGGCCTACC GCCAGGTGCA ACAGCGCAAC GAGCTGGGCG TCTACCTGAT GAACCTCAGC 180
ATCGCCGACC TGCTGTACAT CTGCACGCTG CCGCTGTGGG TGGACTACTT CCTGCACCAC 240
GACAACTGGA TCCACGGCCC CGGGTCCTGC AAGCTCTTTG GGTTCATCTT CTACACCAAT 300
ATCTACATCA GCATCGCCTT CCTGTGCTGC ATCTCGGTGG ACCGCTACCT GGCTGTGGCC 360
CACCCACTCC GCTTCGCCCC CCTGCGCCGC GTCAAGACCG CCGTGGCCGT GAGCTCCGTG 420
GTCTGGGCCA CGGAGCTGGG CGCCAACTCG GCGCCCCGTG TCCATGACGA GCTCTTCCGA 480
GACCGCTACA ACCACACCTT CTGCTTTGAG AAGTTCCCCA TGAAGGCTG GGTGGCCTGG 540
ATGAACCTCT ATCGGGTGTT CGTGGGCTTC CTCTTCCCGT GGGCGCTCAT GCTGCTGTCTG 600
TACCGGGGCA TCCTGCGGGC CGTGCGGGGC AGCGTGTCCA CCGAGCGCCA GGAGAAGGCC 660
AAGATCAAGC GGCTGGCCCT CAGCCTCATC GCCATCGTGC TGGTCTGCTT TGCGCCCTAT 720
CACGTGCTCT TGCTGTCCCG CAGCGCCATC TACCTGGGCC GCCCCTGGGA CTGCGGCTTC 780
GAGGAGCGCG TCTTTTCTGC ATACCACAGC TCACTGGCTT TCACCAGCCT CAACTGTGTG 840
GCGGACCCCA TCCTCTACTG CCTGGTCAAC GAGGGCGCCC GCAGCGATGT GGCCAAGGCC 900
CTGCACAACC TGCTCCGCTT TCTGGCCAGC GACAAGCCCC AGGAGATGGC CAATGCCTCG 960
CTCACCTTGG AGACCCCACT CACCTCCAAG AGGAACAGCA CAGCCAAAGC CATGACTGGC 1020
AGCTGGGCGG CCACTCCGCC TTCCCAGGGG GACCAGGTGC AGCTGAAGAT GCTGCCGCCA 1080
GCACAATGA 1089

20

30

(5) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 362 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```

Met Gly Asn His Thr Trp Glu Gly Cys His Val Asp Ser Arg Val Asp
1           5           10           15
His Leu Phe Pro Pro Ser Leu Tyr Ile Phe Val Ile Gly Val Gly Leu
20          25          30
Pro Thr Asn Cys Leu Ala Leu Trp Ala Ala Tyr Arg Gln Val Gln Gln
35          40          45
Arg Asn Glu Leu Gly Val Tyr Leu Met Asn Leu Ser Ile Ala Asp Leu
50          55          60
Leu Tyr Ile Cys Thr Leu Pro Leu Trp Val Asp Tyr Phe Leu His His
65          70          75          80
Asp Asn Trp Ile His Gly Pro Gly Ser Cys Lys Leu Phe Gly Phe Ile
85          90          95
Phe Tyr Thr Asn Ile Tyr Ile Ser Ile Ala Phe Leu Cys Cys Ile Ser
100         105         110
Val Asp Arg Tyr Leu Ala Val Ala His Pro Leu Arg Phe Ala Arg Leu
115         120         125
Arg Arg Val Lys Thr Ala Val Ala Val Ser Ser Val Val Trp Ala Thr
130         135         140
Glu Leu Gly Ala Asn Ser Ala Pro Leu Phe His Asp Glu Leu Phe Arg
145         150         155         160
Asp Arg Tyr Asn His Thr Phe Cys Phe Glu Lys Phe Pro Met Glu Gly
165         170         175
Trp Val Ala Trp Met Asn Leu Tyr Arg Val Phe Val Gly Phe Leu Phe
180         185         190
Pro Trp Ala Leu Met Leu Leu Ser Tyr Arg Gly Ile Leu Arg Ala Val
195         200         205
Arg Gly Ser Val Ser Thr Glu Arg Gln Glu Lys Ala Lys Ile Lys Arg
210         215         220
Leu Ala Leu Ser Leu Ile Ala Ile Val Leu Val Cys Phe Ala Pro Tyr

```

10

20

30

225 230 235 240
 His Val Leu Leu Leu Ser Arg Ser Ala Ile Tyr Leu Gly Arg Pro Trp
 245 250 255
 Asp Cys Gly Phe Glu Glu Arg Val Phe Ser Ala Tyr His Ser Ser Leu
 260 265 270
 Ala Phe Thr Ser Leu Asn Cys Val Ala Asp Pro Ile Leu Tyr Cys Leu
 275 280 285
 Val Asn Glu Gly Ala Arg Ser Asp Val Ala Lys Ala Leu His Asn Leu
 290 295 300
 Leu Arg Phe Leu Ala Ser Asp Lys Pro Gln Glu Met Ala Asn Ala Ser
 305 310 315 320
 Leu Thr Leu Glu Thr Pro Leu Thr Ser Lys Arg Asn Ser Thr Ala Lys
 325 330 335
 Ala Met Thr Gly Ser Trp Ala Ala Thr Pro Pro Ser Gln Gly Asp Gln
 340 345 350
 Val Gln Leu Lys Met Leu Pro Pro Ala Gln
 355 360

10

(6) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TATGAATTCA GATGCTCTAA ACGTCCCTGC

30

(7) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TCCGGATCCA CCTGCACCTG CGCCTGCACC

30

(8) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1002 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

ATGGAGTCCT CAGGCAACCC AGAGAGCACC ACCTTTTTTT ACTATGACCT TCAGAGCCAG	60
CCGTGTGAGA ACCAGGCCTG GGTCTTTGCT ACCCTCGCCA CCACTGTCCT GTACTGCCTG	120
GTGTTTCTCC TCAGCCTAGT GGGCAACAGC CTGGTCCTGT GGGTCCTGGT GAAGTATGAG	180
AGCCTGGAGT CCCTCACCAA CATCTTCATC CTCAACCTGT GCCTCTCAGA CCTGGTGTTC	240
GCCTGCTTGT TGCCTGTGTG GATCTCCCCA TACCACTGGG GCTGGGTGCT GGGAGACTTC	300
CTCTGCAAAC TCCTCAATAT GATCTTCTCC ATCAGCCTCT ACAGCAGCAT CTTCTTCCTG	360
ACCATCATGA CCATCCACCG CTACCTGTCG GTAGTGAGCC CCCTCTCCAC CCTGCGCGTC	420
CCCACCTCC GCTGCCGGT GCTGGTGACC ATGGCTGTGT GGGTAGCCAG CATCCTGTCC	480
TCCATCCTCG ACACCATCTT CCACAAGGTG CTTTCTTCGG GCTGTGATTA TTCCGAACTC	540
ACGTGGTACC TCACCTCCGT CTACCAGCAC AACCTCTTCT TCCTGCTGTC CCTGGGGATT	600
ATCCTGTTCT GCTACGTGGA GATCCTCAGG ACCCTGTTCC GCTCACGCTC CAAGCGGCGC	660
CACCGCACGG TCAAGCTCAT CTTGCCATC GTGGTGGCCT ACTTCTCAG CTGGGGTCCC	720
TACAACTTCA CCCTGTTTCT GCAGACGCTG TTTCCGACCC AGATCATCCG GAGCTGCGAG	780
GCCAAACAGC AGCTAGAATA CGCCCTGCTC ATCTGCCGCA ACCTCGCCTT CTCCCCTGC	840
TGCTTTAACC CGGTGCTCTA TGTCTTCGTG GGGGTCAAGT TCCGCACACA CCTGAAACAT	900
GTTCTCCGGC AGTTCTGGTT CTGCCGGCTG CAGGCACCCA GCCCAGCCTC GATCCCCCAC	960
TCCCCTGGTG CTTTCGCCTA TGAGGGCGCC TCCTTCTACT GA	1002

10

20

30

(9) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 333 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

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

10

20

30

Phe Val Gly Val Lys Phe Arg Thr His Leu Lys His Val Leu Arg Gln
 290 295 300

Phe Trp Phe Cys Arg Leu Gln Ala Pro Ser Pro Ala Ser Ile Pro His
 305 310 315 320

Ser Pro Gly Ala Phe Ala Tyr Glu Gly Ala Ser Phe Tyr
 325 330

(10) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GCAAGCTTGG GGGACGCCAG GTCGCCGGCT 30

(11) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GCGGATCCGG ACGCTGGGGG AGTCAGGCTG C 31

(12) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 987 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ATGGACAACG CCTCGTTCTC GGAGCCCTGG CCCGCCAACG CATCGGGCCC GGACCCGGCG 60

CTGAGCTGCT CCAACGCGTC GACTCTGGCG CCGCTGCCGG CGCCGCTGGC GGTGGCTGTA 120

CCAGTTGTCT ACGCGGTGAT CTGCGCCGTG GGTCTGGCGG GCAACTCCGC CGTGCTGTAC 180

GTGTTGCTGC GGGCGCCCCG CATGAAGACC GTCACCAACC TGTTCATCCT CAACCTGGCC 240
 ATCGCCGACG AGCTCTTCAC GCTGGTGCTG CCCATCAACA TCGCCGACTT CCTGCTGCGG 300
 CAGTGGCCCT TCGGGGAGCT CATGTGCAAG CTCATCGTGG CTATCGACCA GTACAACACC 360
 TTCTCCAGCC TCTACTTCCT CACCGTCATG AGCGCCGACC GCTACCTGGT GGTGTTGGCC 420
 ACTGCGGAGT CGCGCCGGGT GGCCGGCCGC ACCTACAGCG CCGCGCGCGC GGTGAGCCTG 480
 GCCGTGTGGG GGATCGTCAC ACTCGTCGTG CTGCCCTTCG CAGTCTTCGC CCGGCTAGAC 540
 GACGAGCAGG GCCGGCGCCA GTGCGTGCTA GTCTTTCCGC AGCCCGAGGC CTTCTGGTGG 600
 CGCGCGAGCC GCCTCTACAC GCTCGTGCTG GGCTTCGCCA TCCCGGTGTC CACCATCTGT 660
 GTCCTCTATA CCACCTGCT GTGCCGGCTG CATGCCATGC GGCTGGACAG CCACGCCAAG 720
 GCCCTGGAGC GCGCCAAGAA GCGGGTGACC TTCCTGGTGG TGGCAATCCT GCGGTGTGC 780
 CTCCTCTGCT GGACGCCCTA CCACCTGAGC ACCGTGGTGG CGCTCACCAC CGACCTCCCG 840
 CAGACGCCGC TGGTCATCGC TATCTCCTAC TTCATCACCA GCCTGACGTA CGCCAACAGC 900
 TGCCTCAACC CTTCTCTCTA CGCCTTCCTG GACGCCAGCT TCCGCAGGAA CCTCCGCCAG 960
 CTGATAACTT GCCGCGCGGC AGCCTGA 987

(13) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 328 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

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

10

20

30

Ile Ala Asp Glu Leu Phe Thr Leu Val Leu Pro Ile Asn Ile Ala Asp
 85 90 95
 Phe Leu Leu Arg Gln Trp Pro Phe Gly Glu Leu Met Cys Lys Leu Ile
 100 105 110
 Val Ala Ile Asp Gln Tyr Asn Thr Phe Ser Ser Leu Tyr Phe Leu Thr
 115 120 125
 Val Met Ser Ala Asp Arg Tyr Leu Val Val Leu Ala Thr Ala Glu Ser
 130 135 140
 Arg Arg Val Ala Gly Arg Thr Tyr Ser Ala Ala Arg Ala Val Ser Leu
 145 150 155 160
 Ala Val Trp Gly Ile Val Thr Leu Val Val Leu Pro Phe Ala Val Phe
 165 170 175
 Ala Arg Leu Asp Asp Glu Gln Gly Arg Arg Gln Cys Val Leu Val Phe
 180 185 190
 Pro Gln Pro Glu Ala Phe Trp Trp Arg Ala Ser Arg Leu Tyr Thr Leu
 195 200 205
 Val Leu Gly Phe Ala Ile Pro Val Ser Thr Ile Cys Val Leu Tyr Thr
 210 215 220
 Thr Leu Leu Cys Arg Leu His Ala Met Arg Leu Asp Ser His Ala Lys
 225 230 235 240
 Ala Leu Glu Arg Ala Lys Lys Arg Val Thr Phe Leu Val Val Ala Ile
 245 250 255
 Leu Ala Val Cys Leu Leu Cys Trp Thr Pro Tyr His Leu Ser Thr Val
 260 265 270
 Val Ala Leu Thr Thr Asp Leu Pro Gln Thr Pro Leu Val Ile Ala Ile
 275 280 285
 Ser Tyr Phe Ile Thr Ser Leu Thr Tyr Ala Asn Ser Cys Leu Asn Pro
 290 295 300
 Phe Leu Tyr Ala Phe Leu Asp Ala Ser Phe Arg Arg Asn Leu Arg Gln
 305 310 315 320
 Leu Ile Thr Cys Arg Ala Ala Ala
 325

10

20

30

(14) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CGGAATTCGT CAACGGTCCC AGCTACAATG

30

(15) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

ATGGATCCCA GGCCCTTCAG CACCGCAATA T

31

(16) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1002 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

ATGCAGGCCG CTGGGCACCC AGAGCCCCTT GACAGCAGGG GCTCCTTCTC CCTCCCCACG 60
ATGGGTGCCA ACGTCTCTCA GGACAATGGC ACTGGCCACA ATGCCACCTT CTCCGAGCCA 120
CTGCCGTTCC TCTATGTGCT CCTGCCCCGCC GTGTACTCCG GGATCTGTGC TGTGGGGCTG 180
ACTGGCAACA CGGCCGTCAT CCTTGTAATC CTAAGGGCGC CCAAGATGAA GACGGTGACC 240
AACGTGTTCA TCCTGAACCT GGCCGTCGCC GACGGGCTCT TCACGCTGGT ACTGCCCCGC 300
AACATCGCGG AGCACCTGCT GCAGTACTGG CCCTTCGGGG AGCTGCTCTG CAAGCTGGTG 360
CTGGCCGTCG ACCACTACAA CATCTTCTCC AGCATCTACT TCCTAGCCGT GATGAGCGTG 420
GACCGATACC TGGTGGTGCT GGCCACCGTG AGGTCCCGCC ACATGCCCTG GCGCACCTAC 480
CGGGGGGCGA AGGTGCGCCAG CCTGTGTGTC TGGCTGGGCG TCACGGTCCT GGTTCGCCCC 540
TTCTTCTCTT TCGCTGGCGT CTACAGCAAC GAGCTGCAGG TCCCAAGCTG TGGGCTGAGC 600
TTCCCGTGGC CCGAGCGGGT CTGGTTCAAG GCCAGCCGTG TCTACACTTT GGTCTTGGGC 660
TTCGTGCTGC CCGTGTGCAC CATCTGTGTG CTCTACACAG ACCTCCTGCG CAGGCTGCGG 720

10

20

30

```

GCCGTGCGGC TCCGCTCTGG AGCCAAGGCT CTAGGCAAGG CCAGGCGGAA GGTGACCGTC   780
CTGGTCCTCG TCGTGCTGGC CGTGTGCCTC CTCTGCTGGA CGCCCTTCCA CCTGGCCTCT   840
GTCGTGGGCC TGACCACGGA CCTGCCCCAG ACCCCACTGG TCATCAGTAT GTCCTACGTC   900
ATCACCAGCC TCACGTACGC CAACTCGTGC CTGAACCCCT TCCTCTACGC CTTTCTAGAT   960
GACAACTTCC GGAAGAACTT CCGCAGCATA TTGCGGTGCT GA                      1002

```

(17) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 333 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

10

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

```

Met Gln Ala Ala Gly His Pro Glu Pro Leu Asp Ser Arg Gly Ser Phe
1           5           10           15

Ser Leu Pro Thr Met Gly Ala Asn Val Ser Gln Asp Asn Gly Thr Gly
          20           25           30

His Asn Ala Thr Phe Ser Glu Pro Leu Pro Phe Leu Tyr Val Leu Leu
          35           40           45

Pro Ala Val Tyr Ser Gly Ile Cys Ala Val Gly Leu Thr Gly Asn Thr
          50           55           60

Ala Val Ile Leu Val Ile Leu Arg Ala Pro Lys Met Lys Thr Val Thr
          65           70           75           80

Asn Val Phe Ile Leu Asn Leu Ala Val Ala Asp Gly Leu Phe Thr Leu
          85           90           95

Val Leu Pro Val Asn Ile Ala Glu His Leu Leu Gln Tyr Trp Pro Phe
          100          105          110

Gly Glu Leu Leu Cys Lys Leu Val Leu Ala Val Asp His Tyr Asn Ile
          115          120          125

Phe Ser Ser Ile Tyr Phe Leu Ala Val Met Ser Val Asp Arg Tyr Leu
          130          135          140

Val Val Leu Ala Thr Val Arg Ser Arg His Met Pro Trp Arg Thr Tyr
          145          150          155          160

Arg Gly Ala Lys Val Ala Ser Leu Cys Val Trp Leu Gly Val Thr Val
          165          170          175

```

20

30

Leu Val Leu Pro Phe Phe Ser Phe Ala Gly Val Tyr Ser Asn Glu Leu
 180 185 190
 Gln Val Pro Ser Cys Gly Leu Ser Phe Pro Trp Pro Glu Arg Val Trp
 195 200 205
 Phe Lys Ala Ser Arg Val Tyr Thr Leu Val Leu Gly Phe Val Leu Pro
 210 215 220
 Val Cys Thr Ile Cys Val Leu Tyr Thr Asp Leu Leu Arg Arg Leu Arg
 225 230 235 240
 Ala Val Arg Leu Arg Ser Gly Ala Lys Ala Leu Gly Lys Ala Arg Arg
 245 250 255
 Lys Val Thr Val Leu Val Leu Val Val Leu Ala Val Cys Leu Leu Cys
 260 265 270
 Trp Thr Pro Phe His Leu Ala Ser Val Val Ala Leu Thr Thr Asp Leu
 275 280 285
 Pro Gln Thr Pro Leu Val Ile Ser Met Ser Tyr Val Ile Thr Ser Leu
 290 295 300
 Thr Tyr Ala Asn Ser Cys Leu Asn Pro Phe Leu Tyr Ala Phe Leu Asp
 305 310 315 320
 Asp Asn Phe Arg Lys Asn Phe Arg Ser Ile Leu Arg Cys
 325 330

10

20

(18) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 48 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

ACGAATTCAG CCATGGTCCT TGAGGTGAGT GACCACCAAG TGCTAAAT

48

30

(19) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 27 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GAGGATCCTG GAATGCGGGG AAGTCAG

27

(20) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1107 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

ATGGTCCTTG AGGTGAGTGA CCACCAAGTG CTAAATGACG CCGAGGTTGC CGCCCTCCTG 60
 GAGAACTTCA GCTCTTCCTA TGAATATGGA GAAAACGAGA GTGACTCGTG CTGTACCTCC 120
 CCGCCCTGCC CACAGGACTT CAGCCTGAAC TTCGACCGGG CCTTCCTGCC AGCCCTCTAC 180
 AGCCTCCTCT TTCTGCTGGG GCTGCTGGGC AACGGCGCGG TGGCAGCCGT GCTGCTGAGC 240
 CGGCGGACAG CCCTGAGCAG CACCGACACC TTCCTGCTCC ACCTAGCTGT AGCAGACACG 300
 CTGCTGGTGC TGACACTGCC GCTCTGGGCA GTGGACGCTG CCGTCCAGTG GGTCTTTGGC 360
 TCTGGCCTCT GCAAAGTGGC AGGTGCCCTC TTCAACATCA ACTTCTACGC AGGAGCCCTC 420
 CTGCTGGCCT GCATCAGCTT TGACCGCTAC CTGAACATAG TTCATGCCAC CCAGCTCTAC 480
 CGCCGGGGGC CCGCGGCCCG CGTGACCCTC ACCTGCCTGG CTGTCTGGGG GCTCTGCCTG 540
 CTTTTCGCCC TCCAGACTT CATCTTCCTG TCGGCCACAC ACGACGAGCG CCTCAACGCC 600
 ACCCACTGCC AATACAACCT CCCACAGGTG GGCCGCACGG CTCTGCGGGT GCTGCAGCTG 660
 GTGGCTGGCT TTCTGCTGCC CCTGCTGGTC ATGGCCTACT GCTATGCCCA CATCCTGGCC 720
 GTGCTGCTGG TTTCCAGGGG CCAGCGGCGC CTGCGGGCCA TGCGGCTGGT GGTGGTGGTC 780
 GTGGTGGCCT TTGCCCTCTG CTGGACCCCC TATCACCTGG TGGTGTGGT GGACATCCTC 840
 ATGGACCTGG GCGCTTTGGC CCGCAACTGT GGCCGAGAAA GCAGGGTAGA CGTGGCCAAG 900
 TCGGTCACCT CAGGCTGGG CTACATGCAC TGCTGCCTCA ACCCGCTGCT CTATGCCTTT 960
 GTAGGGGTCA AGTTCCGGGA GCGGATGTGG ATGCTGCTCT TGCGCCTGGG CTGCCCCAAC 1020
 CAGAGAGGGC TCCAGAGGCA GCCATCGTCT TCCCGCCGGG ATTCATCCTG GTCTGAGACC 1080
 TCAGAGGCCT CCTACTCGGG CTTGTGA 1107

10

20

30

(21) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 368 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

```

Met Val Leu Glu Val Ser Asp His Gln Val Leu Asn Asp Ala Glu Val
1           5           10           15

Ala Ala Leu Leu Glu Asn Phe Ser Ser Ser Tyr Asp Tyr Gly Glu Asn
           20           25           30

Glu Ser Asp Ser Cys Cys Thr Ser Pro Pro Cys Pro Gln Asp Phe Ser
           35           40           45

Leu Asn Phe Asp Arg Ala Phe Leu Pro Ala Leu Tyr Ser Leu Leu Phe
           50           55           60

Leu Leu Gly Leu Leu Gly Asn Gly Ala Val Ala Ala Val Leu Leu Ser
           65           70           75           80

Arg Arg Thr Ala Leu Ser Ser Thr Asp Thr Phe Leu Leu His Leu Ala
           85           90           95

Val Ala Asp Thr Leu Leu Val Leu Thr Leu Pro Leu Trp Ala Val Asp
           100          105          110

Ala Ala Val Gln Trp Val Phe Gly Ser Gly Leu Cys Lys Val Ala Gly
           115          120          125

Ala Leu Phe Asn Ile Asn Phe Tyr Ala Gly Ala Leu Leu Leu Ala Cys
           130          135          140

Ile Ser Phe Asp Arg Tyr Leu Asn Ile Val His Ala Thr Gln Leu Tyr
           145          150          155          160

Arg Arg Gly Pro Pro Ala Arg Val Thr Leu Thr Cys Leu Ala Val Trp
           165          170          175

Gly Leu Cys Leu Leu Phe Ala Leu Pro Asp Phe Ile Phe Leu Ser Ala
           180          185          190

His His Asp Glu Arg Leu Asn Ala Thr His Cys Gln Tyr Asn Phe Pro
           195          200          205

Gln Val Gly Arg Thr Ala Leu Arg Val Leu Gln Leu Val Ala Gly Phe
           210          215          220

Leu Leu Pro Leu Leu Val Met Ala Tyr Cys Tyr Ala His Ile Leu Ala
           225          230          235          240

```

10

20

30

Val Leu Leu Val Ser Arg Gly Gln Arg Arg Leu Arg Ala Met Arg Leu
 245 250 255

Val Val Val Val Val Val Ala Phe Ala Leu Cys Trp Thr Pro Tyr His
 260 265 270

Leu Val Val Leu Val Asp Ile Leu Met Asp Leu Gly Ala Leu Ala Arg
 275 280 285

Asn Cys Gly Arg Glu Ser Arg Val Asp Val Ala Lys Ser Val Thr Ser
 290 295 300

Gly Leu Gly Tyr Met His Cys Cys Leu Asn Pro Leu Leu Tyr Ala Phe
 305 310 315 320

Val Gly Val Lys Phe Arg Glu Arg Met Trp Met Leu Leu Leu Arg Leu
 325 330 335

Gly Cys Pro Asn Gln Arg Gly Leu Gln Arg Gln Pro Ser Ser Ser Arg
 340 345 350

Arg Asp Ser Ser Trp Ser Glu Thr Ser Glu Ala Ser Tyr Ser Gly Leu
 355 360 365

10

(22) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TTAAGCTTGA CCTAATGCCA TCTTGTGTCC

30

(23) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TTGGATCCAA AAGAACCATG CACCTCAGAG

30

(24) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1074 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

ATGGCTGATG ACTATGGCTC TGAATCCACA TCTTCCATGG AAGACTACGT TAACTTCAAC	60
TTCACGTACT TCTACTGTGA GAAAAACAAT GTCAGGCAGT TTGCGAGCCA TTTCCTCCCA	120
CCCTTGTA CT GGCTCGTGT CATCGTGGGT GCCTTGGGCA ACAGTCTTGT TATCCTTGTC	180
TACTGGTACT GCACAAGAGT GAAGACCATG ACCGACATGT TCCTTTTGAA TTTGGCAATT	240
GCTGACCTCC TCTTTCTTGT CACTCTTCCC TTCTGGGCCA TTGCTGCTGC TGACCAGTGG	300
AAGTTCAG CTTTCATGTG CAAGGTGGTC AACAGCATGT ACAAGATGAA CTTCTACAGC	360
TGTGTGTG TGATCATGTG CATCAGCCTG GACAGGTACA TTGCCATTGC CCAGGCCATG	420
AGAGCACATA CTTGGAGGGA GAAAAGGCTT TTGTACAGCA AAATGGTTTG CTTTACCATC	480
TGGGTATTGG CAGCTGCTCT CTGCATCCCA GAAATCTTAT ACAGCCAAAT CAAGGAGGAA	540
TCCGGCATTG CTATCTGCAC CATGGTTTAC CCTAGCGATG AGAGCACCAA ACTGAAGTCA	600
GCTGTCTTGA CCCTGAAGGT CATTCTGGGG TTCTTCCTTC CCTTCGTGGT CATGGCTTGC	660
TGCTATACCA TCATCATTCA CACCCTGATA CAAGCCAAGA AGTCTTCCAA GCACAAAGCC	720
CTAAAAGTGA CCATCACTGT CCTGACCGTC TTTGTCTTGT CTCAGTTTCC CTACAACCTG	780
ATTTTGTGG TGCAGACCAT TGACGCCTAT GCCATGTTCA TCTCCAACCTG TGCCGTTTCC	840
ACCAACATTG ACATCTGCTT CCAGGTCACC CAGACCATCG CCTTCTTCCA CAGTTGCCTG	900
AACCCTGTTC TCTATGTTTT TGTGGGTGAG AGATTCCGCC GGGATCTCGT GAAAACCCTG	960
AAGAAGCTGG GTTGCACTAG CCAGGCCAG TGGGTTTCAT TTACAAGGAG AGAGGGAAGC	1020
TTGAAGCTGT CGTCTATGTT GCTGGAGACA ACCTCAGGAG CACTCTCCCT CTGA	1074

10

20

30

(25) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 357 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

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

30

Val Thr Gln Thr Ile Ala Phe Phe His Ser Cys Leu Asn Pro Val Leu
 290 295 300

Tyr Val Phe Val Gly Glu Arg Phe Arg Arg Asp Leu Val Lys Thr Leu
 305 310 315 320

Lys Asn Leu Gly Cys Ile Ser Gln Ala Gln Trp Val Ser Phe Thr Arg
 325 330 335

Arg Glu Gly Ser Leu Lys Leu Ser Ser Met Leu Leu Glu Thr Thr Ser
 340 345 350

Gly Ala Leu Ser Leu
 355

10

(26) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1110 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

ATGGCCTCAT CGACCACTCG GGGCCCCAGG GTTCTGACT TATTTCTGG GCTGCCGCCG 60

GCGGTCACAA CTCCCGCCAA CCAGAGCGCA GAGGCCTCGG CGGGCAACGG GTCGGTGGCT 120

GGCGCGGACG CTCCAGCCGT CACGCCCTTC CAGAGCCTGC AGCTGGTGCA TCAGCTGAAG 180

GGGCTGATCG TGCTGCTCTA CAGCGTCGTG GTGGTCTGGG GGCTGGTGGG CAACTGCCTG 240

CTGGTGCTGG TGATCGCGCG GGTGCCGCGG CTGCACAACG TGACGAACTT CCTCATCGGC 300

AACCTGGCCT TGTCCGACGT GCTCATGTGC ACCGCCTGCG TGCCGCTCAC GCTGGCCTAT 360

GCCTTCGAGC CACGCGGCTG GGTGTTCCGG GCGGGCCTGT GCCACCTGGT CTTCTTCCTG 420

CAGCCGGTCA CCGTCTATGT GTCGGTGTTC ACGCTCACCA CCATCGCAGT GGACCGCTAC 480

GTCGTGCTGG TGCACCCGCT GAGGCGCGCA TCTCGCTGCG CCTCAGCCTA CGCTGTGCTG 540

GCCATCTGGG CGCTGTCCGC GGTGCTGGCG CTGCCGCCCG CCGTGCACAC CTATCACGTG 600

GAGCTCAAGC CGCACGACGT GCGCCTCTGC GAGGAGTTCT GGGGCTCCCA GGAGCGCCAG 660

CGCCAGCTCT ACGCCTGGGG GCTGCTGCTG GTCACCTACC TGCTCCCTCT GCTGGTCATC 720

CTCCTGTCTT ACGTCCGGGT GTCAGTGAAG CTCCGCAACC GCGTGGTGCC GGGCTGCGTG 780

ACCCAGAGCC AGGCCGACTG GGACCGCGCT CGGCGCCGGC GCACCTTCTG CTTGCTGGTG 840

20

30

GTGGTCGTGG TGGTGTTCGC CGTCTGCTGG CTGCCGCTGC ACGTCTTCAA CCTGCTGCGG 900
 GACCTCGACC CCCACGCCAT CGACCCTTAC GCCTTTGGGC TGGTGCAGCT GCTCTGCCAC 960
 TGGCTCGCCA TGAGTTCGGC CTGCTACAAC CCCTTCATCT ACGCTGGCT GCACGACAGC 1020
 TTCCGCGAGG AGCTGCGCAA ACTGTTGGTC GCTTGGCCCC GCAAGATAGC CCCCCATGGC 1080
 CAGAATATGA CCGTCAGCGT GGTCATCTGA 1110

(27) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 369 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

10

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

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

20

30

Tyr Ala Val Leu Ala Ile Trp Ala Leu Ser Ala Val Leu Ala Leu Pro
 180 185 190
 Pro Ala Val His Thr Tyr His Val Glu Leu Lys Pro His Asp Val Arg
 195 200 205
 Leu Cys Glu Glu Phe Trp Gly Ser Gln Glu Arg Gln Arg Gln Leu Tyr
 210 215 220
 Ala Trp Gly Leu Leu Leu Val Thr Tyr Leu Leu Pro Leu Leu Val Ile
 225 230 235 240
 Leu Leu Ser Tyr Val Arg Val Ser Val Lys Leu Arg Asn Arg Val Val
 245 250 255
 Pro Gly Cys Val Thr Gln Ser Gln Ala Asp Trp Asp Arg Ala Arg Arg
 260 265 270
 Arg Arg Thr Phe Cys Leu Leu Val Val Val Val Val Phe Ala Val
 275 280 285
 Cys Trp Leu Pro Leu His Val Phe Asn Leu Leu Arg Asp Leu Asp Pro
 290 295 300
 His Ala Ile Asp Pro Tyr Ala Phe Gly Leu Val Gln Leu Leu Cys His
 305 310 315 320
 Trp Leu Ala Met Ser Ser Ala Cys Tyr Asn Pro Phe Ile Tyr Ala Trp
 325 330 335
 Leu His Asp Ser Phe Arg Glu Glu Leu Arg Lys Leu Leu Val Ala Trp
 340 345 350
 Pro Arg Lys Ile Ala Pro His Gly Gln Asn Met Thr Val Ser Val Val
 355 360 365
 Ile

10

20

(28) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1083 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

ATGGACCCAG AAGAACTTC AGTTTATTG GATTATTACT ATGCTACGAG CCCAACTCT 60
 GACATCAGGG AGACCCACTC CCATGTTCTT TACACCTCTG TCTTCCTTCC AGTCTTTTAC 120

ACAGCTGTGT TCCTGACTGG AGTGCTGGGG AACCTTGTTT TCATGGGAGC GTTGCATTTT 180
 AAACCCGGCA GCCGAAGACT GATCGACATC TTTATCATCA ATCTGGCTGC CTCTGACTTC 240
 ATTTTCTTGT TCACATTGCC TCTCTGGGTG GATAAAGAAG CATCTCTAGG ACTGTGGAGG 300
 ACGGGCTCCT TCCTGTGCAA AGGGAGCTCC TACATGATCT CCGTCAATAT GCACTGCAGT 360
 GTCCTCCTGC TCACTTGCAAT GAGTGTGAC CGCTACCTGG CCATTGTGTG GCCAGTCGTA 420
 TCCAGGAAAT TCAGAAGGAC AGACTGTGCA TATGTAGTCT GTGCCAGCAT CTGGTTTATC 480
 TCCTGCCTGC TGGGGTTGCC TACTCTTCTG TCCAGGGAGC TCACGCTGAT TGATGATAAG 540
 CCATACTGTG CAGAGAAAAA GGCAACTCCA ATTAACTCA TATGGTCCCT GGTGGCCTTA 600
 ATTTTCACCT TTTTGTCCC TTTGTTGAGC ATTGTGACCT GCTACTGTTG CATTGCAAGG 660
 AAGCTGTGTG CCCATTACCA GCAATCAGGA AAGCACAACA AAAAGCTGAA GAAATCTATA 720
 AAGATCATCT TTATTGTCGT GGCAGCCTTT CTTGTCTCCT GGCTGCCCTT CAATACTTTC 780
 AAGTTCCTGG CCATTGTCTC TGGGTTGCGG CAAGAACACT ATTTACCCTC AGCTATTCTT 840
 CAGCTTGGA TGGAGGTGAG TGGACCCTTG GCATTTGCCA ACAGCTGTGT CAACCCTTTC 900
 ATTTACTATA TCTTCGACAG CTACATCCGC CGGGCCATTG TCCACTGCTT GTGCCCTTGC 960
 CTGAAAAACT ATGACTTTGG GAGTAGCACT GAGACATCAG ATAGTCACCT CACTAAGGCT 1020
 CTCTCCACCT TCATTCATGC AGAAGATTTT GCCAGGAGGA GGAAGAGGTC TGTGTCACTC 1080
 TAA 1083

10

(29) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 360 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Met Asp Pro Glu Glu Thr Ser Val Tyr Leu Asp Tyr Tyr Tyr Ala Thr
 1 5 10 15
 Ser Pro Asn Ser Asp Ile Arg Glu Thr His Ser His Val Pro Tyr Thr
 20 25 30
 Ser Val Phe Leu Pro Val Phe Tyr Thr Ala Val Phe Leu Thr Gly Val
 35 40 45

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

10

20

30

340

345

350

Arg Arg Lys Arg Ser Val Ser Leu
355 360

(30) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CTAGAATTCT GACTCCAGCC AAAGCATGAA T

31

(31) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GCTGGATCCT AAACAGTCTG CGCTCGGCCT

30

(32) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1020 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

30

ATGAATGGCC TTGAAGTGGC TCCCCCAGGT CTGATCACCA ACTTCTCCCT GGCCACGGCA 60

GAGCAATGTG GCCAGGAGAC GCCACTGGAG AACATGCTGT TCGCCTCCTT CTACCTTCTG 120

GATTTTATCC TGGCTTTAGT TGGCAATACC CTGGCTCTGT GGCTTTTCAT CCGAGACCAC 180

AAGTCCGGGA CCCCAGCCAA CGTGTTCTG ATGCATCTGG CCGTGGCCGA CTTGTCGTGC 240

GTGCTGGTCC TGCCCAACCCG CCTGGTCTAC CACTTCTCTG GGAACCACTG GCCATTTGGG 300

GAAATCGCAT GCCGTCTCAC CGGCTTCCTC TTCTACCTCA ACATGTACGC CAGCATCTAC 360
 TTCCTCACCT GCATCAGCGC CGACCGTTTC CTGGCCATTG TGCACCCGGT CAAGTCCCTC 420
 AAGCTCCGCA GGCCCCCTCTA CGCACACCTG GCCTGTGCCT TCCTGTGGGT GGTGGTGGCT 480
 GTGGCCATGG CCCCCTGCT GGTGAGCCCA CAGACCGTGC AGACCAACCA CACGGTGGTC 540
 TGCCTGCAGC TGTACCGGGA GAAGGCCTCC CACCATGCCC TGGTGTCCCT GGCAGTGGCC 600
 TTCACCTTCC CGTTCATCAC CACGGTCACC TGCTACCTGC TGATCATCCG CAGCCTGCGG 660
 CAGGGCCTGC GTGTGGAGAA GCGCCTCAAG ACCAAGGCAG TGCGCATGAT CGCCATAGTG 720
 CTGGCCATCT TCCTGGTCTG CTTCGTGCCC TACCACGTCA ACCGCTCCGT CTACGTGCTG 780
 CACTACCGCA GCCATGGGGC CTCCTGCGCC ACCCAGCGCA TCCTGGCCCT GGCAAACCGC 840
 ATCACCTCCT GCCTCACCAG CCTCAACGGG GCACTCGACC CCATCATGTA TTTCTCGTG 900
 GCTGAGAAGT TCCGCCACGC CCTGTGCAAC TTGCTCTGTG GCAAAAGGCT CAAGGGCCCCG 960
 CCCCCAGCT TCGAAGGGAA AACCAACGAG AGCTCGCTGA GTGCCAAGTC AGAGCTGTGA 1020

(33) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 339 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

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

10

20

30

100	105	110
Leu Asn Met Tyr Ala Ser Ile Tyr Phe Leu Thr Cys Ile Ser Ala Asp		
115	120	125
Arg Phe Leu Ala Ile Val His Pro Val Lys Ser Leu Lys Leu Arg Arg		
130	135	140
Pro Leu Tyr Ala His Leu Ala Cys Ala Phe Leu Trp Val Val Val Ala		
145	150	155 160
Val Ala Met Ala Pro Leu Leu Val Ser Pro Gln Thr Val Gln Thr Asn		
165	170	175
His Thr Val Val Cys Leu Gln Leu Tyr Arg Glu Lys Ala Ser His His		
180	185	190
Ala Leu Val Ser Leu Ala Val Ala Phe Thr Phe Pro Phe Ile Thr Thr		
195	200	205
Val Thr Cys Tyr Leu Leu Ile Ile Arg Ser Leu Arg Gln Gly Leu Arg		
210	215	220
Val Glu Lys Arg Leu Lys Thr Lys Ala Val Arg Met Ile Ala Ile Val		
225	230	235 240
Leu Ala Ile Phe Leu Val Cys Phe Val Pro Tyr His Val Asn Arg Ser		
245	250	255
Val Tyr Val Leu His Tyr Arg Ser His Gly Ala Ser Cys Ala Thr Gln		
260	265	270
Arg Ile Leu Ala Leu Ala Asn Arg Ile Thr Ser Cys Leu Thr Ser Leu		
275	280	285
Asn Gly Ala Leu Asp Pro Ile Met Tyr Phe Phe Val Ala Glu Lys Phe		
290	295	300
Arg His Ala Leu Cys Asn Leu Leu Cys Gly Lys Arg Leu Lys Gly Pro		
305	310	315 320
Pro Pro Ser Phe Glu Gly Lys Thr Asn Glu Ser Ser Leu Ser Ala Lys		
325	330	335
Ser Glu Leu		

10

20

30

(34) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

ATAAGATGAT CACCCTGAAC AATCAAGAT

29

(35) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 33 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

TCCGAATTCA TAACATTTCA CTGTTTATAT TGC

33

(36) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 996 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

ATGATCACCC TGAACAATCA AGATCAACCT GTCACCTTTTA ACAGCTCACA TCCAGATGAA 60
TACAAAATTG CAGCCCTTGT CTCTATAGC TGTATCTTCA TAATTGGATT ATTTGTAAAC 120
ATCACTGCAT TATGGGTTTT CAGTTGTACC ACCAAGAAGA GAACACGGT AACCATCTAT 180
ATGATGAATG TGGCATTAGT GGAATTGATA TTTATAATGA CTTTACCCTT TCGAATGTTT 240
TATTATGCAA AAGATGCATG GCCATTTGGA GAGTACTTCT GCCAGATTAT TGGAGCTCTC 300
ACAGTGTTTT ACCCAAGCAT TGCTTTATGG CTCTTGCCCT TTATTAGTGC TGACAGATAC 360
ATGGCCATTG TACAGCCGAA GTACGCCAAA GAACTTAAAA ACACGTGCAA AGCCGTGCTG 420
GCGTGTGTGG GAGTCTGGAT AATGACCCTG ACCACGACCA CCCCTCTGCT ACTGCTCTAT 480
AAAGACCCAG ATAAAGACTC CACTCCCGCC ACCTGCCTCA AGATTCTGA CATCATCTAT 540
CTAAAAGCTG TGAACGTGCT GAACCTCACT CGACTGACAT TTTTTTCTT GATTCCTTTG 600
TTCATCATGA TTGGGTGCTA CTTGGTCATT ATTCATAATC TCCTTCACGG CAGGACGTCT 660
AAGCTGAAAC CCAAAGTCAA GGAGAAGTCC ATAAGGATCA TCATCACGCT GCTGGTGACG 720

10

20

30

```

GTGCTCGTCT GCTTTATGCC CTTCCACATC TGTTCGCTT TCCTGATGCT GGAACGGGG 780
GAGAACAGTT ACAATCCCTG GGGAGCCTTT ACCACCTTCC TCATGAACCT CAGCACGTGT 840
CTGGATGTGA TTCTCTACTA CATCGTTTCA AAACAATTC AGGCTCGAGT CATTAGTGTC 900
ATGCTATAACC GTAATTACCT TCGAAGCCTG CGCAGAAAAA GTTCCGATC TGGTAGTCTA 960
AGGTCATAAA GCAATATAAA CAGTGAAATG TTATGA 996

```

(37) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 331 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: not relevant

10

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

```

Met Ile Thr Leu Asn Asn Gln Asp Gln Pro Val Thr Phe Asn Ser Ser
1           5           10           15

His Pro Asp Glu Tyr Lys Ile Ala Ala Leu Val Phe Tyr Ser Cys Ile
20          25          30

Phe Ile Ile Gly Leu Phe Val Asn Ile Thr Ala Leu Trp Val Phe Ser
35          40          45

Cys Thr Thr Lys Lys Arg Thr Thr Val Thr Ile Tyr Met Met Asn Val
50          55          60

Ala Leu Val Asp Leu Ile Phe Ile Met Thr Leu Pro Phe Arg Met Phe
65          70          75          80

Tyr Tyr Ala Lys Asp Ala Trp Pro Phe Gly Glu Tyr Phe Cys Gln Ile
85          90          95

Ile Gly Ala Leu Thr Val Phe Tyr Pro Ser Ile Ala Leu Trp Leu Leu
100         105         110

Ala Phe Ile Ser Ala Asp Arg Tyr Met Ala Ile Val Gln Pro Lys Tyr
115         120         125

Ala Lys Glu Leu Lys Asn Thr Cys Lys Ala Val Leu Ala Cys Val Gly
130         135         140

Val Trp Ile Met Thr Leu Thr Thr Thr Thr Pro Leu Leu Leu Tyr
145         150         155         160

Lys Asp Pro Asp Lys Asp Ser Thr Pro Ala Thr Cys Leu Lys Ile Ser
165         170         175

```

20

30

Asp Ile Ile Tyr Leu Lys Ala Val Asn Val Leu Asn Leu Thr Arg Leu
 180 185 190
 Thr Phe Phe Phe Leu Ile Pro Leu Phe Ile Met Ile Gly Cys Tyr Leu
 195 200 205
 Val Ile Ile His Asn Leu Leu His Gly Arg Thr Ser Lys Leu Lys Pro
 210 215 220
 Lys Val Lys Glu Lys Ser Ile Arg Ile Ile Ile Thr Leu Leu Val Gln
 225 230 235 240
 Val Leu Val Cys Phe Met Pro Phe His Ile Cys Phe Ala Phe Leu Met
 245 250 255
 Leu Gly Thr Gly Glu Asn Ser Tyr Asn Pro Trp Gly Ala Phe Thr Thr
 260 265 270
 Phe Leu Met Asn Leu Ser Thr Cys Leu Asp Val Ile Leu Tyr Tyr Ile
 275 280 285
 Val Ser Lys Gln Phe Gln Ala Arg Val Ile Ser Val Met Leu Tyr Arg
 290 295 300
 Asn Tyr Leu Arg Ser Leu Arg Arg Lys Ser Phe Arg Ser Gly Ser Leu
 305 310 315 320
 Arg Ser Leu Ser Asn Ile Asn Ser Glu Met Leu
 325 330

10

20

(38) INFORMATION FOR SEQ ID NO:37:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 28 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

CCAAGCTTCC AGGCCTGGGG TGTGCTGG

28

30

(39) INFORMATION FOR SEQ ID NO:38:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 29 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

ATGGATCCTG ACCTTCGGCC CCTGGCAGA

29

(40) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1077 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

ATGCCCTCTG TGTCTCCAGC GGGGCCCTCG GCCGGGGCAG TCCCCAATGC CACCGCAGTG 60
 ACAACAGTGC GGACCAATGC CAGCGGGCTG GAGGTGCCCC TGTTCACCT GTTTGCCCGG 120
 CTGGACGAGG AGCTGCATGG CACCTTCCCA GGCTGTGCG TGGCGCTGAT GGCGGTGCAC 180
 GGAGCCATCT TCCTGGCAGG GCTGGTGCTC AACGGGCTGG CGCTGTACGT CTTCTGCTGC 240
 CGCACCCGGG CCAAGACACC CTCAGTCATC TACACCATCA ACCTGGTGGT GACCGATCTA 300
 CTGGTAGGGC TGTCCTGCG CACGCGCTTC GCTGTGTACT ACGGCGCCAG GGGCTGCCTG 360
 CGCTGTGCCT TCCCGCACGT CCTCGGTAC TTCCTCAACA TGCACTGCTC CATCCTCTTC 420
 CTCACCTGCA TCTGCGTGGA CCGCTACCTG GCCATCGTGC GGCCGAAGG CTCCCGCCGC 480
 TGCCGCCAGC CTGCCTGTGC CAGGGCCGTG TCGCCTTCG TGTGGCTGGC CGCCGCTGCC 540
 GTCACCCTGT CGGTGCTGGG CGTGACAGGC AGCCGGCCCT GCTGCCGTGT CTTTGCGCTG 600
 ACTGTCCTGG AGTTCCTGCT GCCCCTGCTG GTCATCAGCG TGTTTACCGG CCGCATCATG 660
 TGTGCACTGT CGCGGCCGGG TCTGCTCCAC CAGGGTCGCC AGCGCCGCGT GCGGGCCATG 720
 CAGCTCCTGC TCACGGTGCT CATCATCTTT CTCGTCTGCT TCACGCCCTT CCACGCCCGC 780
 CAAGTGGCCG TGGCGCTGTG GCCCACATG CCACACCACA CGAGCCTCGT GGTCTACCAC 840
 GTGGCCGTGA CCCTCAGCAG CCTCAACAGC TGCATGGACC CCATCGTCTA CTGCTTCGTC 900
 ACCAGTGGCT TCCAGGCCAC CGTCCGAGGC CTCTTCGGCC AGCACGGAGA GCGTGAGCCC 960
 AGCAGCGGTG ACGTGGTCAG CATGCACAGG AGCTCCAAGG GCTCAGGCCG TCATCACATC 1020
 CTCAGTGCCG GCCCTCACGC CCTACCCAG GCCCTGGCTA ATGGGCCCGA GGCTTAG 1077

10

20

30

(41) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 358 amino acids

(B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

```

Met Pro Ser Val Ser Pro Ala Gly Pro Ser Ala Gly Ala Val Pro Asn
1           5           10           15
Ala Thr Ala Val Thr Thr Val Arg Thr Asn Ala Ser Gly Leu Glu Val
          20           25           30
Pro Leu Phe His Leu Phe Ala Arg Leu Asp Glu Glu Leu His Gly Thr
          35           40           45
Phe Pro Gly Leu Cys Val Ala Leu Met Ala Val His Gly Ala Ile Phe
          50           55           60
Leu Ala Gly Leu Val Leu Asn Gly Leu Ala Leu Tyr Val Phe Cys Cys
65           70           75           80
Arg Thr Arg Ala Lys Thr Pro Ser Val Ile Tyr Thr Ile Asn Leu Val
          85           90           95
Val Thr Asp Leu Leu Val Gly Leu Ser Leu Pro Thr Arg Phe Ala Val
          100          105          110
Tyr Tyr Gly Ala Arg Gly Cys Leu Arg Cys Ala Phe Pro His Val Leu
          115          120          125
Gly Tyr Phe Leu Asn Met His Cys Ser Ile Leu Phe Leu Thr Cys Ile
          130          135          140
Cys Val Asp Arg Tyr Leu Ala Ile Val Arg Pro Glu Ala Pro Ala Ala
145          150          155          160
Cys Arg Gln Pro Ala Cys Ala Arg Ala Val Cys Ala Phe Val Trp Leu
          165          170          175
Ala Ala Gly Ala Val Thr Leu Ser Val Leu Gly Val Thr Gly Ser Arg
          180          185          190
Pro Cys Cys Arg Val Phe Ala Leu Thr Val Leu Glu Phe Leu Leu Pro
          195          200          205
Leu Leu Val Ile Ser Val Phe Thr Gly Arg Ile Met Cys Ala Leu Ser
          210          215          220
Arg Pro Gly Leu Leu His Gln Gly Arg Gln Arg Arg Val Arg Ala Met
225          230          235          240
Gln Leu Leu Leu Thr Val Leu Ile Ile Phe Leu Val Cys Phe Thr Pro
          245          250          255

```

10

20

30

Phe His Ala Arg Gln Val Ala Val Ala Leu Trp Pro Asp Met Pro His
 260 265 270
 His Thr Ser Leu Val Val Tyr His Val Ala Val Thr Leu Ser Ser Leu
 275 280 285
 Asn Ser Cys Met Asp Pro Ile Val Tyr Cys Phe Val Thr Ser Gly Phe
 290 295 300
 Gln Ala Thr Val Arg Gly Leu Phe Gly Gln His Gly Glu Arg Glu Pro
 305 310 315 320
 Ser Ser Gly Asp Val Val Ser Met His Arg Ser Ser Lys Gly Ser Gly
 325 330 335
 Arg His His Ile Leu Ser Ala Gly Pro His Ala Leu Thr Gln Ala Leu
 340 345 350
 Ala Asn Gly Pro Glu Ala
 355

10

(42) INFORMATION FOR SEQ ID NO:41:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

GAGAATTCAC TCCTGAGCTC AAGATGAACT

30

(43) INFORMATION FOR SEQ ID NO:42:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

CGGGATCCCC GTAAGTGGC CACTTCAGAT

30

(44) INFORMATION FOR SEQ ID NO:43:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1050 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single

30

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

```

ATGAACTCCA CCTTGGATGG TAATCAGAGC AGCCACCCTT TTGCGCTCTT GGCATTGGGC    60
TATTTGGAAA CTGTCAATTT TTGCCTTTTG GAAGTATTGA TTATTGTCTT TCTAACTGTA    120
TTGATTATTT CTGGCAACAT CATTGTGATT TTTGTATTTC ACTGTGCACC TTTGTTGAAC    180
CATCACACTA CAAGTTATTT TATCCAGACT ATGGCATATG CTGACCTTTT TGTGGGGTG    240
AGCTGCGTGG TCCCTTCTTT ATCACTCCTC CATCACCCCC TTCCAGTAGA GGAGTCCTTG    300
ACTTGCCAGA TATTTGGTTT TGTAGTATCA GTTCTGAAGA GCGTCTCCAT GGCTTCTCTG    360
GCCTGTATCA GCATTGATAG ATACATTGCC ATTACTAAAC CTTTAACCTA TAATACTCTG    420
GTTACACCCT GGAGACTACG CCTGTGTATT TTCCTGATTI GGCTATACTC GACCCTGGTC    480
TTCCTGCCTT CTTTTTCCA CTGGGGCAAA CCTGGATATC ATGGAGATGT GTTTCAGTGG    540
TGTGCGGAGT CCTGGCACAC CGACTCCTAC TTCACCCTGT TCATCGTGAT GATGTTATAT    600
GCCCCAGCAG CCCTTATTGT CTGCTTCACC TATTTCAACA TCTTCCGCAT CTGCCAACAG    660
CACACAAAGG ATATCAGCGA AAGGCAAGCC CGCTTCAGCA GCCAGAGTGG GGAGACTGGG    720
GAAGTGCAGG CCTGTCTGA TAAGCGCTAT GCCATGGTCC TGTTCGAAT CACTAGTGTA    780
TTTTACATCC TCTGGTTGCC ATATATCATC TACTTCTTGT TGGAAAGCTC CACTGGCCAC    840
AGCAACCGCT TCGCATCCTT CTGACCACC TGGCTTGCTA TTAGTAACAG TTTCTGCAAC    900
TGTGTAATTT ATAGTCTCTC CAACAGTGTA TTCCAAAGAG GACTAAAGCG CCTCTCAGGG    960
GCTATGTGTA CTTCTTGTGC AAGTCAGACT ACAGCCAACG ACCCTTACAC AGTTAGAAGC   1020
AAAGGCCCTC TTAATGGATG TCATATCTGA                               1050

```

10

20

(45) INFORMATION FOR SEQ ID NO:44:

30

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 349 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

Met Asn Ser Thr Leu Asp Gly Asn Gln Ser Ser His Pro Phe Cys Leu

30

Ser Leu Ser Asn Ser Val Phe Gln Arg Gly Leu Lys Arg Leu Ser Gly
 305 310 315 320

Ala Met Cys Thr Ser Cys Ala Ser Gln Thr Thr Ala Asn Asp Pro Tyr
 325 330 335

Thr Val Arg Ser Lys Gly Pro Leu Asn Gly Cys His Ile
 340 345

(46) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

TCCCCCGGGA AAAAAACCAA CTGCTCCAAA 30

(47) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

TAGGATCCAT TTGAATGTGG ATTTGGTGAA A 31

(48) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1302 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

ATGTGTTTTT CTCCATTCT GGAAATCAAC ATGCAGTCTG AATCTAACAT TACAGTGCGA 60

GATGACATTG ATGACATCAA CACCAATATG TACCAACCAC TATCATATCC GTTAAGCTTT 120

CAAGTGTCTC TCACCGGATT TCTTATGTTA GAAATTGTGT TGGGACTTGG CAGCAACCTC 180

ACTGTATTGG TACTTTACTG CATGAAATCC AACTTAATCA ACTCTGTCAG TAACATTATT 240
 ACAATGAATC TTCATGTACT TGATGTAATA ATTTGTGTGG GATGTATTCC TCTAACTATA 300
 GTTATCCTTC TGCTTTCACT GGAGAGTAAC ACTGCTCTCA TTTGCTGTTT CCATGAGGCT 360
 TGTGTATCTT TTGCAAGTGT CTCAACAGCA ATCAACGTTT TTGCTATCAC TTTGGACAGA 420
 TATGACATCT CTGTAAAACC TGCAAACCGA ATTCTGACAA TGGGCAGAGC TGTAATGTTA 480
 ATGATATCCA TTTGGATTTT TTCTTTTTTC TCTTTCCTGA TTCCTTTTAT TGAGGTAAAT 540
 TTTTTCAGTC TTCAAAGTGG AAATACCTGG GAAAACAAGA CACTTTTATG TGTCAGTACA 600
 AATGAATACT AACTGAACT GGAATGTAT TATCACCTGT TAGTACAGAT CCCAATATTC 660
 TTTTTCAGTC TTGTAGTAAT GTTAATCACA TACACCAAAA TACTTCAGGC TCTTAATATT 720
 CGAATAGGCA CAAGATTTTC AACAGGGCAG AAGAAGAAAG CAAGAAAGAA AAAGACAATT 780
 TCTCTAACCA CACAACATGA GGCTACAGAC ATGTCACAAA GCAGTGGTGG GAGAAATGTA 840
 GTCTTTGGTG TAAGAACTTC AGTTTCTGTA ATAATTGCCC TCCGGCGAGC TGTGAAACGA 900
 CACCGTGAAC GACGAGAAAG ACAAAGAGA GTCTTCAGGA TGTCTTTATT GATTATTTCT 960
 ACATTTCTTC TCTGCTGGAC ACCAATTCTT GTTTTAAATA CCACCATTTT ATGTTTAGGC 1020
 CCAAGTGACC TTTTAGTAAA ATTAAGATTG TGTTTTTTAG TCATGGCTTA TGGAACAAC 1080
 ATATTTCAAC CTCTATTATA TGCATTCAC 1140
 AAAATGAAAA AGCGAGTTGT TTCTATAGTA GAAGCTGATC CCCTGCCTAA TAATGCTGTA 1200
 ATACACAAC CTTGGATAGA TCCCAAAAGA AACAAAAAAA TTACCTTTGA AGATAGTGAA 1260
 ATAAGAGAAA AACGTTTAGT GCCTCAGGTT GTCACAGACT AG 1302

(49) INFORMATION FOR SEQ ID NO:48:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 433 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

Met	Cys	Phe	Ser	Pro	Ile	Leu	Glu	Ile	Asn	Met	Gln	Ser	Glu	Ser	Asn
1				5					10					15	
Ile	Thr	Val	Arg	Asp	Asp	Ile	Asp	Asp	Ile	Asn	Thr	Asn	Met	Tyr	Gln
			20					25					30		

10

20

30

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

10

20

30

(B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

ATGTTGTGTC	CTTCCAAGAC	AGATGGCTCA	GGGCACTCTG	GTAGGATTCA	CCAGGAAACT	60
CATGGAGAAG	GGAAAAGGGA	CAAGATTAGC	AACAGTGAAG	GGAGGGAGAA	TGGTGGGAGA	120
GGATTCCAGA	TGAACGGTGG	GTCGCTGGAG	GCTGAGCATG	CCAGCAGGAT	GTCAGTTCTC	180
AGAGCAAAGC	CCATGTCAAA	CAGCCAACGC	TTGCTCCTTC	TGTCCCCAGG	ATCACCTCCT	240
CGCACGGGGA	GCATCTCCTA	CATCAACATC	ATCATGCCTT	CGGTGTTGGG	CACCATCTGC	300
CTCCTGGGCA	TCATCGGGAA	CTCCACGGTC	ATCTTCGCGG	TCGTGAAGAA	GTCCAAGCTG	360
CACTGGTGCA	ACAACGTCCC	CGACATCTTC	ATCATCAACC	TCTCGGTAGT	AGATCTCCTC	420
TTTCTCCTGG	GCATGCCCTT	CATGATCCAC	CAGCTCATGG	GCAATGGGGT	GTGGCACTTT	480
GGGGAGACCA	TGTGCACCCT	CATCACGGCC	ATGGATGCCA	ATAGTCAGTT	CACCAGCACC	540
TACATCCTGA	CCGCCATGGC	CATTGACCGC	TACCTGGCCA	CTGTCCACCC	CATCTCTTCC	600
ACGAAGTTCC	GGAAGCCCTC	TGTGGCCACC	CTGGTGATCT	GCCTCCTGTG	GGCCCTCTCC	660
TTCATCAGCA	TCACCCCTGT	GTGGCTGTAT	GCCAGACTCA	TCCCCTTCCC	AGGAGGTGCA	720
GTGGGCTGCG	GCATACGCCT	GCCCAACCCA	GACACTGACC	TCTACTGGTT	CACCCTGTAC	780
CAGTTTTTCC	TGGCCTTTGC	CCTGCCTTTT	GTGGTCATCA	CAGCCGCATA	CGTGAGGATC	840
CTGCAGCGCA	TGACGTCTTC	AGTGGCCCCC	GCCTCCCAGC	GCAGCATCCG	GCTGCGGACA	900
AAGAGGGTGA	CCCGCACAGC	CATCGCCATC	TGTCTGGTCT	TCTTTGTGTG	CTGGGCACCC	960
TACTATGTGC	TACAGCTGAC	CCAGTTGTCC	ATCAGCCGCC	CGACCCTCAC	CTTTGTCTAC	1020
TTATACAATG	CGGCCATCAG	CTTGGGCTAT	GCCAACAGCT	GCCTCAACCC	CTTTGTGTAC	1080
ATCGTGCTCT	GTGAGACGTT	CCGCAAACGC	TTGGTCCTGT	CGGTGAAGCC	TGCAGCCCAG	1140
GGGCAGCTTC	GCGCTGTCAG	CAACGCTCAG	ACGGCTGACG	AGGAGAGGAC	AGAAAGCAAA	1200
GGCACCTGA						1209

(53) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 402 amino acids
 (B) TYPE: amino acid

10

20

30

(C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

```

Met Leu Cys Pro Ser Lys Thr Asp Gly Ser Gly His Ser Gly Arg Ile
1           5           10           15

His Gln Glu Thr His Gly Glu Gly Lys Arg Asp Lys Ile Ser Asn Ser
          20           25           30

Glu Gly Arg Glu Asn Gly Gly Arg Gly Phe Gln Met Asn Gly Gly Ser
          35           40           45

Leu Glu Ala Glu His Ala Ser Arg Met Ser Val Leu Arg Ala Lys Pro
50           55           60

Met Ser Asn Ser Gln Arg Leu Leu Leu Leu Ser Pro Gly Ser Pro Pro
65           70           75           80

Arg Thr Gly Ser Ile Ser Tyr Ile Asn Ile Ile Met Pro Ser Val Phe
          85           90           95

Gly Thr Ile Cys Leu Leu Gly Ile Ile Gly Asn Ser Thr Val Ile Phe
          100          105          110

Ala Val Val Lys Lys Ser Lys Leu His Trp Cys Asn Asn Val Pro Asp
          115          120          125

Ile Phe Ile Ile Asn Leu Ser Val Val Asp Leu Leu Phe Leu Leu Gly
          130          135          140

Met Pro Phe Met Ile His Gln Leu Met Gly Asn Gly Val Trp His Phe
145           150           155           160

Gly Glu Thr Met Cys Thr Leu Ile Thr Ala Met Asp Ala Asn Ser Gln
          165          170          175

Phe Thr Ser Thr Tyr Ile Leu Thr Ala Met Ala Ile Asp Arg Tyr Leu
          180          185          190

Ala Thr Val His Pro Ile Ser Ser Thr Lys Phe Arg Lys Pro Ser Val
          195          200          205

Ala Thr Leu Val Ile Cys Leu Leu Trp Ala Leu Ser Phe Ile Ser Ile
          210          215          220

Thr Pro Val Trp Leu Tyr Ala Arg Leu Ile Pro Phe Pro Gly Gly Ala
225           230           235           240

Val Gly Cys Gly Ile Arg Leu Pro Asn Pro Asp Thr Asp Leu Tyr Trp
          245          250          255

```

10

20

30

Phe Thr Leu Tyr Gln Phe Phe Leu Ala Phe Ala Leu Pro Phe Val Val
 260 265 270
 Ile Thr Ala Ala Tyr Val Arg Ile Leu Gln Arg Met Thr Ser Ser Val
 275 280 285
 Ala Pro Ala Ser Gln Arg Ser Ile Arg Leu Arg Thr Lys Arg Val Thr
 290 295 300
 Arg Thr Ala Ile Ala Ile Cys Leu Val Phe Phe Val Cys Trp Ala Pro
 305 310 315 320
 Tyr Tyr Val Leu Gln Leu Thr Gln Leu Ser Ile Ser Arg Pro Thr Leu
 325 330 335
 Thr Phe Val Tyr Leu Tyr Asn Ala Ala Ile Ser Leu Gly Tyr Ala Asn
 340 345 350
 Ser Cys Leu Asn Pro Phe Val Tyr Ile Val Leu Cys Glu Thr Phe Arg
 355 360 365
 Lys Arg Leu Val Leu Ser Val Lys Pro Ala Ala Gln Gly Gln Leu Arg
 370 375 380
 Ala Val Ser Asn Ala Gln Thr Ala Asp Glu Glu Arg Thr Glu Ser Lys
 385 390 395 400

Gly Thr

10

(54) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 27 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

GGCGGATCCA TGGATGTGAC TTCCCAA

27

(55) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 27 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

20

30

GGCGGATCCC TACACGGCAC TGCTGAA

27

(56) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1128 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

ATGGATGTGA CTTCCCAAGC CCGGGGCGTG GGCCTGGAGA TGTACCCAGG CACCGCGCAC 60
 GCTGCGGCCC CCAACACCAC CTCCTCCGAG CTCACCTGT CCCACCCGCT CCTGGGCACC 120
 GCCCTGGCCA ATGGGACAGG TGAGCTCTCG GAGCACCAGC AGTACGTGAT CGGCCTGTTC 180
 CTCTCGTGCC TCTACACCAT CTTCCTCTTC CCCATCGGCT TTGTGGGCAA CATCTGATC 240
 CTGGTGGTGA ACATCAGCTT CCGCGAGAAG ATGACCATCC CCGACCTGTA CTTTCATCAAC 300
 CTGGCGGTGG CGGACCTCAT CCTGGTGGCC GACTCCCTCA TTGAGGTGTT CAACCTGCAC 360
 GAGCGGTACT ACGACATCGC CGTCCTGTGC ACCTTCATGT CGCTCTTCCT GCAGGTCAAC 420
 ATGTACAGCA GCGTCTTCTT CCTCACCTGG ATGAGCTTCG ACCGCTACAT CGCCCTGGCC 480
 AGGGCCATGC GCTGCAGCCT GTTCCGCACC AAGCACCACG CCCGGCTGAG CTGTGGCCTC 540
 ATCTGGATGG CATCCGTGTC AGCCACGCTG GTGCCCTTCA CCGCCGTGCA CCTGCAGCAC 600
 ACCGACGAGG CCTGCTTCTG TTTCGCGGAT GTCCGGGAGG TGCAGTGGCT CGAGGTACAG 660
 CTGGGCTTCA TCGTGCCCTT CGCCATCATC GGCCTGTGCT ACTCCCTCAT TGTCCGGGTG 720
 CTGGTCAGGG CGCACCAGCA CCGTGGGCTG CGGCCCGGC GGCAGAAGGC GCTCCGCATG 780
 ATCCTCGCGG TGGTGTGGT CTTCTTCGTC TGCTGGCTGC CGGAGAACGT CTTTCATCAGC 840
 GTGCACCTCC TGCAGCGGAC GCAGCCTGGG GCCGCTCCCT GCAAGCAGTC TTTCCGCCAT 900
 GCCCACCCCC TCACGGGCCA CATTGTCAAC CTCGCCGCCT TCTCCAACAG CTGCCTAAAC 960
 CCCCTCATCT ACAGCTTTCT CGGGGAGACC TTCAGGGACA AGCTGAGGCT GTACATTGAG 1020
 CAGAAAACAA ATTTGCCGGC CCTGAACCGC TTCTGTACG CTGCCCTGAA GGCCGTCATT 1080
 CCAGACAGCA CCGAGCAGTC GGATGTGAGG TTCAGCAGTG CCGTGTGA 1128

10

20

30

(57) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 375 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

```

Met Asp Val Thr Ser Gln Ala Arg Gly Val Gly Leu Glu Met Tyr Pro
1           5           10           15

Gly Thr Ala His Ala Ala Ala Pro Asn Thr Thr Ser Pro Glu Leu Asn
           20           25           30

Leu Ser His Pro Leu Leu Gly Thr Ala Leu Ala Asn Gly Thr Gly Glu
           35           40           45

Leu Ser Glu His Gln Gln Tyr Val Ile Gly Leu Phe Leu Ser Cys Leu
           50           55           60

Tyr Thr Ile Phe Leu Phe Pro Ile Gly Phe Val Gly Asn Ile Leu Ile
65           70           75           80

Leu Val Val Asn Ile Ser Phe Arg Glu Lys Met Thr Ile Pro Asp Leu
           85           90           95

Tyr Phe Ile Asn Leu Ala Val Ala Asp Leu Ile Leu Val Ala Asp Ser
           100          105          110

Leu Ile Glu Val Phe Asn Leu His Glu Arg Tyr Tyr Asp Ile Ala Val
           115          120          125

Leu Cys Thr Phe Met Ser Leu Phe Leu Gln Val Asn Met Tyr Ser Ser
           130          135          140

Val Phe Phe Leu Thr Trp Met Ser Phe Asp Arg Tyr Ile Ala Leu Ala
145          150          155          160

Arg Ala Met Arg Cys Ser Leu Phe Arg Thr Lys His His Ala Arg Leu
           165          170          175

Ser Cys Gly Leu Ile Trp Met Ala Ser Val Ser Ala Thr Leu Val Pro
           180          185          190

Phe Thr Ala Val His Leu Gln His Thr Asp Glu Ala Cys Phe Cys Phe
           195          200          205

Ala Asp Val Arg Glu Val Gln Trp Leu Glu Val Thr Leu Gly Phe Ile
           210          215          220

Val Pro Phe Ala Ile Ile Gly Leu Cys Tyr Ser Leu Ile Val Arg Val
225          230          235          240

Leu Val Arg Ala His Arg His Arg Gly Leu Arg Pro Arg Arg Gln Lys

```

10

20

30

245	250	255
Ala Leu Arg Met Ile Leu Ala Val	Val Leu Val Phe Phe Val Cys Trp	
260	265	270
Leu Pro Glu Asn Val Phe Ile Ser Val His Leu Leu Gln Arg Thr Gln		
275	280	285
Pro Gly Ala Ala Pro Cys Lys Gln Ser Phe Arg His Ala His Pro Leu		
290	295	300
Thr Gly His Ile Val Asn Leu Ala Ala Phe Ser Asn Ser Cys Leu Asn		
305	310	315
Pro Leu Ile Tyr Ser Phe Leu Gly Glu Thr Phe Arg Asp Lys Leu Arg		
325	330	335
Leu Tyr Ile Glu Gln Lys Thr Asn Leu Pro Ala Leu Asn Arg Phe Cys		
340	345	350
His Ala Ala Leu Lys Ala Val Ile Pro Asp Ser Thr Glu Gln Ser Asp		
355	360	365
Val Arg Phe Ser Ser Ala Val		
370	375	

(58) INFORMATION FOR SEQ ID NO:57:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 31 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

AAGGAATTCA CGGCCGGGTG ATGCCATTCC C

31

(59) INFORMATION FOR SEQ ID NO:58:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

GGTGGATCCA TAAACACGGG CGTTGAGGAC

30

(60) INFORMATION FOR SEQ ID NO:59:

10

20

30

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 960 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

ATGCCATTCC CAAACTGCTC AGCCCCCAGC ACTGTGGTGG CCACAGCTGT GGGTGTCTTG	60
CTGGGGCTGG AGTGTGGGCT GGGTCTGCTG GGCAACGCGG TGGCGCTGTG GACCTTCCTG	120
TTCCGGGTCA GGGTGTGGAA GCCGTACGCT GTCTACCTGC TCAACCTGGC CCTGGCTGAC	180
CTGCTGTTGG CTGCGTGCCT GCCTTTCCCTG GCCGCCTTCT ACCTGAGCCT CCAGGCTTGG	240
CATCTGGGCC GTGTGGGCTG CTGGGCCCTG CGCTTCCTGC TGGACCTCAG CCGCAGCGTG	300
GGGATGGCCT TCCTGGCCGC CGTGGCTTTG GACCGGTACC TCCGTGTGGT CCACCCTCGG	360
CTTAAGGTCA ACCTGCTGTC TCCTCAGGCG GCCCTGGGGG TCTCGGGCCT CGTCTGGCTC	420
CTGATGGTCG CCCTCACCTG CCCGGGCTTG CTCATCTCTG AGGCCGCCCA GAACTCCACC	480
AGGTGCCACA GTTTCTACTC CAGGGCAGAC GGCTCCTTCA GCATCATCTG GCAGGAAGCA	540
CTCTCCTGCC TTCAGTTTGT CCTCCCCTTT GGCCTCATCG TGTCTGCAA TGCAGGCATC	600
ATCAGGGCTC TCCAGAAAAG ACTCCGGGAG CCTGAGAAAC AGCCCAAGCT TCAGCGGGCC	660
CAGGCACTGG TCACCTTGGT GGTGGTGCTG TTTGCTCTGT GCTTTCTGCC CTGCTTCCTG	720
GCCAGAGTCC TGATGCACAT CTTCCAGAAAT CTGGGGAGCT GCAGGGCCCT TTGTGCAGTG	780
GCTCATACCT CGGATGTAC GGGCAGCCTC ACCTACCTGC ACAGTGTGCT CAACCCCGTG	840
GTATACTGCT TCTCCAGCCC CACCTTCAGG AGCTCCTATC GGAGGGTCTT CCACACCCTC	900
CGAGGCAAAG GGCAGGCAGC AGAGCCCCCA GATTTCAACC CCAGAGACTC CTATTCTTGA	960

(61) INFORMATION FOR SEQ ID NO:60:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 319 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

Met Pro Phe Pro Asn Cys Ser Ala Pro Ser Thr Val Val Ala Thr Ala

10

20

30

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

10

20

30

Gln Ala Ala Glu Pro Pro Asp Phe Asn Pro Arg Asp Ser Tyr Ser
 305 310 315

(62) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1143 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

ATGGAGGAAG GTGGTGATTT TGACAACACTAC TATGGGGCAG ACAACCAGTC TGAGTGTGAG 60
 TACACAGACT GGAAATCCTC GGGGGCCCTC ATCCCTGCCA TCTACATGTT GGTCTTCCTC 120
 CTGGGCACCA CGGGAAACGG TCTGGTGCTC TGGACCGTGT TTCGGAGCAG CCGGGAGAAG 180
 AGGCGCTCAG CTGATATCTT CATTTGCTAGC CTGGCGGTGG CTGACCTGAC CTTCTGTGGTG 240
 ACGCTGCCCC TGTGGGCTAC CTACACGTAC CGGGACTATG ACTGGCCCTT TGGGACCTTC 300
 TTCTGCAAGC TCAGCAGCTA CCTCATCTTC GTCAACATGT ACGCCAGCGT CTTCTGCCTC 360
 ACCGGCCTCA GCTTCGACCG CTACCTGGCC ATCGTGAGGC CAGTGGCCAA TGCTCGGCTG 420
 AGGCTGCGGG TCAGCGGGGC CGTGGCCACG GCAGTTCTTT GGGTGCTGGC CGCCCTCCTG 480
 GCCATGCCTG TCATGGTGTT ACGCACCACC GGGGACTTGG AGAACACCAC TAAGGTGCAG 540
 TGCTACATGG ACTACTCCAT GGTGGCCACT GTGAGCTCAG AGTGGGCCTG GGAGGTGGGC 600
 CTTGGGGTCT CGTCCACCAC CGTGGGCTTT GTGGTGCCCT TCACCATCAT GCTGACCTGT 660
 TACTTCTTCA TCGCCCAAAC CATCGCTGGC CACTTCCGCA AGGAACGCAT CGAGGGCCTG 720
 CGGAAGCGGC GCCGGCTGCT CAGCATCATC GTGGTGCTGG TGGTGACCTT TGCCCTGTGC 780
 TGGATGCCCT ACCACCTGGT GAAGACGCTG TACATGCTGG GCAGCCTGCT GCACTGGCCC 840
 TGTGACTTTG ACCTCTTCCT CATGAACATC TTCCCCTACT GCACCTGCAT CAGCTACGTC 900
 AACAGCTGCC TCAACCCCTT CCTCTATGCC TTTTTCGACC CCCGCTTCCG CCAGGCCTGC 960
 ACCTCCATGC TCTGCTGTGG CCAGAGCAGG TGCGCAGGCA CCTCCCACAG CAGCAGTGGG 1020
 GAGAAGTCAG CCAGCTACTC TTCGGGGCAC AGCCAGGGGC CCGGCCCCAA CATGGGCAAG 1080
 GGTGGAGAAC AGATGCACGA GAAATCCATC CCCTACAGCC AGGAGACCCT TGTGGTTGAC 1140
 TAG 1143

10

20

30

(63) INFORMATION FOR SEQ ID NO:62:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 380 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

```

Met Glu Glu Gly Gly Asp Phe Asp Asn Tyr Tyr Gly Ala Asp Asn Gln
1           5           10           15
Ser Glu Cys Glu Tyr Thr Asp Trp Lys Ser Ser Gly Ala Leu Ile Pro
20           25           30
Ala Ile Tyr Met Leu Val Phe Leu Leu Gly Thr Thr Gly Asn Gly Leu
35           40           45
Val Leu Trp Thr Val Phe Arg Ser Ser Arg Glu Lys Arg Arg Ser Ala
50           55           60
Asp Ile Phe Ile Ala Ser Leu Ala Val Ala Asp Leu Thr Phe Val Val
65           70           75           80
Thr Leu Pro Leu Trp Ala Thr Tyr Thr Tyr Arg Asp Tyr Asp Trp Pro
85           90           95
Phe Gly Thr Phe Phe Cys Lys Leu Ser Ser Tyr Leu Ile Phe Val Asn
100          105          110
Met Tyr Ala Ser Val Phe Cys Leu Thr Gly Leu Ser Phe Asp Arg Tyr
115          120          125
Leu Ala Ile Val Arg Pro Val Ala Asn Ala Arg Leu Arg Leu Arg Val
130          135          140
Ser Gly Ala Val Ala Thr Ala Val Leu Trp Val Leu Ala Ala Leu Leu
145          150          155          160
Ala Met Pro Val Met Val Leu Arg Thr Thr Gly Asp Leu Glu Asn Thr
165          170          175
Thr Lys Val Gln Cys Tyr Met Asp Tyr Ser Met Val Ala Thr Val Ser
180          185          190
Ser Glu Trp Ala Trp Glu Val Gly Leu Gly Val Ser Ser Thr Thr Val
195          200          205
Gly Phe Val Val Pro Phe Thr Ile Met Leu Thr Cys Tyr Phe Phe Ile
210          215          220
Ala Gln Thr Ile Ala Gly His Phe Arg Lys Glu Arg Ile Glu Gly Leu

```

10

20

30

225 230 235 240
 Arg Lys Arg Arg Arg Leu Leu Ser Ile Ile Val Val Leu Val Val Thr
 245 250 255
 Phe Ala Leu Cys Trp Met Pro Tyr His Leu Val Lys Thr Leu Tyr Met
 260 265 270
 Leu Gly Ser Leu Leu His Trp Pro Cys Asp Phe Asp Leu Phe Leu Met
 275 280 285
 Asn Ile Phe Pro Tyr Cys Thr Cys Ile Ser Tyr Val Asn Ser Cys Leu
 290 295 300
 Asn Pro Phe Leu Tyr Ala Phe Phe Asp Pro Arg Phe Arg Gln Ala Cys
 305 310 315 320
 Thr Ser Met Leu Cys Cys Gly Gln Ser Arg Cys Ala Gly Thr Ser His
 325 330 335
 Ser Ser Ser Gly Glu Lys Ser Ala Ser Tyr Ser Ser Gly His Ser Gln
 340 345 350
 Gly Pro Gly Pro Asn Met Gly Lys Gly Gly Glu Gln Met His Glu Lys
 355 360 365
 Ser Ile Pro Tyr Ser Gln Glu Thr Leu Val Val Asp
 370 375 380

10

(64) INFORMATION FOR SEQ ID NO:63:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 31 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

TGAGAATTCT GGTGACTCAC AGCCGGCACA G

31

(65) INFORMATION FOR SEQ ID NO:64:

30

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 31 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

GCCGGATCCA AGGAAAAGCA GCAATAAAAG G

31

(66) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1119 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

ATGAACTACC CGCTAACGCT GGAAATGGAC CTCGAGAACC TGGAGGACCT GTTCTGGGAA 60
 CTGGACAGAT TGGACAACCTA TAACGACACC TCCCTGGTGG AAAATCATCT CTGCCCTGCC 120
 ACAGAGGGTC CCCTCATGGC CTCCTTCAAG GCCGTGTTTCG TGCCCGTGGC CTACAGCCTC 180
 ATCTTCCTCC TGGGCGTGAT CGGCAACGTC CTGGTGCTGG TGATCCTGGA GCGGCACCGG 240
 CAGACACGCA GTTCCACGGA GACCTTCCTG TTCCACCTGG CCGTGGCCGA CCTCCTGCTG 300
 GTCTTCATCT TGCCCTTTGC CGTGGCCGAG GGCTCTGTGG GCTGGGTCCT GGGGACCTTC 360
 CTCTGCAAAA CTGTGATTGC CCTGCACAAA GTCAACTTCT ACTGCAGCAG CCTGCTCCTG 420
 GCCTGCATCG CCGTGGACCG CTACCTGGCC ATTGTCCACG CCGTCCATGC CTACCGCCAC 480
 CGCCGCCTCC TCTCCATCCA CATCACCTGT GGGACCATCT GGCTGGTGGG CTTCTCCTTT 540
 GCCTTGCCAG AGATTCTCTT CGCCAAAGTC AGCCAAGGCC ATCACAACAA CTCCCTGCCA 600
 CGTTGCACCT TCTCCAAGA GAACCAAGCA GAAACGCATG CCTGGTTTAC CTCCCGATTG 660
 CTCTACCATG TGGCGGGATT CCTGCTGCCC ATGCTGGTGA TGGGCTGGTG CTACGTGGGG 720
 GTAGTGACACA GGTTCGCCCA GGCCAGCGG CGCCCTCAGC GGCAGAAGGC AGTCAGGGTG 780
 GCCATCCTGG TGACAAGCAT CTTCTTCCTC TGCTGGTCAC CCTACCACAT CGTCATCTTC 840
 CTGGACACCC TGGCGAGGCT GAAGCCCGTG GACAATACCT GCAAGCTGAA TGGCTCTCTC 900
 CCCGTGGCCA TCACCATGTG TGAGTTCTCTG GGCCTGGCCC ACTGCTGCCT CAACCCCATG 960
 CTCTACACTT TCGCCGGCGT GAAGTTCCGC AGTGACCTGT CGCGGCTCCT GACCAAGCTG 1020
 GGCTGTACCG GCCCTGCCTC CCTGTGCCAG CTCTTCCCTA GCTGGCGCAG GAGCAGTCTC 1080
 TCTGAGTCAG AGAATGCCAC CTCTCTCACC ACGTTCTAG 1119

10

20

30

(67) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 372 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

```

Met Asn Tyr Pro Leu Thr Leu Glu Met Asp Leu Glu Asn Leu Glu Asp
1           5           10           15
Leu Phe Trp Glu Leu Asp Arg Leu Asp Asn Tyr Asn Asp Thr Ser Leu
20           25           30
Val Glu Asn His Leu Cys Pro Ala Thr Glu Gly Pro Leu Met Ala Ser
35           40           45
Phe Lys Ala Val Phe Val Pro Val Ala Tyr Ser Leu Ile Phe Leu Leu
50           55           60
Gly Val Ile Gly Asn Val Leu Val Leu Val Ile Leu Glu Arg His Arg
65           70           75           80
Gln Thr Arg Ser Ser Thr Glu Thr Phe Leu Phe His Leu Ala Val Ala
85           90           95
Asp Leu Leu Leu Val Phe Ile Leu Pro Phe Ala Val Ala Glu Gly Ser
100          105          110
Val Gly Trp Val Leu Gly Thr Phe Leu Cys Lys Thr Val Ile Ala Leu
115          120          125
His Lys Val Asn Phe Tyr Cys Ser Ser Leu Leu Leu Ala Cys Ile Ala
130          135          140
Val Asp Arg Tyr Leu Ala Ile Val His Ala Val His Ala Tyr Arg His
145          150          155          160
Arg Arg Leu Leu Ser Ile His Ile Thr Cys Gly Thr Ile Trp Leu Val
165          170          175
Gly Phe Leu Leu Ala Leu Pro Glu Ile Leu Phe Ala Lys Val Ser Gln
180          185          190
Gly His His Asn Asn Ser Leu Pro Arg Cys Thr Phe Ser Gln Glu Asn
195          200          205
Gln Ala Glu Thr His Ala Trp Phe Thr Ser Arg Phe Leu Tyr His Val
210          215          220
Ala Gly Phe Leu Leu Pro Met Leu Val Met Gly Trp Cys Tyr Val Gly
225          230          235          240
Val Val His Arg Leu Arg Gln Ala Gln Arg Arg Pro Gln Arg Gln Lys

```

10

20

30

[illegible]

(68) INFORMATION FOR SEQ ID NO:67:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

CAAAGCTTGA AAGCTGCACG GTGCAGAGAC 30

(69) INFORMATION FOR SEQ ID NO:68:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

GCGGATCCCG AGTCACACCC TGGCTGGGCC 30

(70) INFORMATION FOR SEQ ID NO:69:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1128 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

ATGGATGTGA CTTCCCAAGC CCGGGGCGTG GGCCTGGAGA TGTACCCAGG CACCGCGCAG	60
CCTGCGGCCC CCAACACCAC CTCCCCGAG CTCAACCTGT CCCACCCGCT CCTGGGCACC	120
GCCCTGGCCA ATGGGACAGG TGAGCTCTCG GAGCACCAGC AGTACGTGAT CGGCCTGTTC	180
CTCTCGTGCC TCTACACCAT CTTCCTCTTC CCCATCGGCT TTGTGGGCAA CATCCTGATC	240
CTGGTGGTGA ACATCAGCTT CCGCGAGAAG ATGACCATCC CCGACCTGTA CTTTCATCAAC	300
CTGGCGGTGG CGGACCTCAT CCTGGTGGCC GACTCCCTCA TTGAGGTGTT CAACCTGCAC	360
GAGCGGTACT ACGACATCGC CGTCCTGTGC ACCTTCATGT CGCTCTTCCT GCAGGTCAAC	420
ATGTACAGCA GCGTCTTCTT CCTCACCTGG ATGAGCTTCG ACCGCTACAT CGCCCTGGCC	480
AGGGCCATGC GCTGCAGCCT GTTCCGCACC AAGCACCAG CCCGCTGAG CTGTGGCCTC	540
ATCTGGATGG CATCCGTGTC AGCCACGCTG GTGCCCTTCA CCGCCGTGCA CCTGCAGCAC	600
ACCGACGAGG CCTGCTTCTG TTTCGCGGAT GTCCGGGAGG TGCAGTGGCT CGAGGTCACG	660
CTGGGCTTCA TCGTGCCCTT CGCCATCATC GGCCTGTGCT ACTCCCTCAT TGTCCGGGTG	720
CTGGTCAGGG CGCACCGGCA CCGTGGGCTG CGGCCCGGC GGCAGAAGGC GCTCCGCATG	780
ATCCTCGCGG TGGTGCTGGT CTTCTTCGTC TGCTGGCTGC CGGAGAACGT CTTTCATCAGC	840
GTGCACCTCC TGCAGCGGAC GCAGCCTGGG GCGCTCCCT GCAAGCAGTC TTTCCGCCAT	900
GCCCACCCCC TCACGGGCCA CATTGTCAAC CTCACCGCCT TCTCCAACAG CTGCCTAAAC	960
CCCCTCATCT ACAGCTTTCT CGGGGAGACC TTCAGGGACA AGCTGAGGCT GTACATTGAG	1020
CAGAAAACAA ATTTGCCGGC CCTGAACCGC TTCTGTCACG CTGCCCTGAA GGCCGTCATT	1080
CCAGACAGCA CCGAGCAGTC GGATGTGAGG TTCAGCAGTG CCGTGTAG	1128

(71) INFORMATION FOR SEQ ID NO:70:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 375 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

10

20

30

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

```

Met Asp Val Thr Ser Gln Ala Arg Gly Val Gly Leu Glu Met Tyr Pro
1          5          10          15

Gly Thr Ala Gln Pro Ala Ala Pro Asn Thr Thr Ser Pro Glu Leu Asn
          20          25          30

Leu Ser His Pro Leu Leu Gly Thr Ala Leu Ala Asn Gly Thr Gly Glu
          35          40          45

Leu Ser Glu His Gln Gln Tyr Val Ile Gly Leu Phe Leu Ser Cys Leu
          50          55          60

Tyr Thr Ile Phe Leu Phe Pro Ile Gly Phe Val Gly Asn Ile Leu Ile
65          70          75          80

Leu Val Val Asn Ile Ser Phe Arg Glu Lys Met Thr Ile Pro Asp Leu
          85          90          95

Tyr Phe Ile Asn Leu Ala Val Ala Asp Leu Ile Leu Val Ala Asp Ser
          100          105          110

Leu Ile Glu Val Phe Asn Leu His Glu Arg Tyr Tyr Asp Ile Ala Val
          115          120          125

Leu Cys Thr Phe Met Ser Leu Phe Leu Gln Val Asn Met Tyr Ser Ser
          130          135          140

Val Phe Phe Leu Thr Trp Met Ser Phe Asp Arg Tyr Ile Ala Leu Ala
145          150          155          160

Arg Ala Met Arg Cys Ser Leu Phe Arg Thr Lys His His Ala Arg Leu
          165          170          175

Ser Cys Gly Leu Ile Trp Met Ala Ser Val Ser Ala Thr Leu Val Pro
          180          185          190

Phe Thr Ala Val His Leu Gln His Thr Asp Glu Ala Cys Phe Cys Phe
          195          200          205

Ala Asp Val Arg Glu Val Gln Trp Leu Glu Val Thr Leu Gly Phe Ile
          210          215          220

Val Pro Phe Ala Ile Ile Gly Leu Cys Tyr Ser Leu Ile Val Arg Val
225          230          235          240

Leu Val Arg Ala His Arg His Arg Gly Leu Arg Pro Arg Arg Gln Lys
          245          250          255

Ala Leu Arg Met Ile Leu Ala Val Val Leu Val Phe Phe Val Cys Trp
          260          265          270

```

10

20

30

Leu Pro Glu Asn Val Phe Ile Ser Val His Leu Leu Gln Arg Thr Gln
 275 280 285
 Pro Gly Ala Ala Pro Cys Lys Gln Ser Phe Arg His Ala His Pro Leu
 290 295 300
 Thr Gly His Ile Val Asn Leu Thr Ala Phe Ser Asn Ser Cys Leu Asn
 305 310 315 320
 Pro Leu Ile Tyr Ser Phe Leu Gly Glu Thr Phe Arg Asp Lys Leu Arg
 325 330 335
 Leu Tyr Ile Glu Gln Lys Thr Asn Leu Pro Ala Leu Asn Arg Phe Cys
 340 345 350
 His Ala Ala Leu Lys Ala Val Ile Pro Asp Ser Thr Glu Gln Ser Asp
 355 360 365
 Val Arg Phe Ser Ser Ala Val
 370 375

10

(72) INFORMATION FOR SEQ ID NO:71:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

20

- (ii) MOLECULE TYPE: DNA (genomic)

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

ACAGAATTCC TGTGTGGTTT TACCGCCCAG

30

(73) INFORMATION FOR SEQ ID NO:72:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

30

- (ii) MOLECULE TYPE: DNA (genomic)

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

CTCGGATCCA GGCAGAAGAG TCGCCTATGG

30

(74) INFORMATION FOR SEQ ID NO:73:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1137 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

```
ATGGACCTGG GGAACCAAT GAAAAGCGTG CTGGTGGTGG CTCTCCTTGT CATTTTCCAG    60
GTATGCCTGT GTCAAGATGA GGTACGGAC GATTACATCG GAGACAACAC CACAGTGGAC    120
TACACTTTGT TCGAGTCTTT GTGCTCCAAG AAGGACGTGC GGAACTTTAA AGCCTGGTTC    180
CTCCCTATCA TGTACTCCAT CATTTGTTTC GTGGGCCTAC TGGGCAATGG GCTGGTCGTG    240
TTGACCTATA TCTATTTCAG GAGGCTCAAG ACCATGACCG ATACCTACCT GCTCAACCTG    300
GCGGTGGCAG ACATCCTCTT CCTCCTGACC CTTCCTTCTT GGGCCTACAG CGCGGCCAAG    360
TCCTGGGTCT TCGGTGTCCA CTTTGTGCAAG CTCATCTTTG CCATCTACAA GATGAGCTTC    420
TTCAGTGGCA TGCTCCTACT TCTTTGCATC AGCATTGACC GCTACGTGGC CATCGTCCAG    480
GCTGTCTCAG CTCACCGCCA CCGTGCCCGC GTCCTTCTCA TCAGCAAGCT GTCCTGTGTG    540
GGCATCTGGA TACTAGCCAC AGTGCTCTCC ATCCCAGAGC TCCTGTACAG TGACCTCCAG    600
AGGAGCAGCA GTGAGCAAGC GATGCGATGC TCTCTCATCA CAGAGCATGT GGAGGCCTTT    660
ATCACCATCC AGGTGGCCCA GATGGTGATC GGCTTTCTGG TCCCCTGCTT GGCCATGAGC    720
TTCTGTTACC TTGTATCAT CCGCACCTCT CTCCAGGCAC GCAACTTTGA GCGCAACAAG    780
GCCATCAAGG TGATCATCGC TGTGGTCGTG GTCTTCATAG TCTTCCAGCT GCCCTACAAT    840
GGGGTGGTCC TGGCCCAGAC GGTGGCCAAC TTCAACATCA CCAGTAGCAC CTGTGAGCTC    900
AGTAAGCAAC TCAACATCGC CTACGACGTC ACCTACAGCC TGGCCTGCGT CCGCTGCTGC    960
GTCAACCCTT TCTTGTACGC CTTTCATCGC GTCAAGTTCC GCAACGATCT CTTCAAGCTC   1020
TTCAAGGACC TGGGCTGCCT CAGCCAGGAG CAGCTCCGGC AGTGGTCTTC CTGTCGGCAC   1080
ATCCGGCGCT CCTCCATGAG TGTGGAGGCC GAGACCACCA CCACCTTCTC CCCATAG    1137
```

(75) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 378 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

10

20

30

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

10

20

30

Ala Asn Phe Asn Ile Thr Ser Ser Thr Cys Glu Leu Ser Lys Gln Leu
 290 295 300

Asn Ile Ala Tyr Asp Val Thr Tyr Ser Leu Ala Cys Val Arg Cys Cys
 305 310 315 320

Val Asn Pro Phe Leu Tyr Ala Phe Ile Gly Val Lys Phe Arg Asn Asp
 325 330 335

Leu Phe Lys Leu Phe Lys Asp Leu Gly Cys Leu Ser Gln Glu Gln Leu
 340 345 350

Arg Gln Trp Ser Ser Cys Arg His Ile Arg Arg Ser Ser Met Ser Val
 355 360 365

Glu Ala Glu Thr Thr Thr Thr Phe Ser Pro
 370 375

10

(76) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 32 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

20

CTGGAATTCA CCTGGACCAC CACCAATGGA TA

32

(77) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

30

CTCGGATCCT GCAAAGTTTG TCATACAGTT

30

(78) INFORMATION FOR SEQ ID NO:77:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1085 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

```

ATGGATATAC AAATGGCAAA CAATTTTACT CCGCCCTCTG CAACTCCTCA GGGAAATGAC    60
TGTGACCTCT ATGCACATCA CAGCACGGCC AGGATAGTAA TGCCTCTGCA TTACAGCCTC   120
GTCTTCATCA TTGGGCTCGT GGGAACTTA CTAGCCTTGG TCGTCATTGT TCAAAACAGG   180
AAAAAATCA ACTCTACCAC CCTCTATTCA ACAAATTTGG TGATTTCTGA TATACTTTT    240
ACCACGGCTT TGCCTACACG AATAGCCTAC TATGCAATGG GCTTTGACTG GAGAATCGGA   300
GATGCCTTGT GTAGGATAAC TGCCTAGTGT TTTTACATCA ACACATATGC AGGTGTGAAC   360
TTTATGACCT GCCTGAGTAT TGACCGCTTC ATTGCTGTGG TGCACCTCT ACGCTACAAC   420
AAGATAAAAA GGATTGAACA TGCAAAAGGC GTGTGCATAT TTGTCTGGAT TCTAGTATTT   480
GCTCAGACAC TCCCACTCCT CATCAACCCT ATGTCAAAGC AGGAGGCTGA AAGGATTACA   540
TGCATGGAGT ATCCAAACTT TGAAGAACT AAATCTCTTC CCTGGATTCT GCTTGGGGCA   600
TGTTTCATAG GATATGTACT TCCACTTATA ATCATTCTCA TCTGCTATTC TCAGATCTGC   660
TGCAAACTCT TCAGAACTGC CAAACAAAAC CCACTCACTG AGAAATCTGG TGTAACAAA   720
AAGGCTCTCA ACACAATTAT TCTTATTATT GTTGTGTTTG TTCTCTGTTT CACACCTTAC   780
CATGTTGCAA TTATTCAACA TATGATTAAG AAGCTTCGTT TCTCTAATTT CCTGGAATGT   840
AGCCAAAGAC ATTCGTTCCA GATTTCTCTG CACTTTACAG TATGCCTGAT GAACTTCAAT   900
TGCTGCATGG ACCCTTTTAT CTAATTCTTT GCATGTAAAG GGTATAAGAG AAAGGTTATG   960
AGGATGCTGA AACGGCAAGT CAGTGTATCG ATTTCTAGTG CTGTGAAGTC AGCCCCTGAA  1020
GAAAATTCAC GTGAAATGAC AGAAACGCAG ATGATGATAC ATTCCAAGTC TTCAAATGGA  1080
AAGTGA                                           1086

```

10

20

(79) INFORMATION FOR SEQ ID NO:78:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 361 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

30

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

```

Met Asp Ile Gln Met Ala Asn Asn Phe Thr Pro Pro Ser Ala Thr Pro
1           5           10           15

```

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

10

20

30

305	310	315	320
Arg Met Leu Lys Arg Gln Val Ser Val Ser Ile Ser Ser Ala Val Lys			
325	330	335	
Ser Ala Pro Glu Glu Asn Ser Arg Glu Met Thr Glu Thr Gln Met Met			
340	345	350	
Ile His Ser Lys Ser Ser Asn Gly Lys			
355	360		

(80) INFORMATION FOR SEQ ID NO:79:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 31 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

CTGGAATTCT CCTGCTCATC CAGCCATGCG G 31

(81) INFORMATION FOR SEQ ID NO:80:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

CCTGGATCCC CACCCCTACT GGGGCCTCAG 30

(82) INFORMATION FOR SEQ ID NO:81:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1446 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

ATGCGGTGGC TGTGGCCCTT GGCTGTCTCT CTGCTGTGA TTTTGGCTGT GGGGCTAAGC 60

AGGGTCTCTG GGGGTGCCCC CCTGCACCTG GGCAGGCACA GAGCCGAGAC CCAGGAGCAG 120

CAGAGCCGAT CCAAGAGGGG CACCGAGGAT GAGGAGGCCA AGGGCGTGCA GCAGTATGTG 180
 CCTGAGGAGT GGGCGGAGTA CCCCCGGCCC ATTCACCCTG CTGGCCTGCA GCCAACCAAG 240
 CCCTTGGTGG CCACCAGCCC TAACCCCGAC AAGGATGGGG GCACCCAGA CAGTGGGCAg 300
 GAACTGAGGG GCAATCTGAC AGGGGCACCA GGGCAGAGGC TACAGATCCA GAACCCCTG 360
 TATCCGGTGA CCGAGAGCTC CTACAGTGCC TATGCCATCA TGCTTCTGGC GCTGGTGGTG 420
 TTTGCGGTGG GCATTGTGGG CAACCTGTCG GTCATGTGCA TCGTGTGGCA CAGCTACTAC 480
 CTGAAGAGCG CCTGGAATC CATCCTTGCC AGCCTGGCCC TCTGGGATTT TCTGGTCCTC 540
 TTTTCTGCC TCCCTATTGT CATCTTCAAC GAGATCACCA AGCAGAGGCT ACTGGGTGAC 600
 GTTCTTGTC GTGCCGTGCC CTTTCATGGAG GTCTCCTCTC TGGGAGTCAC GACTTTCAGC 660
 CTCTGTGCCC TGGGCATTGA CCGCTTCCAC GTGGCCACCA GCACCCTGCC CAAGGTGAGG 720
 CCCATCGAGC GGTGCCAATC CATCCTGGCC AAGTTGGCTG TCATCTGGGT GGGCTCCATG 780
 ACGCTGGCTG TGCCTGAGCT CCTGCTGTGG CAGCTGGCAC AGGAGCCTGC CCCACCATG 840
 GGCACCCTGG ACTCATGCAT CATGAAACCC TCAGCCAGCC TGCCCGAGTC CCTGTATTCA 900
 CTGGTGATGA CCTACCAGAA CGCCCGCATG TGGTGGTACT TTGGCTGCTA CTTCTGCCTG 960
 CCCATCCTCT TCACAGTCAC CTGCCAGCTG GTGACATGGC GGGTGCGAGG CCCTCCAGGG 1020
 AGGAAGTCAG AGTGCAGGGC CAGCAAGCAC GAGCAGTGTG AGAGCCAGCT CAACAGCACC 1080
 GTGGTGGGCC TGACCGTGGT CTACGCCTTC TGACCCCTCC CAGAGAACGT CTGCAACATC 1140
 GTGGTGGCCT ACCTCTCCAC CGAGCTGACC CGCCAGACCC TGGACCTCCT GGGCCTCATC 1200
 AACCAGTTCT CCACCTTCTT CAAGGGCGCC ATCACCCAG TGCTGCTCCT TTGCATCTGC 1260
 AGGCCGCTGG GCCAGGCCTT CCTGGACTGC TGCTGCTGCT GCTGCTGTGA GGAGTGCGGC 1320
 GGGGCTTCGG AGGCCTCTGC TGCCAATGGG TCGGACAACA AGCTCAAGAC CGAGGTGTCC 1380
 TCTTCCATCT ACTTCCACAA GCCCAGGGAG TCACCCCCAC TCCTGCCCCT GGGCACACCT 1440
 TGCTGA 1446

10

20

30

(83) INFORMATION FOR SEQ ID NO:82:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 481 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant
- (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

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

10

20

30

275 280 285
 Lys Pro Ser Ala Ser Leu Pro Glu Ser Leu Tyr Ser Leu Val Met Thr
 290 295 300
 Tyr Gln Asn Ala Arg Met Trp Trp Tyr Phe Gly Cys Tyr Phe Cys Leu
 305 310 315 320
 Pro Ile Leu Phe Thr Val Thr Cys Gln Leu Val Thr Trp Arg Val Arg
 325 330 335
 Gly Pro Pro Gly Arg Lys Ser Glu Cys Arg Ala Ser Lys His Glu Gln
 340 345 350
 Cys Glu Ser Gln Leu Asn Ser Thr Val Val Gly Leu Thr Val Val Tyr
 355 360 365
 Ala Phe Cys Thr Leu Pro Glu Asn Val Cys Asn Ile Val Val Ala Tyr
 370 375 380
 Leu Ser Thr Glu Leu Thr Arg Gln Thr Leu Asp Leu Leu Gly Leu Ile
 385 390 395 400
 Asn Gln Phe Ser Thr Phe Phe Lys Gly Ala Ile Thr Pro Val Leu Leu
 405 410 415
 Leu Cys Ile Cys Arg Pro Leu Gly Gln Ala Phe Leu Asp Cys Cys Cys
 420 425 430
 Cys Cys Cys Cys Glu Glu Cys Gly Gly Ala Ser Glu Ala Ser Ala Ala
 435 440 445
 Asn Gly Ser Asp Asn Lys Leu Lys Thr Glu Val Ser Ser Ser Ile Tyr
 450 455 460
 Phe His Lys Pro Arg Glu Ser Pro Pro Leu Leu Pro Leu Gly Thr Pro
 465 470 475 480
 Cys

10

20

(84) INFORMATION FOR SEQ ID NO:83:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

ATGTGGAACG CGACGCCAG CG

22

(85) INFORMATION FOR SEQ ID NO:84:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

TCATGTATTA ATACTAGATT CT

22

(86) INFORMATION FOR SEQ ID NO:85:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 38 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

TACCATGTGG AACGCGACGC CCAGCGAAGA GCCGGGGT

38

(87) INFORMATION FOR SEQ ID NO:86:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 39 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

CGGAATTCAT GTATTAATAC TAGATTCTGT CCAGGCCCG

39

(88) INFORMATION FOR SEQ ID NO:87:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1101 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

10

20

30

ATGTGGAACG CGACGCCCAG CGAAGAGCCG GGGTTCAACC TCACACTGGC CGACCTGGAC 60
 TGGGATGCTT CCCCCGGCAA CGACTCGCTG GGCACGAGC TGCTGCAGCT CTTCCCCGCG 120
 CCGCTGCTGG CCGGCGTCAC AGCCACCTGC GTGGCACTCT TCGTGGTGGG TATCGCTGGC 180
 AACCTGCTCA CCATGCTGGT GGTGTCGCGC TTCCGCGAGC TGCGCACCAC CACCAACCTC 240
 TACCTGTCCA GCATGGCCTT CTCCGATCTG CTCATCTTCC TCTGCATGCC CCTGGACCTC 300
 GTTCGCCTCT GGCAGTACCG GCCTGGAAC TTCGGCGACC TCCTCTGCAA ACTCTTCCAA 360
 TTCGTAGTG AGAGCTGCAC CTACGCCACG GTGCTACCA TCACAGCGCT GAGCGTCGAG 420
 CGCTACTTCG CCATCTGCTT CCCACTCCGG GCCAAGGTGG TGGTCACCAA GGGGCGGGTG 480
 AAGCTGGTCA TCTTCGTCAT CTGGGCCGTG GCCTTCTGCA GCGCCGGGCC CATCTTCGTG 540
 CTAGTCGGGG TGGAGCACGA GAACGGCACC GACCCTTGGG ACACCAACGA GTGCCGCCCC 600
 ACCGAGTTTG CGGTGCGCTC TGGACTGCTC ACGGTCATGG TGTGGGTGTC CAGCATCTTC 660
 TTCTTCCTTC CTGTCTTCTG TCTACGGTC CTCTACAGTC TCATCGGCAG GAAGCTGTGG 720
 CGGAGGAGGC GCGGCGATGC TGTCGTGGGT GCCTCGCTCA GGGACCAGAA CCACAAGCAA 780
 ACCGTGAAAA TGCTGGCTGT AGTGGTGTTC GCCTTCATCC TCTGCTGGCT CCCCTTCCAC 840
 GTAGGCGCAT ATTTATTTTC CAAATCCTTT GAGCCTGGCT CCTTGGAGAT TGCTCAGATC 900
 AGCCAGTACT GCAACCTCGT GTCCTTTGTC CTCTTCTACC TCAGTGCTGC CATCAACCCC 960
 ATTCTGTACA ACATCATGTC CAAGAAGTAC CGGGTGGCAG TGTTCAGACT TCTGGGATTC 1020
 GAACCCCTTCT CCCAGAGAAA GCTCTCCACT CTGAAAGATG AAAGTTCTCG GGCCTGGACA 1080
 GAATCTAGTA TTAATACATG A 1101

(89) INFORMATION FOR SEQ ID NO:88:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 366 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

Met Trp Asn Ala Thr Pro Ser Glu Glu Pro Gly Phe Asn Leu Thr Leu
 1 5 10 15
 Ala Asp Leu Asp Trp Asp Ala Ser Pro Gly Asn Asp Ser Leu Gly Asp
 20 25 30

10

20

30

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

10

20

30

325 330 335
Leu Leu Gly Phe Glu Pro Phe Ser Gln Arg Lys Leu Ser Thr Leu Lys
340 345 350
Asp Glu Ser Ser Arg Ala Trp Thr Glu Ser Ser Ile Asn Thr
355 360 365

(90) INFORMATION FOR SEQ ID NO:89:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

GCAAGCTTGT GCCCTCACCA AGCCATGCCA GCC 33

(91) INFORMATION FOR SEQ ID NO:90:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

CGGAATTCAG CAATGAGTTC CGACAGAAGC 30

(92) INFORMATION FOR SEQ ID NO:91:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1842 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

ATGCGAGCCC CGGGCGCGCT TCTCGCCCGC ATGTCGCGGC TACTGCTTCT GCTACTGCTC 60
AAGGTGTCTG CCTCTTCTGC CCTCGGGGTC GCCCCTGCGT CCAGAAACGA AACTTGTCTG 120
GGGAGAGCT GTGCACCTAC AGTGATCCAG CGCCGCGGCA GGGACGCCTG GGGACCGGGA 180
AATTCTGCAA GAGACGTTCT GCGAGCCCGA GCACCCAGGG AGGAGCAGGG GGCAGCGTTT 240

CTTGCGGGAC CCTCCTGGGA CCTGCCGGCG GCGCCGGGCC GTGACCCGGC TGCAGGCAGA 300
 GGGGCGGAGG CGTCGGCAGC CGGACCCCCG GGACCTCCAA CCAGGCCACC TGGCCCCTGG 360
 AGGTGGAAAG GTGCTCGGGG TCAGGAGCCT TCTGAAACTT TGGGGAGAGG GAACCCACG 420
 GCCCTCCAGC TCTTCCTTCA GATCTCAGAG GAGGAAGAGA AGGGTCCCAG AGGCGCTGGC 480
 ATTTCCGGGC GTAGCCAGGA GCAGAGTGTG AAGACAGTCC CCGGAGCCAG CGATCTTTTT 540
 TACTGGCCAA GGAGAGCCGG GAAACTCCAG GGTTCACC ACCAAGCCCT GTCCAAGACG 600
 GCCAATGGAC TGGCGGGGCA CGAAGGGTGG ACAATTGCAC TCCCGGGCCG GCGCTGGCC 660
 CAGAAATGGAT CCTTGGGTGA AGGAATCCAT GAGCCTGGGG GTCCCCGCCG GGGAAACAGC 720
 ACGAACCAGC GTGTGAGACT GAAGAACCCC TTCTACCCGC TGACCCAGGA GTCCTATGGA 780
 GCCTACGCGG TCATGTGTCT GTCCGTGGTG ATCTTCGGGA CCGGCATCAT TGGCAACCTG 840
 GCGGTGATGA GCATCGTGTG CCACAACCTAC TACATGCGGA GCATCTCCAA CTCCCTCTTG 900
 GCCAACCTGG CCTTCTGGGA CTTTCTCATC ATCTTCTTCT GCCTTCGCT GGTCTCTTC 960
 CACGAGCTGA CCAAGAAGTG GCTGCTGGAG GACTTCTCCT GCAAGATCGT GCCCTATATA 1020
 GAGGTCGCTT CTCTGGGAGT CACCCTTTC ACCTTATGTG CTCTGTGCAT AGACCGCTTC 1080
 CGTGCTGCCA CCAACGTACA GATGTACTAC GAAATGATCG AAACTGTTC CTCAACAACT 1140
 GCCAACTTG CTGTTATATG GGTGGGAGCT CTATTGTTAG CACTTCAGA AGTTGTTCTC 1200
 CGCCAGCTGA GCAAGGAGGA TTTGGGTTT AGTGGCCGAG CTCCGCAGA AAGGTGCATT 1260
 ATTAAGATCT CTCCTGATTT ACCAGACACC ATCTATGTTT TAGCCCTCAC CTACGACAGT 1320
 GCGAGACTGT GGTGGTATTT TGGCTGTAC TTTTGTGTC CCACGCTTTT CACCATCACC 1380
 TGCTCTCTAG TGAAGTGCAG GAAATCCGC AAAGCAGAGA AAGCCTGTAC CCGAGGGAAT 1440
 AAACGGCAGA TTCAACTAGA GAGTCAGATG AACTGTACAG TAGTGGCACT GACCATTTTA 1500
 TATGGATTTT GCATTATTCC TGAAATATC TGCAACATTG TTAGTGCTTA CATGGCTACA 1560
 GGGGTTTCAC AGCAGACAAT GGACCTCCTT AATATCATCA GCCAGTTCCT TTTGTTCTTT 1620
 AAGTCCTGTG TCACCCAGT CCTCCTTTTC TGTCTCTGCA AACCTTCAG TCGGGCCTTC 1680
 ATGGAGTGCT GCTGCTGTTG CTGTGAGGAA TGCAATCAGA AGTCTTCAAC GGTGACCAGT 1740
 GATGACAATG ACAACGAGTA CACCACGGAA CTCGAACTCT CGCCTTTCAG TACCATACGC 1800
 CGTGAAATGT CCACTTTTGC TTCTGTGCGA ACTCATTGCT GA 1842

10

20

30

(93) INFORMATION FOR SEQ ID NO:92:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 613 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

```

Met Arg Ala Pro Gly Ala Leu Leu Ala Arg Met Ser Arg Leu Leu Leu
1           5           10           15

Leu Leu Leu Leu Lys Val Ser Ala Ser Ser Ala Leu Gly Val Ala Pro
           20           25           30

Ala Ser Arg Asn Glu Thr Cys Leu Gly Glu Ser Cys Ala Pro Thr Val
           35           40           45

Ile Gln Arg Arg Gly Arg Asp Ala Trp Gly Pro Gly Asn Ser Ala Arg
           50           55           60

Asp Val Leu Arg Ala Arg Ala Pro Arg Glu Glu Gln Gly Ala Ala Phe
65           70           75           80

Leu Ala Gly Pro Ser Trp Asp Leu Pro Ala Ala Pro Gly Arg Asp Pro
           85           90           95

Ala Ala Gly Arg Gly Ala Glu Ala Ser Ala Ala Gly Pro Pro Gly Pro
           100          105          110

Pro Thr Arg Pro Pro Gly Pro Trp Arg Trp Lys Gly Ala Arg Gly Gln
           115          120          125

Glu Pro Ser Glu Thr Leu Gly Arg Gly Asn Pro Thr Ala Leu Gln Leu
           130          135          140

Phe Leu Gln Ile Ser Glu Glu Glu Glu Lys Gly Pro Arg Gly Ala Gly
145          150          155          160

Ile Ser Gly Arg Ser Gln Glu Gln Ser Val Lys Thr Val Pro Gly Ala
           165          170          175

Ser Asp Leu Phe Tyr Trp Pro Arg Arg Ala Gly Lys Leu Gln Gly Ser
           180          185          190

His His Lys Pro Leu Ser Lys Thr Ala Asn Gly Leu Ala Gly His Glu
           195          200          205

Gly Trp Thr Ile Ala Leu Pro Gly Arg Ala Leu Ala Gln Asn Gly Ser
           210          215          220

Leu Gly Glu Gly Ile His Glu Pro Gly Gly Pro Arg Arg Gly Asn Ser
225          230          235          240

```

10

20

30

Thr Asn Arg Arg Val Arg Leu Lys Asn Pro Phe Tyr Pro Leu Thr Gln
 245 250 255
 Glu Ser Tyr Gly Ala Tyr Ala Val Met Cys Leu Ser Val Val Ile Phe
 260 265 270
 Gly Thr Gly Ile Ile Gly Asn Leu Ala Val Met Ser Ile Val Cys His
 275 280 285
 Asn Tyr Tyr Met Arg Ser Ile Ser Asn Ser Leu Leu Ala Asn Leu Ala
 290 295 300
 Phe Trp Asp Phe Leu Ile Ile Phe Phe Cys Leu Pro Leu Val Ile Phe
 305 310 315 320
 His Glu Leu Thr Lys Lys Trp Leu Leu Glu Asp Phe Ser Cys Lys Ile
 325 330 335
 Val Pro Tyr Ile Glu Val Ala Ser Leu Gly Val Thr Thr Phe Thr Leu
 340 345 350
 Cys Ala Leu Cys Ile Asp Arg Phe Arg Ala Ala Thr Asn Val Gln Met
 355 360 365
 Tyr Tyr Glu Met Ile Glu Asn Cys Ser Ser Thr Thr Ala Lys Leu Ala
 370 375 380
 Val Ile Trp Val Gly Ala Leu Leu Leu Ala Leu Pro Glu Val Val Leu
 385 390 395 400
 Arg Gln Leu Ser Lys Glu Asp Leu Gly Phe Ser Gly Arg Ala Pro Ala
 405 410 415
 Glu Arg Cys Ile Ile Lys Ile Ser Pro Asp Leu Pro Asp Thr Ile Tyr
 420 425 430
 Val Leu Ala Leu Thr Tyr Asp Ser Ala Arg Leu Trp Trp Tyr Phe Gly
 435 440 445
 Cys Tyr Phe Cys Leu Pro Thr Leu Phe Thr Ile Thr Cys Ser Leu Val
 450 455 460
 Thr Ala Arg Lys Ile Arg Lys Ala Glu Lys Ala Cys Thr Arg Gly Asn
 465 470 475 480
 Lys Arg Gln Ile Gln Leu Glu Ser Gln Met Asn Cys Thr Val Val Ala
 485 490 495
 Leu Thr Ile Leu Tyr Gly Phe Cys Ile Ile Pro Glu Asn Ile Cys Asn
 500 505 510
 Ile Val Thr Ala Tyr Met Ala Thr Gly Val Ser Gln Gln Thr Met Asp
 515 520 525
 Leu Leu Asn Ile Ile Ser Gln Phe Leu Leu Phe Phe Lys Ser Cys Val

10

20

30

530 535 540
 Thr Pro Val Leu Leu Phe Cys Leu Cys Lys Pro Phe Ser Arg Ala Phe
 545 550 555 560
 Met Glu Cys Cys Cys Cys Cys Glu Glu Cys Ile Gln Lys Ser Ser
 565 570 575
 Thr Val Thr Ser Asp Asp Asn Asp Asn Glu Tyr Thr Thr Glu Leu Glu
 580 585 590
 Leu Ser Pro Phe Ser Thr Ile Arg Arg Glu Met Ser Thr Phe Ala Ser
 595 600 605
 Val Gly Thr His Cys
 610

10

(94) INFORMATION FOR SEQ ID NO:93:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

CAGAATTCAG AGAAAAAAG TGAATATGGT TTTT

34

20

(95) INFORMATION FOR SEQ ID NO:94:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 32 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

TTGGATCCCT GGTGCATAAC AATTGAAAGA AT

32

30

(96) INFORMATION FOR SEQ ID NO:95:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1248 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

```

ATGGTTTTTTG CTCACAGAAT GGATAACAGC AAGCCACATT TGATTATTCC TACTTCTCTG    60
GTGCCCCCTCC AAAACCGCAG CTGCACTGAA ACAGCCACAC CTCTGCCAAG CCAATACCTG    120
ATGGAATTAA GTGAGGAGCA CAGTTGGATG AGCAACCAAA CAGACCTTCA CTATGTGCTG    180
AAACCCGGGG AAGTGGCCAC AGCCAGCATC TTCTTTGGGA TTCTGTGGTT GTTTTCTATC    240
TTCGGCAATT CCCTGGTTTG TTTGGTCATC CATAGGAGTA GGAGGACTCA GTCTACCACC    300
AACTACTTTG TGGTCTCCAT GGCATGTGCT GACCTTCTCA TCAGCGTTGC CAGCACGCCCT    360
TTCGTCTCTG TCCAGTTCAC CACTGGAAGG TGGACGCTGG GTAGTGCAAC GTGCAAGGTT    420
GTGCGATATT TTCAATATCT CACTCCAGGT GTCCAGATCT ACGTTCTCCT CTCCATCTGC    480
ATAGACCGGT TCTACACCAT CGTCTATCCT CTGAGCTTCA AGGTGTCCAG AGAAAAAGCC    540
AAGAAAATGA TTGCGGCATC GTGGATCTTT GATGCAGGCT TTGTGACCCC TGTGCTCTTT    600
TTCTATGGCT CCAACTGGGA CAGTCATTGT AACTATTTCC TCCCCTCCTC TTGGGAAGGC    660
ACTGCCTACA CTGTCATCCA CTTCTTGGTG GGCTTTGTGA TTCCATCTGT CCTCATAATT    720
TTATTTTACC AAAAGGTCAT AAAATATATT TGGAGAATAG GCACAGATGG CCGAACGGTG    780
AGGAGGACAA TGAACATTGT CCCTCGGACA AAAGTGAAAA CTATCAAGAT GTTCCTCATT    840
TTAAATCTGT TGTTTTTGCT CTCCTGGCTG CCTTTTCATG TAGCTCAGCT ATGGCACCCC    900
CATGAACAAG ACTATAAGAA AAGTTCCCTT GTTTTCACAG CTATCACATG GATATCCTTT    960
AGTTCTTCAG CCTCTAAACC TACTCTGTAT TCAATTTATA ATGCCAATTT TCGGAGAGGG    1020
ATGAAAGAGA CTTTTTGCAT GTCCTCTATG AAATGTTACC GAAGCAATGC CTATACTATC    1080
ACAACAAGTT CAAGGATGGC CAAAAAAAAC TACGTTGGCA TTTCAGAAAT CCCTTCCATG    1140
GCCAAAATA TTACCAAAGA CTCGATCTAT GACTCATTG ACAGAGAAGC CAAGGAAAAA    1200
AAGCTTGCTT GGCCCATTA CTCAAATCCA CCAAATACTT TTGTCTAA    1248

```

10

20

30

(97) INFORMATION FOR SEQ ID NO:96:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 415 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

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

10

20

30

Trp Leu Pro Phe His Val Ala Gln Leu Trp His Pro His Glu Gln Asp
 290 295 300
 Tyr Lys Lys Ser Ser Leu Val Phe Thr Ala Ile Thr Trp Ile Ser Phe
 305 310 315 320
 Ser Ser Ser Ala Ser Lys Pro Thr Leu Tyr Ser Ile Tyr Asn Ala Asn
 325 330 335
 Phe Arg Arg Gly Met Lys Glu Thr Phe Cys Met Ser Ser Met Lys Cys
 340 345 350
 Tyr Arg Ser Asn Ala Tyr Thr Ile Thr Thr Ser Ser Arg Met Ala Lys
 355 360 365
 Lys Asn Tyr Val Gly Ile Ser Glu Ile Pro Ser Met Ala Lys Thr Ile
 370 375 380
 Thr Lys Asp Ser Ile Tyr Asp Ser Phe Asp Arg Glu Ala Lys Glu Lys
 385 390 395 400
 Lys Leu Ala Trp Pro Ile Asn Ser Asn Pro Pro Asn Thr Phe Val
 405 410 415

10

(98) INFORMATION FOR SEQ ID NO:97:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

GGAAAGCTTA ACGATCCCCA GGAGCAACAT

30

(99) INFORMATION FOR SEQ ID NO:98:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

CTGGGATCCT ACGAGAGCAT TTTTCACACA G

31

(100) INFORMATION FOR SEQ ID NO:99:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1842 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

ATGGGGCCCA CCCTAGCGGT TCCCACCCCC TATGGCTGTA TTGGCTGTAA GCTACCCCAG	60
CCAGAATACC CACCGGCTCT AATCATCTTT ATGTTCTGCG CGATGGTTAT CACCATCGTT	120
GTAGACCTAA TCGGCAACTC CATGGTCATT TTGGCTGTGA CGAAGAACAA GAAGCTCCGG	180
AATCTGGCA ACATCTTCGT GGTCACTCTC TCTGTGGCCG ATATGCTGGT GGCCATCTAC	240
CCATACCCCT TGATGCTGCA TGCCATGTCC ATTGGGGGCT GGGATCTGAG CCAGTTACAG	300
TGCCAGATGG TCGGGTTCAT CACAGGGCTG AGTGTGGTCG GCTCCATCTT CAACATCGTG	360
GCAATCGCTA TCAACCGTTA CTGCTACATC TGCCACAGCC TCCAGTACGA ACGGATCTTC	420
AGTGTGCGCA ATACCTGCAT CTACCTGGTC ATCACCTGGA TCATGACCGT CCTGGCTGTC	480
CTGCCCAACA TGTACATTGG CACCATCGAG TACGATCCTC GCACCTACAC CTGCATCTTC	540
AACTATCTGA ACAACCCTGT CTTCACTGTT ACCATCGTCT GCATCCACTT CGTCTCCCT	600
CTCCTCATCG TGGGTTTCTG CTACGTGAGG ATCTGGACCA AAGTGCTGGC GGCCCGTGAC	660
CCTGCAGGGC AGAATCCTGA CAACCAACTT GCTGAGGTTC GCAATTTTCT AACCATGTTT	720
GTGATCTTCC TCCTCTTTCG AGTGTGCTGG TGCCCTATCA ACGTGCTCAC TGTCTTGGTG	780
GCTGTCAGTC CGAAGGAGAT GGCAGGCAAG ATCCCCAACT GGCTTTATCT TGCAGCCTAC	840
TTCATAGCCT ACTTCAACAG CTGCCTCAAC GCTGTGATCT ACGGGCTCCT CAATGAGAAT	900
TTCCGAAGAG AATACTGGAC CATCTTCCAT GCTATGCGGC ACCCTATCAT ATTCTTCCCT	960
GGCCTCATCA GTGATATTCG TGAGATGCAG GAGGCCCCGT CCCTGGCCCCG CGCCCGTGCC	1020
CATGCTCGCG ACCAAGCTCG TGAACAAGAC CGTGCCCATG CCTGTCCTGC TGTGGAGGAA	1080
ACCCCGATGA ATGTCCGGAA TGTTCCATTA CCTGGTGATG CTGCAGCTGG CCACCCCGAC	1140
CGTGCCCTCTG GCCACCCTAA GCCCCATTCC AGATCCTCCT CTGCCTATCG CAAATCTGCC	1200
TCTACCCACC ACAAGTCTGT CTTTAGCCAC TCCAAGGCTG CCTCTGGTCA CCTCAAGCCT	1260
GTCTCTGGCC ACTCCAAGCC TGCCTCTGGT CACCCCAAGT CTGCCACTGT CTACCCTAAG	1320
CCTGCCTCTG TCCATTTCAA GGGTGACTCT GTCCATTTC AAGGTGACTC TGTCCATTTT	1380

10

20

30

AAGCCTGACT CTGTTTCATTT CAAGCCTGCT TCCAGCAACC CCAAGCCCAT CACTGGCCAC 1440
 CATGTCTCTG CTGGCAGCCA CTCCAAGTCT GCCTTCAGTG CTGCCACCAG CCACCCTAAA 1500
 CCCATCAAGC CAGCTACCAG CCATGCTGAG CCCACCACTG CTGACTATCC CAAGCCTGCC 1560
 ACTACCAGCC ACCCTAAGCC CGCTGCTGCT GACAACCCTG AGCTCTCTGC CTCCCATTGC 1620
 CCGGAGATCC CTGCCATTGC CCACCCTGTG TCTGACGACA GTGACCTCCC TGAGTCGGCC 1680
 TCTAGCCCTG CCGCTGGGGCC CACCAAGCCT GCTGCCAGCC AGCTGGAGTC TGACACCATC 1740
 GCTGACCTTC CTGACCCTAC TGTAGTCACT ACCAGTACCA ATGATTACCA TGATGTCGTG 1800
 GTTGTGATG TTGAAGATGA TCCTGATGAA ATGGCTGTGT GA 1842

10

(101) INFORMATION FOR SEQ ID NO:100:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 613 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

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

20

30

Thr Cys Ile Tyr Leu Val Ile Thr Trp Ile Met Thr Val Leu Ala Val
 145 150 155 160
 Leu Pro Asn Met Tyr Ile Gly Thr Ile Glu Tyr Asp Pro Arg Thr Tyr
 165 170 175
 Thr Cys Ile Phe Asn Tyr Leu Asn Asn Pro Val Phe Thr Val Thr Ile
 180 185 190
 Val Cys Ile His Phe Val Leu Pro Leu Leu Ile Val Gly Phe Cys Tyr
 195 200 205
 Val Arg Ile Trp Thr Lys Val Leu Ala Ala Arg Asp Pro Ala Gly Gln
 210 215 220
 Asn Pro Asp Asn Gln Leu Ala Glu Val Arg Asn Phe Leu Thr Met Phe
 225 230 235 240
 Val Ile Phe Leu Leu Phe Ala Val Cys Trp Cys Pro Ile Asn Val Leu
 245 250 255
 Thr Val Leu Val Ala Val Ser Pro Lys Glu Met Ala Gly Lys Ile Pro
 260 265 270
 Asn Trp Leu Tyr Leu Ala Ala Tyr Phe Ile Ala Tyr Phe Asn Ser Cys
 275 280 285
 Leu Asn Ala Val Ile Tyr Gly Leu Leu Asn Glu Asn Phe Arg Arg Glu
 290 295 300
 Tyr Trp Thr Ile Phe His Ala Met Arg His Pro Ile Ile Phe Phe Pro
 305 310 315 320
 Gly Leu Ile Ser Asp Ile Arg Glu Met Gln Glu Ala Arg Thr Leu Ala
 325 330 335
 Arg Ala Arg Ala His Ala Arg Asp Gln Ala Arg Glu Gln Asp Arg Ala
 340 345 350
 His Ala Cys Pro Ala Val Glu Glu Thr Pro Met Asn Val Arg Asn Val
 355 360 365
 Pro Leu Pro Gly Asp Ala Ala Ala Gly His Pro Asp Arg Ala Ser Gly
 370 375 380
 His Pro Lys Pro His Ser Arg Ser Ser Ser Ala Tyr Arg Lys Ser Ala
 385 390 395 400
 Ser Thr His His Lys Ser Val Phe Ser His Ser Lys Ala Ala Ser Gly
 405 410 415
 His Leu Lys Pro Val Ser Gly His Ser Lys Pro Ala Ser Gly His Pro
 420 425 430
 Lys Ser Ala Thr Val Tyr Pro Lys Pro Ala Ser Val His Phe Lys Gly

10

20

30

435 440 445
 Asp Ser Val His Phe Lys Gly Asp Ser Val His Phe Lys Pro Asp Ser
 450 455 460
 Val His Phe Lys Pro Ala Ser Ser Asn Pro Lys Pro Ile Thr Gly His
 465 470 475 480
 His Val Ser Ala Gly Ser His Ser Lys Ser Ala Phe Ser Ala Ala Thr
 485 490 495
 Ser His Pro Lys Pro Ile Lys Pro Ala Thr Ser His Ala Glu Pro Thr
 500 505 510
 Thr Ala Asp Tyr Pro Lys Pro Ala Thr Thr Ser His Pro Lys Pro Ala
 515 520 525
 Ala Ala Asp Asn Pro Glu Leu Ser Ala Ser His Cys Pro Glu Ile Pro
 530 535 540
 Ala Ile Ala His Pro Val Ser Asp Asp Ser Asp Leu Pro Glu Ser Ala
 545 550 555 560
 Ser Ser Pro Ala Ala Gly Pro Thr Lys Pro Ala Ala Ser Gln Leu Glu
 565 570 575
 Ser Asp Thr Ile Ala Asp Leu Pro Asp Pro Thr Val Val Thr Thr Ser
 580 585 590
 Thr Asn Asp Tyr His Asp Val Val Val Val Asp Val Glu Asp Asp Pro
 595 600 605
 Asp Glu Met Ala Val
 610

10

20

(102) INFORMATION FOR SEQ ID NO:101:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 32 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

TCCAAGCTTC GCCATGGGAC ATAACGGGAG CT

32

(103) INFORMATION FOR SEQ ID NO:102:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

CGTGAATTCC AAGAATTTAC AATCCTTGCT 30

(104) INFORMATION FOR SEQ ID NO:103:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1548 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

ATGGGACATA ACGGGAGCTG GATCTCTCCA AATGCCAGCG AGCCGCACAA CGCGTCCGGC 60
 GCCGAGGCTG CGGGTGTGAA CCGCAGCGCG CTCGGGGAGT TCGGCGAGGC GCAGCTGTAC 120
 CGCCAGTTCA CCACCACCGT GCAGGTCGTC ATCTTCATAG GCTCGCTGCT CGGAAACTTC 180
 ATGGTGTAT GGTCAACTTG CCGCACAACC GTGTCAAAT CTGTCACCAA CAGGTCATT 240
 AAAAACCTGG CCTGCTCGGG GATTGTGCC AGCCTGGTCT GTGTGCCCTT CGACATCATC 300
 CTCAGCACCA GTCCTCACTG TTGCTGGTGG ATCTACACCA TGCTCTTCTG CAAGGTCGTC 360
 AAATTTTTC ACAAAGTATT CTGCTCTGTG ACCATCCTCA GCTTCCCTGC TATTGCTTTG 420
 GACAGGTA CTCTAGTCTT CTATCCACTG GAGAGGAAAA TATCTGATGC CAAGTCCCGT 480
 GAACTGGTGA TGTACATCTG GGCCCATGCA GTGGTGGCCA GTGTCCCTGT GTTTGCAGTA 540
 ACCAATGTGG CTGACATCTA TGCCACGTCC ACCTGCACGG AAGTCTGGAG CAACTCCTTG 600
 GGCCACCTGG TGTACGTTCT GGTGTATAAC ATCACCACGG TCATTGTGCC TGTGGTGGTG 660
 GTGTTCTCT TCTTGATACT GATCCGACGG GCCCTGAGTG CCAGCCAGAA GAAGAAGGTC 720
 ATCATAGCAG CGCTCCGGAC CCCACAGAAC ACCATCTCTA TTCCCTATGC CTCCAGCGG 780
 GAGGCCGAGC TGCACGCCAC CCTGCTCTCC ATGGTGATGG TCTTCATCTT GTGTAGCGTG 840
 CCCTATGCCA CCCTGGTCGT CTACCAGACT GTGCTCAATG TCCCTGACAC TTCCGTCTTC 900
 TTGCTGCTCA CTGCTGTTTG GCTGCCCAAA GTCTCCCTGC TGGCAAACCC TGTTCCTTTT 960
 CTTACTGTGA ACAAATCTGT CCGCAAGTGC TTGATAGGGA CCCTGGTGCA ACTACACCAC 1020
 CGGTACAGTC GCCGTAATGT GGTCAGTACA GGGAGTGGCA TGGCTGAGGC CAGCCTGGAA 1080

10

20

30

CCCAGCATAC GCTCGGGTAG CCAGCTCCTG GAGATGTTCC ACATTGGGCA GCAGCAGATC 1140
 TTTAAGCCCA CAGAGGATGA GGAAGAGAGT GAGGCCAAGT ACATTGGCTC AGCTGACTTC 1200
 CAGGCCAAGG AGATATTTAG CACCTGCCTC GAGGGAGAGC AGGGGCCACA GTTTGCGCCC 1260
 TCTGCCCCAC CCCTGAGCAC AGTGGACTCT GTATCCCAGG TGGCACCAGC AGCCCTGTG 1320
 GAACCTGAAA CATTCCCTGA TAAGTATTCC CTGCAGTTTG GCTTTGGGCC TTTTGAGTTG 1380
 CCTCCTCAGT GGCTCTCAGA GACCCGAAAC AGCAAGAAGC GGCTGCTTCC CCCCTTGGGC 1440
 AACACCCAG AAGAGCTGAT CCAGACAAAG GTGCCCAAGG TAGGCAGGGT GGAGCGGAAG 1500
 ATGAGCAGAA ACAATAAAGT GAGCATT TTTT CCAAAGGTGG ATTCCTAG 1548

10

(105) INFORMATION FOR SEQ ID NO:104:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 515 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

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

20

30

Ser Val Leu Tyr Pro Leu Glu Arg Lys Ile Ser Asp Ala Lys Ser Arg
 145 150 155 160
 Glu Leu Val Met Tyr Ile Trp Ala His Ala Val Val Ala Ser Val Pro
 165 170 175
 Val Phe Ala Val Thr Asn Val Ala Asp Ile Tyr Ala Thr Ser Thr Cys
 180 185 190
 Thr Glu Val Trp Ser Asn Ser Leu Gly His Leu Val Tyr Val Leu Val
 195 200 205
 Tyr Asn Ile Thr Thr Val Ile Val Pro Val Val Val Val Phe Leu Phe
 210 215 220
 Leu Ile Leu Ile Arg Arg Ala Leu Ser Ala Ser Gln Lys Lys Lys Val
 225 230 235 240
 Ile Ile Ala Ala Leu Arg Thr Pro Gln Asn Thr Ile Ser Ile Pro Tyr
 245 250 255
 Ala Ser Gln Arg Glu Ala Glu Leu His Ala Thr Leu Leu Ser Met Val
 260 265 270
 Met Val Phe Ile Leu Cys Ser Val Pro Tyr Ala Thr Leu Val Val Tyr
 275 280 285
 Gln Thr Val Leu Asn Val Pro Asp Thr Ser Val Phe Leu Leu Leu Thr
 290 295 300
 Ala Val Trp Leu Pro Lys Val Ser Leu Leu Ala Asn Pro Val Leu Phe
 305 310 315 320
 Leu Thr Val Asn Lys Ser Val Arg Lys Cys Leu Ile Gly Thr Leu Val
 325 330 335
 Gln Leu His His Arg Tyr Ser Arg Arg Asn Val Val Ser Thr Gly Ser
 340 345 350
 Gly Met Ala Glu Ala Ser Leu Glu Pro Ser Ile Arg Ser Gly Ser Gln
 355 360 365
 Leu Leu Glu Met Phe His Ile Gly Gln Gln Gln Ile Phe Lys Pro Thr
 370 375 380
 Glu Asp Glu Glu Glu Ser Glu Ala Lys Tyr Ile Gly Ser Ala Asp Phe
 385 390 395 400
 Gln Ala Lys Glu Ile Phe Ser Thr Cys Leu Glu Gly Glu Gln Gly Pro
 405 410 415
 Gln Phe Ala Pro Ser Ala Pro Pro Leu Ser Thr Val Asp Ser Val Ser
 420 425 430
 Gln Val Ala Pro Ala Ala Pro Val Glu Pro Glu Thr Phe Pro Asp Lys

10

20

30

435 440 445
 Tyr Ser Leu Gln Phe Gly Phe Gly Pro Phe Glu Leu Pro Pro Gln Trp
 450 455 460
 Leu Ser Glu Thr Arg Asn Ser Lys Lys Arg Leu Leu Pro Pro Leu Gly
 465 470 475 480
 Asn Thr Pro Glu Glu Leu Ile Gln Thr Lys Val Pro Lys Val Gly Arg
 485 490 495
 Val Glu Arg Lys Met Ser Arg Asn Asn Lys Val Ser Ile Phe Pro Lys
 500 505 510
 Val Asp Ser
 515

10

(106) INFORMATION FOR SEQ ID NO:105:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:

GGAGAATTCA CTAGGCGAGG CGCTCCATC

29

20

(107) INFORMATION FOR SEQ ID NO:106:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:

GGAGGATCCA GGAAACCTTA GGCCGAGTCC

30

30

(108) INFORMATION FOR SEQ ID NO:107:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1164 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:107:

```

ATGAATCGGC ACCATCTGCA GGATCACTTT CTGGAAATAG ACAAGAAGAA CTGCTGTGTG    60
TTCCGAGATG ACTTCATTGC CAAGGTGTTG CCGCCGGTGT TGGGGCTGGA GTTATCTTT    120
GGGCTTCTGG GCAATGGCCT TGCCCTGTGG ATTTTCTGTT TCCACCTCAA GTCCTGGAAA    180
TCCAGCCGGA TTTTCCTGTT CAACCTGGCA GTAGCTGACT TTCTACTGAT CATCTGCCTG    240
CCGTTCTGTA TGGACTACTA TGTGCGGCGT TCAGACTGGA ACTTTGGGGA CATCCCTTGC    300
CGGCTGGTGC TCTTCATGTT TGCCATGAAC CGCCAGGGCA GCATCATCTT CCTCACGGTG    360
GTGGCGGTAG ACAGGTATTT CCGGGTGGTC CATCCCCACC ACGCCCTGAA CAAGATCTCC    420
AATTGGACAG CAGCCATCAT CTCTGCTT CTGTGGGGCA TCACTGTTGG CCTAACAGTC    480
CACCTCCTGA AGAAGAAGTT GCTGATCCAG AATGGCCCTG CAAATGTGTG CATCAGCTTC    540
AGCATCTGCC ATACCTTCGG GTGGCACGAA GCTATGTTCC TCCTGGAGTT CCTCCTGCCC    600
CTGGGCATCA TCCTGTTCTG CTCAGCCAGA ATTATCTGGA GCCTGCGGCA GAGACAAATG    660
GACCGGCATG CCAAGATCAA GAGAGCCATC ACCTTCATCA TGGTGGTGGC CATCGTCTTT    720
GTCATCTGCT TCCTTCCCAG CGTGGTTGTG CGGATCCGCA TCTTCTGGCT CCTGCACACT    780
TCGGGCACGC AGAATTGTGA AGTGTACCGC TCGGTGGACC TGGCGTTCTT TACTACTCTC    840
AGCTTCACCT ACATGAACAG CATGCTGGAC CCCGTGGTGT ACTACTTCTC CAGCCCATCC    900
TTTCCCAACT TCTTCTCCAC TTGATCAAC CGCTGCCTCC AGAGGAAGAT GACAGGTGAG    960
CCAGATAATA ACCGCAGCAC GAGCGTCGAG CTCACAGGGG ACCCCAACAA AACCAGAGGC   1020
GCTCCAGAGG CGTTAATGGC CAACTCCGGT GAGCCATGGA GCCCCTCTTA TCTGGGCCCA   1080
ACCTCAAATA ACCATTCCAA GAAGGGACAT TGTCACCAAG AACCAGCATC TCTGGAGAAA   1140
CAGTTGGGCT GTTGCATCGA GTAA                                           1164

```

(109) INFORMATION FOR SEQ ID NO:108:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 387 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:108:

Met Asn Arg His His Leu Gln Asp His Phe Leu Glu Ile Asp Lys Lys

10

20

30

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

10

20

30

Phe Ser Thr Leu Ile Asn Arg Cys Leu Gln Arg Lys Met Thr Gly Glu
 305 310 315 320
 Pro Asp Asn Asn Arg Ser Thr Ser Val Glu Leu Thr Gly Asp Pro Asn
 325 330 335
 Lys Thr Arg Gly Ala Pro Glu Ala Leu Met Ala Asn Ser Gly Glu Pro
 340 345 350
 Trp Ser Pro Ser Tyr Leu Gly Pro Thr Ser Asn Asn His Ser Lys Lys
 355 360 365
 Gly His Cys His Gln Glu Pro Ala Ser Leu Glu Lys Gln Leu Gly Cys
 370 375 380
 Cys Ile Glu
 385

10

(110) INFORMATION FOR SEQ ID NO:109:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 37 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: NO

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:109:

ACCATGGCTT GCAATGGCAG TCGGCCAGG GGGCACT

37

(111) INFORMATION FOR SEQ ID NO:110:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 39 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:

CGACCAGGAC AAACAGCATC TTGGTCACTT GTCTCCGGC

39

(112) INFORMATION FOR SEQ ID NO:111:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 39 base pairs
 - (B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:

GACCAAGATG CTGTTTGTCC TGGTCGTGGT GTTGGCAT 39

(113) INFORMATION FOR SEQ ID NO:112:

(i) SEQUENCE CHARACTERISTICS: 10
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:112:

CGGAATTCAG GATGGATCGG TCTCTTGCTG CGCCT 35

(114) INFORMATION FOR SEQ ID NO:113: 20

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1212 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:113:

ATGGCTTGCA ATGGCAGTGC GGCCAGGGGG CACTTTGACC CTGAGGACTT GAACCTGACT 60

GACGAGGCAC TGAGACTCAA GTACCTGGGG CCCAGCAGA CAGAGCTGTT CATGCCCATC 120

TGTGCCACAT ACCTGCTGAT CTTGCTGGTG GCGCTGTGG GCAATGGGCT GACCTGTCTG 180

GTCATCCTGC GCCACAAGGC CATGCGCACG CCTACCAACT ACTACCTCTT CAGCCTGGCC 240

GTGTCGGACC TGCTGGTGCT GCTGGTGGGC CTGCCCTGG AGCTCTATGA GATGTGGCAC 300

AACTACCCCT TCCTGCTGGG CGTTGGTGGC TGCTATTTCC GCACGCTACT GTTTGAGATG 360

GTCTGCCTGG CCTCAGTGCT CAACGTCCT GCCCTGAGCG TGAACGCTA TGTGGCCGTG 420

GTGCACCCAC TCCAGGCCAG GTCCATGGTG ACGCGGGCCC ATGTGCGCCG AGTGCTTGGG 480

30

GCCGTCTGGG GTCTTGCCAT GCTCTGCTCC CTGCCCAACA CCAGCCTGCA CGGCATCCGG 540
 CAGCTGCACG TGCCCTGCCG GGGCCCAGTG CCAGACTCAG CTGTTTGCAT GCTGGTCCGC 600
 CCACGGGGCC TCTACAACAT GGTAAGTCAG ACCACCGCGC TGCTCTTCTT CTGCCTGCCC 660
 ATGGCCATCA TGAGCGTGCT CTACCTGCTC ATTGGGCTGC GACTGCGGCG GGAGAGGCTG 720
 CTGCTCATGC AGGAGGCCAA GGGCAGGGGC TCTGCAGCAG CCAGGTCCAG ATACACCTGC 780
 AGGCTCCAGC AGCAGGATCG GGGCCGGAGA CAAGTGACCA AGATGCTGTT TGTCTGGTC 840
 GTGGTGTGTT GCATCTGCTG GGCCCCGTTC CACGCCGACC GCGTCATGTG GAGCGTCGTG 900
 TCACAGTGGA CAGATGGCCT GCACCTGGCC TTCCAGCACG TGCACGTCAT CTCCGGCATC 960
 TTCTTCTACC TGGGCTCGGC GGCCAACCCC GTGCTCTATA GCCTCATGTC CAGCCGCTTC 1020
 CGAGAGACCT TCCAGGAGGC CCTGTGCCTC GGGGCCTGCT GCCATCGCCT CAGACCCCGC 1080
 CACAGCTCCC ACAGCCTCAG CAGGATGACC ACAGGCAGCA CCCTGTGTGA TGTGGGCTCC 1140
 CTGGGCAGCT GGGTCCACCC CCTGGCTGGG AACGATGGCC CAGAGGCGCA GCAAGAGACC 1200
 GATCCATCCT GA 1212

10

(115) INFORMATION FOR SEQ ID NO:114:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 403 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

20

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:114:

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

30

Glu Met Trp His Asn Tyr Pro Phe Leu Leu Gly Val Gly Gly Cys Tyr
 100 105 110
 Phe Arg Thr Leu Leu Phe Glu Met Val Cys Leu Ala Ser Val Leu Asn
 115 120 125
 Val Thr Ala Leu Ser Val Glu Arg Tyr Val Ala Val Val His Pro Leu
 130 135 140
 Gln Ala Arg Ser Met Val Thr Arg Ala His Val Arg Arg Val Leu Gly
 145 150 155 160
 Ala Val Trp Gly Leu Ala Met Leu Cys Ser Leu Pro Asn Thr Ser Leu
 165 170 175
 His Gly Ile Arg Gln Leu His Val Pro Cys Arg Gly Pro Val Pro Asp
 180 185 190
 Ser Ala Val Cys Met Leu Val Arg Pro Arg Ala Leu Tyr Asn Met Val
 195 200 205
 Val Gln Thr Thr Ala Leu Leu Phe Phe Cys Leu Pro Met Ala Ile Met
 210 215 220
 Ser Val Leu Tyr Leu Leu Ile Gly Leu Arg Leu Arg Arg Glu Arg Leu
 225 230 235 240
 Leu Leu Met Gln Glu Ala Lys Gly Arg Gly Ser Ala Ala Ala Arg Ser
 245 250 255
 Arg Tyr Thr Cys Arg Leu Gln Gln His Asp Arg Gly Arg Arg Gln Val
 260 265 270
 Thr Lys Met Leu Phe Val Leu Val Val Val Phe Gly Ile Cys Trp Ala
 275 280 285
 Pro Phe His Ala Asp Arg Val Met Trp Ser Val Val Ser Gln Trp Thr
 290 295 300
 Asp Gly Leu His Leu Ala Phe Gln His Val His Val Ile Ser Gly Ile
 305 310 315 320
 Phe Phe Tyr Leu Gly Ser Ala Ala Asn Pro Val Leu Tyr Ser Leu Met
 325 330 335
 Ser Ser Arg Phe Arg Glu Thr Phe Gln Glu Ala Leu Cys Leu Gly Ala
 340 345 350
 Cys Cys His Arg Leu Arg Pro Arg His Ser Ser His Ser Leu Ser Arg
 355 360 365
 Met Thr Thr Gly Ser Thr Leu Cys Asp Val Gly Ser Leu Gly Ser Trp
 370 375 380
 Val His Pro Leu Ala Gly Asn Asp Gly Pro Glu Ala Gln Gln Glu Thr

10

20

30

385

390

395

400

Asp Pro Ser

(116) INFORMATION FOR SEQ ID NO:115:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:115:

GGAAGCTTCA GGCCCAAAGA TGGGGAACAT

30

(117) INFORMATION FOR SEQ ID NO:116:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:116:

GTGGATCCAC CCGCGGAGGA CCCAGGCTAG

30

(118) INFORMATION FOR SEQ ID NO:117:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1098 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:117:

ATGGGGAACA TCACTGCAGA CAACTCCTCG ATGAGCTGTA CCATCGACCA TACCATCCAC

60

CAGACGCTGG CCCCSTGGT CTATGTTACC GTGCTGGTGG TGGGCTTCCC GGCCAACTGC

120

CTGTCCCTCT ACTTCGGCTA CCTGCAGATC AAGGCCCGGA ACGAGCTGGG CGTGACCTG

180

TGCAACCTGA CGGTGGCCGA CCTCTTCTAC ATCTGCTCGC TGCCCTTCTG GCTGCAGTAC

240

GTGCTGCAGC ACGACAACTG GTCTCACGGC GACCTGTCCT GCCAGGTGTG CGGCATCCTC

300

CTGTACGAGA ACATCTACAT CAGCGTGGGC TTCCTCTGCT GCATCTCCGT GGACCGCTAC

360

10

20

30

CTGGCTGTGG CCCATCCCTT CCGCTTCCAC CAGTTCCGGA CCCTGAAGGC GGCCGTCGGC 420
 GTCAGCGTGG TCATCTGGGC CAAGGAGCTG CTGACCAGCA TCTACTTCCT GATGCACGAG 480
 GAGGTCATCG AGGACGAGAA CCAGCACCGC GTGTGCTTTG AGCACTACCC CATCCAGGCA 540
 TGGCAGCGCG CCATCAACTA CTACCGCTTC CTGGTGGGCT TCCTCTTCCC CATCTGCCTG 600
 CTGCTGGCGT CCTACCAGGG CATCCTGCGC GCCGTGCGCC GGAGCCACGG CACCCAGAAG 660
 AGCCGCAAGG ACCAGATCCA GCGGCTGGTG CTCAGCACCG TGGTCATCTT CCTGGCCTGC 720
 TTCCTGCCCT ACCACGTGTT GCTGCTGGTG CGCAGCGTCT GGGAGGCCAG CTGCGACTTC 780
 GCCAAGGGCG TTTTCAACGC CTACCACTTC TCCCTCCTGC TCACCAGCTT CAACTGCGTC 840
 GCCGACCCCG TGCTCTACTG CTTCGTGAGC GAGACCACCC ACCGGGACCT GGCCCGCCTC 900
 CGCGGGGCCT GCCTGGCCTT CCTCACCTGC TCCAGGACCG GCCGGGCCAG GGAGGCCTAC 960
 CCGCTGGGTG CCCCCGAGGC CTCCGGGAAA AGCGGGGCCC AGGGTGAGGA GCCCGAGCTG 1020
 TTGACCAAGC TCCACCCGGC CTTCCAGACC CCTAACTCGC CAGGGTCGGG CGGGTTCCCC 1080
 ACGGGCAGGT TGGCCTAG 1098

(119) INFORMATION FOR SEQ ID NO:118:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 365 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:118:

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

Cys Gly Ile Leu Leu Tyr Glu Asn Ile Tyr Ile Ser Val Gly Phe Leu
 100 105 110
 Cys Cys Ile Ser Val Asp Arg Tyr Leu Ala Val Ala His Pro Phe Arg
 115 120 125
 Phe His Gln Phe Arg Thr Leu Lys Ala Ala Val Gly Val Ser Val Val
 130 135 140
 Ile Trp Ala Lys Glu Leu Leu Thr Ser Ile Tyr Phe Leu Met His Glu
 145 150 155 160
 Glu Val Ile Glu Asp Glu Asn Gln His Arg Val Cys Phe Glu His Tyr
 165 170 175
 Pro Ile Gln Ala Trp Gln Arg Ala Ile Asn Tyr Tyr Arg Phe Leu Val
 180 185 190
 Gly Phe Leu Phe Pro Ile Cys Leu Leu Leu Ala Ser Tyr Gln Gly Ile
 195 200 205
 Leu Arg Ala Val Arg Arg Ser His Gly Thr Gln Lys Ser Arg Lys Asp
 210 215 220
 Gln Ile Gln Arg Leu Val Leu Ser Thr Val Val Ile Phe Leu Ala Cys
 225 230 235 240
 Phe Leu Pro Tyr His Val Leu Leu Leu Val Arg Ser Val Trp Glu Ala
 245 250 255
 Ser Cys Asp Phe Ala Lys Gly Val Phe Asn Ala Tyr His Phe Ser Leu
 260 265 270
 Leu Leu Thr Ser Phe Asn Cys Val Ala Asp Pro Val Leu Tyr Cys Phe
 275 280 285
 Val Ser Glu Thr Thr His Arg Asp Leu Ala Arg Leu Arg Gly Ala Cys
 290 295 300
 Leu Ala Phe Leu Thr Cys Ser Arg Thr Gly Arg Ala Arg Glu Ala Tyr
 305 310 315 320
 Pro Leu Gly Ala Pro Glu Ala Ser Gly Lys Ser Gly Ala Gln Gly Glu
 325 330 335
 Glu Pro Glu Leu Leu Thr Lys Leu His Pro Ala Phe Gln Thr Pro Asn
 340 345 350
 Ser Pro Gly Ser Gly Gly Phe Pro Thr Gly Arg Leu Ala
 355 360 365

10

20

30

(120) INFORMATION FOR SEQ ID NO:119:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 base pairs

(B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:119:

GACCTCGAGT CCTTCTACAC CTCATC

26

(121) INFORMATION FOR SEQ ID NO:120:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:120:

TGCTCTAGAT TCCAGATAGG TGAAACTTG

30

(122) INFORMATION FOR SEQ ID NO:121:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1416 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:121:

ATGGATATTC TTTGTGAAGA AAATACTTCT TTGAGCTCAA CTACGAACTC CCTAATGCAA 60
 TTAAATGATG ACAACAGGCT CTACAGTAAT GACTTTAACT CCGGAGAAGC TAACACTTCT 120
 GATGCATTTA ACTGGACAGT CGACTCTGAA AATCGAACCA ACCTTTCCTG TGAAGGGTGC 180
 CTCTCACCGT CGTGTCTCTC CTTACTTCAT CTCCAGGAAA AAAACTGGTC TGCTTTACTG 240
 ACAGCCGTAG TGATTATTCT AACTATTGCT GGAAACATAC TCGTCATCAT GGCAGTGTCC 300
 CTAGAGAAAA AGCTGCAGAA TGCCACCAAC TATTTCTCTGA TGTCACCTGC CATAGCTGAT 360
 ATGCTGCTGG GTTTCCTTGT CATGCCCGTG TCCATGTTAA CCATCCTGTA TGGGTACCGG 420
 TGGCCTCTGC CGAGCAAGCT TTGTGCAGTC TGGATTTACC TGGACGTGCT CTTCTCCACG 480
 GCCTCCATCA TGCACCTCTG CGCCATCTCG CTGGACCGCT ACGTCGCCAT CCAGAATCCC 540
 ATCCACCACA GCCGCTTCAA CTCCAGAACT AAGGCATTTC TGAAAATCAT TGCTGTTTGG 600

30

ACCATATCAG TAGGTATATC CATGCCAATA CCAGTCTTTG GGCTACAGGA CGATTCTGAAG 660
 GTCTTTAAGG AGGGGAGTTG CTTACTCGCC GATGATAACT TTGTCCTGAT CGGCTCTTTT 720
 GTGTCAATTTT TCATTCCCTT AACCATCATG GTGATCACCT ACTTTCTAAC TATCAAGTCA 780
 CTCCAGAAAG AAGCTACTTT GTGTGTAAGT GATCTTGGCA CACGGGCCAA ATTAGCTTCT 840
 TTCAGCTTCC TCCCTCAGAG TTCTTTGTCT TCAGAAAAGC TCTTCCAGCG GTCGATCCAT 900
 AGGGAGCCAG GGTCTACAC AGGCAGGAGG ACTATGCAGT CCATCAGCAA TGAGCAAAAG 960
 GCATGCAAGG TGCTGGGCAT CGTCTTCTTC CTGTTTGTGG TGATGTGGTG CCCTTTCTTC 1020
 ATCACAAACA TCATGGCCGT CATCTGCAA GAGTCCTGCA ATGAGGATGT CATTGGGGCC 1080
 CTGCTCAATG TGTTTGTTTG GATCGGTTAT CTCTCTTCAG CAGTCAACCC ACTAGTCTAC 1140
 ACACTGTTCA ACAAGACCTA TAGGTCAGCC TTTTCACGGT ATATTCAGTG TCAGTACAAG 1200
 GAAAACAAAA AACCATTGCA GTTAATTTTA GTGAACACAA TACCGGCTTT GGCCTACAAG 1260
 TCTAGCCAAC TTCAAATGGG ACAAAAAAAG AATTCAAAGC AAGATGCCAA GACAACAGAT 1320
 AATGACTGCT CAATGGTTC TCTAGGAAAG CAGTATTCTG AAGAGGCTTC TAAAGACAAT 1380
 AGCGACGGAG TGAATGAAAA GGTGAGCTGT GTGTGA 1416

(123) INFORMATION FOR SEQ ID NO:122:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 471 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:122:

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

Thr Ala Val Val Ile Ile Leu Thr Ile Ala Gly Asn Ile Leu Val Ile
 85 90 95
 Met Ala Val Ser Leu Glu Lys Lys Leu Gln Asn Ala Thr Asn Tyr Phe
 100 105 110
 Leu Met Ser Leu Ala Ile Ala Asp Met Leu Leu Gly Phe Leu Val Met
 115 120 125
 Pro Val Ser Met Leu Thr Ile Leu Tyr Gly Tyr Arg Trp Pro Leu Pro
 130 135 140
 Ser Lys Leu Cys Ala Val Trp Ile Tyr Leu Asp Val Leu Phe Ser Thr
 145 150 155 160
 Ala Ser Ile Met His Leu Cys Ala Ile Ser Leu Asp Arg Tyr Val Ala
 165 170 175
 Ile Gln Asn Pro Ile His His Ser Arg Phe Asn Ser Arg Thr Lys Ala
 180 185 190
 Phe Leu Lys Ile Ile Ala Val Trp Thr Ile Ser Val Gly Ile Ser Met
 195 200 205
 Pro Ile Pro Val Phe Gly Leu Gln Asp Asp Ser Lys Val Phe Lys Glu
 210 215 220
 Gly Ser Cys Leu Leu Ala Asp Asp Asn Phe Val Leu Ile Gly Ser Phe
 225 230 235 240
 Val Ser Phe Phe Ile Pro Leu Thr Ile Met Val Ile Thr Tyr Phe Leu
 245 250 255
 Thr Ile Lys Ser Leu Gln Lys Glu Ala Thr Leu Cys Val Ser Asp Leu
 260 265 270
 Gly Thr Arg Ala Lys Leu Ala Ser Phe Ser Phe Leu Pro Gln Ser Ser
 275 280 285
 Leu Ser Ser Glu Lys Leu Phe Gln Arg Ser Ile His Arg Glu Pro Gly
 290 295 300
 Ser Tyr Thr Gly Arg Arg Thr Met Gln Ser Ile Ser Asn Glu Gln Lys
 305 310 315 320
 Ala Cys Lys Val Leu Gly Ile Val Phe Phe Leu Phe Val Val Met Trp
 325 330 335
 Cys Pro Phe Phe Ile Thr Asn Ile Met Ala Val Ile Cys Lys Glu Ser
 340 345 350
 Cys Asn Glu Asp Val Ile Gly Ala Leu Leu Asn Val Phe Val Trp Ile
 355 360 365
 Gly Tyr Leu Ser Ser Ala Val Asn Pro Leu Val Tyr Thr Leu Phe Asn

10

20

30

370	375	380
Lys Thr Tyr Arg Ser Ala Phe Ser Arg Tyr Ile Gln Cys Gln Tyr Lys		
385	390	395 400
Glu Asn Lys Lys Pro Leu Gln Leu Ile Leu Val Asn Thr Ile Pro Ala		
	405	410 415
Leu Ala Tyr Lys Ser Ser Gln Leu Gln Met Gly Gln Lys Lys Asn Ser		
	420	425 430
Lys Gln Asp Ala Lys Thr Thr Asp Asn Asp Cys Ser Met Val Ala Leu		
	435	440 445
Gly Lys Gln Tyr Ser Glu Glu Ala Ser Lys Asp Asn Ser Asp Gly Val		
	450	455 460
Asn Glu Lys Val Ser Cys Val		
465	470	

10

(124) INFORMATION FOR SEQ ID NO:123:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:123:

GACCTCGAGG TTGCTTAAGA CTGAAGC

27

(125) INFORMATION FOR SEQ ID NO:124:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:124:

ATTCTAGAC ATATGTAGCT TGTACCG

27

(126) INFORMATION FOR SEQ ID NO:125:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1377 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:125:

ATGGTGAACC TGAGGAATGC GGTGCATTCA TTCCTTGTGC ACCTAATTGG CCTATTGGTT 60
 TGGCAATGTG ATATTTCTGT GAGCCCAGTA GCAGCTATAG TAACTGACAT TTTCAATACC 120
 TCCGATGGTG GACGCTTCAA ATTCCCAGAC GGGGTACAAA ACTGGCCAGC ACTTTCAATC 180
 GTCATCATAA TAATCATGAC AATAGGTGGC AACATCCTTG TGATCATGGC AGTAAGCATG 240
 GAAAAGAAAC TGCACAATGC CACCAATTAC TTCTTAATGT CCCTAGCCAT TGCTGATATG 300
 CTAGTGGGAC TACTTGTCAT GCCCCTGTCT CTCCTGGCAA TCCTTTATGA TTATGTCTGG 360
 CCACTACCTA GATATTTGTG CCCCCTCTGG ATTTCTTTAG ATGTTTATT TTCAACAGCG 420
 TCCATCATGC ACCTCTGCGC TATATCGCTG GATCGGTATG TAGCAATACG TAATCCTATT 480
 GAGCATAGCC GTTTC AATTC GCGGACTAAG GCCATCATGA AGATTGCTAT TGT TTGGGCA 540
 ATTTCTATAG GTGTATCAGT TCCTATCCCT GTGATTGGAC TGAGGGACGA AGAAAAGGTG 600
 TTCGTGAACA ACACGACGTG CGTGCTCAAC GACCCAAATT TCGTTCCTTAT TGGGTCCTTC 660
 GTAGCTTTCT TCATACCGCT GACGATTATG GTGATTACGT ATTGCCTGAC CATCTACGTT 720
 CTGCGCCGAC AAGCTTTGAT GTTACTGCAC GGCCACACCG AGGAACCGCC TGGACTAAGT 780
 CTGGATTTC TGAAGTGCTG CAAGAGGAAT ACGGCCGAGG AAGAGAACTC TGCAAACCTT 840
 AACCAAGACC AGAACGACG CCGAAGAAAG AAGAAGGAGA GACGTCCTAG GGGCACCATG 900
 CAGGCTATCA ACAATGAAAG AAAAGCTTCG AAAGTCCTTG GGATTGTTTT CTTTGTGTTT 960
 CTGATCATGT GGTGCCCATT TTTCATTACC AATATTCTGT CTGTTCTTTG TGAGAAGTCC 1020
 TGTAAACAAA AGCTCATGGA AAAGCTTCTG AATGTGTTTG TTTGGATTGG CTATGTTTGT 1080
 TCAGGAATCA ATCCTCTGGT GTATACTCTG TTCAACAAAA TTTACCGAAG GGCATTCTCC 1140
 AACTATTTGC GTTGCAATTA TAAGGTAGAG AAAAAGCCTC CTGTCAGGCA GATTCCAAGA 1200
 GTTGCCGCCA CTGCTTTGTC TGGGAGGGAG CTTAATGTTA ACATTTATCG GCATACCAAT 1260
 GAACCGGTGA TCGAGAAAGC CAGTGACAAT GAGCCCGGTA TAGAGATGCA AGTTGAGAAT 1320
 TTAGAGTTAC CAGTAAATCC CTCCAGTGTG GTTAGCGAAA GGATTAGCAG TGTGTGA 1377

(127) INFORMATION FOR SEQ ID NO:126:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 458 amino acids

(B) TYPE: amino acid

10

20

30

(C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:126:

```

Met Val Asn Leu Arg Asn Ala Val His Ser Phe Leu Val His Leu Ile
1           5           10           15
Gly Leu Leu Val Trp Gln Cys Asp Ile Ser Val Ser Pro Val Ala Ala
          20           25           30
Ile Val Thr Asp Ile Phe Asn Thr Ser Asp Gly Gly Arg Phe Lys Phe
          35           40           45
Pro Asp Gly Val Gln Asn Trp Pro Ala Leu Ser Ile Val Ile Ile Ile
          50           55           60
Ile Met Thr Ile Gly Gly Asn Ile Leu Val Ile Met Ala Val Ser Met
65           70           75           80
Glu Lys Lys Leu His Asn Ala Thr Asn Tyr Phe Leu Met Ser Leu Ala
          85           90           95
Ile Ala Asp Met Leu Val Gly Leu Leu Val Met Pro Leu Ser Leu Leu
          100          105          110
Ala Ile Leu Tyr Asp Tyr Val Trp Pro Leu Pro Arg Tyr Leu Cys Pro
          115          120          125
Val Trp Ile Ser Leu Asp Val Leu Phe Ser Thr Ala Ser Ile Met His
          130          135          140
Leu Cys Ala Ile Ser Leu Asp Arg Tyr Val Ala Ile Arg Asn Pro Ile
145          150          155          160
Glu His Ser Arg Phe Asn Ser Arg Thr Lys Ala Ile Met Lys Ile Ala
          165          170          175
Ile Val Trp Ala Ile Ser Ile Gly Val Ser Val Pro Ile Pro Val Ile
          180          185          190
Gly Leu Arg Asp Glu Glu Lys Val Phe Val Asn Asn Thr Thr Cys Val
          195          200          205
Leu Asn Asp Pro Asn Phe Val Leu Ile Gly Ser Phe Val Ala Phe Phe
          210          215          220
Ile Pro Leu Thr Ile Met Val Ile Thr Tyr Cys Leu Thr Ile Tyr Val
225          230          235          240
Leu Arg Arg Gln Ala Leu Met Leu Leu His Gly His Thr Glu Glu Pro
          245          250          255

```

10

20

30

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

10

20

(128) INFORMATION FOR SEQ ID NO:127:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:127:

GGTAAGCTTG GCAGTCCACG CCAGGCCTTC

30

(129) INFORMATION FOR SEQ ID NO:128:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:128:

TCCGAATTCT CTGTAGACAC AAGGCTTTGG 30

(130) INFORMATION FOR SEQ ID NO:129:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1068 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:129:

ATGGATCAGT TCCCTGAATC AGTGACAGAA AACTTTGAGT ACGATGATTT GGCTGAGGCC 60
 TGTATATATG GGGACATCGT GGTCTTTGGG ACTGTGTTCC TGTCCATATT CTACTCCGTC 120
 ATCTTTGCCA TTGGCCTGGT GGGAAATTG TTGGTAGTGT TTGCCCTCAC CAACAGCAAG 180
 AAGCCCAAGA GTGTCACCGA CATTACCTC CTGAACCTGG CCTTGTCTGA TCTGCTGTTT 240
 GTAGCCACTT TGCCCTTCTG GACTCACTAT TTGATAAATG AAAAGGGCCT CCACAATGCC 300
 ATGTGCAAAAT TCACTACCGC CTCTTCTTC ATCGGCTTTT TTGGAAGCAT ATTCTTCATC 360
 ACCGTCAATCA GCATTGATAG GTACCTGGCC ATCGTCCTGG CCGCCAACTC CATGAACAAC 420
 CGGACCGTGC AGCATGGCGT CACCATCAGC CTAGGCGTCT GGGCAGCAGC CATTTTGGTG 480
 GCAGCACCCC AGTTCATGTT CACAAAGCAG AAAGAAAATG AATGCCTTGG TGAATACCCC 540
 GAGGTCCCTC AGGAAATCTG GCCCGTGCTC CGCAATGTTG AAACAAATTT TCTTGGCTTC 600
 CTACTCCCCC TGCTCATTAT GAGTTATTGC TACTTCAGAA TCATCCAGAC GCTGTTTTC 660
 TGCAAGAACC ACAAGAAAGC CAAAGCCATT AACTGATCC TTCTGGTGGT CATCGTGTTT 720
 TTCCTCTTCT GGACACCTA CAACGTTATG ATTTTCCTGG AGACGCTTAA GCTCTATGAC 780
 TTCTTTCCCA GTTGTGACAT GAGGAAGGAT CTGAGGCTGG CCCTCAGTGT GACTGAGACG 840
 GTTGCAATTTA GCCATGTTG CCTGAATCCT CTCATCTATG CATTGCTGG GGAGAAGTTC 900
 AGAAGATACC TTTACCACCT GTATGGGAAA TGCCTGGCTG TCCTGTGTGG GCGCTCAGTC 960

10

20

30

CACGTTGATT TCTCCTCATC TGAATCACAA AGGAGCAGGC ATGGAAGTGT TCTGAGCAGC 1020

AATTTTACTT ACCACACGAG TGATGGAGAT GCATTGCTCC TTCTCTGA 1068

(131) INFORMATION FOR SEQ ID NO:130:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 355 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:130:

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

10

20

30

Tyr Cys Tyr Phe Arg Ile Ile Gln Thr Leu Phe Ser Cys Lys Asn His
 210 215 220
 Lys Lys Ala Lys Ala Ile Lys Leu Ile Leu Leu Val Val Ile Val Phe
 225 230 235 240
 Phe Leu Phe Trp Thr Pro Tyr Asn Val Met Ile Phe Leu Glu Thr Leu
 245 250 255
 Lys Leu Tyr Asp Phe Phe Pro Ser Cys Asp Met Arg Lys Asp Leu Arg
 260 265 270
 Leu Ala Leu Ser Val Thr Glu Thr Val Ala Phe Ser His Cys Cys Leu
 275 280 285
 Asn Pro Leu Ile Tyr Ala Phe Ala Gly Glu Lys Phe Arg Arg Tyr Leu
 290 295 300
 Tyr His Leu Tyr Gly Lys Cys Leu Ala Val Leu Cys Gly Arg Ser Val
 305 310 315 320
 His Val Asp Phe Ser Ser Ser Glu Ser Gln Arg Ser Arg His Gly Ser
 325 330 335
 Val Leu Ser Ser Asn Phe Thr Tyr His Thr Ser Asp Gly Asp Ala Leu
 340 345 350
 Leu Leu Leu
 355

10

20

(132) INFORMATION FOR SEQ ID NO:131:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:131:

GATCTCCAGT AGGCATAAGT GGACAATTCT GG

32

30

(133) INFORMATION FOR SEQ ID NO:132:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:132:

CTCCTTCGGT CCTCCTATCG TTGTCAGAAG

30

(134) INFORMATION FOR SEQ ID NO:133:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:133:

AGAAGGCCAA GATCGCGCGG CTGGCCCTCA

30

10

(135) INFORMATION FOR SEQ ID NO:134:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:134:

CGGCGCCACC GCACGAAAAA GCTCATCTTC

30

20

(136) INFORMATION FOR SEQ ID NO:135:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:135:

GCCAAGAAGC GGGTGAAGTT CCTGGTGGTG GCA

33

30

(137) INFORMATION FOR SEQ ID NO:136:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:136:

CAGGCGGAAG GTGAAAGTCC TGGTCCTCGT 30

(138) INFORMATION FOR SEQ ID NO:137:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic) 10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:137:

CGGCGCCTGC GGGCCAAGCG GCTGGTGGTG GTG 33

(139) INFORMATION FOR SEQ ID NO:138:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:138: 20

CCAAGCACAA AGCCAAGAAA GTGACCATCA C 31

(140) INFORMATION FOR SEQ ID NO:139:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:139: 30

GCGCCGGCGC ACCAAATGCT TGCTGGTGGT 30

(141) INFORMATION FOR SEQ ID NO:140:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:140:

CAAAAAGCTG AAGAAATCTA AGAAGATCAT CTTTATTGTC G 41

(142) INFORMATION FOR SEQ ID NO:141:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic) 10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:141:

CAAGACCAAG GCAAAACGCA TGATCGCCAT 30

(143) INFORMATION FOR SEQ ID NO:142:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic) 20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:142:

GTCAAGGAGA AGTCCAAAAG GATCATCATC 30

(144) INFORMATION FOR SEQ ID NO:143:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic) 30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:143:

CGCCGCGTGC GGGCCAAGCA GCTCCTGCTC 30

(145) INFORMATION FOR SEQ ID NO:144:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:144:

CCTGATAAGC GCTATAAAAT GGTCCTGTTT CGA 33

(146) INFORMATION FOR SEQ ID NO:145:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic) 10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:145:

GAAAGACAAA AGAGAGTCAA GAGGATGTCT TTATTG 36

(147) INFORMATION FOR SEQ ID NO:146:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic) 20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:146:

CGGAGAAAGA GGGTGAAACG CACAGCCATC GCC 33

(148) INFORMATION FOR SEQ ID NO:147:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic) 30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:147:

AAGCTTCAGC GGGCCAAGGC ACTGGTCACC 30

(149) INFORMATION FOR SEQ ID NO:148:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:148:

CAGCGGCAGA AGGC^AAAAG GGTGGCCATC 30

(150) INFORMATION FOR SEQ ID NO:149:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic) 10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:149:

CGGCAGAAGG CGAAGCGCAT GATCCTCGCG 30

(151) INFORMATION FOR SEQ ID NO:150:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic) 20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:150:

GAGCGCAACA AGGCCAAAAA GGTGATCATC 30

(152) INFORMATION FOR SEQ ID NO:151:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic) 30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:151:

GGTGTAACA AAAAGGCTAA AAACACAATT ATTCTTATT 39

(153) INFORMATION FOR SEQ ID NO:152:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:152:

GAGAGCCAGC TCAAGAGCAC CGTGGTG 27

(154) INFORMATION FOR SEQ ID NO:153:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic) 10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:153:

CCACAAGCAA ACCAAGAAAA TGCTGGCTGT 30

(155) INFORMATION FOR SEQ ID NO:154:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic) 20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:154:

CATCAAGTGT ATCATGTGCC AAGTACGCCC 30

(156) INFORMATION FOR SEQ ID NO:155:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic) 30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:155:

CTAGAGAGTC AGATGAAGTG TACAGTAGTG GCAC 34

(157) INFORMATION FOR SEQ ID NO:156:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:156:

CGGACAAAAG TGAAAACTAA AAAGATGTTT CTCATT 36

(158) INFORMATION FOR SEQ ID NO:157:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic) 10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:157:

GCTGAGGTTC GCAATAAACT AACCATGTTT GTG 33

(159) INFORMATION FOR SEQ ID NO:158:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic) 20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:158:

GGGAGGCCGA GCTGAAAGCC ACCCTGCTC 29

(160) INFORMATION FOR SEQ ID NO:159:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic) 30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:159:

CAAGATCAAG AGAGCCAAA CCTTCATCAT G 31

(161) INFORMATION FOR SEQ ID NO:160:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:160:

CCGGAGACAA GTGAAGAAGA TGCTGTTTGT C 31

(162) INFORMATION FOR SEQ ID NO:161:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic) 10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:161:

GCAAGGACCA GATCAAGCGG CTGGTGCTCA 30

(163) INFORMATION FOR SEQ ID NO:162:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic) 20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:162:

CAAGAAAGCC AAAGCCAAGA AACTGATCCT TCTG 34

(164) INFORMATION FOR SEQ ID NO:163:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1068 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic) 30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:163:

ATGGAAGATT TGGAGGAAAC ATTATTTGAA GAATTTGAAA ACTATTCCTA TGACCTAGAC 60

TATTACTCTC TGGAGTCTGA TTTGGAGGAG AAAGTCCAGC TGGGAGTTGT TCACTGGGTC 120

TCCCTGGTGT TATATTGTTT GGCTTTTGTT CTGGGAATTC CAGGAAATGC CATCGTCATT 180

TGGTTCACGG GGCTCAAGTG GAAGAAGACA GTCACCACTC TGTGGTTCCT CAATCTAGCC 240

ATTGCGGATT TCATTTTCT TCTCTTCTG CCCCTGTACA TCTCCTATGT GGCCATGAAT 300

TTCCACTGGC CCTTTGGCAT CTGGCTGTGC AAAGCCAATT CCTTCACTGC CCAGTTGAAC 360
 ATGTTTGCCA GTGTTTTTTT CCTGACAGTG ATCAGCCTGG ACCACTATAT CCACTTGATC 420
 CATCCTGTCT TATCTCATCG GCATCGAACC CTCAAGAACT CTCTGATIGT CATTATATTC 480
 ATCTGGCTTT TGGCTTCTCT AATTGGCGGT CCTGCCCTGT ACTTCCGGGA CACTGTGGAG 540
 TTCAATAATC ATACTCTTTG CTATAACAAT TTTCAGAAGC ATGATCCTGA CCTCACTTTG 600
 ATCAGGCACC ATGTTCTGAC TTGGGTGAAA TTTATCATTG GCTATCTCTT CCCTTTGCTA 660
 ACAATGAGTA TTTGCTACTT GTGTCTCATC TTCAAGGIGA AGAAGCGAAC AGTCCTGATC 720
 TCCAGTAGGC ATAAGTGGAC AATTCTGGTT GTGGTTGTGG CCTTTGTGGT TTGCTGGACT 780
 CCTTATCACC TGTTTAGCAT TTGGGAGCTC ACCATTCAACC ACAATAGCTA TTCCCACCAT 840
 GTGATGCAGG CTGGAATCCC CCTCTCCACT GGTTTGGCAT TCCTCAATAG TTGCTTGAAC 900
 CCCATCCTTT ATGTCCTAAT TAGTAAGAAG TTCCAAGCTC GCTTCCGGTC CTCAGTTGCT 960
 GAGATACTCA AGTACACACT GTGGGAAGTC AGCTGTTCTG GCACAGTGAG TGAACAGCTC 1020
 AGGAACTCAG AAACCAAGAA TCTGTGTCTC CTGGAAACAG CTCAATAA 1068

(165) INFORMATION FOR SEQ ID NO:164:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 355 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:164:

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

Val Ala Met Asn Phe His Trp Pro Phe Gly Ile Trp Leu Cys Lys Ala
 100 105 110
 Asn Ser Phe Thr Ala Gln Leu Asn Met Phe Ala Ser Val Phe Phe Leu
 115 120 125
 Thr Val Ile Ser Leu Asp His Tyr Ile His Leu Ile His Pro Val Leu
 130 135 140
 Ser His Arg His Arg Thr Leu Lys Asn Ser Leu Ile Val Ile Ile Phe
 145 150 155 160
 Ile Trp Leu Leu Ala Ser Leu Ile Gly Gly Pro Ala Leu Tyr Phe Arg
 165 170 175
 Asp Thr Val Glu Phe Asn Asn His Thr Leu Cys Tyr Asn Asn Phe Gln
 180 185 190
 Lys His Asp Pro Asp Leu Thr Leu Ile Arg His His Val Leu Thr Trp
 195 200 205
 Val Lys Phe Ile Ile Gly Tyr Leu Phe Pro Leu Leu Thr Met Ser Ile
 210 215 220
 Cys Tyr Leu Cys Leu Ile Phe Lys Val Lys Lys Arg Thr Val Leu Ile
 225 230 235 240
 Ser Ser Arg His Lys Trp Thr Ile Leu Val Val Val Val Ala Phe Val
 245 250 255
 Val Cys Trp Thr Pro Tyr His Leu Phe Ser Ile Trp Glu Leu Thr Ile
 260 265 270
 His His Asn Ser Tyr Ser His His Val Met Gln Ala Gly Ile Pro Leu
 275 280 285
 Ser Thr Gly Leu Ala Phe Leu Asn Ser Cys Leu Asn Pro Ile Leu Tyr
 290 295 300
 Val Leu Ile Ser Lys Lys Phe Gln Ala Arg Phe Arg Ser Ser Val Ala
 305 310 315 320
 Glu Ile Leu Lys Tyr Thr Leu Trp Glu Val Ser Cys Ser Gly Thr Val
 325 330 335
 Ser Glu Gln Leu Arg Asn Ser Glu Thr Lys Asn Leu Cys Leu Leu Glu
 340 345 350
 Thr Ala Gln
 355

10

20

30

(166) INFORMATION FOR SEQ ID NO:165:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1089 base pairs

(B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:165:

ATGGGCAACC ACACGTGGGA GGGCTGCCAC GTGGACTCGC GCGTGGACCA CCTCTTTCCG	60
CCATCCCTCT ACATCTTTGT CATCGGCGTG GGGCTGCCCA CCAACTGCCT GGCTCTGTGG	120
GCGGCCTACC GCCAGGTGCA ACAGCGCAAC GAGCTGGGCG TCTACCTGAT GAACCTCAGC	180
ATCGCCGACC TGCTGTACAT CTGCACGCTG CCGCTGTGGG TGGACTACTT CCTGCACCAC	240
GACAACTGGA TCCACGGCCC CGGGTCCTGC AAGCTCTTTG GGTTCATCTT CTACACCAAT	300
ATCTACATCA GCATCGCCTT CCTGTGCTGC ATCTCGGTGG ACCGCTACCT GGCTGTGGCC	360
CACCCACTCC GCTTCGCCCC CCTGCGCCGC GTCAAGACCG CCGTGGCCGT GAGCTCCGTG	420
GTCTGGGCCA CGGAGCTGGG CGCCAACTCG GCGCCCTGT TCCATGACGA GCTCTCCGA	480
GACCGCTACA ACCACACCTT CTGCTTTGAG AAGTTCCCA TGAAGGCTG GGTGGCCTGG	540
ATGAACCTCT ATCGGGTGTT CGTGGGCTTC CTCTTCCCGT GGGCGCTCAT GCTGCTGTCG	600
TACCGGGGCA TCCTGCGGGC CGTGGGGGC AGCGTGTCCA CCGAGCGCCA GGAGAAGGCC	660
AAGATCGGCG GGCTGGCCCT CAGCCTCATC GCCATCGTGC TGGTCTGCTT TCGCCCTAT	720
CACGTGCTCT TGCTGTCCCG CAGCGCCATC TACCTGGGCC GCCCTGGGA CTGCGGCTTC	780
GAGGAGCGCG TCTTTTCTGC ATACCACAGC TCACTGGCTT TCACCAGCCT CAACTGTGTG	840
GCGGACCCCA TCCTCTACTG OCTGGTCAAC GAGGGCGCCC GCAGCGATGT GGCCAAGGCC	900
CTGCACAACC TGCTCCGCTT TCTGGCCAGC GACAAGCCCC AGGAGATGGC CAATGCCTCG	960
CTCACCTTGG AGACCCCACT CACCTCCAAG AGGAACAGCA CAGCCAAAGC CATGACTGGC	1020
AGCTGGGCGG CCACTCCGCC TTCCAGGGG GACCAGGTGC AGCTGAAGAT GCTGCCGCCA	1080
GCACAATGA	1089

10

20

30

(167) INFORMATION FOR SEQ ID NO:166:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 362 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:166:

Met Gly Asn His Thr Trp Glu Gly Cys His Val Asp Ser Arg Val Asp
 1 5 10 15

His Leu Phe Pro Pro Ser Leu Tyr Ile Phe Val Ile Gly Val Gly Leu
 20 25 30

Pro Thr Asn Cys Leu Ala Leu Trp Ala Ala Tyr Arg Gln Val Gln Gln
 35 40 45

Arg Asn Glu Leu Gly Val Tyr Leu Met Asn Leu Ser Ile Ala Asp Leu
 50 55 60

Leu Tyr Ile Cys Thr Leu Pro Leu Trp Val Asp Tyr Phe Leu His His
 65 70 75 80

Asp Asn Trp Ile His Gly Pro Gly Ser Cys Lys Leu Phe Gly Phe Ile
 85 90 95

Phe Tyr Thr Asn Ile Tyr Ile Ser Ile Ala Phe Leu Cys Cys Ile Ser
 100 105 110

Val Asp Arg Tyr Leu Ala Val Ala His Pro Leu Arg Phe Ala Arg Leu
 115 120 125

Arg Arg Val Lys Thr Ala Val Ala Val Ser Ser Val Val Trp Ala Thr
 130 135 140

Glu Leu Gly Ala Asn Ser Ala Pro Leu Phe His Asp Glu Leu Phe Arg
 145 150 155 160

Asp Arg Tyr Asn His Thr Phe Cys Phe Glu Lys Phe Pro Met Glu Gly
 165 170 175

Trp Val Ala Trp Met Asn Leu Tyr Arg Val Phe Val Gly Phe Leu Phe
 180 185 190

Pro Trp Ala Leu Met Leu Leu Ser Tyr Arg Gly Ile Leu Arg Ala Val
 195 200 205

Arg Gly Ser Val Ser Thr Glu Arg Gln Glu Lys Ala Lys Ile Ala Arg
 210 215 220

Leu Ala Leu Ser Leu Ile Ala Ile Val Leu Val Cys Phe Ala Pro Tyr
 225 230 235 240

His Val Leu Leu Leu Ser Arg Ser Ala Ile Tyr Leu Gly Arg Pro Trp
 245 250 255

Asp Cys Gly Phe Glu Glu Arg Val Phe Ser Ala Tyr His Ser Ser Leu
 260 265 270

Ala Phe Thr Ser Leu Asn Cys Val Ala Asp Pro Ile Leu Tyr Cys Leu

10

20

30

275 280 285
 Val Asn Glu Gly Ala Arg Ser Asp Val Ala Lys Ala Leu His Asn Leu
 290 295 300
 Leu Arg Phe Leu Ala Ser Asp Lys Pro Gln Glu Met Ala Asn Ala Ser
 305 310 315 320
 Leu Thr Leu Glu Thr Pro Leu Thr Ser Lys Arg Asn Ser Thr Ala Lys
 325 330 335
 Ala Met Thr Gly Ser Trp Ala Ala Thr Pro Pro Ser Gln Gly Asp Gln
 340 345 350
 Val Gln Leu Lys Met Leu Pro Pro Ala Gln
 355 360

10

(168) INFORMATION FOR SEQ ID NO:167:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1002 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:167:

20

ATGGAGTCCT CAGGCAACCC AGAGAGCACC ACCTTTTTTT ACTATGACCT TCAGAGCCAG 60
 CCGTGTGAGA ACCAGGCCTG GGTCTTTGCT ACCCTGCGCA CCACTGTCCT GTACTGCCTG 120
 GTGTTCTCTC TCAGCCTAGT GGGCAACAGC CTGGTCCTGT GGGTCCTGGT GAAGTATGAG 180
 AGCCTGGAGT CCCTCACCAA CATCTTCATC CTCAACCTGT GCCTCTCAGA CCTGGTGTTT 240
 GCCTGCTTGT TGCCTGTGTG GATCTCCCCA TACCACTGGG GCTGGGTGCT GGGAGACTTC 300
 CTCTGCAAAC TCCTCAATAT GATCTTCTCC ATCAGCCTCT ACAGCAGCAT CTTCTTCCTG 360
 ACCATCATGA CCATCCACCG CTACCTGTGG GTAGTGAGCC CCCTCTCCAC CCTGCGCGTC 420
 CCCACCCTCC GCTGCCGGGT GCTGGTGACC ATGGCTGTGT GGGTAGCCAG CATCCTGTCC 480
 TCCATCCTCG ACACCATCTT CCACAAGGTG CTTTCTTCGG GCTGTGATTA TTCCGAATC 540
 ACGTGGTACC TCACCTCCGT CTACCAGCAC AACCTCTTCT TCCTGCTGTC CCTGGGGATT 600
 ATCCTGTTCT GCTACGTGGA GATCCTCAGG ACCCTGTTCC GCTCACGCTC CAAGCGGCGC 660
 CACCGCACGA AAAAGCTCAT CTTGCCATC GTGGTGGCCT ACTTCCTCAG CTGGGGTCCC 720
 TACAACTTCA CCCTGTTTCT GCAGACGCTG TTTCGGACCC AGATCATCCG GAGCTGCGAG 780

30

GCCAAACAGC AGCTAGAATA CGCCCTGCTC ATCTGCCGCA ACCTCGCCTT CTCCCCTGTC 840
 TGCTTTAACC CGGTGCTCTA TGTCTTCGTG GGGGTCAAGT TCCGCACACA CCTGAAACAT 900
 GTTCTCCGGC AGTTCTGGTT CTGCCGGCTG CAGGCACCCA GCCCAGCCTC GATCCCCCAC 960
 TCCCCTGGTG CCTTCGCCTA TGAGGGCGCC TCCTTCTACT GA 1002

(169) INFORMATION FOR SEQ ID NO:168:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 333 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: not relevant

10

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:168:

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

20

30

Phe Phe Leu Leu Ser Leu Gly Ile Ile Leu Phe Cys Tyr Val Glu Ile
 195 200 205
 Leu Arg Thr Leu Phe Arg Ser Arg Ser Lys Arg Arg His Arg Thr Lys
 210 215 220
 Lys Leu Ile Phe Ala Ile Val Val Ala Tyr Phe Leu Ser Trp Gly Pro
 225 230 235 240
 Tyr Asn Phe Thr Leu Phe Leu Gln Thr Leu Phe Arg Thr Gln Ile Ile
 245 250 255
 Arg Ser Cys Glu Ala Lys Gln Gln Leu Glu Tyr Ala Leu Leu Ile Cys
 260 265 270
 Arg Asn Leu Ala Phe Ser His Cys Cys Phe Asn Pro Val Leu Tyr Val
 275 280 285
 Phe Val Gly Val Lys Phe Arg Thr His Leu Lys His Val Leu Arg Gln
 290 295 300
 Phe Trp Phe Cys Arg Leu Gln Ala Pro Ser Pro Ala Ser Ile Pro His
 305 310 315 320
 Ser Pro Gly Ala Phe Ala Tyr Glu Gly Ala Ser Phe Tyr
 325 330

10

(170) INFORMATION FOR SEQ ID NO:169:

20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 987 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:169:

ATGGACAACG CCTCGTTCTC GGAGCCCTGG CCCGCCAACG CATCGGGCCC GGACCCGGCG 60
 CTGAGCTGCT CCAACGCGTC GACTCTGGCG CCGCTGCCGG CGCCGCTGGC GGTGGCTGTA 120
 CCAGTTGTCT ACGCGGTGAT CTGCGCCGTG GGTCTGGCGG GCAACTCCGC CGTGCTGTAC 180
 GTGTTGCTGC GGGCGCCCCG CATGAAGACC GTCACCAACC TGTTTCATCCT CAACCTGGCC 240
 ATCGCCGACG AGCTCTTCAC GCTGGTGCTG CCCATCAACA TCGCCGACTT CCTGCTGCGG 300
 CAGTGGCCCT TCGGGGAGCT CATGTGCAAG CTCATCGTGG CTATCGACCA GTACAACACC 360
 TTCTCCAGCC TCTACTTCCT CACCGTCATG AGCGCCGACC GCTACCTGGT GGTGTTGGCC 420
 ACTGCGGAGT CGCGCCGGGT GGCCGGCCGC ACCTACAGCG CCGCGCGCGC GGTGAGCCTG 480

30

GCCGTGTGGG GGATCGTCAC ACTCGTCGTG CTGCCCTTCG CAGTCTTCGC CCGGCTAGAC 540
 GACGAGCAGG GCCGGCGCCA GTGCGTGCTA GTCTTTCCGC AGCCCGAGGC CTTCTGGTGG 600
 CGCGCGAGCC GCCTCTACAC GCTCGTGCTG GGCTTCGCCA TCCCCGTGTC CACCATCTGT 660
 GTCCTCTATA CCACCTGCT GTGCCGGCTG CATGCCATGC GGCTGGACAG CCACGCCAAG 720
 GCCCTGGAGC GCGCCAAGAA GCGGGTGAAG TTCCTGGTGG TGGCAATCCT GCGGGTGTGC 780
 CTCCTCTGCT GGACGCCCTA CCACCTGAGC ACCGTGGTGG CGCTCACCAC CGACCTCCCG 840
 CAGACGCCGC TGGTCATCGC TATCTCCTAC TTCATCACCA GCCTGACGTA CGCCAACAGC 900
 TGCCTCAACC CCTTCCTCTA CGCCTTCCTG GACGCCAGCT TCCGCAGGAA CCTCCGCCAG 960
 CTGATAACTT GCCGCGCGGC AGCCTGA 987

(171) INFORMATION FOR SEQ ID NO:170:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 328 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:170:

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

10

20

30

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

10

20

(172) INFORMATION FOR SEQ ID NO:171:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1002 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:171:

ATGCAGGCCG CTGGGCACCC AGAGCCCCTT GACAGCAGGG GCTCCTTCTC CCTCCCCACG 60
 ATGGGTGCCA ACGTCTCTCA GGACAATGGC ACTGGCCACA ATGCCACCTT CTCCGAGCCA 120
 CTGCCGTTCC TCTATGTGCT CCTGCCCGCC GTGTACTCCG GGAICTGTGC TGTGGGGCTG 180

ACTGGCAACA CGGCCGTCAT CCTTGTATC CTAAGGGCGC CCAAGATGAA GACGGTGACC 240
 AACGTGTTCA TCCTGAACCT GGCCGTCGCC GACGGGCTCT TCACGCTGGT ACTGCCTGTC 300
 AACATCGCGG AGCACCTGCT GCAGTACTGG CCCTTCGGGG AGCTGCTCTG CAAGCTGGTG 360
 CTGGCCGTCG ACCACTACAA CATCTTCTCC AGCATCTACT TCCTAGCCGT GATGAGCGTG 420
 GACCGATACC TGGTGGTGCT GGCCACCGTG AGGTCCCGCC ACATGCCCTG GCGCACCTAC 480
 CGGGGGGCGA AGGTCGCCAG CCTGTGTGTC TGGCTGGGCG TCACGGTCTT GGTCTGCCC 540
 TTCTTCTCTT TCGCTGGCGT CTACAGCAAC GAGCTGCAGG TCCCAAGCTG TGGGCTGAGC 600
 TTCCCGTGGC CCGAGCAGGT CTGGTICAAG GCCAGCCGTG TCTACACGTT GGTCTGGGC 660
 TTCGTGCTGC CCGTGTGCAC CATCTGTGTG CTCTACACAG ACCTCTGCG CAGGCTGCGG 720
 GCCGTGCGGC TCCGCTCTGG AGCCAAGGCT CTAGGCAAGG CCAGGCGGAA GGTGAAAGTC 780
 CTGGTCTCTG TCGTGTGGC CGTGTGCCTC CTCTGCTGGA CGCCCTTCCA CCTGGCCTCT 840
 GTCGTGGCCC TGACCACGGA CCTGCCCCAG ACCCCACTGG TCATCAGTAT GTCCTACGTC 900
 ATCACCAGCC TCACGTACGC CAACTCGTGC CTGAACCCCT TCCTCTACGC CTTTCTAGAT 960
 GACAACTTCC GGAAGAACTT CCGCAGCATA TTGCGGTGCT GA 1002

(173) INFORMATION FOR SEQ ID NO:172:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 333 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:172:

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

10

20

30

Asn Val Phe Ile Leu Asn Leu Ala Val Ala Asp Gly Leu Phe Thr Leu
 85 90 95
 Val Leu Pro Val Asn Ile Ala Glu His Leu Leu Gln Tyr Trp Pro Phe
 100 105 110
 Gly Glu Leu Leu Cys Lys Leu Val Leu Ala Val Asp His Tyr Asn Ile
 115 120 125
 Phe Ser Ser Ile Tyr Phe Leu Ala Val Met Ser Val Asp Arg Tyr Leu
 130 135 140
 Val Val Leu Ala Thr Val Arg Ser Arg His Met Pro Trp Arg Thr Tyr
 145 150 155 160
 Arg Gly Ala Lys Val Ala Ser Leu Cys Val Trp Leu Gly Val Thr Val
 165 170 175
 Leu Val Leu Pro Phe Phe Ser Phe Ala Gly Val Tyr Ser Asn Glu Leu
 180 185 190
 Gln Val Pro Ser Cys Gly Leu Ser Phe Pro Trp Pro Glu Gln Val Trp
 195 200 205
 Phe Lys Ala Ser Arg Val Tyr Thr Leu Val Leu Gly Phe Val Leu Pro
 210 215 220
 Val Cys Thr Ile Cys Val Leu Tyr Thr Asp Leu Leu Arg Arg Leu Arg
 225 230 235 240
 Ala Val Arg Leu Arg Ser Gly Ala Lys Ala Leu Gly Lys Ala Arg Arg
 245 250 255
 Lys Val Lys Val Leu Val Leu Val Val Leu Ala Val Cys Leu Leu Cys
 260 265 270
 Trp Thr Pro Phe His Leu Ala Ser Val Val Ala Leu Thr Thr Asp Leu
 275 280 285
 Pro Gln Thr Pro Leu Val Ile Ser Met Ser Tyr Val Ile Thr Ser Leu
 290 295 300
 Thr Tyr Ala Asn Ser Cys Leu Asn Pro Phe Leu Tyr Ala Phe Leu Asp
 305 310 315 320
 Asp Asn Phe Arg Lys Asn Phe Arg Ser Ile Leu Arg Cys
 325 330

10

20

30

(174) INFORMATION FOR SEQ ID NO:173:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1107 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:173:

```
ATGGTCCTTG AGGTGAGTGA CCACCAAGTG CTAAATGACG CCGAGGTTGC CGCCCTCCTG    60
GAGAACTTCA GCTCTTCCTA TGACTATGGA GAAAACGAGA GTGACTCGTG CTGTACCTCC    120
CGCCCTGCC CACAGGACTT CAGCCTGAAC TTCGACCGGG CCTTCCTGCC AGCCCTCTAC    180
AGCCTCCTCT TTCTGCTGGG GCTGCTGGGC AACGGCGCGG TGGCAGCCGT GCTGCTGAGC    240
CGGCGGACAG CCCTGAGCAG CACCGACACC TTCCTGCTCC ACCTAGCTGT AGCAGACACG    300
CTGCTGGTGC TGACACTGCC GCTCTGGGCA GTGGACGCTG CCGTCCAGTG GGTCTTTGGC    360
TCTGGCCTCT GCAAAGTGGC AGGTGCCCTC TTCAACATCA ACTTCTACGC AGGAGCCCTC    420
CTGCTGGCCT GCATCAGCTT TGACCGCTAC CTGAACATAG TTCATGCCAC CCAGCTCTAC    480
CGCCGGGGGC CCCC GGCCCG CGTGACCTC ACCTGCCTGG CTGTCTGGGG GCTCTGCCTG    540
CTTTTCGCCC TCCAGACTT CATCTTCCTG TCGGCCACC ACGACGAGCG CCTCAACGCC    600
ACCCACTGCC AATACAACTT CCCACAGGTG GGCCGCACGG CTCTGCGGGT GCTGCAGCTG    660
GTGGCTGGCT TTCTGCTGCC CCTGCTGGTC ATGGCCTACT GCTATGCCCA CATCCTGGCC    720
GTGCTGCTGG TTTCCAGGGG CCAGCGGCGC CTGCGGGCCA AGCGGCTGGT GGTGGTGGTC    780
GTGGTGGCCT TTGCCCTCTG CTGGACCCCC TATCACCTGG TGGTGCTGGT GGACATCCTC    840
ATGGACCTGG GCGCTTTGGC CCGCAACTGT GGCCGAGAAA GCAGGGTAGA CGTGGCCAAG    900
TCGGTCACCT CAGGCCTGGG CTACATGCAC TGCTGCCTCA ACCCGCTGCT CTATGCCTTT    960
GTAGGGGTCA AGTTCCGGGA GCGGATGTGG ATGCTGCTCT TGCGCCTGGG CTGCCCCAAC   1020
CAGAGAGGGC TCCAGAGGCA GCCATCGTCT TCCCGCCGGG ATTCATCCTG GTCTGAGACC   1080
TCAGAGGCCT CCTACTCGGG CTTGTGA                                     1107
```

10

20

(175) INFORMATION FOR SEQ ID NO:174:

30

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 368 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:174:

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

10

20

30

290	295	300
Gly Leu Gly Tyr Met His Cys Cys Leu Asn Pro Leu Leu Tyr Ala Phe		
305	310	315 320
Val Gly Val Lys Phe Arg Glu Arg Met Trp Met Leu Leu Leu Arg Leu		
	325	330 335
Gly Cys Pro Asn Gln Arg Gly Leu Gln Arg Gln Pro Ser Ser Ser Arg		
	340	345 350
Arg Asp Ser Ser Trp Ser Glu Thr Ser Glu Ala Ser Tyr Ser Gly Leu		
	355	360 365

10

(176) INFORMATION FOR SEQ ID NO:175:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1074 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:175:

ATGGCTGATG ACTATGGCTC TGAATCCACA TCTTCCATGG AAGACTACGT TAACTTCAAC	60
TTCACCTGACT TCTACTGTGA GAAAAACAAT GTCAGGCAGT TTGCGAGCCA TTTCTCTCCA	120
CCCTTGTAAT GGCTCGTGT CATCGTGGGT GCCTTGGGCA ACAGTCTTGT TATCCTTGTC	180
TACTGGTACT GCACAAGAGT GAAGACCATG ACCGACATGT TCCTTTTGAA TTTGGCAATT	240
GCTGACCTCC TCTTTCTTGT CACTCTTCCC TTCTGGGCCA TTGCTGCTGC TGACCAGTGG	300
AAGTTCCAGA CCTTCATGTG CAAGGTGGTC AACAGCATGT ACAAGATGAA CTTCTACAGC	360
TGTGTGTTGC TGATCATGTG CATCAGCGTG GACAGGTACA TTGCCATTGC CCAGGCCATG	420
AGAGCACATA CTTGGAGGGA GAAAAGGCTT TTGTACAGCA AAATGGTTTG CTTTACCATC	480
TGGGTATTGG CAGCTGCTCT CTGCATCCCA GAAATCTTAT ACAGCCAAAT CAAGGAGGAA	540
TCCGGCATTG CTATCTGCAC CATGGTTTAC CCTAGCGATG AGAGCACCAA ACTGAAGTCA	600
GCTGTCTTGA CCCTGAAGGT CATTCTGGGG TTCTTCCTTC CCTTCGTGGT CATGGCTTGC	660
TGCTATACCA TCATCATTCA CACCCTGATA CAAGCCAAGA AGTCTTCCAA GCACAAAGCC	720
AAGAAAGTGA CCATCACTGT CCTGACCGTC TTTGTCTTGT CTCAGTTTCC CTACAACTGC	780
ATTTTGTGG TGCAGACCAT TGACGCCTAT GCCATGTTCA TCTCCAAC TGCCGTTTCC	840

20

30

ACCAACATTG ACATCTGCTT CCAGGTCACC CAGACCATCG CCTTCTTCCA CAGTTGCCTG 900
 AACCCGTGTC TCTATGTTTT TGTGGGTGAG AGATTCCGCC GGGATCTCGT GAAAACCCTG 960
 AAGAACTTGG GTTGCATCAG CCAGGCCAG TGGGTTTCAT TTACAAGGAG AGAGGGAAGC 1020
 TTGAAGCTGT CGTCTATGTT GCTGGAGACA ACCTCAGGAG CACTCTCCCT CTGA 1074

(177) INFORMATION FOR SEQ ID NO:176:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 357 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: not relevant

10

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:176:

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

20

30

Asp Glu Ser Thr Lys Leu Lys Ser Ala Val Leu Thr Leu Lys Val Ile
 195 200 205
 Leu Gly Phe Phe Leu Pro Phe Val Val Met Ala Cys Cys Tyr Thr Ile
 210 215 220
 Ile Ile His Thr Leu Ile Gln Ala Lys Lys Ser Ser Lys His Lys Ala
 225 230 235 240
 Lys Lys Val Thr Ile Thr Val Leu Thr Val Phe Val Leu Ser Gln Phe
 245 250 255
 Pro Tyr Asn Cys Ile Leu Leu Val Gln Thr Ile Asp Ala Tyr Ala Met
 260 265 270
 Phe Ile Ser Asn Cys Ala Val Ser Thr Asn Ile Asp Ile Cys Phe Gln
 275 280 285
 Val Thr Gln Thr Ile Ala Phe Phe His Ser Cys Leu Asn Pro Val Leu
 290 295 300
 Tyr Val Phe Val Gly Glu Arg Phe Arg Arg Asp Leu Val Lys Thr Leu
 305 310 315 320
 Lys Asn Leu Gly Cys Ile Ser Gln Ala Gln Trp Val Ser Phe Thr Arg
 325 330 335
 Arg Glu Gly Ser Leu Lys Leu Ser Ser Met Leu Leu Glu Thr Thr Ser
 340 345 350
 Gly Ala Leu Ser Leu
 355

10

20

(178) INFORMATION FOR SEQ ID NO:177:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1110 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:177:

30

ATGGCCTCAT CGACCACTCG GGGCCCCAGG GTTTCTGACT TATTTTCTGG GCTGCCGCCG 60
 GCGGTCACAA CTCCCCCAA CCAGAGCGCA GAGGCCTCGG CGGGCAACGG GTCGGTGGCT 120
 GCGCGGGACG CTCCAGCCGT CACGCCCTTC CAGAGCCTGC AGCTGGTGCA TCAGCTGAAG 180
 GGGCTGATCG TGCTGCTCTA CAGCGTCGTG GTGGTCGTGG GGCTGGTGGG CAACTGCCTG 240
 CTGGTGCTGG TGATCGCGCG GGTGCCGCGG CTGCACAACG TGACGAACTT CCTCATCGGC 300

AACCTGGCCT TGTCCGACGT GTCATGTGC ACCGCCTGCG TGCCGCTCAC GCTGGCCTAT 360
 GCCTTCGAGC CACGCGGCTG GGTGTTGCGC GCGGCGCTGT GCCACCTGGT CTTCTTCCTG 420
 CAGCCGGTCA CCGTCTATGT GTCGGTGTTC ACGCTCACCA CCATCGCAGT GGACCGCTAC 480
 GTCGTGCTGG TGCACCGCT GAGGCGCGCA TCTCGCTGCG CCTCAGCCTA CGCTGTGCTG 540
 GCCATCTGGG CGCTGTCCGC GGTGCTGGCG CTGCCGCCCG CCGTGACAC CTATCACGTG 600
 GAGCTCAAGC CGCAGACGT GCGCCTCTGC GAGGAGTTCT GGGGCTCCCA GGAGCGCCAG 660
 CGCCAGCTCT ACGCCTGGGG GCTGCTGCTG GTCACCTACC TGCTCCCTCT GCTGGTCATC 720
 CTCCTGTCTT ACGTCCGGGT GTCAGTGAAG CTCGCAACC GCGTGGTGCC GGGCTGCGTG 780
 ACCCAGAGCC AGGCCGACTG GGACCGCGCT CGGCGCCGGC GCACCAAATG CTTGCTGGTG 840
 GTGGTCGTGG TGGTGTTCGC CGTCTGCTGG CTGCCGCTGC ACGTCTTCAA CCTGCTGCGG 900
 GACCTCGACC CCCACCCCAT CGACCTTAC GCCTTTGGGC TGGTGCAGCT GCTCTGCCAC 960
 TGGCTCGCCA TGAGTTCGGC CTGCTACAAC CCCTTCATCT ACGCCTGGCT GCACGACAGC 1020
 TTCCGCGAGG AGCTGCGCAA ACTGTTGGTC GCTTGGCCCC GCAAGATAGC CCCCCATGGC 1080
 CAGAATATGA CCGTCAGCGT GGTCACTCTA 1110

(179) INFORMATION FOR SEQ ID NO:178:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 369 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:178:

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

10

20

30

Leu Val Leu Val Ile Ala Arg Val Pro Arg Leu His Asn Val Thr Asn
 85 90 95
 Phe Leu Ile Gly Asn Leu Ala Leu Ser Asp Val Leu Met Cys Thr Ala
 100 105 110
 Cys Val Pro Leu Thr Leu Ala Tyr Ala Phe Glu Pro Arg Gly Trp Val
 115 120 125
 Phe Gly Gly Gly Leu Cys His Leu Val Phe Phe Leu Gln Pro Val Thr
 130 135 140
 Val Tyr Val Ser Val Phe Thr Leu Thr Thr Ile Ala Val Asp Arg Tyr
 145 150 155 160
 Val Val Leu Val His Pro Leu Arg Arg Ala Ser Arg Cys Ala Ser Ala
 165 170 175
 Tyr Ala Val Leu Ala Ile Trp Ala Leu Ser Ala Val Leu Ala Leu Pro
 180 185 190
 Pro Ala Val His Thr Tyr His Val Glu Leu Lys Pro His Asp Val Arg
 195 200 205
 Leu Cys Glu Glu Phe Trp Gly Ser Gln Glu Arg Gln Arg Gln Leu Tyr
 210 215 220
 Ala Trp Gly Leu Leu Leu Val Thr Tyr Leu Leu Pro Leu Leu Val Ile
 225 230 235 240
 Leu Leu Ser Tyr Val Arg Val Ser Val Lys Leu Arg Asn Arg Val Val
 245 250 255
 Pro Gly Cys Val Thr Gln Ser Gln Ala Asp Trp Asp Arg Ala Arg Arg
 260 265 270
 Arg Arg Thr Lys Cys Leu Leu Val Val Val Val Val Phe Ala Val
 275 280 285
 Cys Trp Leu Pro Leu His Val Phe Asn Leu Leu Arg Asp Leu Asp Pro
 290 295 300
 His Ala Ile Asp Pro Tyr Ala Phe Gly Leu Val Gln Leu Leu Cys His
 305 310 315 320
 Trp Leu Ala Met Ser Ser Ala Cys Tyr Asn Pro Phe Ile Tyr Ala Trp
 325 330 335
 Leu His Asp Ser Phe Arg Glu Glu Leu Arg Lys Leu Leu Val Ala Trp
 340 345 350
 Pro Arg Lys Ile Ala Pro His Gly Gln Asn Met Thr Val Ser Val Val
 355 360 365
 Ile

10

20

30

(180) INFORMATION FOR SEQ ID NO:179:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1083 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:179:

ATGGACCCAG AAGAACTTC AGTTTATTIG GATTATTACT ATGCTACGAG CCCAACTCT	60
GACATCAGGG AGACCCACTC CCATGTTCC TACACCTCTG TCTTCCTTCC AGTCTTTTAC	120
ACAGCTGTGT TCCTGACTGG AGTGCTGGGG AACCTTGTTT TCATGGGAGC GTTGCAATTC	180
AAACCCGGCA GCCGAAGACT GATCGACATC TTTATCATCA ATCTGGCTGC CTCTGACTTC	240
ATTTTCTTIG TCACATTGCC TCTCTGGGTG GATAAAGAAG CATCTCTAGG ACTGTGGAGG	300
ACGGGCTCCT TCCTGTGCAA AGGGAGCTCC TACATGATCT CCGTCAATAT GCACTGCAGT	360
GTCCTCCTGC TCACTTGCA T GAGTGTGAC CGCTACCTGG CCATTGTGTG GCCAGTCGTA	420
TCCAGGAAAT TCAGAAGGAC AGACTGTGCA TATGTAGTCT GTGCCAGCAT CTGGTTTATC	480
TCCTGCCTGC TGGGGTTGCC TACTCTTCTG TCCAGGGAGC TCACGCTGAT TGATGATAAG	540
CCATACTGTG CAGAGAAAAA GGCAACTCCA ATTAACTCA TATGGTCCCT GGTGGCCTTA	600
ATTTTCACCT TTTTGTCCC TTGTGTGAGC ATTGTGACCT GCTACTGTTG CATTGCAAGG	660
AAGCTGTGTG CCCATTACCA GCAATCAGGA AAGCACAACA AAAAGCTGAA GAAATCTAAG	720
AAGATCATCT TTATTGTGCT GGCAGCCTTT CTGTCTCCTT GGCTGCCCTT CAATACTTTC	780
AAGTTCCTGG CCATTGTCTC TGGGTTGCGG CAAGAACACT ATTTACCTC AGCTATTCTT	840
CAGCTTGGTA TGGAGGTGAG TGGACCCTTG GCATTTGCCA ACAGCTGTGT CAACCCTTTC	900
ATTACTATA TCTTCGACAG CTACATCCGC CGGGCCATTG TCCACTGCTT GTGCCCTTGC	960
CTGAAAAACT ATGACTTTGG GAGTAGCACT GAGACATCAG ATAGTCACCT CACTAAGGCT	1020
CTCTCCACCT TCATTCATGC AGAAGATTTT GCCAGGAGGA GGAAGAGGTC TGTGTCACTC	1080
TAA	1083

(181) INFORMATION FOR SEQ ID NO:180:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 360 amino acids

10

20

30

(B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:180:

```

Met Asp Pro Glu Glu Thr Ser Val Tyr Leu Asp Tyr Tyr Tyr Ala Thr
1          5          10          15
Ser Pro Asn Ser Asp Ile Arg Glu Thr His Ser His Val Pro Tyr Thr
          20          25          30
Ser Val Phe Leu Pro Val Phe Tyr Thr Ala Val Phe Leu Thr Gly Val
          35          40          45
Leu Gly Asn Leu Val Leu Met Gly Ala Leu His Phe Lys Pro Gly Ser
          50          55          60
Arg Arg Leu Ile Asp Ile Phe Ile Ile Asn Leu Ala Ala Ser Asp Phe
65          70          75          80
Ile Phe Leu Val Thr Leu Pro Leu Trp Val Asp Lys Glu Ala Ser Leu
          85          90          95
Gly Leu Trp Arg Thr Gly Ser Phe Leu Cys Lys Gly Ser Ser Tyr Met
          100          105          110
Ile Ser Val Asn Met His Cys Ser Val Leu Leu Leu Thr Cys Met Ser
          115          120          125
Val Asp Arg Tyr Leu Ala Ile Val Trp Pro Val Val Ser Arg Lys Phe
          130          135          140
Arg Arg Thr Asp Cys Ala Tyr Val Val Cys Ala Ser Ile Trp Phe Ile
145          150          155          160
Ser Cys Leu Leu Gly Leu Pro Thr Leu Leu Ser Arg Glu Leu Thr Leu
          165          170          175
Ile Asp Asp Lys Pro Tyr Cys Ala Glu Lys Lys Ala Thr Pro Ile Lys
          180          185          190
Leu Ile Trp Ser Leu Val Ala Leu Ile Phe Thr Phe Phe Val Pro Leu
          195          200          205
Leu Ser Ile Val Thr Cys Tyr Cys Cys Ile Ala Arg Lys Leu Cys Ala
          210          215          220
His Tyr Gln Gln Ser Gly Lys His Asn Lys Lys Leu Lys Lys Ser Lys
225          230          235          240
Lys Ile Ile Phe Ile Val Val Ala Ala Phe Leu Val Ser Trp Leu Pro
          245          250          255

```

10

20

30

Phe Asn Thr Phe Lys Phe Leu Ala Ile Val Ser Gly Leu Arg Gln Glu
 260 265 270
 His Tyr Leu Pro Ser Ala Ile Leu Gln Leu Gly Met Glu Val Ser Gly
 275 280 285
 Pro Leu Ala Phe Ala Asn Ser Cys Val Asn Pro Phe Ile Tyr Tyr Ile
 290 295 300
 Phe Asp Ser Tyr Ile Arg Arg Ala Ile Val His Cys Leu Cys Pro Cys
 305 310 315 320
 Leu Lys Asn Tyr Asp Phe Gly Ser Ser Thr Glu Thr Ser Asp Ser His
 325 330 335
 Leu Thr Lys Ala Leu Ser Thr Phe Ile His Ala Glu Asp Phe Ala Arg
 340 345 350
 Arg Arg Lys Arg Ser Val Ser Leu
 355 360

10

(182) INFORMATION FOR SEQ ID NO:181:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1020 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

20

- (ii) MOLECULE TYPE: DNA (genomic)

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:181:

ATGAATGGCC TTGAAGTGGC TCCCCAGGT CTGATCACCA ACTTCTCCCT GGCCACGGCA 60
 GAGCAATGTG GCCAGGAGAC GCCACTGGAG AACATGCTGT TCGCCTCCTT CTACCTTCTG 120
 GATTTTATCC TGGCTTTAGT TGGCAATACC CTGGCTCTGT GGCTTTTCAT CCGAGACCAC 180
 AAGTCCGGGA CCCCAGCCAA CGTGTTCTGT ATGCATCTGG CCGTGGCCGA CTTGTCTGTG 240
 GTGCTGGTCC TGCCACCCCG CCTGGTCTAC CACTTCTCTG GGAACCACTG GCCATTGGGG 300
 GAAATCGCAT GCCGTCTCAC CGGCTTCCTC TTCTACCTCA ACATGTACGC CAGCATCTAC 360
 TTCCTCACCT GCATCAGCGC CGACCGTTTC CTGGCCATTG TGCACCCGGT CAAGTCCCTC 420
 AAGCTCCGCA GGCCCTCTA CGCACACCTG GCCTGTGCCT TCCTGTGGGT GGTGGTGGCT 480
 GTGGCCATGG CCCCCTGTCT GGTGAGCCCA CAGACCGTGC AGACCAACCA CACGGTGGTC 540
 TGCCTGCAGC TGTACCGGGA GAAGGCCTCC CACCATGCCC TGGTGTCCCT GGCAGTGGCC 600
 TTCACCTTCC CGTTCATCAC CACGGTCACC TGCTACCTGC TGATCATCCG CAGCCTGCGG 660

30

CAGGGCCTGC GTGTGGAGAA GCGCCTCAAG ACCAAGGCAA AACGCATGAT CGCCATAGTG 720
 CTGGCCATCT TCCTGGTCTG CTTCGTGCCC TACCACGTCA ACCGCTCCGT CTACGTGCTG 780
 CACTACCGCA GCCATGGGGC CTCCTGCGCC ACCCAGCGCA TCCTGGCCCT GGCAAACCGC 840
 ATCACCTCCT GCCTCACCAG CCTCAACGGG GCACTCGACC CCATCATGTA TTTCTTCGTG 900
 GCTGAGAAGT TCCGCCACGC CCTGTGCAAC TTGCTCTGTG GCAAAAGGCT CAAGGGCCCG 960
 CCCCCAGCT TCGAAGGGAA AACCAACGAG AGCTCGCTGA GTGCCAAGTC AGAGCTGTGA 1020

(183) INFORMATION FOR SEQ ID NO:182:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 339 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:182:

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

10

20

30

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

10

20

(183) INFORMATION FOR SEQ ID NO:183:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 996 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:183:

ATGATCACCC TGAACAATCA AGATCAACCT GTCCCTTTTIA ACAGCTCACA TCCAGATGAA	60
TACAAAATTG CAGCCCTTGT CTTCTATAGC TGTATCTTCA TAATTGGATT ATTTGTTAAC	120
ATCACTGCAT TATGGGTTTT CAGTTGTACC ACCAAGAAGA GAACCACGGT AACCATCTAT	180
ATGATGAATG TGGCATTAGT GGAATTGATA TTTATAATGA CTTTACCCTT TCGAATGTTT	240

TATTATGCAA AAGATGAATG GCCATTTGGA GAGTACTTCT GCCAGATTCT TGGAGCTCTC 300
 ACAGTGT TTTT ACCCAAGCAT TGCTTTATGG CTTCTTGCCCT TTATTAGTGC TGACAGATAC 360
 ATGGCCATTG TACAGCCGAA GTACGCCAAA GAACTTAAAA ACACGTGCAA AGCCGTGCTG 420
 GCGTGTGTGG GAGTCTGGAT AATGACCCCTG ACCACGACCA CCCCTCTGCT ACTGCTCTAT 480
 AAAGACCCAG ATAAAGACTC CACTCCCGCC ACCTGCCTCA AGATTTCTGA CATCATCTAT 540
 CTAAAAGCTG TGAACGTGCT GAACCTCACT CGACTGACAT TTTTTTCTT GATTCCTTTG 600
 TTCATCATGA TTGGGTGCTA CTTGGTCATT ATTCATAATC TCCTTCACGG CAGGACGTCT 660
 AAGCTGAAAC CCAAAGTCAA GGAGAAGTCC AAAAGGATCA TCATCACGCT GCTGGTGCAG 720
 GTGCTCGTCT GCTTTATGCC CTTCCACATC TGTTCGCTT TCCTGATGCT GGGAACGGGG 780
 GAGAATAGTT ACAATCCCTG GGGAGCCTTT ACCACCTTCC TCATGAACCT CAGCACGTGT 840
 CTGGATGTGA TTCTCTACTA CATCGTTTCA AAACAATTTC AGGCTCGAGT CATTAGTGTC 900
 ATGCTATACC GTAATTACCT TCGAAGCATG CGCAGAAAAA GTTCCGATC TGGTAGTCTA 960
 AGGTCACTAA GCAATATAAA CAGTGAAATG TTATGA 996

10

(185) INFORMATION FOR SEQ ID NO:184:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 331 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

20

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:184:

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

30

Leu Gly Ala Leu Thr Val Phe Tyr Pro Ser Ile Ala Leu Trp Leu Leu
 100 105 110
 Ala Phe Ile Ser Ala Asp Arg Tyr Met Ala Ile Val Gln Pro Lys Tyr
 115 120 125
 Ala Lys Glu Leu Lys Asn Thr Cys Lys Ala Val Leu Ala Cys Val Gly
 130 135 140
 Val Trp Ile Met Thr Leu Thr Thr Thr Thr Pro Leu Leu Leu Leu Tyr
 145 150 155 160
 Lys Asp Pro Asp Lys Asp Ser Thr Pro Ala Thr Cys Leu Lys Ile Ser
 165 170 175
 Asp Ile Ile Tyr Leu Lys Ala Val Asn Val Leu Asn Leu Thr Arg Leu
 180 185 190
 Thr Phe Phe Phe Leu Ile Pro Leu Phe Ile Met Ile Gly Cys Tyr Leu
 195 200 205
 Val Ile Ile His Asn Leu Leu His Gly Arg Thr Ser Lys Leu Lys Pro
 210 215 220
 Lys Val Lys Glu Lys Ser Lys Arg Ile Ile Ile Thr Leu Leu Val Gln
 225 230 235 240
 Val Leu Val Cys Phe Met Pro Phe His Ile Cys Phe Ala Phe Leu Met
 245 250 255
 Leu Gly Thr Gly Glu Asn Ser Tyr Asn Pro Trp Gly Ala Phe Thr Thr
 260 265 270
 Phe Leu Met Asn Leu Ser Thr Cys Leu Asp Val Ile Leu Tyr Tyr Ile
 275 280 285
 Val Ser Lys Gln Phe Gln Ala Arg Val Ile Ser Val Met Leu Tyr Arg
 290 295 300
 Asn Tyr Leu Arg Ser Met Arg Arg Lys Ser Phe Arg Ser Gly Ser Leu
 305 310 315 320
 Arg Ser Leu Ser Asn Ile Asn Ser Glu Met Leu
 325 330

10

20

30

(186) INFORMATION FOR SEQ ID NO:185:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1077 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:185:

```

ATGCCCTCTG TGTCTCCAGC GGGGCCCTCG GCCGGGGCAG TCCCCAATGC CACCGCAGTG    60
ACAACAGTGC GGACCAATGC CAGCGGGCTG GAGGTGCCCC TGTTCCACCT GTTTGCCCCG    120
CTGGACGAGG AGCTGCATGG CACCTTCCCA GGCCTGTGCG TGGCGTGAT GCGGGTGAC    180
GGAGCCATCT TCCTGGCAGG GCTGGTGCTC AACGGGCTGG CGCTGTACGT CTTCTGCTGC    240
CGCACCCGGG CCAAGACACC CTCAGTCATC TACACCATCA ACCTGGTGGT GACCGATCTA    300
CTGGTAGGGC TGTCCCTGCC CACGCGCTTC GCTGTGTACT ACGGCGCCAG GGGCTGCCTG    360
CGCTGTGCCT TCCCGCACGT CCTCGGTTAC TTCCTCAACA TGCACTGCTC CATCCTCTTC    420
CTCACCTGCA TCTGCGTGGA CCGCTACCTG GCCATCGTGC GGCCCGAAGG CTCCCGCCGC    480
TGCCGCCAGC CTGCCTGTGC CAGGGCCGTG TGCGCCTTCG TGTGGCTGGC CGCCGGTGCC    540
GTCACCTGT CGGTGCTGGG CGTGACAGGC AGCCGGCCCT GCTGCCGTGT CTTTGCGCTG    600
ACTGTCCTGG AGTTCCTGCT GCCCCTGCTG GTCATCAGCG TGTTTACCGG CCGCATCATG    660
TGTGCACTGT CGCGGCCGGG TCTGCTCCAC CAGGGTCGCC AGCGCCGCGT GCGGGCCAAG    720
CAGCTCCTGC TCACGGTGCT CATCATCTTT CTCGTCTGCT TCACGCCCTT CCACGCCCGC    780
CAAGTGGCCG TGGCGCTGTG GCCCGACATG CCACACCACA CGAGCCTCGT GGTCTACCAC    840
GTGGCCGTGA CCCTCAGCAG CCTCAACAGC TGCATGGACC CCATCGTCTA CTGCTTCGTC    900
ACCAGTGGCT TCCAGGCCAC CGTCCGAGGC CTCTTCGGCC AGCACGGAGA GCGTGAGCCC    960
AGCAGCGGTG ACGTGGTCAG CATGCACAGG AGCTCCAAGG GCTCAGGCCG TCATCACATC   1020
CTCAGTGCCG GCCCTCACGC CCTCACCAGG GCCCTGGCTA ATGGGCCCGA GGCTTAG    1077

```

10

20

(187) INFORMATION FOR SEQ ID NO:186:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 358 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

30

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:186:

```

Met Pro Ser Val Ser Pro Ala Gly Pro Ser Ala Gly Ala Val Pro Asn
1           5           10           15
Ala Thr Ala Val Thr Thr Val Arg Thr Asn Ala Ser Gly Leu Glu Val
          20           25           30

```

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

10

20

30

325 330 335
 Arg His His Ile Leu Ser Ala Gly Pro His Ala Leu Thr Gln Ala Leu
 340 345 350
 Ala Asn Gly Pro Glu Ala
 355

(188) INFORMATION FOR SEQ ID NO:187:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1050 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:187:

ATGAACTCCA CCTGGATGG TAATCAGAGC AGCCACCCTT TTTGCCTCTT GGCATTGGGC	60
TATTTGGAAA CTGTCAATTT TTGCCTTTTG GAAGTATTGA TTATTGTCTT TCTAACTGTA	120
TTGATTATTT CTGGCAACAT CATTGTGATT TTTGTATTTT ACTGTGCACC TTTGTTGAAC	180
CATCACACTA CAAGTTATTT TATCCAGACT ATGGCATATG CTGACCTTTT TGTGGGGTG	240
AGCTGCGTGG TCCCTTCTTT ATCACTCCTC CATCACCCCC TTCCAGTAGA GGAGTCCTTG	300
ACTTGCCAGA TATTTGGTTT TGTAGTATCA GTTCTGAAGA GCGTCTCCAT GGCTTCTCTG	360
GCCTGTATCA GCATTGATAG ATACATTGCC ATTACTAAAC CTTTAACCTA TAATACTCTG	420
GTTACACCCT GGAGACTACG CCTGTGTATT TTCCTGATTT GGCTATACTC GACCCTGGTC	480
TTCCTGCCTT CCTTTTTCCTA CTGGGGCAAA CCTGGATATC ATGGAGATGT GTTTCAGTGG	540
TGTGCGGAGT CCTGGCACAC CGACTCCTAC TTCACCCTGT TCATCGTGAT GATGTTATAT	600
GCCCCAGCAG CCCTTATTGT CTGCTTCACC TATTTCAACA TCTTCCGCAT CTGCCAACAG	660
CACACAAAGG ATATCAGCGA AAGGCAAGCC CGCTTCAGCA GCCAGAGTGG GGAGACTGGG	720
GAAGTGCAGG CCTGTCCTGA TAAGCGCTAT AAAATGGTCC TGTTCGAAT CACTAGTGTA	780
TTTTACATCC TCTGGTTGCC ATATATCATC TACTTCTTGT TGGAAAGCTC CACTGGCCAC	840
AGCAACCGCT TCGCATCCTT CTGACCACC TGGCTTGCTA TTAGTAACAG TTTCTGCAAC	900
TGTGTAATTT ATAGTCTCTC CAACAGTGTA TTCCAAAGAG GACTAAAGCG CCTCTCAGGG	960
GCTATGTGTA CTTCTTGTGC AAGTCAGACT ACAGCCAACG ACCCTTACAC AGTTAGAAGC	1020
AAAGGCCCTC TTAATGGATG TCATATCTGA	1050

20

30

(189) INFORMATION FOR SEQ ID NO:188:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 349 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:188:

```

Met Asn Ser Thr Leu Asp Gly Asn Gln Ser Ser His Pro Phe Cys Leu
1           5           10           15
Leu Ala Phe Gly Tyr Leu Glu Thr Val Asn Phe Cys Leu Leu Glu Val
20          25          30
Leu Ile Ile Val Phe Leu Thr Val Leu Ile Ile Ser Gly Asn Ile Ile
35          40          45
Val Ile Phe Val Phe His Cys Ala Pro Leu Leu Asn His His Thr Thr
50          55          60
Ser Tyr Phe Ile Gln Thr Met Ala Tyr Ala Asp Leu Phe Val Gly Val
65          70          75          80
Ser Cys Val Val Pro Ser Leu Ser Leu Leu His His Pro Leu Pro Val
85          90          95
Glu Glu Ser Leu Thr Cys Gln Ile Phe Gly Phe Val Val Ser Val Leu
100         105         110
Lys Ser Val Ser Met Ala Ser Leu Ala Cys Ile Ser Ile Asp Arg Tyr
115         120         125
Ile Ala Ile Thr Lys Pro Leu Thr Tyr Asn Thr Leu Val Thr Pro Trp
130         135         140
Arg Leu Arg Leu Cys Ile Phe Leu Ile Trp Leu Tyr Ser Thr Leu Val
145         150         155         160
Phe Leu Pro Ser Phe Phe His Trp Gly Lys Pro Gly Tyr His Gly Asp
165         170         175
Val Phe Gln Trp Cys Ala Glu Ser Trp His Thr Asp Ser Tyr Phe Thr
180         185         190
Leu Phe Ile Val Met Met Leu Tyr Ala Pro Ala Ala Leu Ile Val Cys
195         200         205
Phe Thr Tyr Phe Asn Ile Phe Arg Ile Cys Gln Gln His Thr Lys Asp
210         215         220

```

10

20

30

Ile Ser Glu Arg Gln Ala Arg Phe Ser Ser Gln Ser Gly Glu Thr Gly
 225 230 235 240

Glu Val Gln Ala Cys Pro Asp Lys Arg Tyr Lys Met Val Leu Phe Arg
 245 250 255

Ile Thr Ser Val Phe Tyr Ile Leu Trp Leu Pro Tyr Ile Ile Tyr Phe
 260 265 270

Leu Leu Glu Ser Ser Thr Gly His Ser Asn Arg Phe Ala Ser Phe Leu
 275 280 285

Thr Thr Trp Leu Ala Ile Ser Asn Ser Phe Cys Asn Cys Val Ile Tyr
 290 295 300

Ser Leu Ser Asn Ser Val Phe Gln Arg Gly Leu Lys Arg Leu Ser Gly
 305 310 315 320

Ala Met Cys Thr Ser Cys Ala Ser Gln Thr Thr Ala Asn Asp Pro Tyr
 325 330 335

Thr Val Arg Ser Lys Gly Pro Leu Asn Gly Cys His Ile
 340 345

10

(190) INFORMATION FOR SEQ ID NO:189:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1302 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:189:

ATGTGTTTTT CTCCCATCTT GGAAATCAAC ATGCAGTCTG AATCTAACAT TACAGTGCGA 60

GATGACATTG ATGACATCAA CACCAATATG TACCAACCAC TATCATATCC GTTAAGCTTT 120

CAAGTGTCTC TCACCGGATT TCTTATGTTA GAAATTGTGT TGGGACTTGG CAGCAACCTC 180

ACTGTATTGG TACTTTACTG CATGAAATCC AACTTAATCA ACTCTGTCAG TAACATTATT 240

ACAATGAATC TTCATGTACT TGATGTAATA APTTGTGTGG GATGTATTCC TCTAACTATA 300

GTTATCCTTC TGCTTTCACCT GGAGAGTAAC ACTGCTCTCA TTTGCTGTTT CCATGAGGCT 360

TGTGTATCTT TTGCAAGTGT CTCAACAGCA ATCAACGTTT TTGCTATCAC FTTGGACAGA 420

TATGACATCT CTGTAAAACC TGCAAACCGA ATTCTGACAA TGGGCAGAGC TGTAATGTTA 480

ATGATATCCA TTTGGATTTT TTCTTTTTC TCTTTCCTGA TTCCTTTTAT TGAGGTAAAT 540

TTTTTCAGTC TTCAAAGTGG AAATACCTGG GAAAACAAGA CACTTTTATG TGTCAGTACA 600

30

AATGAATACT AACTGAACT GGAATGTAT TATCACCTGT TAGTACAGAT CCCAATATTC 660
 TTTTTCCTG TTGTAGTAAT GTTAATCACA TACACCAAAA TACTTCAGGC TCTTAATATT 720
 CGAATAGGCA CAAGATTTTC AACAGGGCAG AAGAAGAAAG CAAGAAAGAA AAAGACAATT 780
 TCTCTAACCA CACAACATGA GGCTACAGAC ATGTCACAAA GCAGTGCTGG GAGAAATGTA 840
 GTCTTTGGTG TAAGAACTTC AGTTTCTGTA ATAATTGCCC TCCGGCGAGC TGTGAAACGA 900
 CACCGTGAAC GACGAGAAAG ACAAAGAGA GTCAAGAGGA TGTCTTTATT GATTATTTCT 960
 ACATTTCTTC TCTGCTGGAC ACCAATTTCT GTTTTAAATA CCACCATTTT ATGTTTAGGC 1020
 CCAAGTGACC TTTTAGTAAA ATTAAGATTG TGTTTTTTAG TCATGGCTTA TGAACAACCT 1080
 ATATTTCAAC CTCTATTATA TGCATTCCT AGACAAAAAT TTCAAAAGGT CTTGAAAAGT 1140
 AAAATGAAAA AGCGAGTTGT TTCTATAGTA GAAGCTGATC CCCTGCCTAA TAATGCTGTA 1200
 ATACACAACCT CTTGGATAGA TCCCAAAAGA AACAAAAAAA TTACCTTTGA AGATAGTGAA 1260
 ATAAGAGAAA AACGTTTAGT GCCTCAGGTT GTCACAGACT AG 1302

(191) INFORMATION FOR SEQ ID NO:190:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 433 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:190:

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

100	105	110
Leu Ile Cys Cys Phe His Glu Ala Cys Val Ser Phe Ala Ser Val Ser 115 120 125		
Thr Ala Ile Asn Val Phe Ala Ile Thr Leu Asp Arg Tyr Asp Ile Ser 130 135 140		
Val Lys Pro Ala Asn Arg Ile Leu Thr Met Gly Arg Ala Val Met Leu 145 150 155 160		
Met Ile Ser Ile Trp Ile Phe Ser Phe Phe Ser Phe Leu Ile Pro Phe 165 170 175		
Ile Glu Val Asn Phe Phe Ser Leu Gln Ser Gly Asn Thr Trp Glu Asn 180 185 190		
Lys Thr Leu Leu Cys Val Ser Thr Asn Glu Tyr Tyr Thr Glu Leu Gly 195 200 205		
Met Tyr Tyr His Leu Leu Val Gln Ile Pro Ile Phe Phe Phe Thr Val 210 215 220		
Val Val Met Leu Ile Thr Tyr Thr Lys Ile Leu Gln Ala Leu Asn Ile 225 230 235 240		
Arg Ile Gly Thr Arg Phe Ser Thr Gly Gln Lys Lys Lys Ala Arg Lys 245 250 255		
Lys Lys Thr Ile Ser Leu Thr Thr Gln His Glu Ala Thr Asp Met Ser 260 265 270		
Gln Ser Ser Gly Gly Arg Asn Val Val Phe Gly Val Arg Thr Ser Val 275 280 285		
Ser Val Ile Ile Ala Leu Arg Arg Ala Val Lys Arg His Arg Glu Arg 290 295 300		
Arg Glu Arg Gln Lys Arg Val Lys Arg Met Ser Leu Leu Ile Ile Ser 305 310 315 320		
Thr Phe Leu Leu Cys Trp Thr Pro Ile Ser Val Leu Asn Thr Thr Ile 325 330 335		
Leu Cys Leu Gly Pro Ser Asp Leu Leu Val Lys Leu Arg Leu Cys Phe 340 345 350		
Leu Val Met Ala Tyr Gly Thr Thr Ile Phe His Pro Leu Leu Tyr Ala 355 360 365		
Phe Thr Arg Gln Lys Phe Gln Lys Val Leu Lys Ser Lys Met Lys Lys 370 375 380		
Arg Val Val Ser Ile Val Glu Ala Asp Pro Leu Pro Asn Asn Ala Val 385 390 395 400		

10

20

30

Ile His Asn Ser Trp Ile Asp Pro Lys Arg Asn Lys Lys Ile Thr Phe
 405 410 415
 Glu Asp Ser Glu Ile Arg Glu Lys Arg Leu Val Pro Gln Val Val Thr
 420 425 430

Asp

(192) INFORMATION FOR SEQ ID NO:191:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1209 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:191:

ATGTTGTGTC CTTCCAAGAC AGATGGCTCA GGGCACTCTG GTAGGATTCA CCAGGAAACT	60
CATGGAGAAG GGAAAAGGGA CAAGATTAGC AACAGTGAAG GGAGGGAGAA TGGTGGGAGA	120
GGATTCCAGA TGAACGGTGG GTCGCTGGAG GCTGAGCATG CCAGCAGGAT GTCAGTTCTC	180
AGAGCAAAGC CCATGTCAAA CAGCCAACGC TTGCTCCTTC TGTCCCCAGG ATCACCTCCT	240
CGCACGGGGA GCATCTCCTA CATCAACATC ATCATGCCTT CGGTGTTTCG CACCATCTGC	300
CTCCTGGGCA TCATCGGGAA CTCCACGGTC ATCTTCGCGG TCGTGAAGAA GTCCAAGCTG	360
CACTGGTGCA ACAACGTCCC CGACATCTTC ATCATCAACC TCTCGGTAGT AGATCTCCTC	420
TTTCTCCTGG GCATGCCCTT CATGATCCAC CAGCTCATGG GCAATGGGGT GTGGCACTTT	480
GGGGAGACCA TGTGCACCCT CATCACGGCC ATGGATGCCA ATAGTCAGTT CACCAGCACC	540
TACATCCTGA CCGCCATGGC CATTGACCGC TACCTGGCCA CTGTCCACCC CATCTCTTCC	600
ACGAAGTTCC GGAAGCCCTC TGTGGCCACC CTGGTGATCT GCCTCCTGTG GGCCCTCTCC	660
TTCATCAGCA TCACCCCTGT GTGGCTGTAT GCCAGACTCA TCCCCTTCCC AGGAGGTGCA	720
GTGGGCTGCG GCATACGCCT GCCCAACCCA GACACTGACC TCTACTGGTT CACCCTGTAC	780
CAGTTTTTCC TGGCCTTTGC CCTGCCTTTT GTGGTCATCA CAGCCGCATA CGTGAGGATC	840
CTGCAGCGCA TGACGTCTC AGTGGCCCCC GCCTCCCAGC GCAGCATCCG GCTGCGGACA	900
AAGAGGGTGA AACGCACAGC CATCGCCATC TGTCTGGTCT TCTTTGTGTG CTGGGCACCC	960
TACTATGTGC TACAGCTGAC CCAGTTGTCC ATCAGCCGCC CGACCCTCAC CTTTGTCTAC	1020
TTATACAATG CGGCCATCAG CTTGGGCTAT GCCAACAGCT GCCTCAACCC CTTTGTGTAC	1080

20

30

ATCGTGCTCT GTGAGACGTT CCGCAAACGC TTGGTCCTGT CCGTGAAGCC TGCAGCCCAG 1140
 GGGCAGCTTC GCGCTGTCAG CAACGCTCAG ACGGCTGACG AGGAGAGGAC AGAAAGCAAA 1200
 GGCACCTCA 1209

(193) INFORMATION FOR SEQ ID NO:192:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 402 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:192:

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

10

20

30

195 200 205
 Ala Thr Leu Val Ile Cys Leu Leu Trp Ala Leu Ser Phe Ile Ser Ile
 210 215 220
 Thr Pro Val Trp Leu Tyr Ala Arg Leu Ile Pro Phe Pro Gly Gly Ala
 225 230 235 240
 Val Gly Cys Gly Ile Arg Leu Pro Asn Pro Asp Thr Asp Leu Tyr Trp
 245 250 255
 Phe Thr Leu Tyr Gln Phe Phe Leu Ala Phe Ala Leu Pro Phe Val Val
 260 265 270
 Ile Thr Ala Ala Tyr Val Arg Ile Leu Gln Arg Met Thr Ser Ser Val
 275 280 285
 Ala Pro Ala Ser Gln Arg Ser Ile Arg Leu Arg Thr Lys Arg Val Lys
 290 295 300
 Arg Thr Ala Ile Ala Ile Cys Leu Val Phe Phe Val Cys Trp Ala Pro
 305 310 315 320
 Tyr Tyr Val Leu Gln Leu Thr Gln Leu Ser Ile Ser Arg Pro Thr Leu
 325 330 335
 Thr Phe Val Tyr Leu Tyr Asn Ala Ala Ile Ser Leu Gly Tyr Ala Asn
 340 345 350
 Ser Cys Leu Asn Pro Phe Val Tyr Ile Val Leu Cys Glu Thr Phe Arg
 355 360 365
 Lys Arg Leu Val Leu Ser Val Lys Pro Ala Ala Gln Gly Gln Leu Arg
 370 375 380
 Ala Val Ser Asn Ala Gln Thr Ala Asp Glu Glu Arg Thr Glu Ser Lys
 385 390 395 400
 Gly Thr

10

20

(194) INFORMATION FOR SEQ ID NO:193:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1128 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:193:

ATGGATGTGA CTTCCCAAGC CCGGGGCGTG GGCCTGGAGA TGTACCCAGG CACCGCGCAC 60
 GCTGCGGCCC CCAACACCAC CTCCCCGAG CTCAACCTGT CCCACCCGCT CCTGGGCACC 120

```

GCCCTGGCCA ATGGGACAGG TGAGCTCTCG GAGCACCAGC AGTACGTGAT CGGCCTGTTC 180
CTCTCGTGCC TCTACACCAT CTTCTCTTTC CCCATCGGCT TTGTGGGCAA CATCCTGATC 240
CTGGTGGTGA ACATCAGCTT CCGCGAGAAG ATGACCATCC CCGACCTGTA CTTCATCAAC 300
CTGGCGGTGG CGGACCTCAT CCTGGTGGCC GACTCCCTCA TTGAGGTGTT CAACCTGCAC 360
GAGCGGTACT ACGACATCGC CGTCCTGTGC ACCTTCATGT CGCTCTTCCT GCAGGTCAAC 420
ATGTACAGCA GCGTCTTCTT CCTCACCTGG ATGAGCTTCG ACCGCTACAT CGCCCTGGCC 480
AGGGCCATGC GCTGCAGCCT GTTCGACACC AAGCACCACG CCCGGCTGAG CTGTGGCCTC 540
ATCTGGATGG CATCCGTGTC AGCCACGCTG GTGCCCTTCA CCGCCGTGCA CCTGCAGCAC 600
ACCGACGAGG CCTGCTTCTG TTTCGCGGAT GTCCGGGAGG TGCAGTGGCT CGAGGTCACG 660
CTGGGCTTCA TCGTGCCCTT CGCCATCATC GGCCTGTGCT ACTCCCTCAT TGTCCGGGTG 720
CTGGTCAGGG CGCACCAGCA CCGTGGGCTG CGGCCCCGAC GGCAGAAGGC GAAACGCATG 780
ATCCTCGCGG TGGTGTGGT CTTCTTCGTC TGCTGGCTGC CGGAGAACGT CTTTCATCAGC 840
GTGCACCTCC TGCAGCGGAC GCAGCCTGGG GCCGCTCCCT GCAAGCAGTC TTTCCGCCAT 900
GCCCCCCCC TCACGGGCCA CATTGTCAAC CTCGCCGCT TCTCCAACAG CTGCCTAAAC 960
CCCCTCATCT ACAGCTTTCT CGGGGAGACC TTCAGGGACA AGCTGAGGCT GTACATTGAG 1020
CAGAAAACAA ATTTGCCGGC CCTGAACCGC TTCTGTCACG CTGCCCTGAA GGCCGTCATT 1080
CCAGACAGCA CCGAGCAGTC GGATGTGAGG TTCAGCAGTG CCGTGTGA 1128

```

(195) INFORMATION FOR SEQ ID NO:194:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 375 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:194:

```

Met Asp Val Thr Ser Gln Ala Arg Gly Val Gly Leu Glu Met Tyr Pro
1           5           10           15
Gly Thr Ala His Ala Ala Ala Pro Asn Thr Thr Ser Pro Glu Leu Asn
20          25          30
Leu Ser His Pro Leu Leu Gly Thr Ala Leu Ala Asn Gly Thr Gly Glu
35          40          45

```

10

20

30

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

10

20

30

340 345 350
His Ala Ala Leu Lys Ala Val Ile Pro Asp Ser Thr Glu Gln Ser Asp
355 360 365
Val Arg Phe Ser Ser Ala Val
370 375

(196) INFORMATION FOR SEQ ID NO:195:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 960 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:195:

ATGCCATTCC CAAACTGCTC AGCCCCCAGC ACTGTGGTGG CCACAGCTGT GGGTGTCTTG 60
CTGGGGCTGG AGTGTGGGCT GGGTCTGCTG GGCAACGCGG TGGCGCTGTG GACCTTCCTG 120
TTCCGGGTCA GGGTGTGGAA GCCGTACGCT GTCTACCTGC TCAACCTGGC CCTGGCTGAC 180
CTGCTGTTGG CTGCGTGCCT GCCTTTCTCTG GCGCCTTCT ACCTGAGCCT CCAGGCTTGG 240
CATCTGGGCC GTGTGGGCTG CTGGGCCCTG CGCTTCCTGC TGGACCTCAG CCGCAGCGTG 300
GGGATGGCCT TCCTGGCCGC CGTGGCTTTG GACCGGTACC TCCGTGTGGT CCACCCTCGG 360
CTTAAGGTCA ACCTGCTGTC TCCTCAGGCG GCCCTGGGGG TCTCGGGCCT CGTCTGGCTC 420
CTGATGGTCG CCCTCACCTG CCCGGGCTTG CTCATCTCTG AGGCCGCCCA GAACTCCACC 480
AGGTGCCACA GTTTCTACTC CAGGGCAGAC GGCTCCTTCA GCATCATCTG GCAGGAAGCA 540
CTCTCCTGCC TTCAGTTTGT CCTCCCCTTT GGCTCATCG TGTTCGCAA TGCAGGCATC 600
ATCAGGGCTC TCCAGAAAAG ACTCCGGGAG CCTGAGAAAC AGCCCAAGCT TCAGCGGGCC 660
AAGGCACTGG TCACCTTGGT GGTGGTGCTG TTTGCTCTGT GCTTCTGCC CTGCTTCCTG 720
GCCAGAGTCC TGATGCACAT CTTCAGAAAT CTGGGGAGCT GCAGGGCCCT TTGTGCAGTG 780
GCTCATACCT CGGATGTCAC GGGCAGCCTC ACCTACCTGC ACAGTGTCGT CAACCCCGTG 840
GTATACTGCT TCTCCAGCCC CACCTTCAGG AGCTCCTATC GGAGGGTCTT CCACACCCTC 900
CGAGGCAAAG GGCAGGCAGC AGAGCCCCCA GATTTCAACC CCAGAGACTC CTATTCCTGA 960

20

30

(197) INFORMATION FOR SEQ ID NO:196:

- (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 319 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:196:

```

Met Pro Phe Pro Asn Cys Ser Ala Pro Ser Thr Val Val Ala Thr Ala
1           5           10           15
Val Gly Val Leu Leu Gly Leu Glu Cys Gly Leu Gly Leu Leu Gly Asn
                20           25           30
Ala Val Ala Leu Trp Thr Phe Leu Phe Arg Val Arg Val Trp Lys Pro
        35           40           45
Tyr Ala Val Tyr Leu Leu Asn Leu Ala Leu Ala Asp Leu Leu Leu Ala
        50           55           60
Ala Cys Leu Pro Phe Leu Ala Ala Phe Tyr Leu Ser Leu Gln Ala Trp
65           70           75           80
His Leu Gly Arg Val Gly Cys Trp Ala Leu Arg Phe Leu Leu Asp Leu
        85           90           95
Ser Arg Ser Val Gly Met Ala Phe Leu Ala Ala Val Ala Leu Asp Arg
        100          105          110
Tyr Leu Arg Val Val His Pro Arg Leu Lys Val Asn Leu Leu Ser Pro
        115          120          125
Gln Ala Ala Leu Gly Val Ser Gly Leu Val Trp Leu Leu Met Val Ala
        130          135          140
Leu Thr Cys Pro Gly Leu Leu Ile Ser Glu Ala Ala Gln Asn Ser Thr
145          150          155          160
Arg Cys His Ser Phe Tyr Ser Arg Ala Asp Gly Ser Phe Ser Ile Ile
        165          170          175
Trp Gln Glu Ala Leu Ser Cys Leu Gln Phe Val Leu Pro Phe Gly Leu
        180          185          190
Ile Val Phe Cys Asn Ala Gly Ile Ile Arg Ala Leu Gln Lys Arg Leu
        195          200          205
Arg Glu Pro Glu Lys Gln Pro Lys Leu Gln Arg Ala Lys Ala Leu Val
        210          215          220
Thr Leu Val Val Val Leu Phe Ala Leu Cys Phe Leu Pro Cys Phe Leu
225          230          235          240
Ala Arg Val Leu Met His Ile Phe Gln Asn Leu Gly Ser Cys Arg Ala

```

10

20

30

245	250	255
Leu Cys Ala Val Ala His Thr Ser Asp Val Thr Gly Ser Leu Thr Tyr		
260	265	270
Leu His Ser Val Val Asn Pro Val Val Tyr Cys Phe Ser Ser Pro Thr		
275	280	285
Phe Arg Ser Ser Tyr Arg Arg Val Phe His Thr Leu Arg Gly Lys Gly		
290	295	300
Gln Ala Ala Glu Pro Pro Asp Phe Asn Pro Arg Asp Ser Tyr Ser		
305	310	315

(198) INFORMATION FOR SEQ ID NO:197:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1143 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:197:

ATGGAGGAAG GTGGTGATTT TGACAACTAC TATGGGGCAG ACAACCAGTC TGAGTGTGAG	60
TACACAGACT GGAAATCCTC GGGGGCCCTC ATCCCTGCCA TCTACATGTT GGTCTTCCTC	120
CTGGGCACCA CGGGAAACGG TCTGGTGCTC TGGACCGTGT TTCGGAGCAG CCGGGAGAAG	180
AGGCGCTCAG CTGATATCTT CATTGCTAGC CTGGCGGTGG CTGACCTGAC CTTCTGTGGT	240
ACGCTGCCCC TGTGGGCTAC CTACACGTAC CGGGACTATG ACTGGCCCTT TGGGACCTTC	300
TTCTGCAAGC TCAGCAGCTA CCTCATCTTC GTCAACATGT ACGCCAGCGT CTTCTGCCTC	360
ACCGGCCTCA GCTTCGACCG CTACCTGGCC ATCGTGAGGC CAGTGGCCAA TGCTCGGCTG	420
AGGCTGCGGG TCAGCGGGGC CGTGGCCACG GCAGTTCTTT GGGTGCTGGC CGCCCTCCTG	480
GCCATGCCTG TCATGGTGTT ACGCACCACC GGGGACTTGG AGAACACCAC TAAGGTGCAG	540
TGCTACATGG ACTACTCCAT GGTGGCCACT GTGAGCTCAG AGTGGGCCTG GGAGGTGGGC	600
CTTGGGGTCT CGTCCACCAC CGTGGGCTTT GTGGTGCCCT TCACCATCAT GCTGACCTGT	660
TACTTCTTCA TCGCCCAAAC CATCGCTGGC CACTTCCGCA AGGAACGCAT CGAGGGCCTG	720
CGGAAGCGGC GCCGGCTTAA GAGCATCATC GTGGTGCTGG TGGTGACCTT TGCCCTGTGC	780
TGGATGCCCT ACCACCTGGT GAAGACGCTG TACATGCTGG GCAGCCTGCT GCACTGGCCC	840
TGTGACTTTG ACCTCTTCCT CATGAACATC TTCCCCTACT GCACCTGCAT CAGCTACGTC	900

10

20

30

AACAGCTGCC TCAACCCCTT CCTCTATGCC TTTTTCGACC CCCGCTTCCG CCAGGCCTGC 960
 ACCTCCATGC TCTGCTGTGG CCAGAGCAGG TGCGCAGGCA CCTCCCACAG CAGCAGTGGG 1020
 GAGAAGTCAG CCAGCTACTC TTCGGGGCAC AGCCAGGGGC CCGGCCCAA CATCGGCAAG 1080
 GGTGGAGAAC AGATGCACGA GAAATCCATC CCCTACAGCC AGGAGACCCT TGTGGTTGAC 1140
 TAG 1143

(199) INFORMATION FOR SEQ ID NO:198:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 380 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:198:

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

10

20

30

Thr Lys Val Gln Cys Tyr Met Asp Tyr Ser Met Val Ala Thr Val Ser
 180 185 190
 Ser Glu Trp Ala Trp Glu Val Gly Leu Gly Val Ser Ser Thr Thr Val
 195 200 205
 Gly Phe Val Val Pro Phe Thr Ile Met Leu Thr Cys Tyr Phe Phe Ile
 210 215 220
 Ala Gln Thr Ile Ala Gly His Phe Arg Lys Glu Arg Ile Glu Gly Leu
 225 230 235 240
 Arg Lys Arg Arg Arg Leu Lys Ser Ile Ile Val Val Leu Val Val Thr
 245 250 255
 Phe Ala Leu Cys Trp Met Pro Tyr His Leu Val Lys Thr Leu Tyr Met
 260 265 270
 Leu Gly Ser Leu Leu His Trp Pro Cys Asp Phe Asp Leu Phe Leu Met
 275 280 285
 Asn Ile Phe Pro Tyr Cys Thr Cys Ile Ser Tyr Val Asn Ser Cys Leu
 290 295 300
 Asn Pro Phe Leu Tyr Ala Phe Phe Asp Pro Arg Phe Arg Gln Ala Cys
 305 310 315 320
 Thr Ser Met Leu Cys Cys Gly Gln Ser Arg Cys Ala Gly Thr Ser His
 325 330 335
 Ser Ser Ser Gly Glu Lys Ser Ala Ser Tyr Ser Ser Gly His Ser Gln
 340 345 350
 Gly Pro Gly Pro Asn Met Gly Lys Gly Gly Glu Gln Met His Glu Lys
 355 360 365
 Ser Ile Pro Tyr Ser Gln Glu Thr Leu Val Val Asp
 370 375 380

10

20

(200) INFORMATION FOR SEQ ID NO:199:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1119 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:199:

ATGAACTACC CGCTAACGCT GGAAATGGAC CTCGAGAACC TGGAGGACCT GTTCTGGGAA 60
 CTGGACAGAT TGGACAACTA TAACGACACC TCCCTGGTGG AAAATCATCT CTGCCCTGCC 120

ACAGAGGGTC CCCTCATGGC CTCCTTCAAG GCCGTGTTTCG TGCCCGTGGC CTACAGCCTC 180
 ATCTTCCTCC TGGGCGTGAT CGGCAACGTC CTGGTGCTGG TGATCCTGGA GCGGCACCGG 240
 CAGACACGCA GTTCCACGGA GACCTTCCTG TTCCACCTGG CCGTGGCCGA CTCCTGCTG 300
 GTCTTCATCT TGCCCTTTGC CGTGGCCGAG GGCTCTGTGG GCTGGGTCTT GGGGACCTTC 360
 CTCTGCAAAA CTGTGATTGC CCTGCACAAA GTCAACTTCT ACTGCAGCAG CCTGCTCCTG 420
 GCCTGCATCG CCGTGGACCG CTACCTGGCC ATTGTCCACG CCGTCCATGC CTACCGCCAC 480
 CGCCGCCTCC TCTCCATCCA CATCACCTGT GGGACCATCT GGCTGGTGGG CTTCTCCTT 540
 GCCTTGCCAG AGATTCTCTT CGCCAAAGTC AGCCAAGGCC ATCACAACAA CTCCCTGCCA 600
 CGTTGCACCT TCTCCCAAGA GAACCAAGCA GAAACGCATG CCTGGTTCAC CTCCCGATTC 660
 CTCTACCATG TGGCGGGATT CCTGCTGCCC ATGCTGGTGA TGGGCTGGTG CTACGTGGGG 720
 GTAGTGACACA GGTTCGCGCA GGGCCAGCGG CGCCCTCAGC GGCAGAAGGC AAAAAGGGTG 780
 GCCATCCTGG TGACAAGCAT CTTCTTCCTC TGCTGGTCAC CCTACCACAT CGTCATCTTC 840
 CTGGACACCC TGGCGAGGCT GAAGGCCGTG GACAATACCT GCAAGCTGAA TGGCTCTCTC 900
 CCCGTGGCCA TCACCATGTG TGAGTTCCTG GGCTGGCCC ACTGCTGCCT CAACCCCATG 960
 CTCTACACTT TCGCCGGCGT GAAGTTCGCG AGTGACCTGT CGCGGCTCCT GACCAAGCTG 1020
 GGCTGTACCG GCCCTGCCTC OCTGTGCCAG CTCTTCCCTA GCTGGCGCAG GAGCAGTCTC 1080
 TCTGAGTCAG AGAATGCCAC CTCTCTCACC ACGTTCTAG 1119

(201) INFORMATION FOR SEQ ID NO:200:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 372 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant
- (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:200:

Met Asn Tyr Pro Leu Thr Leu Glu Met Asp Leu Glu Asn Leu Glu Asp
 1 5 10 15
 Leu Phe Trp Glu Leu Asp Arg Leu Asp Asn Tyr Asn Asp Thr Ser Leu
 20 25 30
 Val Glu Asn His Leu Cys Pro Ala Thr Glu Gly Pro Leu Met Ala Ser
 35 40 45
 Phe Lys Ala Val Phe Val Pro Val Ala Tyr Ser Leu Ile Phe Leu Leu

10

20

30

30

Pro Ser Trp Arg Arg Ser Ser Leu Ser Glu Ser Glu Asn Ala Thr Ser
 355 360 365

Leu Thr Thr Phe
 370

(202) INFORMATION FOR SEQ ID NO:201:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1128 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:201:

ATGGATGTGA CTTCCCAAGC CCGGGGCGTG GGCCTGGAGA TGTACCCAGG CACCGCGCAG	60
CCTGCGGCCC CCAACACCAC CTCCCCGAG CTCAACCTGT CCCACCCGCT CCTGGGCACC	120
GCCCTGGCCA ATGGGACAGG TGAGCTCTCG GAGCACCAGC AGTACGTGAT CGGCCTGTTC	180
CTCTCGTGCC TCTACACCAT CTTCCTCTTC CCCATCGGCT TTGTGGGCAA CATCCTGATC	240
CTGGTGGTGA ACATCAGCTT CCGCGAGAAG ATGACCATCC CCGACCTGTA CTTTCATCAAC	300
CTGGCGGTGG CGGACCTCAT CCTGGTGGCC GACTCCCTCA TTGAGGTGTT CAACCTGCAC	360
GAGCGGTACT ACGACATCGC CGTCCTGTGC ACCTTCATGT CGCTCTTCCT GCAGGTCAAC	420
ATGTACAGCA GCGTCTTCTT CCTCACCTGG ATGAGCTTCG ACCGCTACAT CGCCCTGGCC	480
AGGGCCATGC GCTGCAGCCT GTTCCGCACC AAGCACCACG CCCGCTGAG CTGTGGCCTC	540
ATCTGGATGG CATCCGTGTC AGCCACGCTG GTGCCCTTCA CCGCCGTGCA CCTGCAGCAC	600
ACCGACGAGG CCTGCTTCTG TTTCGCGGAT GTCCGGGAGG TGCACTGGCT CGAGGTCACG	660
CTGGGCTTCA TCGTGCCCTT CGCCATCATC GGCCTGTGCT ACTCCCTCAT TGTCCGGGTG	720
CTGGTCAGGG CGCACCAGCA CCGTGGGCTG CGGCCCCGGC GGCAGAAGGC GAAGCGCATG	780
ATCCTCGCGG TGGTGTGGT CTTCTTCGTC TGCTGGCTGC CGGAGAACGT CTTTCATCAGC	840
GTGCACCTCC TGCAGCGGAC GCAGCCTGGG GCCGCTCCCT GCAAGCAGTC TTTCCGCCAT	900
GCCCACCCCC TCACGGGCCA CATTGTCAAC CTCACCGCCT TCTCCAACAG CTGCCTAAAC	960
CCCCTCATCT ACAGCTTTCT CGGGGAGACC TTCAGGGACA AGCTGAGGCT GTACATTGAG	1020
CAGAAAACAA ATTTGCCGGC CCTGAACCGC TTCTGTACAG CTGCCCTGAA GGCCGTCATT	1080
CCAGACAGCA CCGAGCAGTC GGATGTGAGG TTCAGCAGTG CCGTGTAG	1128

10

20

30

(203) INFORMATION FOR SEQ ID NO:202:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 375 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:202:

```

Met Asp Val Thr Ser Gln Ala Arg Gly Val Gly Leu Glu Met Tyr Pro
1           5           10           15
Gly Thr Ala Gln Pro Ala Ala Pro Asn Thr Thr Ser Pro Glu Leu Asn
20           25           30
Leu Ser His Pro Leu Leu Gly Thr Ala Leu Ala Asn Gly Thr Gly Glu
35           40           45
Leu Ser Glu His Gln Gln Tyr Val Ile Gly Leu Phe Leu Ser Cys Leu
50           55           60
Tyr Thr Ile Phe Leu Phe Pro Ile Gly Phe Val Gly Asn Ile Leu Ile
65           70           75           80
Leu Val Val Asn Ile Ser Phe Arg Glu Lys Met Thr Ile Pro Asp Leu
85           90           95
Tyr Phe Ile Asn Leu Ala Val Ala Asp Leu Ile Leu Val Ala Asp Ser
100          105          110
Leu Ile Glu Val Phe Asn Leu His Glu Arg Tyr Tyr Asp Ile Ala Val
115          120          125
Leu Cys Thr Phe Met Ser Leu Phe Leu Gln Val Asn Met Tyr Ser Ser
130          135          140
Val Phe Phe Leu Thr Trp Met Ser Phe Asp Arg Tyr Ile Ala Leu Ala
145          150          155          160
Arg Ala Met Arg Cys Ser Leu Phe Arg Thr Lys His His Ala Arg Leu
165          170          175
Ser Cys Gly Leu Ile Trp Met Ala Ser Val Ser Ala Thr Leu Val Pro
180          185          190
Phe Thr Ala Val His Leu Gln His Thr Asp Glu Ala Cys Phe Cys Phe
195          200          205
Ala Asp Val Arg Glu Val Gln Trp Leu Glu Val Thr Leu Gly Phe Ile
210          215          220
Val Pro Phe Ala Ile Ile Gly Leu Cys Tyr Ser Leu Ile Val Arg Val

```

10

20

30

225 230 235 240
 Leu Val Arg Ala His Arg His Arg Gly Leu Arg Pro Arg Arg Gln Lys
 245 250 255
 Ala Lys Arg Met Ile Leu Ala Val Val Leu Val Phe Phe Val Cys Trp
 260 265 270
 Leu Pro Glu Asn Val Phe Ile Ser Val His Leu Leu Gln Arg Thr Gln
 275 280 285
 Pro Gly Ala Ala Pro Cys Lys Gln Ser Phe Arg His Ala His Pro Leu
 290 295 300
 Thr Gly His Ile Val Asn Leu Thr Ala Phe Ser Asn Ser Cys Leu Asn
 305 310 315 320
 Pro Leu Ile Tyr Ser Phe Leu Gly Glu Thr Phe Arg Asp Lys Leu Arg
 325 330 335
 Leu Tyr Ile Glu Gln Lys Thr Asn Leu Pro Ala Leu Asn Arg Phe Cys
 340 345 350
 His Ala Ala Leu Lys Ala Val Ile Pro Asp Ser Thr Glu Gln Ser Asp
 355 360 365
 Val Arg Phe Ser Ser Ala Val
 370 375

10

(204) INFORMATION FOR SEQ ID NO:203:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1137 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:203:

ATGGACCTGG GGAAACCAAT GAAAAGCGTG CTGGTGGTGG CTCTCCTTGT CATTTTCCAG 60
 GTATGCCTGT GTCAAGATGA GGTACGGAC GATTACATCG GAGACAACAC CACAGTGGAC 120
 TACACTTTGT TCGAGTCTTT GTGCTCCAAG AAGGACGTGC GGAACTTTAA AGCCTGGTTC 180
 CTCCTATCA TGTACTCCAT CATTTGTTTC GTGGGCCTAC TGGGCAATGG GCTGGTCGTG 240
 TTGACCTATA TCTATTTCAG GAGGCTCAAG ACCATGACCG ATACCTACCT GCTCAACCTG 300
 GCGGTGGCAG ACATCCTCTT CCTCCTGACC CTTCCCTTCT GGGCCTACAG CGCGGCCAAG 360
 TCCTGGGTCT TCGGTGTCCA CTTTGTGCAAG CTCATCTTTG CCATCTACAA GATGAGCTTC 420

20

30

TTCAGTGGCA TGCTCCTACT TCTTTGCATC AGCATTGACC GCTACGTGGC CATCGTCCAG 480
 GCTGTCTCAG CTCACCGCCA CCGTGCCCGC GTCCTTCTCA TCAGCAAGCT GTCCTGTGTG 540
 GGCATCTGGA TACTAGCCAC AGTGCTCTCC ATCCCAAGAGC TCCTGTACAG TGACCTCCAG 600
 AGGAGCAGCA GTGAGCAAGC GATGCGATGC TCTCTCATCA CAGAGCATGT GGAGGCCTTT 660
 ATCACCATCC AGGTGGCCCA GATGGTGATC GGCTTTCTGG TCCCCTGCT GCCCATGAGC 720
 TTCTGTTACC TTGTATCAT CCGCACCCCTG CTCCAGGCAC GCAACTTTGA GCGCAACAAG 780
 GCCAAAAGG TGATCATCGC TGTGGTCGTG GTCTTCATAG TCTTCCAGCT GCCCTACAAT 840
 GGGGTGGTCC TGGCCCAGAC GGTGGCCAAC TTCAACATCA CCAGTAGCAC CTGTGAGCTC 900
 AGTAAGCAAC TCAACATCGC CTACGACGTC ACCTACAGCC TGGCCTGCGT CCGCTGCTGC 960
 GTCAACCCTT TCTTGACGC CTTCATCGGC GTCAAGTTCC GCAACGATCT CTTCAAGCTC 1020
 TTCAAGGACC TGGGCTGCCT CAGCCAGGAG CAGCTCCGSC AGTGGTCTTC CTGTGCGCAC 1080
 ATCCGGCGCT CCTCCATGAG TGTGGAGGCC GAGACCACCA CCACCTTCTC CCCATAG 1137

(205) INFORMATION FOR SEQ ID NO:204:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 378 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:204:

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

10

20

30

100	105	110
Phe Trp Ala Tyr Ser Ala Ala Lys Ser Trp Val Phe Gly Val His Phe		
115	120	125
Cys Lys Leu Ile Phe Ala Ile Tyr Lys Met Ser Phe Phe Ser Gly Met		
130	135	140
Leu Leu Leu Leu Cys Ile Ser Ile Asp Arg Tyr Val Ala Ile Val Gln		
145	150	155
Ala Val Ser Ala His Arg His Arg Ala Arg Val Leu Leu Ile Ser Lys		
165	170	175
Leu Ser Cys Val Gly Ile Trp Ile Leu Ala Thr Val Leu Ser Ile Pro		
180	185	190
Glu Leu Leu Tyr Ser Asp Leu Gln Arg Ser Ser Ser Glu Gln Ala Met		
195	200	205
Arg Cys Ser Leu Ile Thr Glu His Val Glu Ala Phe Ile Thr Ile Gln		
210	215	220
Val Ala Gln Met Val Ile Gly Phe Leu Val Pro Leu Leu Ala Met Ser		
225	230	235
Phe Cys Tyr Leu Val Ile Ile Arg Thr Leu Leu Gln Ala Arg Asn Phe		
245	250	255
Glu Arg Asn Lys Ala Lys Lys Val Ile Ile Ala Val Val Val Val Phe		
260	265	270
Ile Val Phe Gln Leu Pro Tyr Asn Gly Val Val Leu Ala Gln Thr Val		
275	280	285
Ala Asn Phe Asn Ile Thr Ser Ser Thr Cys Glu Leu Ser Lys Gln Leu		
290	295	300
Asn Ile Ala Tyr Asp Val Thr Tyr Ser Leu Ala Cys Val Arg Cys Cys		
305	310	315
Val Asn Pro Phe Leu Tyr Ala Phe Ile Gly Val Lys Phe Arg Asn Asp		
325	330	335
Leu Phe Lys Leu Phe Lys Asp Leu Gly Cys Leu Ser Gln Glu Gln Leu		
340	345	350
Arg Gln Trp Ser Ser Cys Arg His Ile Arg Arg Ser Ser Met Ser Val		
355	360	365
Glu Ala Glu Thr Thr Thr Thr Phe Ser Pro		
370	375	

(206) INFORMATION FOR SEQ ID NO:205:

10

20

30

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1086 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:205:

ATGGATATAC AAATGGCAAA CAATTTTACT CCGCCCTCTG CAACTCCTCA GGGAAATGAC	60
TGTGACCTCT ATGCACATCA CAGCACGGCC AGGATAGTAA TGCCTCTGCA TTACAGCCTC	120
GTCTTCATCA TTGGGCTCGT GGGAAACTTA CTAGCCTTGG TCGTCATTGT TCAAAACAGG	180
AAAAAATCA ACTCTACCAC CCTCTATTCA ACAAATTTGG TGATTTCTGA TATACTTTTT	240
ACCACGGCTT TGCCTACACG AATAGCCTAC TATGCAATGG GCTTTGACTG GAGAATCGGA	300
GATGCCTTGT GTAGGATAAC TCGCCTAGTG TTTTACATCA ACACATATGC AGGTGTGAAC	360
TTTATGACCT GCCTGAGTAT TGACCGCTTC ATTGCTGTGG TGCACCTCTC ACGCTACAAC	420
AAGATAAAAA GGATTGAACA TGCAAAAGGC GTGTGCATAT TTGTCTGGAT TCTAGTATTT	480
GCTCAGACAC TCCCACTCCT CATCAACCTT ATGTCAAAGC AGGAGGCTGA AAGGATTACA	540
TGCATGGAGT ATCCAAACTT TGAAGAACT AAATCTCTTC CCTGGATTCT GCTTGGGGCA	600
TGTTTCATAG GATATGTACT TCCACTTATA ATCATTCTCA TCTGCTATTG TCAGATCTGC	660
TGCAAACTCT TCAGAACTGC CAAACAAAAC CCACTCACTG AGAAATCTGG TGTAACAAA	720
AAGGCTAAAA ACACAATTAT TCTTATTATT GTTGTGTTTG TTCTCTGTTT CACACCTTAC	780
CATGTTGCAA TTATTCAACA TATGATTAAG AAGCTTCGTT TCTCTAATTT CCTGGAATGT	840
AGCCAAAGAC ATTCGTTCCA GATTTCTCTG CACTTTACAG TATGCCTGAT GAACTTCAAT	900
TGCTGCATGG ACCCTTTTAT CTACTTCTTT GCATGTAAAG GGTATAAGAG AAAGTTATG	960
AGGATGCTGA AACGGCAAGT CAGTGTATCG ATTTCTAGTG CTGTGAAGTC AGCCCTTGAA	1020
GAAAAATTCAC GTGAAATGAC AGAAACGCAG ATGATGATAC ATTCCAAGTC TTCAAATGGA	1080
AAGTGA	1086

10

20

30

(207) INFORMATION FOR SEQ ID NO:206:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 361 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:206:

```

Met Asp Ile Gln Met Ala Asn Asn Phe Thr Pro Pro Ser Ala Thr Pro
1           5           10           15

Gln Gly Asn Asp Cys Asp Leu Tyr Ala His His Ser Thr Ala Arg Ile
          20           25           30

Val Met Pro Leu His Tyr Ser Leu Val Phe Ile Ile Gly Leu Val Gly
          35           40           45

Asn Leu Leu Ala Leu Val Val Ile Val Gln Asn Arg Lys Lys Ile Asn
50           55           60

Ser Thr Thr Leu Tyr Ser Thr Asn Leu Val Ile Ser Asp Ile Leu Phe
65           70           75           80

Thr Thr Ala Leu Pro Thr Arg Ile Ala Tyr Tyr Ala Met Gly Phe Asp
          85           90           95

Trp Arg Ile Gly Asp Ala Leu Cys Arg Ile Thr Ala Leu Val Phe Tyr
          100          105          110

Ile Asn Thr Tyr Ala Gly Val Asn Phe Met Thr Cys Leu Ser Ile Asp
          115          120          125

Arg Phe Ile Ala Val Val His Pro Leu Arg Tyr Asn Lys Ile Lys Arg
          130          135          140

Ile Glu His Ala Lys Gly Val Cys Ile Phe Val Trp Ile Leu Val Phe
145          150          155          160

Ala Gln Thr Leu Pro Leu Leu Ile Asn Pro Met Ser Lys Gln Glu Ala
          165          170          175

Glu Arg Ile Thr Cys Met Glu Tyr Pro Asn Phe Glu Glu Thr Lys Ser
          180          185          190

Leu Pro Trp Ile Leu Leu Gly Ala Cys Phe Ile Gly Tyr Val Leu Pro
          195          200          205

Leu Ile Ile Ile Leu Ile Cys Tyr Ser Gln Ile Cys Cys Lys Leu Phe
210          215          220

Arg Thr Ala Lys Gln Asn Pro Leu Thr Glu Lys Ser Gly Val Asn Lys
225          230          235          240

Lys Ala Lys Asn Thr Ile Ile Leu Ile Ile Val Val Phe Val Leu Cys
          245          250          255

Phe Thr Pro Tyr His Val Ala Ile Ile Gln His Met Ile Lys Lys Leu
          260          265          270

```

10

20

30

Arg Phe Ser Asn Phe Leu Glu Cys Ser Gln Arg His Ser Phe Gln Ile
 275 280 285

Ser Leu His Phe Thr Val Cys Leu Met Asn Phe Asn Cys Cys Met Asp
 290 295 300

Pro Phe Ile Tyr Phe Phe Ala Cys Lys Gly Tyr Lys Arg Lys Val Met
 305 310 315 320

Arg Met Leu Lys Arg Gln Val Ser Val Ser Ile Ser Ser Ala Val Lys
 325 330 335

Ser Ala Pro Glu Glu Asn Ser Arg Glu Met Thr Glu Thr Gln Met Met
 340 345 350

Ile His Ser Lys Ser Ser Asn Gly Lys
 355 360

10

(208) INFORMATION FOR SEQ ID NO:207:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1446 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:207:

20

ATGCGGTGGC TGTGGCCCCCT GGCTGTCTCT CTTGCTGTGA TTTTGGCTGT GGGGCTAAGC 60

AGGGTCTCTG GGGGTGCCCC CCTGCACCTG GGCAGGCACA GAGCCGAGAC CCAGGAGCAG 120

CAGAGCCGAT CCAAGAGGGG CACCGAGGAT GAGGAGGCCA AGGGCGTGCA GCAGTATGTG 180

CCTGAGGAGT GGGCGGAGTA CCCCCGGCCC ATTCAACCCTG CTGGCCTGCA GCCAACCAAG 240

CCCTTGGTGG CCACCAGCCC TAACCCCGAC AAGGATGGGG GCACCCAGAG CAGTGGGCAG 300

GAAGTGAAGG GCAATCTGAC AGGGGCACCA GGCAGAGGC TACAGATCCA GAACCCCTG 360

TATCCGGTGA CCGAGAGCTC CTACAGTGCC TATGCCATCA TGCTTCTGGC GCTGGTGGTG 420

TTTGCGGTGG GCATTGTGGG CAACCTGTCTG GTCATGTGCA TCGTGTGGCA CAGCTACTAC 480

CTGAAGAGCG CCTGGAATC CATCCTTGCC AGCCTGGCCC TCTGGGATTT TCTGGTCCTC 540

TTTTTCTGCC TCCCTATTGT CATCTTCAAC GAGATCACCA AGCAGAGGCT ACTGGGTGAC 600

GTTTCTTGTC GTGCCGTGCC CTTATGGAG GTCTCCTCTC TGGGAGTCAC GACTTTCAGC 660

CTCTGTGCCC TGGGCATTGA CCGCTTCCAC GTGGCCACCA GCACCCTGCC CAAGGTGAGG 720

CCCATCGAGC GGTGCCAATC CATCCTGGCC AAGTTGGCTG TCATCTGGGT GGGCTCCATG 780

30

ACGCTGGCTG TGCCTGAGCT CCTGCTGTGG CAGCTGGCAC AGGAGCCTGC CCCCACCATG 840
 GGCACCCCTGG ACTCATGCAT CATGAAACCC TCAGCCAGCC TGCCCGAGTC CCTGTATTCA 900
 CTGGTGATGA CCTACCAGAA CGCCCGCATG TGGTGGTACT TTGGCTGCTA CTTCTGCCTG 960
 CCCATCCTCT TCACAGTCAC CTGCCAGCTG GTGACATGGC GGGTGCGAGG CCCTCCAGGG 1020
 AGGAAGTCAG AGTGCAGGGC CAGCAAGCAC GAGCAGTGTG AGAGCCAGCT CAAGAGCACC 1080
 GTGGTGGGCC TGACCGTGGT CTACGCCTTC TGCACCCTCC CAGAGAACGT CTGCAACATC 1140
 GTGGTGGCCT ACCTCTCCAC CGAGCTGACC CGCCAGACCC TGGACCTCCT GGGCCTCATC 1200
 AACCAGTTCT CCACCTTCTT CAAGGGCGCC ATCACCCAG TGCTGCTCCT TTGCATCTGC 1260
 AGGCCGCTGG GCCAGGCCCTT CCTGGACTGC TGCTGCTGCT GCTGCTGTGA GGAGTGCGGC 1320
 GGGGCTTCGG AGGCCTCTGC TGCCAATGGG TCGGACAACA AGCTCAAGAC CGAGGTGTCC 1380
 TCTTCCATCT ACTTCCACAA GCCCAGGGAG TCACCCCCAC TCCTGCCCCCT GGGCACACCT 1440
 TGCTGA 1446

(209) INFORMATION FOR SEQ ID NO:208:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 481 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:208:

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

10

20

30

100	105	110
Arg Leu Gln Ile Gln Asn Pro Leu Tyr Pro Val Thr Glu Ser Ser Tyr 115	120	125
Ser Ala Tyr Ala Ile Met Leu Leu Ala Leu Val Val Phe Ala Val Gly 130	135	140
Ile Val Gly Asn Leu Ser Val Met Cys Ile Val Trp His Ser Tyr Tyr 145	150	155
Leu Lys Ser Ala Trp Asn Ser Ile Leu Ala Ser Leu Ala Leu Trp Asp 165	170	175
Phe Leu Val Leu Phe Phe Cys Leu Pro Ile Val Ile Phe Asn Glu Ile 180	185	190
Thr Lys Gln Arg Leu Leu Gly Asp Val Ser Cys Arg Ala Val Pro Phe 195	200	205
Met Glu Val Ser Ser Leu Gly Val Thr Thr Phe Ser Leu Cys Ala Leu 210	215	220
Gly Ile Asp Arg Phe His Val Ala Thr Ser Thr Leu Pro Lys Val Arg 225	230	235
Pro Ile Glu Arg Cys Gln Ser Ile Leu Ala Lys Leu Ala Val Ile Trp 245	250	255
Val Gly Ser Met Thr Leu Ala Val Pro Glu Leu Leu Leu Trp Gln Leu 260	265	270
Ala Gln Glu Pro Ala Pro Thr Met Gly Thr Leu Asp Ser Cys Ile Met 275	280	285
Lys Pro Ser Ala Ser Leu Pro Glu Ser Leu Tyr Ser Leu Val Met Thr 290	295	300
Tyr Gln Asn Ala Arg Met Trp Trp Tyr Phe Gly Cys Tyr Phe Cys Leu 305	310	315
Pro Ile Leu Phe Thr Val Thr Cys Gln Leu Val Thr Trp Arg Val Arg 325	330	335
Gly Pro Pro Gly Arg Lys Ser Glu Cys Arg Ala Ser Lys His Glu Gln 340	345	350
Cys Glu Ser Gln Leu Lys Ser Thr Val Val Gly Leu Thr Val Val Tyr 355	360	365
Ala Phe Cys Thr Leu Pro Glu Asn Val Cys Asn Ile Val Val Ala Tyr 370	375	380
Leu Ser Thr Glu Leu Thr Arg Gln Thr Leu Asp Leu Leu Gly Leu Ile 385	390	395
		400

10

20

30

Asn Gln Phe Ser Thr Phe Phe Lys Gly Ala Ile Thr Pro Val Leu Leu
 405 410 415

Leu Cys Ile Cys Arg Pro Leu Gly Gln Ala Phe Leu Asp Cys Cys Cys
 420 425 430

Cys Cys Cys Cys Glu Glu Cys Gly Gly Ala Ser Glu Ala Ser Ala Ala
 435 440 445

Asn Gly Ser Asp Asn Lys Leu Lys Thr Glu Val Ser Ser Ser Ile Tyr
 450 455 460

Phe His Lys Pro Arg Glu Ser Pro Pro Leu Leu Pro Leu Gly Thr Pro
 465 470 475 480

Cys

10

(210) INFORMATION FOR SEQ ID NO:209:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1101 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:209:

20

ATGTGGAACG CGACGCCCAG CGAAGAGCCG GGGTTCAACC TCACACTGGC CGACCTGGAC 60

TGGGATGCTT CCCCCGGCAA CGACTCGCTG GCGGACGAGC TGCTGCAGCT CTTCCCCGCG 120

CCGCTGCTGG CGGGCGTCAC AGCCACCTGC GTGGCACTCT TCGTGGTGGG TATCGCTGGC 180

AACCTGCTCA CCATGCTGGT GGTGTCGCGC TTCCGCGAGC TGCGCACCAC CACCAACCTC 240

TACCTGTCCA GCATGGCCTT CTCCGATCTG CTCATCTTCC TCTGCATGCC CCTGGACCTC 300

GTTCGCCTCT GGCAGTACCG GCCCTGGAAC TTCGGCGACC TCCTCTGCAA ACTCTTCCAA 360

TTCGTCAGTG AGAGCTGCAC CTACGCCACG GTGCTCACCA TCACAGCGCT GAGCGTCGAG 420

CGCTACTTCG CCATCTGCTT CCCACTCCGG GCCAAGGTGG TGGTCACCAA GGGGCGGGTG 480

AAGCTGGTCA TCTTCGTCAT CTGGGCCGTG GCCTTCTGCA GCGCCGGGCC CATCTTCGTG 540

CTAGTCGGGG TGGAGCACGA GAACGGCACC GACCCTTGGG ACACCAACGA GTGCCGCCCC 600

ACCGAGTTTG CGGTGCGCTC TGGACTGCTC ACGGTCATGG TGTGGGTGTC CAGCATCTTC 660

TTCTTCCTTC CTGTCTCTG TCTACGGTC CTCTACAGTC TCATCGGCAG GAAGCTGTGG 720

CGGAGGAGGC GCGCGATGC TGTCGTGGGT GCCTCGCTCA GGGACCAGAA CCACAAGCAA 780

30

ACCAAGAAAA TGCTGGCTGT AGTGGTGTTC GCCTTCATCC TCTGCTGGCT CCCCTTCCAC 840
 GTAGGGCGAT ATTTATTTTC CAAATCCTTT GAGCCTGGCT CCTTGGAGAT TGCTCAGATC 900
 AGCCAGTACT GCAACCTCGT GTCCTTTGTC CTCTTCTACC TCAGTGCTGC CATCAACCCC 960
 ATTCTGTACA ACATCATGTC CAAGAAGTAC CGGGTGGCAG TGTTCAGACT TCTGGGATTC 1020
 GAACCCTTCT CCCAGAGAAA GCTCTCCACT CTGAAAGATG AAAGTTCTCG GGCCTGGACA 1080
 GAATCTAGTA TTAATACATG A 1101

(211) INFORMATION FOR SEQ ID NO:210:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 366 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:210:

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

10

20

30

AATTCTGCAA GAGACGTTCT GCGAGCCCGA GCACCCAGGG AGGAGCAGGG GGCAGCGTTT 240
 CTTGCGGGAC CCTCCTGGGA CCTGCCGGCG GCGCCGGGCC GTGACCCGGC TGCAGGCAGA 300
 GGGGCGGAGG CGTCGGCAGC CGGACCCCGG GGACCTCCAA CCAGGCCACC TGGCCCTGG 360
 AGGTGGAAG GTGCTCGGG TCAGGAGCCT TCTGAAACTT TGGGAGAGG GAACCCACG 420
 GCCCTCCAGC TCTTCCTTCA GATCTCAGAG GAGGAAGAGA AGGTCCCAG AGGCGCTGGC 480
 ATTTCCGGGC GTAGCCAGGA GCAGAGTGTG AAGACAGTCC CCGGAGCCAG CGATCTTTT 540
 TACTGGCCAA GGAGAGCCGG GAAACTCCAG GGTCCCACC ACAAGCCCCT GTCCAAGACG 600
 GCCAATGGAC TGGCGGGCA CGAAGGGTGG ACAATTGCAC TCCCGGGCCG GCGCTGGCC 660
 CAGAATGGAT CCTTGGGTGA AGGAATCCAT GAGCCTGGGG GTCCCCCGCG GGGAAACAGC 720
 ACGAACCAGC GTGTGAGACT GAAGAACCC TTCTACCCGC TGACCCAGGA GTCCTATGGA 780
 GCCTACGCGG TCATGTGTCT GTCCGTGGTG ATCTTCGGGA CCGGCATCAT TGGCAACCTG 840
 GCGGTGATGT GCATCGTGTG CCACAACCTAC TACATGCGGA GCATCTCCAA CTCCTCTTG 900
 GCCAACCTGG CCTTCTGGGA CTTTCTCATC ATCTTCTTCT GCCTTCGCT GGTCTCTTC 960
 CACGAGCTGA CCAAGAAGTG GCTGCTGGAG GACTTCTCCT GCAAGATCGT GCCCTATATA 1020
 GAGGTGCGCT CTCTGGGAGT CACCACTTTC ACCTTATGTG CTCTGTGCAT AGACCGCTTC 1080
 CGTGCTGCCA CCAACGTACA GATGTACTAC GAAATGATCG AAAATTGTTC CTCAACAACCT 1140
 GCCAAACTTG CTGTTATATG GGTGGGAGCT CTATTGTTAG CACTTCCAGA AGTTGTTCTC 1200
 CGCCAGCTGA GCAAGGAGGA TTTGGGGTTT AGTGGCCGAG CTCCGGCAGA AAGGTGCATT 1260
 ATTAAGATCT CTCCTGATTT ACCAGACACC ATCTATGTTC TAGCCCTCAC CTACGACAGT 1320
 GCGAGACTGT GGTGATATTT TGGCTGTTAC TTTTGTGTC CCACGCTTTT CACCATCACC 1380
 TGCTCTCTAG TGAAGTGGAG GAAAATCCGC AAAGCAGAGA AAGCCTGTAC CCGAGGGAAT 1440
 AAACGGCAGA TTCAACTAGA GAGTCAGATG AAGTGTACAG TAGTGGCACT GACCATTTTA 1500
 TATGGATTTT GCATTATTCC TGAAAATATC TGCAACATTG TTACTGCCTA CATGGCTACA 1560
 GGGGTTTCAC AGCAGACAAT GGACCTCCTT AATATCATCA GCCAGTTCCT TTTGTTCTTT 1620
 AAGTCCTGTG TCACCCAGT CCTCCTTTTC TGTCTCTGCA AACCCTTCAG TCGGGCCTTC 1680
 ATGGAGTGCT GCTGCTGTTG CTGTGAGGAA TGCATTGAGA AGTCTTCAAC GGTGACCAGT 1740
 GATGACAATG ACAACGAGTA CACCACGGAA CTCGAACTCT CGCCTTTCAG TACCATACGC 1800
 CGTGAAATGT CCACTTTTGC TTCTGTGCGA ACTCATTGCT GA 1842

10

20

30

(213) INFORMATION FOR SEQ ID NO:212:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 613 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:212:

```

Met Arg Ala Pro Gly Ala Leu Leu Ala Arg Met Ser Arg Leu Leu Leu
1           5           10           15
Leu Leu Leu Leu Lys Val Ser Ala Ser Ser Ala Leu Gly Val Ala Pro
20           25           30
Ala Ser Arg Asn Glu Thr Cys Leu Gly Glu Ser Cys Ala Pro Thr Val
35           40           45
Ile Gln Arg Arg Gly Arg Asp Ala Trp Gly Pro Gly Asn Ser Ala Arg
50           55           60
Asp Val Leu Arg Ala Arg Ala Pro Arg Glu Glu Gln Gly Ala Ala Phe
65           70           75           80
Leu Ala Gly Pro Ser Trp Asp Leu Pro Ala Ala Pro Gly Arg Asp Pro
85           90           95
Ala Ala Gly Arg Gly Ala Glu Ala Ser Ala Ala Gly Pro Pro Gly Pro
100          105          110
Pro Thr Arg Pro Pro Gly Pro Trp Arg Trp Lys Gly Ala Arg Gly Gln
115          120          125
Glu Pro Ser Glu Thr Leu Gly Arg Gly Asn Pro Thr Ala Leu Gln Leu
130          135          140
Phe Leu Gln Ile Ser Glu Glu Glu Glu Lys Gly Pro Arg Gly Ala Gly
145          150          155          160
Ile Ser Gly Arg Ser Gln Glu Gln Ser Val Lys Thr Val Pro Gly Ala
165          170          175
Ser Asp Leu Phe Tyr Trp Pro Arg Arg Ala Gly Lys Leu Gln Gly Ser
180          185          190
His His Lys Pro Leu Ser Lys Thr Ala Asn Gly Leu Ala Gly His Glu
195          200          205
Gly Trp Thr Ile Ala Leu Pro Gly Arg Ala Leu Ala Gln Asn Gly Ser
210          215          220
Leu Gly Glu Gly Ile His Glu Pro Gly Gly Pro Arg Arg Gly Asn Ser

```

10

20

30

225 230 235 240
 Thr Asn Arg Arg Val Arg Leu Lys Asn Pro Phe Tyr Pro Leu Thr Gln
 245 250 255
 Glu Ser Tyr Gly Ala Tyr Ala Val Met Cys Leu Ser Val Val Ile Phe
 260 265 270
 Gly Thr Gly Ile Ile Gly Asn Leu Ala Val Met Cys Ile Val Cys His
 275 280 285
 Asn Tyr Tyr Met Arg Ser Ile Ser Asn Ser Leu Leu Ala Asn Leu Ala
 290 295 300
 Phe Trp Asp Phe Leu Ile Ile Phe Phe Cys Leu Pro Leu Val Ile Phe
 305 310 315 320
 His Glu Leu Thr Lys Lys Trp Leu Leu Glu Asp Phe Ser Cys Lys Ile
 325 330 335
 Val Pro Tyr Ile Glu Val Ala Ser Leu Gly Val Thr Thr Phe Thr Leu
 340 345 350
 Cys Ala Leu Cys Ile Asp Arg Phe Arg Ala Ala Thr Asn Val Gln Met
 355 360 365
 Tyr Tyr Glu Met Ile Glu Asn Cys Ser Ser Thr Thr Ala Lys Leu Ala
 370 375 380
 Val Ile Trp Val Gly Ala Leu Leu Leu Ala Leu Pro Glu Val Val Leu
 385 390 395 400
 Arg Gln Leu Ser Lys Glu Asp Leu Gly Phe Ser Gly Arg Ala Pro Ala
 405 410 415
 Glu Arg Cys Ile Ile Lys Ile Ser Pro Asp Leu Pro Asp Thr Ile Tyr
 420 425 430
 Val Leu Ala Leu Thr Tyr Asp Ser Ala Arg Leu Trp Trp Tyr Phe Gly
 435 440 445
 Cys Tyr Phe Cys Leu Pro Thr Leu Phe Thr Ile Thr Cys Ser Leu Val
 450 455 460
 Thr Ala Arg Lys Ile Arg Lys Ala Glu Lys Ala Cys Thr Arg Gly Asn
 465 470 475 480
 Lys Arg Gln Ile Gln Leu Glu Ser Gln Met Lys Cys Thr Val Val Ala
 485 490 495
 Leu Thr Ile Leu Tyr Gly Phe Cys Ile Ile Pro Glu Asn Ile Cys Asn
 500 505 510
 Ile Val Thr Ala Tyr Met Ala Thr Gly Val Ser Gln Gln Thr Met Asp
 515 520 525

10

20

30

Leu Leu Asn Ile Ile Ser Gln Phe Leu Leu Phe Phe Lys Ser Cys Val
 530 535 540
 Thr Pro Val Leu Leu Phe Cys Leu Cys Lys Pro Phe Ser Arg Ala Phe
 545 550 555 560
 Met Glu Cys Cys Cys Cys Cys Cys Glu Glu Cys Ile Gln Lys Ser Ser
 565 570 575
 Thr Val Thr Ser Asp Asp Asn Asp Asn Glu Tyr Thr Thr Glu Leu Glu
 580 585 590
 Leu Ser Pro Phe Ser Thr Ile Arg Arg Glu Met Ser Thr Phe Ala Ser
 595 600 605
 Val Gly Thr His Cys
 610

10

(214) INFORMATION FOR SEQ ID NO:213:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1248 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:213:

20

ATGGTTTTTG CTCACAGAAT GGATAACAGC AAGCCACATT TGATTATTCC TACACTTCTG 60
 GTGCCCTCC AAAACCGCAG CTGCACTGAA ACAGCCACAC CTCTGCCAAG CCAATACCTG 120
 ATGGAATTAA GTGAGGAGCA CAGTTGGATG AGCAACCAAA CAGACCTTCA CTATGTGCTG 180
 AAACCCGGGG AAGTGGCCAC AGCCAGCATC TTCTTTGGGA TTCTGTGGTT GTTTTCTATC 240
 TTCGGCAATT CCCTGGTTTG TTTGGTCATC CATAGGAGTA GGAGGACTCA GTCTACCACC 300
 AACTACTTTG TGGTCTCCAT GGCATGTGCT GACCTTCTCA TCAGCGTTGC CAGCACGCCT 360
 TTCGTCTGCT TCCAGTTCAC CACTGGAAGG TGGACGCTGG GTAGTGCAAC GTGCAAGGTT 420
 GTGCGATATT TTCAATATCT CACTCCAGGT GTCCAGATCT ACGTTCCTCCT CTCCATCTGC 480
 ATAGACCGGT TCTACACCAT CGTCTATCCT CTGAGCTTCA AGGTGTCCAG AGAAAAAGCC 540
 AAGAAAATGA TTGCGGCATC GTGGATCTTT GATGCAGGCT TTGTGACCCC TGTGCTCTTT 600
 TTCTATGGCT CCAACTGGGA CAGTCATTGT AACTATTTC TCCCTCCTC TTGGGAAGGC 660
 ACTGCCTACA CTGTCAATCA CTTCTTGGTG GGCTTTGTGA TTCCATCTGT CCTCATAATT 720
 TTATTTTACC AAAAGGTCAT AAAATATATT TGGAGAATAG GCACAGATGG CCGAACGGTG 780

30

AGGAGGACAA TGAACATTGT CCCTCGGACA AAAGTGAAAA CTAAAAAGAT GTTCCTCATT 840
 TTAAATCTGT TGTTTTTGCT CTCCTGGCTG CCTTTTCATG TAGCTCAGCT ATGGCACCCC 900
 CATGAACAAG ACTATAAGAA AAGTTCCTT GTTTTCACAG CTATCACATG GATATCCTTT 960
 AGTCTTCAG CCTCTAAACC TACTCTGTAT TCAATTTATA ATGCCAATTT TCGGAGAGGG 1020
 ATGAAAGAGA CTTTTTGCAT GTCCTCTATG AAATGTTACC GAAGCAATGC CTATACTATC 1080
 ACAACAAGTT CAAGGATGGC CAAAAAAAAC TACGTTGGCA TTTCAGAAAT CCCTTCCATG 1140
 GCCAAACTA TTACCAAAGA CTCGATCTAT GACTCATTG ACAGAGAAGC CAAGGAAAAA 1200
 AAGCTTGCTT GGCCATTAA CTCAATCCA CCAATACTT TTGTCTAA 1248

10

(215) INFORMATION FOR SEQ ID NO:214:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 415 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:214:

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

30

Gln Tyr Leu Thr Pro Gly Val Gln Ile Tyr Val Leu Leu Ser Ile Cys
 145 150 155 160
 Ile Asp Arg Phe Tyr Thr Ile Val Tyr Pro Leu Ser Phe Lys Val Ser
 165 170 175
 Arg Glu Lys Ala Lys Lys Met Ile Ala Ala Ser Trp Ile Phe Asp Ala
 180 185 190
 Gly Phe Val Thr Pro Val Leu Phe Phe Tyr Gly Ser Asn Trp Asp Ser
 195 200 205
 His Cys Asn Tyr Phe Leu Pro Ser Ser Trp Glu Gly Thr Ala Tyr Thr
 210 215 220
 Val Ile His Phe Leu Val Gly Phe Val Ile Pro Ser Val Leu Ile Ile
 225 230 235 240
 Leu Phe Tyr Gln Lys Val Ile Lys Tyr Ile Trp Arg Ile Gly Thr Asp
 245 250 255
 Gly Arg Thr Val Arg Arg Thr Met Asn Ile Val Pro Arg Thr Lys Val
 260 265 270
 Lys Thr Lys Lys Met Phe Leu Ile Leu Asn Leu Leu Phe Leu Leu Ser
 275 280 285
 Trp Leu Pro Phe His Val Ala Gln Leu Trp His Pro His Glu Gln Asp
 290 295 300
 Tyr Lys Lys Ser Ser Leu Val Phe Thr Ala Ile Thr Trp Ile Ser Phe
 305 310 315 320
 Ser Ser Ser Ala Ser Lys Pro Thr Leu Tyr Ser Ile Tyr Asn Ala Asn
 325 330 335
 Phe Arg Arg Gly Met Lys Glu Thr Phe Cys Met Ser Ser Met Lys Cys
 340 345 350
 Tyr Arg Ser Asn Ala Tyr Thr Ile Thr Thr Ser Ser Arg Met Ala Lys
 355 360 365
 Lys Asn Tyr Val Gly Ile Ser Glu Ile Pro Ser Met Ala Lys Thr Ile
 370 375 380
 Thr Lys Asp Ser Ile Tyr Asp Ser Phe Asp Arg Glu Ala Lys Glu Lys
 385 390 395 400
 Lys Leu Ala Trp Pro Ile Asn Ser Asn Pro Pro Asn Thr Phe Val
 405 410 415

10

20

30

(216) INFORMATION FOR SEQ ID NO:215:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1842 base pairs

(B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:215:

ATGGGGCCCA CCCTAGCGGT TCCCACCCCC TATGGCTGTA TTGGCTGTAA GCTACCCAG	60
CCAGAATACC CACCGGCTCT AATCATCTTT ATGTTCTGCG CGATGGTTAT CACCATCGTT	120
GTAGACCTAA TCGGCAACTC CATGGTCATT TTGGCTGTGA CGAAGAACAA GAAGCTCCGG	180
AATTCTGGCA ACATCTTCGT GGTCACTCTC TCTGTGGCCG ATATGCTGGT GGCCATCTAC	240
CCATACCCCTT TGATGCTGCA TGCCATGTCC ATTGGGGGCT GGGATCTGAG CCAGTTACAG	300
TGCCAGATGG TCGGGTTCAT CACAGGCGTG AGTGTGGTCG GCTCCATCTT CAACATCGTG	360
GCAATCGCTA TCAACCGTTA CTGCTACATC TGCCACAGCC TCCAGTACGA ACGGATCTTC	420
AGTGTGCGCA ATACCTGCAT CTACCTGGTC ATCACCTGGA TCATGACCGT CCTGGCTGTC	480
CTGCCCCACA TGTACATTGG CACCATCGAG TACGATCCTC GCACCTACAC CTGCATCTTC	540
AACIATCTGA ACAACCTGT CTTCCTGTT ACCATCGTCT GCATCCACTT CGTCCTCCCT	600
CTCCTCATCG TGGGTTTCTG CTACGTGAGG ATCTGGACCA AAGTGCTGGC GGCCCGTGAC	660
CCTGCAGGGC AGAATCCTGA CAACCAACTT GCTGAGGTTT GCAATAAACT AACCATGTTT	720
GTGATCTTCC TCCTCTTTCG AGTGTGCTGG TGCCCTATCA ACGTGCTCAC TGTCTTGGTG	780
GCTGTGCTGC CGAAGGAGAT GGCAGGCAAG ATCCCAACT GGCTTTATCT TGCAGCCTAC	840
TTCATAGCCT ACTTCAACAG CTGCCTCAAC GCTGTGATCT ACGGGCTCCT CAATGAGAAT	900
TTCCGAAGAG AATACTGGAC CATCTTCCAT GCTATGCGGC ACCCTATCAT ATTCTTCTCT	960
GGCCTCATCA GTGATATTCG TGAGATGCAG GAGGCCCCGA CCCTGGCCCG CGCCCGTGCC	1020
CATGCTCGCG ACCAAGCTCG TGAACAAGAC CGTGCCCATG CCTGTCTCTG TGTGGAGGAA	1080
ACCCCGATGA ATGTCCGGAA TGTTCATTG CCTGGTGATG CTGCAGCTGG CCACCCCGAC	1140
CGTGCTCTG GCCACCTAA GCCCCATTCC AGATCCTCCT CTGCCTATCG CAAATCTGCC	1200
TCTACCCACC ACAAGTCTGT CTTTAGCCAC TCCAAGGCTG CCTCTGGTCA CCTCAAGCCT	1260
GTCTCTGGCC ACTCCAAGCC TGCTCTGGT CACCCCAAGT CTGCCACTGT CTACCCTAAG	1320
CCTGCCTCTG TCCATTTCAG GGCTGACTCT GTCCATTTCAG AGGGTGACTC TGTCCATTTC	1380
AAGCCTGACT CTGTTCAATT CAAGCCTGCT TCCAGCAACC CCAAGCCCAT CACTGGCCAC	1440

10

20

30

CATGTCTCTG CTGGCAGCCA CTCCAAGTCT GCCTTCAATG CTGCCACCAG CCACCCTAAA 1500
 CCCATCAAGC CAGCTACCAG CCATGCTGAG CCCACCACTG CTGACTATCC CAAGCCTGCC 1560
 ACTACCAGCC ACCCTAAGCC CGCTGCTGCT GACAACCCTG AGCTCTCTGC CTCCCATTGC 1620
 CCCGAGATCC CTGCCATTGC CCACCCTGTG TGTGACGACA GTGACCTCCC TGAGTCGGCC 1680
 TCTAGCCCTG CCGCTGGGCC CACCAAGCCT GCTGCCAGCC AGCTGGAGTC TGACACCATC 1740
 GCTGACCTTC CTGACCCTAC TGTAGTCACT ACCAGTACCA ATGATTACCA TGATGTCGTG 1800
 GTTGTGATG TTGAAGATGA TCCTGATGAA ATGGCTGTGT GA 1842

(217) INFORMATION FOR SEQ ID NO:216:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 613 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:216:

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

10

20

30

Leu Pro Asn Met Tyr Ile Gly Thr Ile Glu Tyr Asp Pro Arg Thr Tyr
 165 170 175
 Thr Cys Ile Phe Asn Tyr Leu Asn Asn Pro Val Phe Thr Val Thr Ile
 180 185 190
 Val Cys Ile His Phe Val Leu Pro Leu Leu Ile Val Gly Phe Cys Tyr
 195 200 205
 Val Arg Ile Trp Thr Lys Val Leu Ala Ala Arg Asp Pro Ala Gly Gln
 210 215 220
 Asn Pro Asp Asn Gln Leu Ala Glu Val Arg Asn Lys Leu Thr Met Phe
 225 230 235 240
 Val Ile Phe Leu Leu Phe Ala Val Cys Trp Cys Pro Ile Asn Val Leu
 245 250 255
 Thr Val Leu Val Ala Val Ser Pro Lys Glu Met Ala Gly Lys Ile Pro
 260 265 270
 Asn Trp Leu Tyr Leu Ala Ala Tyr Phe Ile Ala Tyr Phe Asn Ser Cys
 275 280 285
 Leu Asn Ala Val Ile Tyr Gly Leu Leu Asn Glu Asn Phe Arg Arg Glu
 290 295 300
 Tyr Trp Thr Ile Phe His Ala Met Arg His Pro Ile Ile Phe Phe Ser
 305 310 315 320
 Gly Leu Ile Ser Asp Ile Arg Glu Met Gln Glu Ala Arg Thr Leu Ala
 325 330 335
 Arg Ala Arg Ala His Ala Arg Asp Gln Ala Arg Glu Gln Asp Arg Ala
 340 345 350
 His Ala Cys Pro Ala Val Glu Glu Thr Pro Met Asn Val Arg Asn Val
 355 360 365
 Pro Leu Pro Gly Asp Ala Ala Ala Gly His Pro Asp Arg Ala Ser Gly
 370 375 380
 His Pro Lys Pro His Ser Arg Ser Ser Ser Ala Tyr Arg Lys Ser Ala
 385 390 395 400
 Ser Thr His His Lys Ser Val Phe Ser His Ser Lys Ala Ala Ser Gly
 405 410 415
 His Leu Lys Pro Val Ser Gly His Ser Lys Pro Ala Ser Gly His Pro
 420 425 430
 Lys Ser Ala Thr Val Tyr Pro Lys Pro Ala Ser Val His Phe Lys Ala
 435 440 445
 Asp Ser Val His Phe Lys Gly Asp Ser Val His Phe Lys Pro Asp Ser

10

20

30

450 455 460
 Val His Phe Lys Pro Ala Ser Ser Asn Pro Lys Pro Ile Thr Gly His
 465 470 475 480
 His Val Ser Ala Gly Ser His Ser Lys Ser Ala Phe Asn Ala Ala Thr
 485 490 495
 Ser His Pro Lys Pro Ile Lys Pro Ala Thr Ser His Ala Glu Pro Thr
 500 505 510
 Thr Ala Asp Tyr Pro Lys Pro Ala Thr Thr Ser His Pro Lys Pro Ala
 515 520 525
 Ala Ala Asp Asn Pro Glu Leu Ser Ala Ser His Cys Pro Glu Ile Pro
 530 535 540
 Ala Ile Ala His Pro Val Ser Asp Asp Ser Asp Leu Pro Glu Ser Ala
 545 550 555 560
 Ser Ser Pro Ala Ala Gly Pro Thr Lys Pro Ala Ala Ser Gln Leu Glu
 565 570 575
 Ser Asp Thr Ile Ala Asp Leu Pro Asp Pro Thr Val Val Thr Thr Ser
 580 585 590
 Thr Asn Asp Tyr His Asp Val Val Val Val Asp Val Glu Asp Asp Pro
 595 600 605
 Asp Glu Met Ala Val
 610

10

20

(218) INFORMATION FOR SEQ ID NO:217:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1854 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:217:

ATGGGGCCCA CCCTAGCGGT TCCCACCCCC TATGGCTGTA TTGGCTGTAA GCTACCCAG 60
 CCAGAATACC CACCGGCTCT AATCATCTTT ATGTTCTGCG CGATGGTTAT CACCATCGTT 120
 GTAGACCTAA TCGGCAACTC CATGGTCATT TTGGCTGTGA CGAAGAACAA GAAGCTCCGG 180
 AATTCTGGCA ACATCTTCGT GGTCACTCTC TCTGTGGCCG ATATGCTGGT GGCCATCTAC 240
 CCATACCCTT TGATGCTGCA TGCCATGTCC ATTGGGGGCT GGGATCTGAG CCAGTTACAG 300
 TGCCAGATGG TCGGGTTCAT CACAGGGCTG AGTGTGGTCG GCTCCATCTT CAACATCGTG 360

30

GCAATCGCTA TCAACCGTTA CTGCTACATC TGCCACAGCC TCCAGTACGA ACGGATCTTC 420
 AGTGTGCGCA ATACCTGCAT CTACCTGGTC ATCACCTGGA TCATGACCGT CCTGGCTGTC 480
 CTGCCCCACA TGTACATTGG CACCATCGAG TACGATCCTC GCACCTACAC CTGCATCTTC 540
 AACTATCTGA ACAACCCGTG CTTCACTGTT ACCATCGTCT GCATCCACTT CGTCCTCCCT 600
 CTCCTCATCG TGGGTTTCTG CTACGTGAGG ATCTGGACCA AAGTGCTGGC GGCCCGTGAC 660
 CCTGCAGGGC AGAATCCTGA CAACCAACTT GCTGAGGTTC GCAATAAACT AACCATGTTT 720
 GTGATCTTCC TCCTCTTTGC AGTGTGCTGG TGCCCTATCA ACGTGCTCAC TGTCTTGGTG 780
 GCTGTCAATC CGAAGGAGAT GGCAGGCAAG ATCCCCAACT GGCTTTATCT TGCAGCCTAC 840
 TTCATAGCCT ACTTCAACAG CTGCCTCAAC GCTGTGATCT ACGGGCTCCT CAATGAGAAT 900
 TTCCGAAGAG AATACTGGAC CATCTTCCAT GCTATGCGGC ACCCTATCAT ATTCTTCTCT 960
 GGCCTCATCA GTGATATTCT TGAGATGCAG GAGGCCCCGA CCCTGGCCCCG CGCCCGTGCC 1020
 CATGCTCGCG ACCAAGCTCG TGAACAAGAC CGTGCCCATG CCTGTCTGTC TGTGGAGGAA 1080
 ACCCCGATGA ATGTCCGGAA TGTTCATTA CCTGGTGATG CTGCAGCTGG CCACCCCGAC 1140
 CGTGCTCTG GCCACCCTAA GCCCCATTCC AGATCCTCCT CTGCCTATCG CAAATCTGCC 1200
 TCTACCCACC ACAAGTCTGT CTTAGCCAC TCCAAGGCTG CCTCTGGTCA CCTCAAGCCT 1260
 GTCTCTGGCC ACTCCAAGCC TGCTCTGGT CACCCCAAGT CTGCCACTGT CTACCCTAAG 1320
 CCTGCCTCTG TCCATTCAA GGCTGACTCT GTCCATTTC AAGGTGACTC TGTCCATTTC 1380
 AAGCTGACT CTGTTCAATT CAAGCCTGCT TCCAGCAACC CCAAGCCCAT CACTGGCCAC 1440
 CATGTCTCTG CTGGCAGCCA CTCCAAGTCT GCCTTCAGTG CTGCCACCAG CCACCCTAAA 1500
 CCCACCACTG GCCACATCAA GCCAGCTACC AGCCATGCTG AGCCCACCAC TGCTGACTAT 1560
 CCCAAGCCTG CCACTACCAG CCACCCTAAG CCCACTGCTG CTGACAACCC TGAGCTCTCT 1620
 GCCTCCCAT T GCCCCGAGAT CCCTGCCATT GCCCACCCTG TGTCTGACGA CAGTGACCTC 1680
 CCTGAGTCGG CCTCTAGCCC TGCCGCTGGG CCCACCAAGC CTGCTGCCAG CCAGCTGGAG 1740
 TCTGACACCA TCGCTGACCT TCCTGACCCT ACTGTAGTCA CTACCAGTAC CAATGATTAC 1800
 CATGATGTCG TGGTTGTTGA TGTGAAGAT GATCCTGATG AAATGGCTGT GTGA 1854

10

20

30

(219) INFORMATION FOR SEQ ID NO:218:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 617 amino acids
 - (B) TYPE: amino acid

(C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:218:

```

Met Gly Pro Thr Leu Ala Val Pro Thr Pro Tyr Gly Cys Ile Gly Cys
1           5           10           15

Lys Leu Pro Gln Pro Glu Tyr Pro Pro Ala Leu Ile Ile Phe Met Phe
          20           25           30

Cys Ala Met Val Ile Thr Ile Val Val Asp Leu Ile Gly Asn Ser Met
          35           40           45

Val Ile Leu Ala Val Thr Lys Asn Lys Lys Leu Arg Asn Ser Gly Asn
          50           55           60

Ile Phe Val Val Ser Leu Ser Val Ala Asp Met Leu Val Ala Ile Tyr
          65           70           75           80

Pro Tyr Pro Leu Met Leu His Ala Met Ser Ile Gly Gly Trp Asp Leu
          85           90           95

Ser Gln Leu Gln Cys Gln Met Val Gly Phe Ile Thr Gly Leu Ser Val
          100          105          110

Val Gly Ser Ile Phe Asn Ile Val Ala Ile Ala Ile Asn Arg Tyr Cys
          115          120          125

Tyr Ile Cys His Ser Leu Gln Tyr Glu Arg Ile Phe Ser Val Arg Asn
          130          135          140

Thr Cys Ile Tyr Leu Val Ile Thr Trp Ile Met Thr Val Leu Ala Val
          145          150          155          160

Leu Pro Asn Met Tyr Ile Gly Thr Ile Glu Tyr Asp Pro Arg Thr Tyr
          165          170          175

Thr Cys Ile Phe Asn Tyr Leu Asn Asn Pro Val Phe Thr Val Thr Ile
          180          185          190

Val Cys Ile His Phe Val Leu Pro Leu Leu Ile Val Gly Phe Cys Tyr
          195          200          205

Val Arg Ile Trp Thr Lys Val Leu Ala Ala Arg Asp Pro Ala Gly Gln
          210          215          220

Asn Pro Asp Asn Gln Leu Ala Glu Val Arg Asn Lys Leu Thr Met Phe
          225          230          235          240

Val Ile Phe Leu Leu Phe Ala Val Cys Trp Cys Pro Ile Asn Val Leu
          245          250          255

```

10

20

30

Thr Val Leu Val Ala Val Ser Pro Lys Glu Met Ala Gly Lys Ile Pro
 260 265 270
 Asn Trp Leu Tyr Leu Ala Ala Tyr Phe Ile Ala Tyr Phe Asn Ser Cys
 275 280 285
 Leu Asn Ala Val Ile Tyr Gly Leu Leu Asn Glu Asn Phe Arg Arg Glu
 290 295 300
 Tyr Trp Thr Ile Phe His Ala Met Arg His Pro Ile Ile Phe Phe Ser
 305 310 315 320
 Gly Leu Ile Ser Asp Ile Arg Glu Met Gln Glu Ala Arg Thr Leu Ala
 325 330 335
 Arg Ala Arg Ala His Ala Arg Asp Gln Ala Arg Glu Gln Asp Arg Ala
 340 345 350
 His Ala Cys Pro Ala Val Glu Glu Thr Pro Met Asn Val Arg Asn Val
 355 360 365
 Pro Leu Pro Gly Asp Ala Ala Ala Gly His Pro Asp Arg Ala Ser Gly
 370 375 380
 His Pro Lys Pro His Ser Arg Ser Ser Ser Ala Tyr Arg Lys Ser Ala
 385 390 395 400
 Ser Thr His His Lys Ser Val Phe Ser His Ser Lys Ala Ala Ser Gly
 405 410 415
 His Leu Lys Pro Val Ser Gly His Ser Lys Pro Ala Ser Gly His Pro
 420 425 430
 Lys Ser Ala Thr Val Tyr Pro Lys Pro Ala Ser Val His Phe Lys Ala
 435 440 445
 Asp Ser Val His Phe Lys Gly Asp Ser Val His Phe Lys Pro Asp Ser
 450 455 460
 Val His Phe Lys Pro Ala Ser Ser Asn Pro Lys Pro Ile Thr Gly His
 465 470 475 480
 His Val Ser Ala Gly Ser His Ser Lys Ser Ala Phe Ser Ala Ala Thr
 485 490 495
 Ser His Pro Lys Pro Thr Thr Gly His Ile Lys Pro Ala Thr Ser His
 500 505 510
 Ala Glu Pro Thr Thr Ala Asp Tyr Pro Lys Pro Ala Thr Thr Ser His
 515 520 525
 Pro Lys Pro Thr Ala Ala Asp Asn Pro Glu Leu Ser Ala Ser His Cys
 530 535 540
 Pro Glu Ile Pro Ala Ile Ala His Pro Val Ser Asp Asp Ser Asp Leu

10

20

30

545	550	555	560
Pro Glu Ser Ala Ser Ser Pro Ala Ala Gly Pro Thr Lys Pro Ala Ala			
	565	570	575
Ser Gln Leu Glu Ser Asp Thr Ile Ala Asp Leu Pro Asp Pro Thr Val			
	580	585	590
Val Thr Thr Ser Thr Asn Asp Tyr His Asp Val Val Val Val Asp Val			
	595	600	605
Glu Asp Asp Pro Asp Glu Met Ala Val			
	610	615	

(220) INFORMATION FOR SEQ ID NO:219:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1548 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:219:

ATGGGACATA ACGGGAGCTG GATCTCTCCA AATGCCAGCG AGCCGCACAA CGCGTCCGGC	60
GCCGAGGCTG CGGGTGTGAA CCGCAGCGCG CTCGGGGAGT TCGGCGAGGC GCAGCTGTAC	120
CGCCAGTTCA CCACCACCGT GCAGGTCGTC ATCTTCATAG GCTCGCTGCT CGGAAACTTC	180
ATGGTGTAT GGTCAACTTG CCGCACAAACC GTGTTCAAAT CTGTACCAA CAGGTTTCATT	240
AAAAACCTGG CCTGCTCGGG GATTTGTGCC AGCCTGGTCT GTGTGCCCTT CGACATCATC	300
CTCAGCACCA GTCCTCACTG TTGCTGGTGG ATCTACACCA TGCTCTTCTG CAAGGTCGTC	360
AAATTTTTCG ACAAAGTATT CTGCTCTGTG ACCATCCTCA GCTTCCCTGC TATTGCTTTG	420
GACAGGTACT ACTCAGTCCT CTATCCACTG GAGAGGAAAA TATCTGATGC CAAGTCCCGT	480
GAAGTGGTGA TGTACATCTG GGCCCATGCA GTGGTGGCCA GTGTCCCTGT GTTTGCAGTA	540
ACCAATGTGG CTGACATCTA TGCCACGTCC ACCTGCACGG AAGTCTGGAG CAACTCCTTG	600
GGCCACCTGG TGTACGTTCT GGTGTATAAC ATCACCACGG TCATTGTGCC TGTGGTGGTG	660
GTGTTCTCTT TCTTGATACT GATCCGACGG GCCCTGAGTG CCAGCCAGAA GAAGAAGGTC	720
ATCATAGCAG CGCTCCGGAC CCCACAGAAC ACCATCTCTA TTCCCTATGC CTCCCAGCGG	780
GAGGCCGAGC TGAAAGCCAC CCTGCTCTCC ATGGTGATGG TCTTCATCTT GTGTAGCGTG	840
CCCTATGCCA CCCTGGTCGT CTACCAGACT GTGCTCAATG TCCCTGACAC TTCCGTCTTC	900

10

20

30

```

TTGCTGCTCA CTGCTGTTTG GCTGCCCAAA GTCTCCCTGC TGGCAAACCC TGTCTCTTTT   960
CTTACTGTGA ACAAATCTGT CCGCAAGTGC TTGATAGGGA CCCTGGTGCA ACTACACCAC  1020
CGGTACAGTC GCCGTAATGT GGTCAGTACA GGGAGTGGCA TGGCTGAGGC CAGCCTGGAA  1080
CCCAGCATAC GCTCGGGTAG CCAGCTCCTG GAGATGTTCC ACATTGGGCA GCAGCAGATC  1140
TTTAAGCCCA CAGAGGATGA GGAAGAGAGT GAGGCCAAGT ACATTGGCTC AGCTGACTTC  1200
CAGGCCAAGG AGATATTTAG CACCTGCCTG GAGGGAGAGC AGGGGCCACA GTTTGCGCCC  1260
TCTGCCCCAC CCCTGAGCAC AGTGGACTCT GTATCCCAGG TGGCACCGGC AGCCCCTGTG  1320
GAACCTGAAA CATTCCCTGA TAAGTATTCC CTGCAGTTTG GCTTTGGGCC TTTTGAGTTG  1380
CCTCCTCAGT GGCTCTCAGA GACCCGAAAC AGCAAGAAGC GGCTGCTTCC CCCCTTGGGC  1440
AACACCCCAG AAGAGCTGAT CCAGACAAAG GTGCCCAAGG TAGGCAGGGT GGAGCGGAAG  1500
ATGAGCAGAA ACAATAAAGT GAGCATTTTT CCAAAGGTGG ATTCCTAG                1548

```

(221) INFORMATION FOR SEQ ID NO:220:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 515 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:220:

```

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

```

10

20

30

Thr Met Leu Phe Cys Lys Val Val Lys Phe Leu His Lys Val Phe Cys
 115 120 125
 Ser Val Thr Ile Leu Ser Phe Pro Ala Ile Ala Leu Asp Arg Tyr Tyr
 130 135 140
 Ser Val Leu Tyr Pro Leu Glu Arg Lys Ile Ser Asp Ala Lys Ser Arg
 145 150 155 160
 Glu Leu Val Met Tyr Ile Trp Ala His Ala Val Val Ala Ser Val Pro
 165 170 175
 Val Phe Ala Val Thr Asn Val Ala Asp Ile Tyr Ala Thr Ser Thr Cys
 180 185 190
 Thr Glu Val Trp Ser Asn Ser Leu Gly His Leu Val Tyr Val Leu Val
 195 200 205
 Tyr Asn Ile Thr Thr Val Ile Val Pro Val Val Val Phe Leu Phe
 210 215 220
 Leu Ile Leu Ile Arg Arg Ala Leu Ser Ala Ser Gln Lys Lys Lys Val
 225 230 235 240
 Ile Ile Ala Ala Leu Arg Thr Pro Gln Asn Thr Ile Ser Ile Pro Tyr
 245 250 255
 Ala Ser Gln Arg Glu Ala Glu Leu Lys Ala Thr Leu Leu Ser Met Val
 260 265 270
 Met Val Phe Ile Leu Cys Ser Val Pro Tyr Ala Thr Leu Val Val Tyr
 275 280 285
 Gln Thr Val Leu Asn Val Pro Asp Thr Ser Val Phe Leu Leu Leu Thr
 290 295 300
 Ala Val Trp Leu Pro Lys Val Ser Leu Leu Ala Asn Pro Val Leu Phe
 305 310 315 320
 Leu Thr Val Asn Lys Ser Val Arg Lys Cys Leu Ile Gly Thr Leu Val
 325 330 335
 Gln Leu His His Arg Tyr Ser Arg Arg Asn Val Val Ser Thr Gly Ser
 340 345 350
 Gly Met Ala Glu Ala Ser Leu Glu Pro Ser Ile Arg Ser Gly Ser Gln
 355 360 365
 Leu Leu Glu Met Phe His Ile Gly Gln Gln Gln Ile Phe Lys Pro Thr
 370 375 380
 Glu Asp Glu Glu Glu Ser Glu Ala Lys Tyr Ile Gly Ser Ala Asp Phe
 385 390 395 400
 Gln Ala Lys Glu Ile Phe Ser Thr Cys Leu Glu Gly Glu Gln Gly Pro

10

20

30

405 410 415
 Gln Phe Ala Pro Ser Ala Pro Pro Leu Ser Thr Val Asp Ser Val Ser
 420 425 430
 Gln Val Ala Pro Ala Ala Pro Val Glu Pro Glu Thr Phe Pro Asp Lys
 435 440 445
 Tyr Ser Leu Gln Phe Gly Phe Gly Pro Phe Glu Leu Pro Pro Gln Trp
 450 455 460
 Leu Ser Glu Thr Arg Asn Ser Lys Lys Arg Leu Leu Pro Pro Leu Gly
 465 470 475 480
 Asn Thr Pro Glu Glu Leu Ile Gln Thr Lys Val Pro Lys Val Gly Arg
 485 490 495
 Val Glu Arg Lys Met Ser Arg Asn Asn Lys Val Ser Ile Phe Pro Lys
 500 505 510
 Val Asp Ser
 515

10

(222) INFORMATION FOR SEQ ID NO:221:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1164 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:221:

ATGAATCGGC ACCATCTGCA GGATCACTTT CTGGAAATAG ACAAGAAGAA CTGCTGTGTG 60
 TTCCGAGATG ACTTCATTGC CAAGGTGTTG CCGCCGGTGT TGGGGCTGGA GTTTATCTTT 120
 GGGCTTCTGG GCAATGGCCT TGCCCTGTGG ATTTTCTGTT TCCACCTCAA GTCCTGGAAA 180
 TCCAGCCGGA TTTTCTGTGT CAACCTGGCA GTAGCTGACT TTCTACTGAT CATCTGCCTG 240
 CCGTTCGTGA TGGACTACTA TGTGCGGCGT TCAGACTGGA AGTTTGGGGA CATCCCTTGC 300
 CGGCTGGTGC TCTTCATGTT TGCCATGAAC CGCCAGGGCA GCATCATCTT CCTCACGGTG 360
 GTGGCGGTAG ACAGGTATTT CCGGGTGGTC CATCCCCACC ACGCCCTGAA CAAGATCTCC 420
 AATTGGACAG CAGCCATCAT CTCTTGCCCT CTGTGGGGCA TCACTGTTGG CCTAACAGTC 480
 CACCTCCTGA AGAAGAAGTT GCTGATCCAG AATGGCCCTG CAAATGTGTG CATCAGCTTC 540
 AGCATCTGCC ATACCTTCCG GTGGCACGAA GCTATGTTCC TCCTGGAGTT CCTCCTGCCC 600

30

CTGGGCATCA TCCTGTTCTG CTCAGCCAGA ATTATCTGGA GCCTGCGGCA GAGACAAATG 660
 GACCGGCATG CCAAGATCAA GAGAGCCAAA ACCTTCATCA TGGTGGTGGC CATCGTCTTT 720
 GTCATCTGCT TCCTTCCCAG CGTGTTGTG CGGATCCGCA TCTTCTGGCT CCTGCACACT 780
 TCGGGCACGC AGAATTGTGA AGTGATACGC TCGGTGGACC TGGCGTTCTT TATCACTCTC 840
 AGCTTCACCT ACATGAACAG CATGCTGGAC CCCGTGGTGT ACTACTTCTC CAGCCCATCC 900
 TTCCCAACT TCTTCTCCAC TTGATCAAC CGCTGCCTCC AGAGGAAGAT GACAGGTGAG 960
 CCAGATAATA ACCGCAGCAC GAGCGTCGAG CTCACAGGGG ACCCCAACAA AACCAGAGGC 1020
 GCTCCAGAGG CGTTAATGGC CAACTCCGGT GAGCCATGGA GCCCCTCTTA TCTGGGCCCA 1080
 ACCTCAAATA ACCATTCCAA GAAGGGACAT TGTCACCAAG AACCAGCATC TCTGGAGAAA 1140
 CAGTTGGGCT GTTGCAATCGA GTAA 1164

(223) INFORMATION FOR SEQ ID NO:222:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 387 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:222:

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

10

20

30

Val Val His Pro His His Ala Leu Asn Lys Ile Ser Asn Trp Thr Ala
 130 135 140
 Ala Ile Ile Ser Cys Leu Leu Trp Gly Ile Thr Val Gly Leu Thr Val
 145 150 155 160
 His Leu Leu Lys Lys Lys Leu Leu Ile Gln Asn Gly Pro Ala Asn Val
 165 170 175
 Cys Ile Ser Phe Ser Ile Cys His Thr Phe Arg Trp His Glu Ala Met
 180 185 190
 Phe Leu Leu Glu Phe Leu Leu Pro Leu Gly Ile Ile Leu Phe Cys Ser
 195 200 205
 Ala Arg Ile Ile Trp Ser Leu Arg Gln Arg Gln Met Asp Arg His Ala
 210 215 220
 Lys Ile Lys Arg Ala Lys Thr Phe Ile Met Val Val Ala Ile Val Phe
 225 230 235 240
 Val Ile Cys Phe Leu Pro Ser Val Val Val Arg Ile Arg Ile Phe Trp
 245 250 255
 Leu Leu His Thr Ser Gly Thr Gln Asn Cys Glu Val Tyr Arg Ser Val
 260 265 270
 Asp Leu Ala Phe Phe Ile Thr Leu Ser Phe Thr Tyr Met Asn Ser Met
 275 280 285
 Leu Asp Pro Val Val Tyr Tyr Phe Ser Ser Pro Ser Phe Pro Asn Phe
 290 295 300
 Phe Ser Thr Leu Ile Asn Arg Cys Leu Gln Arg Lys Met Thr Gly Glu
 305 310 315 320
 Pro Asp Asn Asn Arg Ser Thr Ser Val Glu Leu Thr Gly Asp Pro Asn
 325 330 335
 Lys Thr Arg Gly Ala Pro Glu Ala Leu Met Ala Asn Ser Gly Glu Pro
 340 345 350
 Trp Ser Pro Ser Tyr Leu Gly Pro Thr Ser Asn Asn His Ser Lys Lys
 355 360 365
 Gly His Cys His Gln Glu Pro Ala Ser Leu Glu Lys Gln Leu Gly Cys
 370 375 380
 Cys Ile Glu
 385

10

20

30

(224) INFORMATION FOR SEQ ID NO:223:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1212 base pairs

(B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:223:

ATGGCTTGCA ATGGCAGTGC GGCCAGGGGG CACTTTGACC CTGAGGACTT GAACCTGACT	60
GACGAGGCAC TGAGACTCAA GTACCTGGGG CCCAGCAGA CAGAGCTGTT CATGCCCATC	120
TGTGCCACAT ACCTGCTGAT CTTCTGGTG GCGCTGTGG GCAATGGGCT GACCTGTCTG	180
GTCACTCTGC GCCACAAGGC CATGCGCACG CCTACCAACT ACTACCTCTT CAGCCTGGCC	240
GTGTCGGACC TGCTGGTGCT GCTGGTGGGC CTGCCCTGG AGCTCTATGA GATGTGGCAC	300
AACTACCCCT TCCTGCTGGG CGTTGGTGGC TGCTATTTCC GCACGCTACT GTTTGAGATG	360
GTCTGCCTGG CCTCAGTGCT CAACGTCCT GCCCTGAGCG TGGAACGCTA TGTGGCCGTG	420
GTGCACCCAC TCCAGGCCAG GTCCATGGTG ACGCGGGCCC ATGTGCGCCG AGTGCTTGGG	480
GCCGTCTGGG GTCTTGCCAT GCTCTGCTCC CTGCCAACA CCAGCCTGCA CGGCATCCGG	540
CAGCTGCACG TGCCCTGCCG GGGCCAGTG CCAGACTCAG CTGTTTGCA TGTGGTCCGC	600
CCACGGGCCC TCTACAACAT GGTAGTGAG ACCACCGCG TGCTCTCTT CTGCCTGCC	660
ATGGCCATCA TGAGCGTGCT CTACCTGCTC ATTGGGCTGC GACTGCGGCG GGAGAGGCTG	720
CTGCTCATGC AGGAGGCCAA GGGCAGGGGC TCTGCAGCAG CCAGGTCCAG ATACACCTGC	780
AGGCTCCAGC AGCAGATCG GGGCCGAGA CAAGTGAAGA AGATGCTGTT TGTCTGGTC	840
GTGGTGTTTG GCATCTGCTG GGGCCCGTTC CACGCCGACC GCGTCATGTG GAGCGTCGTG	900
TCACAGTGGA CAGATGGCCT GCACCTGGCC TTCCAGCACG TGCACGTCAT CTCCGGCATC	960
TTCTTCTACC TGGGCTCGGC GGCCAACCCC GTGCTCTATA GCCTCATGTC CAGCCGCTTC	1020
CGAGAGACCT TCCAGGAGGC CCTGTGCCTC GGGGCTGTCT GCCATCGCCT CAGACCCGCG	1080
CACAGCTCCC ACAGCCTCAG CAGGATGACC ACAGGCAGCA CCCTGTGTGA TGTGGGCTCC	1140
CTGGGCAGCT GGGTCCACCC CCTGGCTGGG AACGATGGCC CAGAGGCGCA GCAAGAGACC	1200
GATCCATCCT GA	1212

(225) INFORMATION FOR SEQ ID NO:224:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 403 amino acids
 (B) TYPE: amino acid

10

20

30

(C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:224:

```

Met Ala Cys Asn Gly Ser Ala Ala Arg Gly His Phe Asp Pro Glu Asp
1           5           10           15

Leu Asn Leu Thr Asp Glu Ala Leu Arg Leu Lys Tyr Leu Gly Pro Gln
          20           25           30

Gln Thr Glu Leu Phe Met Pro Ile Cys Ala Thr Tyr Leu Leu Ile Phe
          35           40           45

Val Val Gly Ala Val Gly Asn Gly Leu Thr Cys Leu Val Ile Leu Arg
          50           55           60

His Lys Ala Met Arg Thr Pro Thr Asn Tyr Tyr Leu Phe Ser Leu Ala
65           70           75           80

Val Ser Asp Leu Leu Val Leu Leu Val Gly Leu Pro Leu Glu Leu Tyr
          85           90           95

Glu Met Trp His Asn Tyr Pro Phe Leu Leu Gly Val Gly Gly Cys Tyr
          100          105          110

Phe Arg Thr Leu Leu Phe Glu Met Val Cys Leu Ala Ser Val Leu Asn
          115          120          125

Val Thr Ala Leu Ser Val Glu Arg Tyr Val Ala Val Val His Pro Leu
          130          135          140

Gln Ala Arg Ser Met Val Thr Arg Ala His Val Arg Arg Val Leu Gly
145          150          155          160

Ala Val Trp Gly Leu Ala Met Leu Cys Ser Leu Pro Asn Thr Ser Leu
          165          170          175

His Gly Ile Arg Gln Leu His Val Pro Cys Arg Gly Pro Val Pro Asp
          180          185          190

Ser Ala Val Cys Met Leu Val Arg Pro Arg Ala Leu Tyr Asn Met Val
          195          200          205

Val Gln Thr Thr Ala Leu Leu Phe Phe Cys Leu Pro Met Ala Ile Met
          210          215          220

Ser Val Leu Tyr Leu Leu Ile Gly Leu Arg Leu Arg Arg Glu Arg Leu
225          230          235          240

Leu Leu Met Gln Glu Ala Lys Gly Arg Gly Ser Ala Ala Ala Arg Ser
          245          250          255

```

10

20

30

Arg Tyr Thr Cys Arg Leu Gln Gln His Asp Arg Gly Arg Arg Gln Val
 260 265 270
 Lys Lys Met Leu Phe Val Leu Val Val Val Phe Gly Ile Cys Trp Ala
 275 280 285
 Pro Phe His Ala Asp Arg Val Met Trp Ser Val Val Ser Gln Trp Thr
 290 295 300
 Asp Gly Leu His Leu Ala Phe Gln His Val His Val Ile Ser Gly Ile
 305 310 315 320
 Phe Phe Tyr Leu Gly Ser Ala Ala Asn Pro Val Leu Tyr Ser Leu Met
 325 330 335
 Ser Ser Arg Phe Arg Glu Thr Phe Gln Glu Ala Leu Cys Leu Gly Ala
 340 345 350
 Cys Cys His Arg Leu Arg Pro Arg His Ser Ser His Ser Leu Ser Arg
 355 360 365
 Met Thr Thr Gly Ser Thr Leu Cys Asp Val Gly Ser Leu Gly Ser Trp
 370 375 380
 Val His Pro Leu Ala Gly Asn Asp Gly Pro Glu Ala Gln Gln Glu Thr
 385 390 395 400
 Asp Pro Ser

10

20

(226) INFORMATION FOR SEQ ID NO:225:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1098 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:225:

ATGGGGAACA TCACTGCAGA CAACTCCTCG ATGAGCTGTA CCATCGACCA TACCATCCAC 60
 CAGACGCTGG CCCCGTGGT CTATGTTACC GTGCTGGTGG TGGGCTTCCC GGCCAACTGC 120
 CTGTCCCTCT ACTTCGGCTA CCTGCAGATC AAGGCCCGGA ACGAGCTGGG CGTGTACCTG 180
 TGCAACCTGA CGGTGGCCGA CCTCTTCTAC ATCTGCTCGC TGCCCTTCTG GCTGCAGTAC 240
 GTGCTGCAGC ACGACAACTG GTCTCACGGC GACCTGTCTT GCCAGGTGTG CGGCATCCTC 300
 CTGTACGAGA ACATCTACAT CAGCGTGGGC TTCCTCTGCT GCATCTCCGT GGACCGCTAC 360
 CTGGCTGTGG CCCATCCCTT CCGCTTCCAC CAGTTCCGGA CCCTGAAGGC GGCCGTCGGC 420

30

GTCAGCGTGG TCATCTGGGC CAAGGAGCTG CTGACCAGCA TCTACTTCCT GATGCACGAG 480
 GAGGTCATCG AGGACGAGAA CCAGCACCGC GTGTGCTTTG AGCACTACCC CATCCAGGCA 540
 TGGCAGCGCG CCATCAACTA CTACCGCTTC CTGGTGGGCT TCCTCTTCCC CATCTGCCTG 600
 CTGCTGGCGT CCTACCAGGG CATCCTGCGC GCCGTGCGCC GGAGCCACGG CACCCAGAAG 660
 AGCCGCAAGG ACCAGATCAA GCGGCTGGTG CTCAGCACCG TGGTCATCTT CCTGGCCTGC 720
 TTCTTGCCCT ACCACGTGTT GCTGCTGGTG CGCAGCGTCT GGGAGGCCAG CTGCGACTTC 780
 GCCAAGGGCG TTTTCAACGC CTACCACTTC TCCCTCCTGC TCACCAGCTT CAACTGCGTC 840
 GCCGACCCCG TGCTCTACTG CTTCGTCAGC GAGACCACCC ACCGGGACCT GGCCCGCCTC 900
 CGCGGGGCCT GCCTGGCCTT CCTCACCTGC TCCAGGACCG GCCGGGCCAG GGAGGCCTAC 960
 CCGCTGGGTG CCCCCGAGGC CTCCGGGAAA AGCGGGGCC AGGGTGAGGA GCCCGAGCTG 1020
 TTGACCAAGC TCCACCCGGC CTTCAGACC CCTAACTCGC CAGGGTCGGG CGGGTTCCCC 1080
 ACGGGCAGGT TGGCCTAG 1098

(227) INFORMATION FOR SEQ ID NO:226:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 365 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:226:

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

10

20

30

100	105	110
Cys Cys Ile Ser Val Asp Arg Tyr Leu Ala Val Ala His Pro Phe Arg		
115	120	125
Phe His Gln Phe Arg Thr Leu Lys Ala Ala Val Gly Val Ser Val Val		
130	135	140
Ile Trp Ala Lys Glu Leu Leu Thr Ser Ile Tyr Phe Leu Met His Glu		
145	150	155
Glu Val Ile Glu Asp Glu Asn Gln His Arg Val Cys Phe Glu His Tyr		
165	170	175
Pro Ile Gln Ala Trp Gln Arg Ala Ile Asn Tyr Tyr Arg Phe Leu Val		
180	185	190
Gly Phe Leu Phe Pro Ile Cys Leu Leu Leu Ala Ser Tyr Gln Gly Ile		
195	200	205
Leu Arg Ala Val Arg Arg Ser His Gly Thr Gln Lys Ser Arg Lys Asp		
210	215	220
Gln Ile Lys Arg Leu Val Leu Ser Thr Val Val Ile Phe Leu Ala Cys		
225	230	235
Phe Leu Pro Tyr His Val Leu Leu Leu Val Arg Ser Val Trp Glu Ala		
245	250	255
Ser Cys Asp Phe Ala Lys Gly Val Phe Asn Ala Tyr His Phe Ser Leu		
260	265	270
Leu Leu Thr Ser Phe Asn Cys Val Ala Asp Pro Val Leu Tyr Cys Phe		
275	280	285
Val Ser Glu Thr Thr His Arg Asp Leu Ala Arg Leu Arg Gly Ala Cys		
290	295	300
Leu Ala Phe Leu Thr Cys Ser Arg Thr Gly Arg Ala Arg Glu Ala Tyr		
305	310	315
Pro Leu Gly Ala Pro Glu Ala Ser Gly Lys Ser Gly Ala Gln Gly Glu		
325	330	335
Glu Pro Glu Leu Leu Thr Lys Leu His Pro Ala Phe Gln Thr Pro Asn		
340	345	350
Ser Pro Gly Ser Gly Gly Phe Pro Thr Gly Arg Leu Ala		
355	360	365

10

20

30

(228) INFORMATION FOR SEQ ID NO:227:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1416 base pairs
 (B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:227:

```

ATGGATATTC TTTGTGAAGA AAATACTTCT TTGAGCTCAA CTACGAACTC CCTAATGCAA    60
TTAAATGATG ACAACAGGCT CTACAGTAAT GACTTTAACT CCGGAGAAGC TAACACTTCT    120
GATGCATTTA ACTGGACAGT CGACTCTGAA AATCGAACCA ACCTTTCCTG TGAAGGGTGC    180
CTCTCACCGT CGTGTCTCTC CTTACTTCAT CTCCAGGAAA AAAACTGGTC TGCTTTACTG    240
ACAGCCGTAG TGATTATTCT AACTATTGCT GGAAACATAC TCGTCATCAT GGCAGTGTCC    300
CTAGAGAAAA AGCTGCAGAA TGCCACCAAC TATTTCTGTA TGTCACCTGC CATAGCTGAT    360
ATGCTGCTGG GTTTCCTTGT CATGCCCGTG TCCATGTIAA CCATCCTGTA TGGGTACCGG    420
TGGCCTCTGC CGAGCAAGCT TTGTGCAGTC TGGATTACC TGGACGTGCT CTTCTCCACG    480
GCCTCCATCA TGCACCTCTG CGCCATCTCG CTGGACCGCT ACGTCGCCAT CCAGAATCCC    540
ATCCACCACA GCCGCTTCAA CTCCAGAACT AAGGCATTTT TGAAAATCAT TGCTGTTTGG    600
ACCATATCAG TAGGTATATC CATGCCAATA CCAGTCTTTG GGCTACAGGA CGATTGGAAG    660
GTCTTTAAGG AGGGGAGTTG CTTACTCGCC GATGATAACT TTGTCTGAT CGGCTCTTTT    720
GTGTCAATTT TCATTCCCTT AACCATCATG GTGATCACCT ACTTTCTAAC TATCAAGTCA    780
CTCCAGAAAG AAGCTACTTT GTGTGTAAGT GATCTTGGCA CACGGGCCAA ATTAGCTTCT    840
TTCAGCTTCC TCCCTCAGAG TTCTTTGTCT TCAGAAAAGC TCTTCCAGCG GTCGATCCAT    900
AGGGAGCCAG GGTCTTACAC AGGCAGGAGG ACTATGCAGT CCATCAGCAA TGAGCAAAAG    960
GCAAAGAAGG TGCTGGGCAT CGTCTTCTTC CTGTTTGTGG TGATGTGGTG CCCTTTCTTC   1020
ATCACAAACA TCATGGCCGT CATCTGCAA GAGTCCTGCA ATGAGGATGT CATTGGGGCC   1080
CTGCTCAATG TGTTTGTTTG GATCGGTTAT CTCTCTTCAG CAGTCAACCC ACTAGTCTAC   1140
ACACTGTTCA ACAAGACCTA TAGGTACAGC TTTTCACGGT ATATTCAGTG TCAGTACAAG   1200
GAAAACAAAA AACCATTGCA GTTAATTTTA GTGAACACAA TACCGGCTTT GGCCTACAAG   1260
TCTAGCCAAC TTCAAATGGG ACAAAAAAAG AATTCAAAGC AAGATGCCAA GACAACAGAT   1320
AATGACTGCT CAATGGTTGC TCTAGGAAAG CAGTATTCTG AAGAGGCTTC TAAAGACAAT   1380
AGCGACGGAG TGAATGAAAA GGTGAGCTGT GTGTGA                                1416

```

10

20

30

(229) INFORMATION FOR SEQ ID NO:228:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 470 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:228:

1	Met	Asp	Ile	Leu	Cys	Glu	Glu	Asn	Thr	Ser	Leu	Ser	Ser	Thr	Thr	Asn	
			5					10					15				
	Ser	Leu	Met	Gln	Leu	Asn	Asp	Asp	Asn	Arg	Leu	Tyr	Ser	Asn	Asp	Phe	10
		20					25					30					
	Asn	Ser	Gly	Glu	Ala	Asn	Thr	Ser	Asp	Ala	Phe	Asn	Trp	Thr	Val	Asp	
		35					40					45					
	Ser	Glu	Asn	Arg	Thr	Asn	Leu	Ser	Cys	Glu	Gly	Cys	Leu	Ser	Pro	Ser	
		50				55					60						
	Cys	Leu	Ser	Leu	Leu	His	Leu	Gln	Glu	Lys	Asn	Trp	Ser	Ala	Leu	Leu	
	65				70					75					80		
	Thr	Ala	Val	Val	Ile	Ile	Leu	Thr	Ile	Ala	Gly	Asn	Ile	Leu	Val	Ile	
				85				90					95				
	Met	Ala	Val	Ser	Leu	Glu	Lys	Lys	Leu	Gln	Asn	Ala	Thr	Asn	Tyr	Phe	20
		100						105					110				
	Leu	Met	Ser	Leu	Ala	Ile	Ala	Asp	Met	Leu	Leu	Gly	Phe	Leu	Val	Met	
		115					120					125					
	Pro	Val	Ser	Met	Leu	Thr	Ile	Leu	Tyr	Gly	Tyr	Arg	Trp	Pro	Leu	Pro	
		130				135						140					
	Ser	Lys	Leu	Cys	Ala	Val	Trp	Ile	Tyr	Leu	Asp	Val	Leu	Phe	Ser	Thr	
	145				150					155					160		
	Ala	Ser	Ile	Met	His	Leu	Cys	Ala	Ile	Ser	Leu	Asp	Arg	Tyr	Val	Ala	
				165					170					175			
	Ile	Gln	Asn	Pro	Ile	His	His	Ser	Arg	Phe	Asn	Ser	Arg	Thr	Lys	Ala	30
		180						185						190			
	Phe	Leu	Lys	Ile	Ile	Ala	Val	Trp	Thr	Ile	Ser	Val	Gly	Ile	Ser	Met	
		195				200						205					
	Pro	Ile	Pro	Val	Phe	Gly	Leu	Gln	Asp	Asp	Ser	Lys	Val	Phe	Lys	Glu	
		210				215					220						
	Gly	Ser	Cys	Leu	Leu	Ala	Asp	Asp	Asn	Phe	Val	Leu	Ile	Gly	Ser	Phe	
	225				230					235				240			

Val Ser Phe Phe Ile Pro Leu Thr Ile Met Val Ile Thr Tyr Phe Leu
 245 250 255
 Thr Ile Lys Ser Leu Gln Lys Glu Ala Thr Leu Cys Val Ser Asp Leu
 260 265 270
 Gly Thr Arg Ala Lys Leu Ala Ser Phe Ser Phe Leu Pro Gln Ser Ser
 275 280 285
 Leu Ser Ser Glu Lys Leu Phe Gln Arg Ser Ile His Arg Glu Pro Gly
 290 295 300
 Ser Tyr Thr Gly Arg Arg Thr Met Gln Ser Ile Ser Asn Glu Gln Lys
 305 310 315 320
 Ala Lys Lys Val Leu Gly Ile Val Phe Phe Leu Phe Val Val Met Trp
 325 330 335
 Cys Pro Phe Phe Ile Thr Asn Ile Met Ala Val Ile Cys Lys Glu Ser
 340 345 350
 Cys Asn Glu Asp Val Ile Gly Ala Leu Leu Asn Val Phe Val Trp Ile
 355 360 365
 Gly Tyr Leu Ser Ser Ala Val Asn Pro Leu Val Tyr Thr Leu Phe Asn
 370 375 380
 Lys Thr Tyr Arg Ser Ala Phe Ser Arg Tyr Ile Gln Cys Gln Tyr Lys
 385 390 395 400
 Glu Asn Lys Lys Pro Leu Gln Leu Ile Leu Val Asn Thr Ile Pro Ala
 405 410 415
 Leu Ala Tyr Lys Ser Ser Gln Leu Gln Met Gly Gln Lys Lys Asn Ser
 420 425 430
 Lys Gln Asp Ala Lys Thr Thr Asp Asn Asp Cys Ser Met Val Ala Leu
 435 440 445
 Gly Lys Gln Tyr Ser Glu Glu Ala Ser Lys Asp Asn Ser Asp Gly Val
 450 455 460
 Asn Glu Lys Val Ser Cys Val
 465 470

10

20

30

(230) INFORMATION FOR SEQ ID NO:229:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1377 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:229:

ATGGTGAACC	TGAGGAATGC	GGTGCATTCA	TTCCTTGTGC	ACCTAATTGG	CCTATTGGTT	60
TGGCAATGTG	ATATTTCTGT	GAGCCCAGTA	GCAGCTATAG	TAAGTGACAT	TTTCAATACC	120
TCCGATGGTG	GACGCTTCAA	ATTCCCAGAC	GGGGTACAAA	ACTGGCCAGC	ACTTTCAATC	180
GTCATCATAA	TAATCATGAC	AATAGGTGGC	AACATCCTTG	TGATCATGGC	AGTAAGCATG	240
GAAAAGAAAC	TGCACAATGC	CACCAATTAC	TTCTTAATGT	CCCTAGCCAT	TGCTGATATG	300
CTAGTGGGAC	TACTTGTCT	GCCCCTGTCT	CTCCTGGCAA	TCCTTTATGA	TTATGTCTGG	360
CCACTACCTA	GATATTTGTG	CCCCGTCTGG	ATTTCCTTAG	ATGTTTTATT	TTCAACAGCG	420
TCCATCATGC	ACCTCTGCGC	TATATCGCTG	GATCGGTATG	TAGCAATACG	TAATCCTATT	480
GAGCATAGCC	GTTTCAATTC	GCGGACTAAG	GCCATCATGA	AGATTGCTAT	TGTTTGGGCA	540
ATTTCTATAG	GTGTATCAGT	TCCTATCCCT	GTGATTGGAC	TGAGGGACGA	AGAAAAGGTG	600
TTCGTGAACA	ACACGACGTG	CGTGCTCAAC	GACCCAAATT	TCGTTCTTAT	TGGGTCTTTC	660
GTAGCTTTCT	TCATACCGCT	GACGATTATG	GTGATTACGT	ATTGCCTGAC	CATCTACGTT	720
CTGCGCCGAC	AAGCTTTGAT	GTTACTGCAC	GGCCACACCG	AGGAACCGCC	TGACTAAGT	780
CTGGATTTCC	TGAAGTGCTG	CAAGAGGAAT	ACGGCCGAGG	AAGAGAACTC	TGCAAACCTT	840
AACCAAGACC	AGAACGCACG	CCGAAGAAAG	AAGAAGGAGA	GACGTCCTAG	GGGCACCATG	900
CAGGCTATCA	ACAATGAAAG	AAAAGCTAAG	AAAGTCCTTG	GGATTGTTTT	CTTTGTGTTT	960
CTGATCATGT	GGTGCCCAT	TTTCATTACC	AATATTCTGT	CTGTTCTTTG	TGAGAAGTCC	1020
TGTAACCAAA	AGCTCATGGA	AAAGCTTCTG	AATGTGTTTG	TTTGGATTGG	CTATGTTTGT	1080
TCAGGAATCA	ATCCTCTGGT	GTATACTCTG	TTCAACAAAA	TTTACCGAAG	GGCATTCTCC	1140
AACTATTTGC	GTTGCAATTA	TAAGGTAGAG	AAAAAGCCTC	CTGTCAGGCA	GATTCCAAGA	1200
GTTGCCGCCA	CTGCTTTGTC	TGGGAGGGAG	CTTAATGTTA	ACATTTATCG	GCATACCAAT	1260
GAACCGGTGA	TCGAGAAAGC	CAGTGACAAT	GAGCCCGGTA	TAGAGATGCA	AGTTGAGAAT	1320
TTAGAGTTAC	CAGTAAATCC	CTCCAGTGTG	GTTAGCGAAA	GSATTAGCAG	TGTGTGA	1377

(231) INFORMATION FOR SEQ ID NO:230:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 458 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:

10

20

30

(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:230:

```

Met Val Asn Leu Arg Asn Ala Val His Ser Phe Leu Val His Leu Ile
1           5           10           15

Gly Leu Leu Val Trp Gln Cys Asp Ile Ser Val Ser Pro Val Ala Ala
          20           25           30

Ile Val Thr Asp Ile Phe Asn Thr Ser Asp Gly Gly Arg Phe Lys Phe
          35           40           45

Pro Asp Gly Val Gln Asn Trp Pro Ala Leu Ser Ile Val Ile Ile Ile
          50           55           60

Ile Met Thr Ile Gly Gly Asn Ile Leu Val Ile Met Ala Val Ser Met
65           70           75           80

Glu Lys Lys Leu His Asn Ala Thr Asn Tyr Phe Leu Met Ser Leu Ala
          85           90           95

Ile Ala Asp Met Leu Val Gly Leu Leu Val Met Pro Leu Ser Leu Leu
          100          105          110

Ala Ile Leu Tyr Asp Tyr Val Trp Pro Leu Pro Arg Tyr Leu Cys Pro
          115          120          125

Val Trp Ile Ser Leu Asp Val Leu Phe Ser Thr Ala Ser Ile Met His
          130          135          140

Leu Cys Ala Ile Ser Leu Asp Arg Tyr Val Ala Ile Arg Asn Pro Ile
145          150          155          160

Glu His Ser Arg Phe Asn Ser Arg Thr Lys Ala Ile Met Lys Ile Ala
          165          170          175

Ile Val Trp Ala Ile Ser Ile Gly Val Ser Val Pro Ile Pro Val Ile
          180          185          190

Gly Leu Arg Asp Glu Glu Lys Val Phe Val Asn Asn Thr Thr Cys Val
          195          200          205

Leu Asn Asp Pro Asn Phe Val Leu Ile Gly Ser Phe Val Ala Phe Phe
          210          215          220

Ile Pro Leu Thr Ile Met Val Ile Thr Tyr Cys Leu Thr Ile Tyr Val
225          230          235          240

Leu Arg Arg Gln Ala Leu Met Leu Leu His Gly His Thr Glu Glu Pro
          245          250          255

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

```

10

20

30

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

10

20

(232) INFORMATION FOR SEQ ID NO:231:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1068 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:231:

ATGGATCAGT TCCCTGAATC AGTGACAGAA AACTTTGAGT ACGATGATTT GGCTGAGGCC	60
TGTTATATTG GGGACATCGT GGTCTTTGGG ACTGTGTTCC TGTCCATATT CTACTCCGTC	120
ATCTTTGCCA TTGGCCTGGT GGGAAATTG TTGGTAGTGT TTGCCCTCAC CAACAGCAAG	180

AAGCCCAAGA GTGTCACCGA CATTACCTC CTGAACCTGG CCTTGTCTGA TCTGCTGTTT 240
 GTAGCCACTT TGCCCTTCTG GACTCACTAT TTGATAAATG AAAAGGGCCT CCACAATGCC 300
 ATGTGCAAAT TCACTACCGC CTCTTCTTC ATCGGCTTTT TTGAAGCAT ATTCTTCATC 360
 ACCGTCACTA GCATTGATAG GTACCTGGCC ATCGTCCTGG CCGCCAACTC CATGAACAAC 420
 CGGACCGTGC AGCATGGCGT CACCATCAGC CTAGGCGTCT GGGCAGCAGC CATTTTGGTG 480
 GCAGCACCCC AGTTCATGTT CACAAAGCAG AAAGAAAATG AATGCCTTGG TGACTACCCC 540
 GAGGTCTCTC AGGAAATCTG GCCCGTGCTC CGCAATGTGG AAACAAATTT TCTTGGCTTC 600
 CTACTCCCCC TGCTCATTAT GAGTTATTGC TACTTCAGAA TCATCCAGAC GCTGTTTTTC 660
 TGCAAGAACC ACAAGAAAGC CAAAGCCAAG AAACGTATCC TTCTGGTGGT CATCGTGTTT 720
 TTCCTCTTCT GGACACCCTA CAACGTTATG ATTTCTCTGG AGACGCTTAA GCTCTATGAC 780
 TTCTTTCCCA GTTGTGACAT GAGGAAGGAT CTGAGGCTGG CCCTCAGTGT GACTGAGACG 840
 GTTGCAATTA GCCATTGTTG CCTGAATCCT CTCATCTATG CATTGCTGG GGAGAAGTTC 900
 AGAAGATACC TTTACCACCT GTATGGGAAA TGCTGGCTG TCCTGTGTGG GCGCTCAGTC 960
 CACGTTGATT TCTCTCATC TGAATCACA AGGAGCAGGC ATGGAAGTGT TCTGAGCAGC 1020
 AATTTTACTT ACCACACGAG TGATGGAGAT GCATTGCTCC TTCTCTGA 1068

10

20

(233) INFORMATION FOR SEQ ID NO:232:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 355 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:232:

Met Asp Gln Phe Pro Glu Ser Val Thr Glu Asn Phe Glu Tyr Asp Asp
 1 5 10 15
 Leu Ala Glu Ala Cys Tyr Ile Gly Asp Ile Val Val Phe Gly Thr Val
 20 25 30
 Phe Leu Ser Ile Phe Tyr Ser Val Ile Phe Ala Ile Gly Leu Val Gly
 35 40 45
 Asn Leu Leu Val Val Phe Ala Leu Thr Asn Ser Lys Lys Pro Lys Ser
 50 55 60
 Val Thr Asp Ile Tyr Leu Leu Asn Leu Ala Leu Ser Asp Leu Leu Phe

30

[illegible]

(234) INFORMATION FOR SEQ ID NO:233:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 29 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:233:

GGCTTAAGAG CATCATCGTG GTGCTGGTG

29

10

(235) INFORMATION FOR SEQ ID NO:234:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:234:

GTCACCACCA GCACCACGAT GATGCTCTTA AGCC

34

20

(236) INFORMATION FOR SEQ ID NO:235:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 31 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:235:

CAAAGAAAGT ACTGGGCATC GTCTTCTTCC T

31

30

(237) INFORMATION FOR SEQ ID NO:236:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:236:

TGCTCTAGAT TCCAGATAGG TGAAAACTTG 30

(238) INFORMATION FOR SEQ ID NO:237:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: NO 10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:237:

CTAGGGGCAC CATGCAGGCT ATCAACAATG AAAGAAAAGC TAAGAAAGTC 50

(239) INFORMATION FOR SEQ ID NO:238:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic) 20

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:238:

CAAGGACTTT CTTAGCTTTT CTTTCATTGT TGATAGCCTG CATGGTGCCC 50

(240) INFORMATION FOR SEQ ID NO:239:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear 30

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:239:

CGGCGGCAGA AGGCGAAACG CATGATCCTC GCGGT 35

(241) INFORMATION FOR SEQ ID NO:240:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:240:

ACCGCGAGGA TCATGCGTTT CGCCTTCTGC CGCCG 35

(242) INFORMATION FOR SEQ ID NO:241:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid 10
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:241:

GAGACATATT ATCTGCCACG GAGG 24

(243) INFORMATION FOR SEQ ID NO:242:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid 20
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:242:

TTGGCATAGA AACCGGACCC AAGG 24

(244) INFORMATION FOR SEQ ID NO:243:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid 30
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:243:

TAAGAATTCC ATAAAAATTG TGAATGG 28

(245) INFORMATION FOR SEQ ID NO:244:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:244:

CCAGGATCCA GCTGAAGTCT TCCATCATTC

30

(246) INFORMATION FOR SEQ ID NO:245:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1071 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:245:

ATGAATGGGG TCTCGGAGGG GACCAGAGGC TGCAGTGACA GGCAACCTGG GGTCTTGACA 60
 CGTGATCGCT CTTGTTCCAG GAAGATGAAC TCTTCCGGAT GCCTGTCTGA GGAGGTGGGG 120
 TCCCTCCGCC CACTGACTGT GGTATCCTG TCTGCGTCCA TTGTCGTCGG AGTGCTGGGC 180
 AATGGGCTGG TGCTGTGGAT GACTGTCTTC CGTATGGCAC GCACGGTCTC CACCGTCTGC 240
 TTCTTCCACC TGGCCCTTGC CGATTTTCATG CTCTCACTGT CTCTGCCCCAT TGCCATGTAC 300
 TATATTGTCT CCAGGCAGTG GCTCCTCGGA GAGTGGGCCT GCAAACCTCTA CATCACCTTT 360
 GTGTTCTCA GCTACTTTGC CAGTAACTGC CTCCTTGTCT TCATCTCTGT GGACCGTTGC 420
 ATCTCTGTCC TCTACCCCGT CTGGGCCCTG AACCACCGCA CTGTGCAGCG GCGGAGCTGG 480
 CTGGCCTTTG GGGTGTGGCT CCTGGCCGCC GCCTTGTGCT CTGCGCACCT GAAATTCCGG 540
 ACAACCAGAA AATGGAATGG CTGTACGCAC TGCTACTTGG CGTTCAACTC TGACAATGAG 600
 ACTGCCCAGA TTTGGATTGA AGGGGTCGTG GAGGGACACA TTATAGGGAC CATTGGCCAC 660
 TTCCTGCTGG GCTTCTCTGGG GCCCTTAGCA ATCATAGGCA CCTGCGCCCA CCTCATCCGG 720
 GCCAAGCTCT TCGGGGAGGG CTGGGTCCAT GCCAACCAGC CCGCGAGGCT GCTGCTGGTG 780
 CTGGTGAGCG CTTTCTTTAT CTTCTGGTCC CCGTTTAACG TGGTGTCTGT GGTCCATCTG 840
 TGGCGACGGG TGATGCTCAA GGAAATCTAC CACCCCGGA TGCTGCTCAT CCTCCAGGCT 900
 AGCTTTGCCT TGGGCTGTGT CAACAGCAGC CTCACCCCT TCCTCTACGT CTCGTTGGC 960

20

30

AGAGATTTCC AAGAAAAGTT TTTCCAGTCT TTGACTTCTG CCCTGGCGAG GCGGTTTGGA 1020
 GAGGAGGAGT TTCTGTCATC CTGTCCCCGT GGCAACGCCC CCCGGGAATG A 1071

(247) INFORMATION FOR SEQ ID NO:246:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 356 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:246:

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

10

20

30

Val Val Glu Gly His Ile Ile Gly Thr Ile Gly His Phe Leu Leu Gly
 210 215 220

Phe Leu Gly Pro Leu Ala Ile Ile Gly Thr Cys Ala His Leu Ile Arg
 225 230 235 240

Ala Lys Leu Leu Arg Glu Gly Trp Val His Ala Asn Arg Pro Ala Arg
 245 250 255

Leu Leu Leu Val Leu Val Ser Ala Phe Phe Ile Phe Trp Ser Pro Phe
 260 265 270

Asn Val Val Leu Leu Val His Leu Trp Arg Arg Val Met Leu Lys Glu
 275 280 285

Ile Tyr His Pro Arg Met Leu Leu Ile Leu Gln Ala Ser Phe Ala Leu
 290 295 300

Gly Cys Val Asn Ser Ser Leu Asn Pro Phe Leu Tyr Val Phe Val Gly
 305 310 315 320

Arg Asp Phe Gln Glu Lys Phe Phe Gln Ser Leu Thr Ser Ala Leu Ala
 325 330 335

Arg Ala Phe Gly Glu Glu Glu Phe Leu Ser Ser Cys Pro Arg Gly Asn
 340 345 350

Ala Pro Arg Glu
 355

10

20

(248) INFORMATION FOR SEQ ID NO:247:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 32 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:247:

GCAGAATTCG GCGGCCCCAT GGACCTGCCC CC

32

30

(249) INFORMATION FOR SEQ ID NO:248:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:248:

GCTGGATCCC CCGAGCAGTG GCGTTACTTC

30

(250) INFORMATION FOR SEQ ID NO:249:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 903 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:249:

ATGGACCTGC CCCCGCAGCT CTCCTTCGGC CTCTATGTGG CCGCCTTTGC GCTGGGCTTC 60
 CCGCTCAACG TCCTGGCCAT CCGAGGCGCG ACGGCCACG CCCGGCTCCG TCTCACCCCT 120
 AGCCTGGTCT ACGCCCTGAA CCTGGGCTGC TCCGACCTGC TGCTGACAGT CTCTCTGCCC 180
 CTGAAGGCGG TGGAGGCGCT AGCCTCCGGG GCCTGGCCTC TGCCGGCCTC GCTGTGCCCC 240
 GTCTTCGCGG TGGCCCACTT CTTCCCACTC TATGCCGGCG GGGGCTTCCT GGCCGCCCTG 300
 AGTGCAGGCC GCTACCTGGG AGCAGCCTTC CCCTTGGGCT ACCAAGCCTT CCGGAGGCCG 360
 TGCTATTCTT GGGGGGTGTG CGCGGCCATC TGGGCCCTCG TCCTGTGTCA CCTGGGTCTG 420
 GTCTTTGGGT TGGAGGCTCC AGGAGGCTGG CTGGACCACA GCAACACCTC CCTGGGCATC 480
 AACACACCGG TCAACGGCTC TCCGGTCTGC CTGGAGGCCT GGGACCCGGC CTCTGCCGGC 540
 CCGGCCCCGT TCAGCCTCTC TCTCCTGCTC TTTTTCGTG CCTTGGCCAT CACAGCCTTC 600
 TGCTACGTGG GCTGCCTCCG GGCCTGGGCC CGCTCCGGCC TGACGCACAG GCGGAAGCTG 660
 CGGGCCGCCT GGGTGGCCGG CGGGGCCCTC CTCACGCTGC TGCTCTGCGT AGGACCCTAC 720
 AACGCCTCCA ACGTGGCCAG CTCCTGTAC CCCAATCTAG GAGGCTCCTG GCGGAAGCTG 780
 GGGCTCATCA CGGGTGCCCTG GAGTGTGGTG CTTAATCCGC TGGTGACCGG TTAATTGGGA 840
 AGGGGTCTTG GCCTGAAGAC AGTGTGTGCG GCAAGAACGC AAGGGGGCAA GTCCCAGAAG 900
 TAA 903

10

20

30

(251) INFORMATION FOR SEQ ID NO:250:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 300 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:250:

```

Met Asp Leu Pro Pro Gln Leu Ser Phe Gly Leu Tyr Val Ala Ala Phe
1           5           10           15

Ala Leu Gly Phe Pro Leu Asn Val Leu Ala Ile Arg Gly Ala Thr Ala
20           25           30

His Ala Arg Leu Arg Leu Thr Pro Ser Leu Val Tyr Ala Leu Asn Leu
35           40           45

Gly Cys Ser Asp Leu Leu Leu Thr Val Ser Leu Pro Leu Lys Ala Val
50           55           60

Glu Ala Leu Ala Ser Gly Ala Trp Pro Leu Pro Ala Ser Leu Cys Pro
65           70           75           80

Val Phe Ala Val Ala His Phe Phe Pro Leu Tyr Ala Gly Gly Gly Phe
85           90           95

Leu Ala Ala Leu Ser Ala Gly Arg Tyr Leu Gly Ala Ala Phe Pro Leu
100          105          110

Gly Tyr Gln Ala Phe Arg Arg Pro Cys Tyr Ser Trp Gly Val Cys Ala
115          120          125

Ala Ile Trp Ala Leu Val Leu Cys His Leu Gly Leu Val Phe Gly Leu
130          135          140

Glu Ala Pro Gly Gly Trp Leu Asp His Ser Asn Thr Ser Leu Gly Ile
145          150          155          160

Asn Thr Pro Val Asn Gly Ser Pro Val Cys Leu Glu Ala Trp Asp Pro
165          170          175

Ala Ser Ala Gly Pro Ala Arg Phe Ser Leu Ser Leu Leu Leu Phe Phe
180          185          190

Leu Pro Leu Ala Ile Thr Ala Phe Cys Tyr Val Gly Cys Leu Arg Ala
195          200          205

Leu Ala Arg Ser Gly Leu Thr His Arg Arg Lys Leu Arg Ala Ala Trp
210          215          220

Val Ala Gly Gly Ala Leu Leu Thr Leu Leu Leu Cys Val Gly Pro Tyr
225          230          235          240

Asn Ala Ser Asn Val Ala Ser Phe Leu Tyr Pro Asn Leu Gly Gly Ser
245          250          255

Trp Arg Lys Leu Gly Leu Ile Thr Gly Ala Trp Ser Val Val Leu Asn
260          265          270

Pro Leu Val Thr Gly Tyr Leu Gly Arg Gly Pro Gly Leu Lys Thr Val

```

10

20

30

275

280

285

Cys Ala Ala Arg Thr Gln Gly Gly Lys Ser Gln Lys
 290 295 300

(252) INFORMATION FOR SEQ ID NO:251:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:251:

CTCAAGCTTA CTCTCTCTCA CCAGTGGCCA C

31

(253) INFORMATION FOR SEQ ID NO:252:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:252:

CCCTCCTCCC CCGGAGGACC TAGC

24

(254) INFORMATION FOR SEQ ID NO:253:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1041 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:253:

30

ATGGATACAG GCCCCGACCA GTCCTACTTC TCCGGCAATC ACTGGTTCGT CTTCTCGGTG 60
 TACCTTCTCA CTTTCCTGGT GGGGCTCCCC CTCAACCTGC TGGCCCTGGT GGTCTTCGTG 120
 GGCAAGCTGC AGCGCCGCCC GGTGGCCGTG GACGTGCTCC TGCTCAACCT GACCGCCTCG 180
 GACCTGCTCC TGCTGCTGTT CCTGCCTTTC CGCATGGTGG AGGCAGCCAA TGGCATGCAC 240
 TGGCCCCTGC CCTTCATCCT CTGCCCCACTC TCTGGATTCA TCTTCTTCAC CACCATCTAT 300

CTCACCGCCC TCTTCCTGGC AGCTGTGAGC ATTGAACGCT TCCTGAGTGT GGCCACCCCA 360
 CTGTGGTACA AGACCCGGCC GAGGCTGGGG CAGGCAGGTC TGGTGAGTGT GGCCTGCTGG 420
 CTGTTGGCCT CTGCTCACTG CAGCGTGGTC TACGTCATAG AATTCTCAGG GGACATCTCC 480
 CACAGCCAGG GCACCAATGG GACCTGCTAC CTGGAGTTCC GGAAGGACCA GCTAGCCATC 540
 CTCCTGCCCCG TCGGGCTGGA GATGGCTGTG GTCCTCTTTG TGGTCCCGCT GATCATCACC 600
 AGCTACTGCT ACAGCCGCCT GGTGTGGATC CTCGGCAGAG GGGGCAGCCA CCGCCGGCAG 660
 AGGAGGGTGG CGGGGCTGTT GGCGGCCACG CTGCTCAACT TCCTTGCTCTG CTTTGGGCCC 720
 TACAACGTGT CCCATGTCGT GGGCTATATC TCGGGTGAAA GCCCGGCATG GAGGATCTAC 780
 GTGACGCTTC TCAGCACCTT GAACTCCTGT GTCGACCCCT TTGTCTACTA CTTCTCCTCC 840
 TCCGGGTTCC AAGCCGACTT TCATGAGCTG CTGAGGAGGT TGTGTGGGCT CTGGGGCCAG 900
 TGGCAGCAGG AGAGCAGCAT GGAGCTGAAG GAGCAGAAGG GAGGGGAGGA GCAGAGAGCG 960
 GACCGACCAG CTGAAAGAAA GACCAGTGAA CACTCACAGG GCTGTGGAAC TGGTGGCCAG 1020
 GTGGCCTGTG CTGAAAGCTA G 1041

(255) INFORMATION FOR SEQ ID NO:254:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 346 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:254:

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

Thr Thr Ile Tyr Leu Thr Ala Leu Phe Leu Ala Ala Val Ser Ile Glu
 100 105 110
 Arg Phe Leu Ser Val Ala His Pro Leu Trp Tyr Lys Thr Arg Pro Arg
 115 120 125
 Leu Gly Gln Ala Gly Leu Val Ser Val Ala Cys Trp Leu Leu Ala Ser
 130 135 140
 Ala His Cys Ser Val Val Tyr Val Ile Glu Phe Ser Gly Asp Ile Ser
 145 150 155 160
 His Ser Gln Gly Thr Asn Gly Thr Cys Tyr Leu Glu Phe Arg Lys Asp
 165 170 175
 Gln Leu Ala Ile Leu Leu Pro Val Arg Leu Glu Met Ala Val Val Leu
 180 185 190
 Phe Val Val Pro Leu Ile Ile Thr Ser Tyr Cys Tyr Ser Arg Leu Val
 195 200 205
 Trp Ile Leu Gly Arg Gly Gly Ser His Arg Arg Gln Arg Arg Val Ala
 210 215 220
 Gly Leu Leu Ala Ala Thr Leu Leu Asn Phe Leu Val Cys Phe Gly Pro
 225 230 235 240
 Tyr Asn Val Ser His Val Val Gly Tyr Ile Cys Gly Glu Ser Pro Ala
 245 250 255
 Trp Arg Ile Tyr Val Thr Leu Leu Ser Thr Leu Asn Ser Cys Val Asp
 260 265 270
 Pro Phe Val Tyr Tyr Phe Ser Ser Ser Gly Phe Gln Ala Asp Phe His
 275 280 285
 Glu Leu Leu Arg Arg Leu Cys Gly Leu Trp Gly Gln Trp Gln Gln Glu
 290 295 300
 Ser Ser Met Glu Leu Lys Glu Gln Lys Gly Gly Glu Glu Gln Arg Ala
 305 310 315 320
 Asp Arg Pro Ala Glu Arg Lys Thr Ser Glu His Ser Gln Gly Cys Gly
 325 330 335
 Thr Gly Gly Gln Val Ala Cys Ala Glu Ser
 340 345

10

20

30

(256) INFORMATION FOR SEQ ID NO:255:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:255:

TTTAAGCTTC CCCTCCAGGA TGCTGCCGGA C 31

(257) INFORMATION FOR SEQ ID NO:256:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:256:

GGCGAATTCT GAAGGTCCAG GGAAACTGCT A 31

(258) INFORMATION FOR SEQ ID NO:257:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 993 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:257:

ATGCTGCCGG ACTGGAAGAG CTCCTTGATC CTCATGGCTT ACATCATCAT CTCCTCACT 60

GGCCTCCCTG CCAACCTCCT GGCCCTGCGG GCCTTTGTGG GCGGATCCG CCAGCCCCAG 120

CCTGCACCTG TGCACATCCT CCTGCTGAGC CTGACGCTGG CCGACCTCCT CCTGCTGCTG 180

CTGCTGCCCT TCAAGATCAT CGAGGCTGCG TCGAACTTCC GCTGGTACCT GCCCAAGGTC 240

GTCTGCGCCC TCACGAGTTT TGGCTTCTAC AGCAGCATCT ACTGCAGCAC GTGGCTCCTG 300

GCGGGCATCA GCATCGAGCG CTACCTGGGA GTGGCTTTCC CCGTGCAGTA CAAGCTCTCC 360

CGCCGGCCTC TGTATGGAGT GATTGCAGCT CTGGTGGCCT GGGTTATGTC CTTTGGTCAC 420

TGCACCATCG TGATCATCGT TCAATACTTG AACACGACTG AGCAGGTCAG AAGTGGCAAT 480

GAAATTACCT GCTACGAGAA CTTACCGGAT AACCAGTTGG ACGTGGTGCT GCCCGTGCGG 540

CTGGAGCTGT GCCTGGTGCT CTTCTTCATC CCCATGGCAG TCACCATCTT CTGCTACTGG 600

CGTTTTGTGT GGATCATGCT CTCCCAGCCC CTGTGGGGG CCCAGAGGCG GCGCCGAGCC 660

GTGGGGCTGG CTGTGGTGAC GCTGCTCAAT TTCCTGGTGT GCTTCGGACC TTACAACGTG 720

10

20

30

TCCCACCTGG TGGGGTATCA CCAGAGAAAA AGCCCCTGGT GGCGGTCAAT AGCCGTGGTG 780
 TTCAGTTCAC TCAACGCCAG TCTGGACCCC CTGCTCTTCT ATTTCTCTTC TTCAGTGGTG 840
 CGCAGGGCAT TTGGGAGAGG GCTGCAGGTG CTGCGGAATC AGGGCTCCTC CCTGTTGGGA 900
 CGCAGAGGCA AAGACACAGC AGAGGGGACA AATGAGGACA GGGGTGTGGG TCAAGGAGAA 960
 GGGATGCCAA GTTCGGACTT CACTACAGAG TAG 993

(259) INFORMATION FOR SEQ ID NO:258:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 362 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

10

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:258:

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

20

30

Leu Pro Val Arg Leu Glu Leu Cys Leu Val Leu Phe Phe Ile Pro Met
 180 185 190
 Ala Val Thr Ile Phe Cys Tyr Trp Arg Phe Val Trp Ile Met Leu Ser
 195 200 205
 Gln Pro Leu Val Gly Ala Gln Arg Arg Arg Arg Ala Val Gly Leu Ala
 210 215 220
 Val Val Thr Leu Leu Asn Phe Leu Val Cys Phe Gly Pro Tyr Asn Val
 225 230 235 240
 Ser His Leu Val Gly Tyr His Gln Arg Lys Ser Pro Trp Trp Arg Ser
 245 250 255
 Ile Ala Val Val Phe Ser Ser Leu Asn Ala Ser Leu Asp Pro Leu Leu
 260 265 270
 Phe Tyr Phe Ser Ser Ser Val Val Arg Arg Ala Phe Gly Arg Gly Leu
 275 280 285
 Gln Val Leu Arg Asn Gln Gly Ser Ser Leu Leu Gly Arg Arg Gly Lys
 290 295 300
 Asp Thr Ala Glu Gly Thr Asn Glu Asp Arg Gly Val Gly Gln Gly Glu
 305 310 315 320
 Gly Met Pro Ser Ser Asp Phe Thr Thr Glu
 325 330

10

20

(260) INFORMATION FOR SEQ ID NO:259:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:259:

CCCAAGCTTC GGGCACCATG GACACCTCCC

30

30

(261) INFORMATION FOR SEQ ID NO:260:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:260:

ACAGGATCCA AATGCACAGC ACTGGTAAGC

30

(262) INFORMATION FOR SEQ ID NO:261:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:261:

CTATAACTGG GTTACATGGT TTAAC

25

10

(263) INFORMATION FOR SEQ ID NO:262:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:262:

TTTGAATTCA CATATTAATT AGAGACATGG

30

20

(264) INFORMATION FOR SEQ ID NO:263:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2724 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:263:

ATGGACACCT CCCGGCTCGG TGTGCTCCTG TCCTTGCCCTG TGCTGCTGCA GCTGGCGACC 60
GGGGGCAGCT CTCCCAGGTC TGGTGTGTTG CTGAGGGGCT GCCCCACACA CTGTCATTGC 120
GAGCCCGACG GCAGGATGTT GCTCAGGGTG GACTGCTCCG ACCTGGGGCT CTCGGAGCTG 180
CCTTCCAACC TCAGCGTCTT CACCTCCTAC CTAGACCTCA GTATGAACAA CATCAGTCAG 240
CTGCTCCCGA ATCCCCTGCC CAGTCTCCGC TTCTTGAGG AGTTACGTCT TGCGGGAAAC 300
GCTCTGACAT ACATTCCCAA GGGAGCATTC ACTGGCCTTT ACAGTCTTAA AGTTCTTATG 360

30

CTGCAGAATA ATCAGCTAAG ACACGTACCC ACAGAAGCTC TGCAGAATTT GCGAAGCCTT 420
 CAATCCCTGC GTCTGGATGC TAACCACATC AGCTATGTGC CCCCAGCTG TTTCAGTGGC 480
 CTGCATTCCC TGAGGCACCT GTGGCTGGAT GACAATGCGT TAACAGAAAT CCCCGTCCAG 540
 GCTTTTAGAA GTTATCGGC ATTGCAAGCC ATGACCTTGG CCCTGAACAA AATACACCAC 600
 ATACCAGACT ATGCCCTTGG AAACCTCTCC AGCTTGGTAG TTCTACATCT CCATAACAAT 660
 AGAATCCACT CCTGGGAAA GAAATGCTTT GATGGGCTCC ACAGCCTAGA GACTTTAGAT 720
 TTAAATTACA ATAACCTTGA TGAATTCCCC ACTGCAATTA GGACACTCTC CAACCTTAAA 780
 GAACTAGGAT TTCATAGCAA CAATATCAGG TCGATACCTG AGAAAGCATT TGTAGGCAAC 840
 CCTTCTCTTA TTACAATACA TTTCTATGAC AATCCCATCC AATTTGTTGG GAGATCTGCT 900
 TTTCAACATT TACCTGAACT AAGAACACTG ACTCTGAATG GTGCCTCACA AATAACTGAA 960
 TTTCTGATT TAACTGGAAC TGCAAACCTG GAGAGTCTGA CTTTAACTGG AGCACAGATC 1020
 TCATCTCTTC CTCAAACCGT CTGCAATCAG TTACCTAATC TCCAAGTGCT AGATCTGTCT 1080
 TACAACCTAT TAGAAGATTT ACCCAGTTTT TCAGTCTGCC AAAAGCTTCA GAAAATTGAC 1140
 CTAAGACATA ATGAAATCTA CGAAATTAAA GTTGACACTT TCCAGCAGTT GCTTAGCCTC 1200
 CGATCGCTGA ATTTGGCTTG GAACAAAATT GCTATTATTC ACCCCAATGC ATTTTCCACT 1260
 TTGCCATCCC TAATAAAGCT GGACCTATCG TCCAACCTCC TGTCGTCTTT TCCTATAACT 1320
 GGGTTACATG GTTTAACTCA CTTAAAATT AAGGAAATC ATGCCTTACA GAGCTTGATA 1380
 TCATCTGAAA ACTTTCCAGA ACTCAAGGTT ATAGAAATGC CTTATGCTTA CCAGTGCTGT 1440
 GCATTGGAG TGTGTGAGAA TGCCTATAAG ATTTCTAATC AATGGAATAA AGGTGACAAC 1500
 AGCAGTATGG ACGACCTTCA TAAGAAAGAT GCTGGAATGT TTCAGGCTCA AGATGAACGT 1560
 GACCTTGAAG ATTTCTGCT TGAATTTGAG GAAGACCTGA AAGCCCTTCA TTCAGTGCAG 1620
 TGTTACCTT CCCCAGGCCC CTTCAAACCC TGTGAACACC TGCTTGATGG CTGGCTGATC 1680
 AGAATTGGAG TGTGGACCAT AGCAGTTCTG GCACCTACTT GTAATGCTTT GGTGACTTCA 1740
 ACAGTTTTCA GATCCCTCT GTACATTTC CCCATTAAAC TGTTAATTGG GGTGATCGCA 1800
 GCAGTGAACA TGCTCACGGG AGTCTCCAGT GCCGTGCTGG CTGGTGTGGA TGCCTTCACT 1860
 TTTGGCAGCT TTGCACGACA TGGTGCTGG TGGGAGAATG GGTTGGTTG CCATGTCATT 1920
 GGTTTTTTGT CCATTTTTC TTCAGAATCA TCTGTTTTCC TGCTTACTCT GGCAGCCCTG 1980
 GAGCGTGGGT TCTCTGTGAA ATATTCTGCA AAATTGAAA CGAAAGCTCC ATTTTCTAGC 2040

10

20

30

CTGAAAGTAA TCATTTTGCT CTGTGCCCTG CTGGCCTTGA CCATGGCCGC AGTTCCTCTG 2100
 CTGGGTGGCA GCAAGTATGG CGCCTCCCCT CTCTGCCTGC CTTTGCCCTT TGGGGAGCCC 2160
 AGCACCATGG GCTACATGGT CGCTCTCATC TTGCTCAATT CCCTTTGCTT CCTCATGATG 2220
 ACCATTGCCT ACACCAAGCT CTA CTGCAAT TTGGACAAGG GAGACCTGGA GAATATTTGG 2280
 GACTGCTCTA TGGTAAAACA CATTGCCCTG TTGCTCTTCA CCAACTGCAT CCTAAACTGC 2340
 CCTGTGGCTT TCTTGTCTT CTCCTCTTTA ATAAACCTTA CATTATCAG TCCTGAAGTA 2400
 ATTAAGTTTA TCCTTCTGGT GGTAGTCCCA CTCTCTGCAT GTCTCAATCC CCTTCTCTAC 2460
 ATCTTGTTCA ATCCTCACTT TAAGGAGGAT CTGGTGAGCC TGAGAAAGCA AACCTACGTC 2520
 TGGACAAGAT CAAACACCCC AAGCTTGATG TCAATTAACT CTGATGATGT CGAAAAACAG 2580
 TCCTGTGACT CAACTCAAGC CTTGGTAACC TTTACCAGCT CCAGCATCAC TTATGACCTG 2640
 CCTCCCAGTT CCGTGCCATC ACCAGCTTAT CCACTGACTG AGAGCTGCCA TCTTCTCTCT 2700
 GTGGCATTGT TCCCATGTCT CTAA 2724

(265) INFORMATION FOR SEQ ID NO:264:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 907 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:264:

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

100	105	110
Leu Tyr Ser Leu Lys Val Leu Met Leu Gln Asn Asn Gln Leu Arg His 115 120 125		
Val Pro Thr Glu Ala Leu Gln Asn Leu Arg Ser Leu Gln Ser Leu Arg 130 135 140		
Leu Asp Ala Asn His Ile Ser Tyr Val Pro Pro Ser Cys Phe Ser Gly 145 150 155 160		
Leu His Ser Leu Arg His Leu Trp Leu Asp Asp Asn Ala Leu Thr Glu 165 170 175		
Ile Pro Val Gln Ala Phe Arg Ser Leu Ser Ala Leu Gln Ala Met Thr 180 185 190		
Leu Ala Leu Asn Lys Ile His His Ile Pro Asp Tyr Ala Phe Gly Asn 195 200 205		
Leu Ser Ser Leu Val Val Leu His Leu His Asn Asn Arg Ile His Ser 210 215 220		
Leu Gly Lys Lys Cys Phe Asp Gly Leu His Ser Leu Glu Thr Leu Asp 225 230 235 240		
Leu Asn Tyr Asn Asn Leu Asp Glu Phe Pro Thr Ala Ile Arg Thr Leu 245 250 255		
Ser Asn Leu Lys Glu Leu Gly Phe His Ser Asn Asn Ile Arg Ser Ile 260 265 270		
Pro Glu Lys Ala Phe Val Gly Asn Pro Ser Leu Ile Thr Ile His Phe 275 280 285		
Tyr Asp Asn Pro Ile Gln Phe Val Gly Arg Ser Ala Phe Gln His Leu 290 295 300		
Pro Glu Leu Arg Thr Leu Thr Leu Asn Gly Ala Ser Gln Ile Thr Glu 305 310 315 320		
Phe Pro Asp Leu Thr Gly Thr Ala Asn Leu Glu Ser Leu Thr Leu Thr 325 330 335		
Gly Ala Gln Ile Ser Ser Leu Pro Gln Thr Val Cys Asn Gln Leu Pro 340 345 350		
Asn Leu Gln Val Leu Asp Leu Ser Tyr Asn Leu Leu Glu Asp Leu Pro 355 360 365		
Ser Phe Ser Val Cys Gln Lys Leu Gln Lys Ile Asp Leu Arg His Asn 370 375 380		
Glu Ile Tyr Glu Ile Lys Val Asp Thr Phe Gln Gln Leu Leu Ser Leu 385 390 395 400		

10

20

30

Arg Ser Leu Asn Leu Ala Trp Asn Lys Ile Ala Ile Ile His Pro Asn
 405 410 415
 Ala Phe Ser Thr Leu Pro Ser Leu Ile Lys Leu Asp Leu Ser Ser Asn
 420 425 430
 Leu Leu Ser Ser Phe Pro Ile Thr Gly Leu His Gly Leu Thr His Leu
 435 440 445
 Lys Leu Thr Gly Asn His Ala Leu Gln Ser Leu Ile Ser Ser Glu Asn
 450 455 460
 Phe Pro Glu Leu Lys Val Ile Glu Met Pro Tyr Ala Tyr Gln Cys Cys
 465 470 475 480
 Ala Phe Gly Val Cys Glu Asn Ala Tyr Lys Ile Ser Asn Gln Trp Asn
 485 490 495
 Lys Gly Asp Asn Ser Ser Met Asp Asp Leu His Lys Lys Asp Ala Gly
 500 505 510
 Met Phe Gln Ala Gln Asp Glu Arg Asp Leu Glu Asp Phe Leu Leu Asp
 515 520 525
 Phe Glu Glu Asp Leu Lys Ala Leu His Ser Val Gln Cys Ser Pro Ser
 530 535 540
 Pro Gly Pro Phe Lys Pro Cys Glu His Leu Leu Asp Gly Trp Leu Ile
 545 550 555 560
 Arg Ile Gly Val Trp Thr Ile Ala Val Leu Ala Leu Thr Cys Asn Ala
 565 570 575
 Leu Val Thr Ser Thr Val Phe Arg Ser Pro Leu Tyr Ile Ser Pro Ile
 580 585 590
 Lys Leu Leu Ile Gly Val Ile Ala Ala Val Asn Met Leu Thr Gly Val
 595 600 605
 Ser Ser Ala Val Leu Ala Gly Val Asp Ala Phe Thr Phe Gly Ser Phe
 610 615 620
 Ala Arg His Gly Ala Trp Trp Glu Asn Gly Val Gly Cys His Val Ile
 625 630 635 640
 Gly Phe Leu Ser Ile Phe Ala Ser Glu Ser Ser Val Phe Leu Leu Thr
 645 650 655
 Leu Ala Ala Leu Glu Arg Gly Phe Ser Val Lys Tyr Ser Ala Lys Phe
 660 665 670
 Glu Thr Lys Ala Pro Phe Ser Ser Leu Lys Val Ile Ile Leu Leu Cys
 675 680 685
 Ala Leu Leu Ala Leu Thr Met Ala Ala Val Pro Leu Leu Gly Gly Ser

10

20

30

```

        690                695                700
Lys Tyr Gly Ala Ser Pro Leu Cys Leu Pro Leu Pro Phe Gly Glu Pro
705                710                715                720
Ser Thr Met Gly Tyr Met Val Ala Leu Ile Leu Leu Asn Ser Leu Cys
                725                730                735
Phe Leu Met Met Thr Ile Ala Tyr Thr Lys Leu Tyr Cys Asn Leu Asp
                740                745                750
Lys Gly Asp Leu Glu Asn Ile Trp Asp Cys Ser Met Val Lys His Ile
                755                760                765
Ala Leu Leu Leu Phe Thr Asn Cys Ile Leu Asn Cys Pro Val Ala Phe
                770                775                780
Leu Ser Phe Ser Ser Leu Ile Asn Leu Thr Phe Ile Ser Pro Glu Val
785                790                795                800
Ile Lys Phe Ile Leu Leu Val Val Val Pro Leu Pro Ala Cys Leu Asn
                805                810                815
Pro Leu Leu Tyr Ile Leu Phe Asn Pro His Phe Lys Glu Asp Leu Val
                820                825                830
Ser Leu Arg Lys Gln Thr Tyr Val Trp Thr Arg Ser Lys His Pro Ser
                835                840                845
Leu Met Ser Ile Asn Ser Asp Asp Val Glu Lys Gln Ser Cys Asp Ser
                850                855                860
Thr Gln Ala Leu Val Thr Phe Thr Ser Ser Ser Ile Thr Tyr Asp Leu
865                870                875                880
Pro Pro Ser Ser Val Pro Ser Pro Ala Tyr Pro Val Thr Glu Ser Cys
                885                890                895
His Leu Ser Ser Val Ala Phe Val Pro Cys Leu
                900                905

```

10

20

(266) INFORMATION FOR SEQ ID NO:265:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:265:

CGGAAGCTGC GGGCCAAATG GGTGGCCGGC

30

30

(267) INFORMATION FOR SEQ ID NO:266:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:266:

CAGAGGAGGG TGAAGGGCT GTTGCG

27

(268) INFORMATION FOR SEQ ID NO:267:

10

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:267:

GGCGGCGCCG AGCCAAGGG CTGGCTGTGG

30

(269) INFORMATION FOR SEQ ID NO:268:

20

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 32 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:268:

GGGACTGCTC TATGAAAAA CACATTGCCC TG

32

(270) INFORMATION FOR SEQ ID NO:269:

30

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1071 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:269:

ATGAATGGGG TCTCGGAGGG GACCAGAGGC TGCACTGACA GGCAACCTGG GGTCTGACA

60

CGTGATCGCT CTTGTTCCAG GAAGATGAAC TCTTCCGGAT GCCTGTCTGA GGAGGTGGGG 120
 TCCCTCCGCC CACTGACTGT GGTATCCTG TCTGCGTCCA TTGTCGTCGG AGTGCTGGGC 180
 AATGGGCTGG TGCTGTGGAT GACTGTCTTC CGTATGGCAC GCACGGTCTC CACCGTCTGC 240
 TTCTTCCACC TGGCCCTTGC CGATTTCATG CTCTCACTGT CTCTGCCCAT TGCCATGTAC 300
 TATATTGTCT CCAGGCAGTG GCTCCTCGGA GAGTGGGCCT GCAAACCTTA CATCACCTTT 360
 GTGTTCTCA GCTACTTTGC CAGTAACTGC CTCCTTGTCT TCATCTCTGT GGACCGTTGC 420
 ATCTCTGTCC TCTACCCCGT CTGGGCCCTG AACCACCGCA CTGTGCAGCG GGCGAGCTGG 480
 CTGGCCTTTG GGGTGTGGCT CCTGGCCGCC GCCTTGTGCT CTGCGCACCT GAAATTCGG 540
 ACAACCAGAA AATGGAATGG CTGTACGCAC TGCTACTTGG CGTTCAACTC TGACAATGAG 600
 ACTGCCCAGA TTTGGATTGA AGGGGTCGTG GAGGGACACA TTATAGGGAC CATTGGCCAC 660
 TTCCTGCTGG GCTTCTGGG GCCCTTAGCA ATCATAGGCA CCTGCGCCCA CCTCATCCGG 720
 GCCAAGCTCT TCGGGGAGGG CTGGGTCCAT GCCAACCAGC CCAAGAGGCT GCTGCTGGTG 780
 CTGGTGAGCG CTTTCTTTAT CTTCTGGTCC CCGTTTAAAG TGGTGTGTG GGTCCATCTG 840
 TGGCGACGGG TGATGCTCAA GGAAATCTAC CACCCCGGGA TGCTGCTCAT CCTCCAGGCT 900
 AGCTTTGCCT TGGGCTGTGT CAACAGCAGC CTCAACCCCT TCCTCTACGT CTTGTTGGC 960
 AGAGATTTCC AAGAAAAGTT TTTCCAGTCT TTGACTTCTG CCCTGGCGAG GGCCTTTGGA 1020
 GAGGAGGAGT TTCTGTCATC CTGTCCCCGT GGCAACGCCC CCCGGGAATG A 1071

(271) INFORMATION FOR SEQ ID NO:270:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 356 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:270:

Met Asn Gly Val Ser Glu Gly Thr Arg Gly Cys Ser Asp Arg Gln Pro
 1 5 10 15
 Gly Val Leu Thr Arg Asp Arg Ser Cys Ser Arg Lys Met Asn Ser Ser
 20 25 30
 Gly Cys Leu Ser Glu Glu Val Gly Ser Leu Arg Pro Leu Thr Val Val
 35 40 45

10

20

30

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

10

20

30

340

345

350

Ala Pro Arg Glu
355

(272) INFORMATION FOR SEQ ID NO:271:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 903 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:271:

```
ATGGACCTGC CCCCGCAGCT CTCCTTCGGC CTCTATGTGG CCGCCTTTGC GCTGGGCTTC    60
CCGCTCAACG TCCTGGCCAT CCGAGGCGCG ACGGCCACG CCCGGCTCCG TCTACCCCT    120
AGCCTGGTCT ACGCCCTGAA CCTGGGCTGC TCCGACCTGC TGCTGACAGT CTCTCTGCCC    180
CTGAAGGCGG TGGAGGCGCT AGCCTCCGGG GCCTGGCCTC TGCCGGCCTC GCTGTGCCCC    240
GTCTTCGCGG TGGCCCACTT CTTCCTACTC TATGCCGGCG GGGGCTTCCT GGCCGCCCTG    300
AGTGCAGGCC GCTACCTGGG AGCAGCCTTC CCCTTGGGCT ACCAAGCCTT CCGGAGGCCG    360
TGCTATTCTT GGGGGGTGTG CGCGGCCATC TGGGCCCTCG TCCTGTGTCA CCTGGGTCTG    420
GTCTTTGGGT TGGAGGCTCC AGGAGGCTGG CTGGACCACA GCAACACCTC CCTGGGCATC    480
AACACACCGG TCAACGGCTC TCCGGTCTGC CTGGAGGCCT GGGACCCGGC CTCTGCCGGC    540
CCGGCCCGCT TCAGCCTCTC TCTCCTGCTC TTTTCTCTGC CCTTGGCCAT CACAGCCTTC    600
TGCTACGTGG GCTGCCTCCG GGCCTGGGCC CGTCCGGCC TGACGCACAG GCGGAAGCTG    660
CGGGCCAAAT GGGTGGCCGG CGGGGCCCTC CTCACGCTGC TGCTCTGCGT AGGACCCTAC    720
AACGCCTCCA ACGTGGCCAG CTTCTGTAC CCCAATCTAG GAGGCTCCTG GCGGAAGCTG    780
GGGCTCATCA CGGGTGCCCTG GAGTGTGGTG CTTAATCCGC TGGTGACCGG TTAATTGGGA    840
AGGGGTCCTG GCCTGAAGAC AGTGTGTGCG GCAAGAACGC AAGGGGGCAA GTCCCAGAAG    900
TAA                                                                    903
```

20

30

(273) INFORMATION FOR SEQ ID NO:272:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 300 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:

(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:272:

```

Met Asp Leu Pro Pro Gln Leu Ser Phe Gly Leu Tyr Val Ala Ala Phe
1          5          10          15

Ala Leu Gly Phe Pro Leu Asn Val Leu Ala Ile Arg Gly Ala Thr Ala
          20          25          30

His Ala Arg Leu Arg Leu Thr Pro Ser Leu Val Tyr Ala Leu Asn Leu
          35          40          45

Gly Cys Ser Asp Leu Leu Leu Thr Val Ser Leu Pro Leu Lys Ala Val
          50          55          60

Glu Ala Leu Ala Ser Gly Ala Trp Pro Leu Pro Ala Ser Leu Cys Pro
          65          70          75          80

Val Phe Ala Val Ala His Phe Phe Pro Leu Tyr Ala Gly Gly Gly Phe
          85          90          95

Leu Ala Ala Leu Ser Ala Gly Arg Tyr Leu Gly Ala Ala Phe Pro Leu
          100          105          110

Gly Tyr Gln Ala Phe Arg Arg Pro Cys Tyr Ser Trp Gly Val Cys Ala
          115          120          125

Ala Ile Trp Ala Leu Val Leu Cys His Leu Gly Leu Val Phe Gly Leu
          130          135          140

Glu Ala Pro Gly Gly Trp Leu Asp His Ser Asn Thr Ser Leu Gly Ile
          145          150          155          160

Asn Thr Pro Val Asn Gly Ser Pro Val Cys Leu Glu Ala Trp Asp Pro
          165          170          175

Ala Ser Ala Gly Pro Ala Arg Phe Ser Leu Ser Leu Leu Leu Phe Phe
          180          185          190

Leu Pro Leu Ala Ile Thr Ala Phe Cys Tyr Val Gly Cys Leu Arg Ala
          195          200          205

Leu Ala Arg Ser Gly Leu Thr His Arg Arg Lys Leu Arg Ala Lys Trp
          210          215          220

Val Ala Gly Gly Ala Leu Leu Thr Leu Leu Leu Cys Val Gly Pro Tyr
          225          230          235          240

Asn Ala Ser Asn Val Ala Ser Phe Leu Tyr Pro Asn Leu Gly Gly Ser
          245          250          255

Trp Arg Lys Leu Gly Leu Ile Thr Gly Ala Trp Ser Val Val Leu Asn

```

10

20

30

260	265	270
Pro Leu Val Thr Gly Tyr Leu Gly Arg Gly Pro Gly Leu Lys Thr Val		
275	280	285
Cys Ala Ala Arg Thr Gln Gly Gly Lys Ser Gln Lys		
290	295	300

(274) INFORMATION FOR SEQ ID NO:273:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1041 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:273:

ATGGATACAG GCCCCGACCA GTCCTACTTC TCCGGCAATC ACTGGTTCGT CTTCTCGGTG	60
TACCTTCTCA CTTTCCTGGT GGGGCTCCCC CTCAACCTGC TGGCCCTGGT GGTCTTCGTG	120
GGCAAGCTGC AGCGCCGCC GGTGGCCGTG GACGTGCTCC TGCTCAACCT GACCGCCTCG	180
GACCTGCTCC TGCTGCTGTT CCTGCCTTTC CGCATGGTGG AGGCAGCCAA TGGCATGCAC	240
TGGCCCTGTC CTTTCATCCT CTGCCCCTC TCTGGATTCA TCTTCTTCAC CACCATCTAT	300
CTCACCGCCC TCTTCCTGGC AGCTGTGAGC ATTGAACGCT TCCTGAGTGT GGCCACCCCA	360
CTGTGGTACA AGACCCGGCC GAGGCTGGGG CAGGCAGGTC TGGTGAGTGT GGCTGCTGG	420
CTGTTGGCCT CTGCTCACTG CAGCGTGGTC TACGTCATAG AATTCTCAGG GGACATCTCC	480
CACAGCCAGG GCACCAATGG GACCTGCTAC CTGGAGTTCC GGAAGGACCA GCTAGCCATC	540
CTCCTGCCCC TGCGGCTGGA GATGGCTGTG GTCCTCTTTG TGGTCCCGCT GATCATCACC	600
AGCTACTGCT ACAGCCGCCT GGTGTGGATC CTCGGCAGAG GGGGCAGCCA CCGCCGGCAG	660
AGGAGGGTGA AGGGGCTGTT GGCGGCCACG CTGCTCAACT TCCTTGCTCTG CTTTGGGCCC	720
TACAACGTGT CCCATGTCGT GGGCTATATC TCGGGTGAAA GCGCGGCATG GAGGATCTAC	780
GTGACGCTTC TCAGCACCTT GAACTCCTGT GTCGACCCCT TTGTCTACTA CTTCTCCTCC	840
TCCGGGTTCC AAGCCGACTT TCATGAGCTG CTGAGGAGGT TGTGTGGGCT CTGGGGCCAG	900
TGGCAGCAGG AGAGCAGCAT GGAGCTGAAG GAGCAGAAGG GAGGGGAGGA GCAGAGAGCG	960
GACCGACCAG CTGAAAGAAA GACCAGTGAA CACTCACAGG GCTGTGGAAC TGGTGGCCAG	1020
GTGGCCTGTG CTGAAAGCTA G	1041

20

30

(275) INFORMATION FOR SEQ ID NO:274:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 346 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:274:

```

Met Asp Thr Gly Pro Asp Gln Ser Tyr Phe Ser Gly Asn His Trp Phe
1           5           10           15
Val Phe Ser Val Tyr Leu Leu Thr Phe Leu Val Gly Leu Pro Leu Asn
20           25           30
Leu Leu Ala Leu Val Val Phe Val Gly Lys Leu Gln Arg Arg Pro Val
35           40           45
Ala Val Asp Val Leu Leu Leu Asn Leu Thr Ala Ser Asp Leu Leu Leu
50           55           60
Leu Leu Phe Leu Pro Phe Arg Met Val Glu Ala Ala Asn Gly Met His
65           70           75           80
Trp Pro Leu Pro Phe Ile Leu Cys Pro Leu Ser Gly Phe Ile Phe Phe
85           90           95
Thr Thr Ile Tyr Leu Thr Ala Leu Phe Leu Ala Ala Val Ser Ile Glu
100          105          110
Arg Phe Leu Ser Val Ala His Pro Leu Trp Tyr Lys Thr Arg Pro Arg
115          120          125
Leu Gly Gln Ala Gly Leu Val Ser Val Ala Cys Trp Leu Leu Ala Ser
130          135          140
Ala His Cys Ser Val Val Tyr Val Ile Glu Phe Ser Gly Asp Ile Ser
145          150          155          160
His Ser Gln Gly Thr Asn Gly Thr Cys Tyr Leu Glu Phe Arg Lys Asp
165          170          175
Gln Leu Ala Ile Leu Leu Pro Val Arg Leu Glu Met Ala Val Val Leu
180          185          190
Phe Val Val Pro Leu Ile Ile Thr Ser Tyr Cys Tyr Ser Arg Leu Val
195          200          205
Trp Ile Leu Gly Arg Gly Gly Ser His Arg Arg Gln Arg Arg Val Lys
210          215          220
Gly Leu Leu Ala Ala Thr Leu Leu Asn Phe Leu Val Cys Phe Gly Pro

```

10

20

30

225 230 235 240
Tyr Asn Val Ser His Val Val Gly Tyr Ile Cys Gly Glu Ser Pro Ala
 245 250 255
Trp Arg Ile Tyr Val Thr Leu Leu Ser Thr Leu Asn Ser Cys Val Asp
 260 265 270
Pro Phe Val Tyr Tyr Phe Ser Ser Ser Gly Phe Gln Ala Asp Phe His
 275 280 285
Glu Leu Leu Arg Arg Leu Cys Gly Leu Trp Gly Gln Trp Gln Gln Glu
 290 295 300
Ser Ser Met Glu Leu Lys Glu Gln Lys Gly Gly Glu Glu Gln Arg Ala
305 310 315 320
Asp Arg Pro Ala Glu Arg Lys Thr Ser Glu His Ser Gln Gly Cys Gly
 325 330 335
Thr Gly Gly Gln Val Ala Cys Ala Glu Ser
 340 345

(276) INFORMATION FOR SEQ ID NO:275:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 993 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:275:

ATGCTGCCGG ACTGGAAGAG CTCCTTGATC CTCATGGCTT ACATCATCAT CTCCTCACT 60
GGCCTCCCTG CCAACCTCCT GGCCCTGCGG GCCTTTGTGG GCGGGATCCG CCAGCCCCAG 120
CCTGCACCTG TGCACATCCT CCTGCTGAGC CTGACGCTGG CCGACCTCCT CCTGCTGCTG 180
CTGCTGCCCT TCAAGATCAT CGAGGCTGCG TCGAACTTCC GCTGGTACCT GCCCAAGGTC 240
GTCTGCGCCC TCACGAGTTT TGGCTTCTAC AGCAGCATCT ACTGCAGCAC GTGGCTCCTG 300
GCGGGCATCA GCATCGAGCG CTACCTGGGA GTGGCTTTCC CCGTGAGTA CAAGCTCTCC 360
CGCCGGCCTC TGTATGGAGT GATTGCAGCT CTGGTGGCCT GGTTTATGTC CTTTGGTCAC 420
TGCACCATCG TGATCATCGT TCAATACTTG AACACGACTG AGCAGGTCAG AAGTGGCAAT 480
GAAATTACCT GCTACGAGAA CTTACCGAT AACCAGTTGG ACGTGGTGCT GCCCGTCCG 540
CTGGAGCTGT GCCTGGTGCT CTTCTTCATC CCCATGGCAG TCACCATCTT CTGCTACTGG 600

10

20

30

CGTTTGTGT GGATCATGCT CTCCCAGCCC CTTGTGGGGG CCCAGAGGCG GCGCCGAGCC 660
 AAGGGGCTGG CTGTGGTGAC GCTGCTCAAT TTCCTGGTGT GCTTCGGACC TTACAACGTG 720
 TCCCACCTGG TGGGGTATCA CCAGAGAAAA AGCCCCTGGT GGCGGTCAAT AGCCGTGGTG 780
 TTCAGTTCAC TCAACGCCAG TCTGGACCCC CTGCTCTTCT ATTTCTCTTC TTCAGTGGTG 840
 CGCAGGGCAT TTGGGAGAGG GCTGCAGGTG CTGCGGAATC AGGGCTCCTC CCTGTTGGGA 900
 CGCAGAGGCA AAGACACAGC AGAGGGGACA AATGAGGACA GGGGTGTGGG TCAAGGAGAA 960
 GGGATGCCAA GTTCGGAATT CACTACAGAG TAG 993

(277) INFORMATION FOR SEQ ID NO:276:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 330 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant
- (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:276:

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

10

20

30

Glu Ile Thr Cys Tyr Glu Asn Phe Thr Asp Asn Gln Leu Asp Val Val
 165 170 175
 Leu Pro Val Arg Leu Glu Leu Cys Leu Val Leu Phe Phe Ile Pro Met
 180 185 190
 Ala Val Thr Ile Phe Cys Tyr Trp Arg Phe Val Trp Ile Met Leu Ser
 195 200 205
 Gln Pro Leu Val Gly Ala Gln Arg Arg Arg Arg Ala Lys Gly Leu Ala
 210 215 220
 Val Val Thr Leu Leu Asn Phe Leu Val Cys Phe Gly Pro Tyr Asn Val
 225 230 235 240
 Ser His Leu Val Gly Tyr His Gln Arg Lys Ser Pro Trp Trp Arg Ser
 245 250 255
 Ile Ala Val Val Phe Ser Ser Leu Asn Ala Ser Leu Asp Pro Leu Leu
 260 265 270
 Phe Tyr Phe Ser Ser Ser Val Val Arg Arg Ala Phe Gly Arg Gly Leu
 275 280 285
 Gln Val Leu Arg Asn Gln Gly Ser Ser Leu Leu Gly Arg Arg Gly Lys
 290 295 300
 Asp Thr Ala Glu Gly Thr Asn Glu Asp Arg Gly Val Gly Gln Gly Glu
 305 310 315 320
 Gly Met Pro Ser Ser Asp Phe Thr Thr Glu
 325 330

10

20

(278) INFORMATION FOR SEQ ID NO:277:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2724 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:277:

30

ATGGACACCT CCCGGCTCGG TGTGCTCCTG TCCTTGCCCTG TGCTGCTGCA GCTGGCGACC 60
 GGGGGCAGCT CTCCCAGGTC TGGTGTGTGTG CTGAGGGGGCT GCCCCACACA CTGTCATTGC 120
 GAGCCCGACG GCAGGATGTT GCTCAGGGTG GACTGCTCCG ACCTGGGGCT CTCGGAGCTG 180
 CCTTCCAACC TCAGCGTCTT CACCTCCTAC CTAGACCTCA GTATGAACAA CATCAGTCAG 240
 CTGCTCCCGA ATCCCCTGCC CAGTCTCCGC TTCTGGAGG AGTTACGTCT TCGGGGAAAC 300

GCTCTGACAT ACATTCCCAA GGGAGCATTG ACTGGCCTTT ACAGTCTTAA AGTTCTTATG 360
 CTGCAGAATA ATCAGCTAAG ACACGTACCC ACAGAAGCTC TGCAGAATTT GCGAAGCCTT 420
 CAATCCCTGC GTCTGGATGC TAACCACATC AGCTATGTGC CCCCAGCTG TTTCAGTGGC 480
 CTGCATTCCC TGAGGCACCT GTGGCTGGAT GACAATGCGT TAACAGAAAT CCCCCTCCAG 540
 GCTTTTAGAA GTTTATCGGC ATTGCAAGCC ATGACCTTGG CCCTGAACAA AATACACCAC 600
 ATACCAGACT ATGCTTTTGG AAACCTCTCC AGCTTGGTAG TTCTACATCT CCATAACAAT 660
 AGAATCCACT CCCTGGGAAA GAAATGCTTT GATGGGCTCC ACAGCCTAGA GACTTTAGAT 720
 TTAAATTACA ATAACCTTGA TGAATTCCCC ACTGCAATTA GGACACTCTC CAACCTTAAA 780
 GAACTAGGAT TTCATAGCAA CAATATCAGG TCGATACCTG AGAAAGCATT TGTAGGCAAC 840
 CCTTCTCTTA TTACAATACA TTTCTATGAC AATCCCATCC AATTTGTTGG GAGATCTGCT 900
 TTTCAACATT TACCTGAACT AAGAACACTG ACTCTGAATG GTGCCTCACA AATAACTGAA 960
 TTTCTGATT TAACTGGAAC TGCAAACCTG GAGAGTCTGA CTTTAACTGG AGCACAGATC 1020
 TCATCTCTTC CTCAAACCGT CTGCAATCAG TTACCTAATC TCCAAGTGCT AGATCTGTCT 1080
 TACAACCTAT TAGAAGATTT ACCCAGTTTT TCAGTCTGCC AAAAGCTTCA GAAAATTGAC 1140
 CTAAGACATA ATGAAATCTA CGAAATTAAA GTTGACACTT TCCAGCAGTT GCTTAGCCTC 1200
 CGATCGCTGA ATTTGGCTTG GAACAAAATT GCTATTTATC ACCCCAATGC ATTTTCCACT 1260
 TTGCCATCCC TAATAAAGCT GGACCTATCG TCCAACCTCC TGTCGTCTTT TCCTATAACT 1320
 GGGTTACATG GTTTAACTCA CTTAAAATTA ACAGGAAATC ATGCCTTACA GAGCTTGATA 1380
 TCATCTGAAA ACTTTCCAGA ACTCAAGGTT ATAGAAATGC CTTATGCTTA CCAGTGCTGT 1440
 GCATTTGGAG TGTGTGAGAA TGCTTATAAG ATTTCTAATC AATGGAATAA AGGTGACAAC 1500
 AGCAGTATGG ACGACCTTCA TAAGAAAGAT GCTGGAATGT TTCAGGCTCA AGATGAACGT 1560
 GACCTTGAAG ATTTCTTGCT TGACTTTGAG GAAGACCTGA AAGCCCTTCA TTCAGTGCAG 1620
 TGTTCACCTT CCCCAGGCCC CTTCAAACCC TGTGAACACC TGCTTGATGG CTGGCTGATC 1680
 AGAATTGGAG TGTGGACCAT AGCAGTTCTG GCACCTTACT GTAATGCTTT GGTGACTTCA 1740
 ACAGTTTTCA GATCCCCTCT GTACATTTCC CCCATTAAAC TGTTAATTGG GGTGATCGCA 1800
 GCAGTGAACA TGCTCACGGG AGTCTCCAGT GCCGTGCTGG CTGGTGTGGA TCGGTTCACT 1860
 TTTGGCAGCT TTGCACGACA TGGTGCCTGG TGGGAGAATG GGGTTGGTTG CCATGTCATT 1920
 GGTTTTTTGT CCATTTTTCG TTCAGAAATCA TCTGTTTTCC TGCTTACTCT GGCAGCCCTG 1980

10

20

30

GAGCGTGGGT TCTCTGTGAA ATATTCTGCA AAATTTGAAA CGAAAGCTCC ATTTTCTAGC 2040
 CTGAAAGTAA TCATTTTGCT CTGTGCCCTG CTGGCCTTGA CCATGGCCGC AGTTCCCTG 2100
 CTGGGTGGCA GCAAGTATGG CGCTCCCCT CTCTGCCTGC CTTTGCCTTT TGGGGAGCCC 2160
 AGCACCATGG GCTACATGGT CGCTCTCATC TTGCTCAATT CCCTTTGCTT CCTCATGATG 2220
 ACCATTGCCT ACACCAAGCT CTA CTGCAAT TTGGACAAGG GAGACCTGGA GAATATTTGG 2280
 GACTGCTCTA TGA AAAACA CATTGCCCTG TTGCTCTTCA CCAACTGCAT CCTAAACTGC 2340
 CCTGTGGCTT TCTTGTCTT CTCCTCTTTA ATAAACCTTA CATTATCAG TCCTGAAGTA 2400
 ATTAAGTTTA TCCTTCTGGT GGTAGTCCCA CTCCTGCAT GTCTCAATCC CCTTCTCTAC 2460
 ATCTTGTTCA ATCCTCACTT TAAGGAGGAT CTGGTGAGCC TGAGAAAGCA AACCTACGTC 2520
 TGGACAAGAT CAAAACACCC AAGCTTGATG TCAATTA ACT CTGATGATGT CGAAAAACAG 2580
 TCCTGTGACT CAACTCAAGC CTTGGTAACC TTTACCAGCT CCAGCATCAC TTATGACCTG 2640
 CCTCCAGTT CCGTGCCATC ACCAGCTTAT CCAGTGACTG AGAGCTGCCA TCTTCTCTCT 2700
 GTGGCATTG TCCCATGTCT CTAA 2724

(279) INFORMATION FOR SEQ ID NO:278:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 907 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:278:

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

Leu Ala Gly Asn Ala Leu Thr Tyr Ile Pro Lys Gly Ala Phe Thr Gly
 100 105 110
 Leu Tyr Ser Leu Lys Val Leu Met Leu Gln Asn Asn Gln Leu Arg His
 115 120 125
 Val Pro Thr Glu Ala Leu Gln Asn Leu Arg Ser Leu Gln Ser Leu Arg
 130 135 140
 Leu Asp Ala Asn His Ile Ser Tyr Val Pro Pro Ser Cys Phe Ser Gly
 145 150 155 160
 Leu His Ser Leu Arg His Leu Trp Leu Asp Asp Asn Ala Leu Thr Glu
 165 170 175
 Ile Pro Val Gln Ala Phe Arg Ser Leu Ser Ala Leu Gln Ala Met Thr
 180 185 190
 Leu Ala Leu Asn Lys Ile His His Ile Pro Asp Tyr Ala Phe Gly Asn
 195 200 205
 Leu Ser Ser Leu Val Val Leu His Leu His Asn Asn Arg Ile His Ser
 210 215 220
 Leu Gly Lys Lys Cys Phe Asp Gly Leu His Ser Leu Glu Thr Leu Asp
 225 230 235 240
 Leu Asn Tyr Asn Asn Leu Asp Glu Phe Pro Thr Ala Ile Arg Thr Leu
 245 250 255
 Ser Asn Leu Lys Glu Leu Gly Phe His Ser Asn Asn Ile Arg Ser Ile
 260 265 270
 Pro Glu Lys Ala Phe Val Gly Asn Pro Ser Leu Ile Thr Ile His Phe
 275 280 285
 Tyr Asp Asn Pro Ile Gln Phe Val Gly Arg Ser Ala Phe Gln His Leu
 290 295 300
 Pro Glu Leu Arg Thr Leu Thr Leu Asn Gly Ala Ser Gln Ile Thr Glu
 305 310 315 320
 Phe Pro Asp Leu Thr Gly Thr Ala Asn Leu Glu Ser Leu Thr Leu Thr
 325 330 335
 Gly Ala Gln Ile Ser Ser Leu Pro Gln Thr Val Cys Asn Gln Leu Pro
 340 345 350
 Asn Leu Gln Val Leu Asp Leu Ser Tyr Asn Leu Leu Glu Asp Leu Pro
 355 360 365
 Ser Phe Ser Val Cys Gln Lys Leu Gln Lys Ile Asp Leu Arg His Asn
 370 375 380
 Glu Ile Tyr Glu Ile Lys Val Asp Thr Phe Gln Gln Leu Leu Ser Leu

10

20

30

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

10

20

30

Ala Leu Leu Ala Leu Thr Met Ala Ala Val Pro Leu Leu Gly Gly Ser
 690 695 700

Lys Tyr Gly Ala Ser Pro Leu Cys Leu Pro Leu Pro Phe Gly Glu Pro
 705 710 715 720

Ser Thr Met Gly Tyr Met Val Ala Leu Ile Leu Leu Asn Ser Leu Cys
 725 730 735

Phe Leu Met Met Thr Ile Ala Tyr Thr Lys Leu Tyr Cys Asn Leu Asp
 740 745 750

Lys Gly Asp Leu Glu Asn Ile Trp Asp Cys Ser Met Lys Lys His Ile
 755 760 765

Ala Leu Leu Leu Phe Thr Asn Cys Ile Leu Asn Cys Pro Val Ala Phe
 770 775 780

Leu Ser Phe Ser Ser Leu Ile Asn Leu Thr Phe Ile Ser Pro Glu Val
 785 790 795 800

Ile Lys Phe Ile Leu Leu Val Val Val Pro Leu Pro Ala Cys Leu Asn
 805 810 815

Pro Leu Leu Tyr Ile Leu Phe Asn Pro His Phe Lys Glu Asp Leu Val
 820 825 830

Ser Leu Arg Lys Gln Thr Tyr Val Trp Thr Arg Ser Lys His Pro Ser
 835 840 845

Leu Met Ser Ile Asn Ser Asp Asp Val Glu Lys Gln Ser Cys Asp Ser
 850 855 860

Thr Gln Ala Leu Val Thr Phe Thr Ser Ser Ser Ile Thr Tyr Asp Leu
 865 870 875 880

Pro Pro Ser Ser Val Pro Ser Pro Ala Tyr Pro Val Thr Glu Ser Cys
 885 890 895

His Leu Ser Ser Val Ala Phe Val Pro Cys Leu
 900 905

10

20

(280) INFORMATION FOR SEQ ID NO:279:

30

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 32 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:279:

CATGCCAACC GGCCCGCGAG GCTGCTGCTG GT

32

(281) INFORMATION FOR SEQ ID NO:280:

40

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 32 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

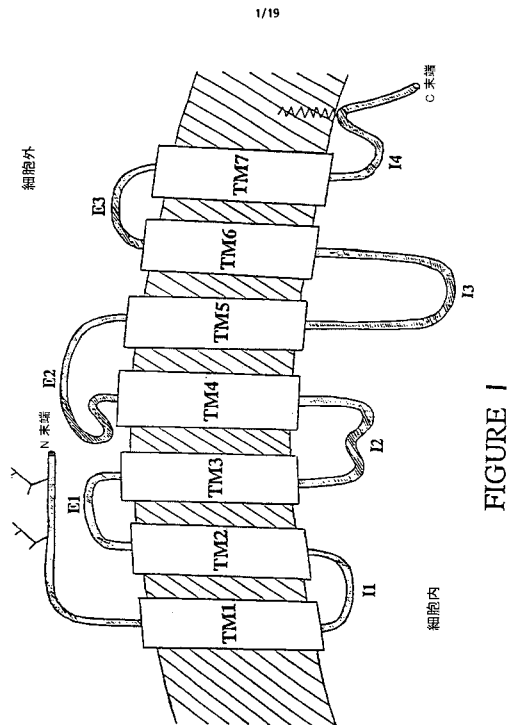
(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:280:

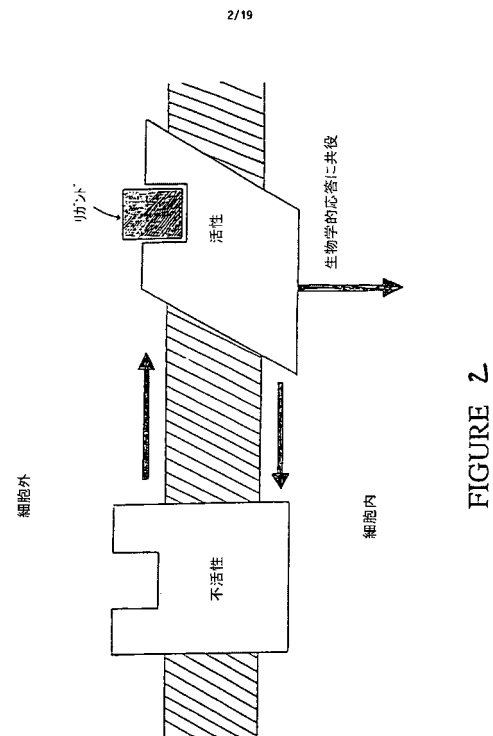
ACCAGCAGCA GCCTCGCGGG CCGGTTGGCA TG

32

【図 1】



【図 2】



【図 3 A】

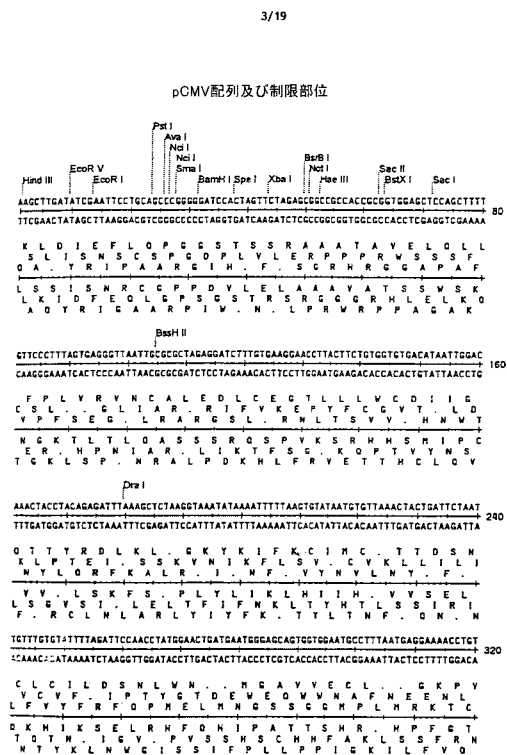


FIGURE 3A

【図 3 B】

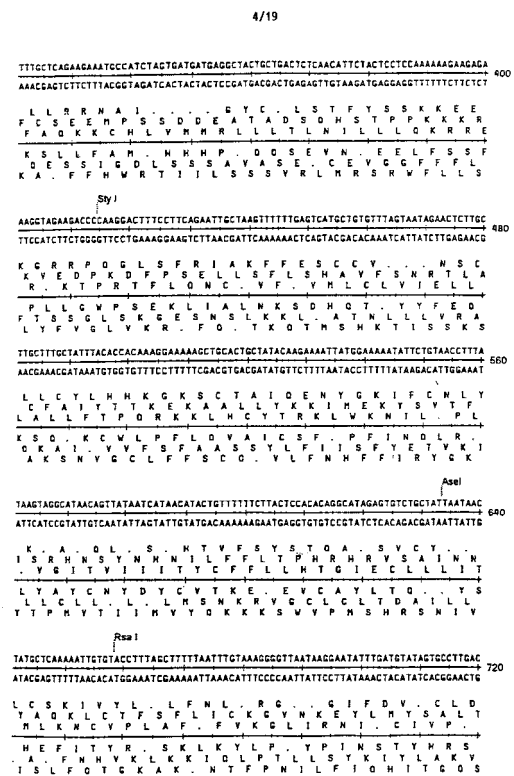


FIGURE 3B

【図 3 C】

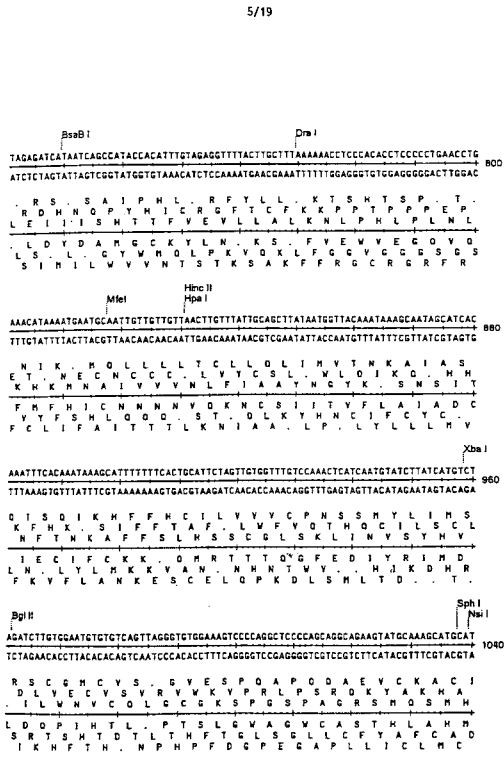


FIGURE 3C

【図 3 D】

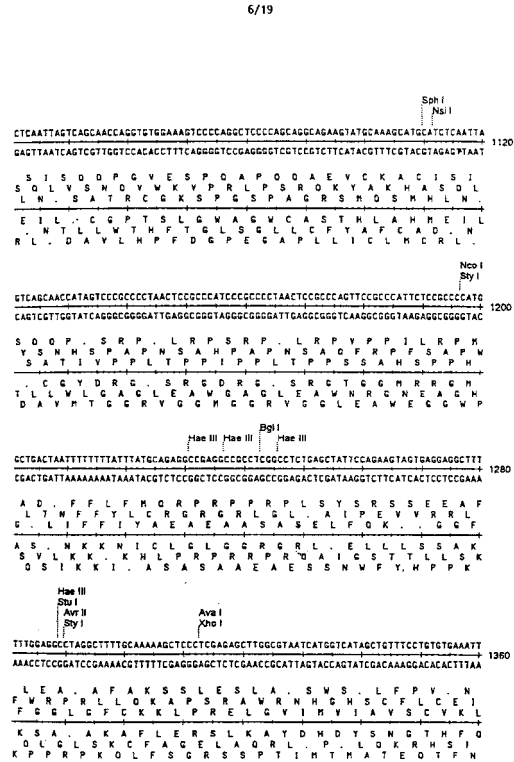


FIGURE 3D

【図 3 E】

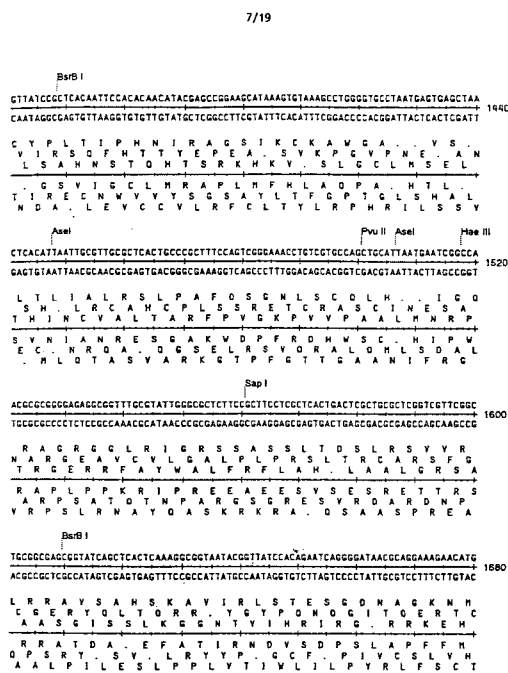
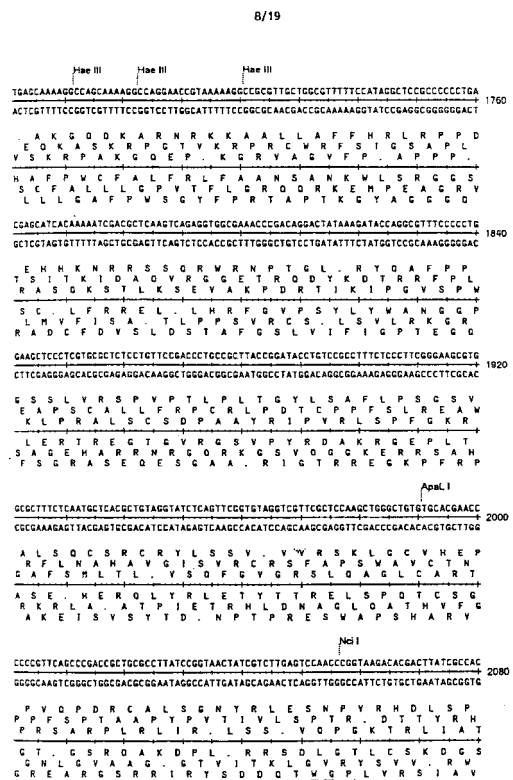


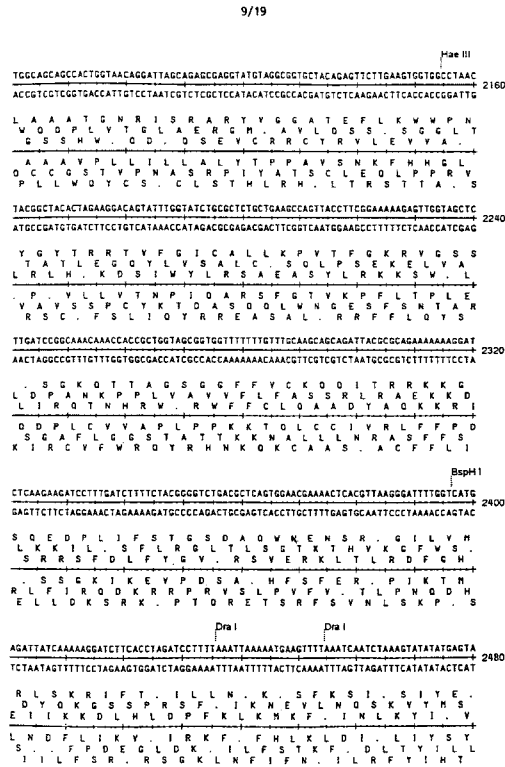
FIGURE 3E

【図 3 F】

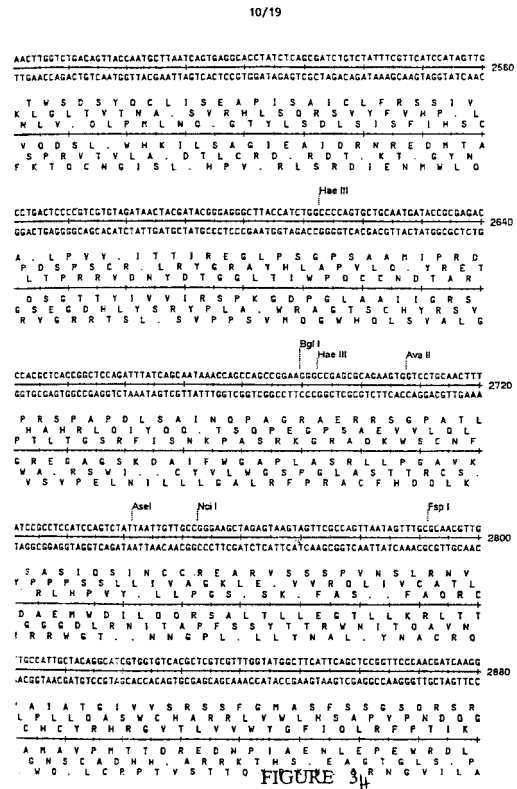


D O T W G P L
FIGURE 3F

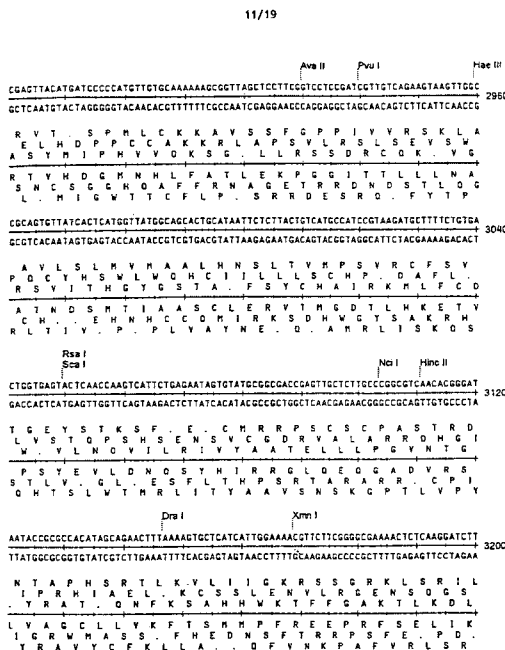
【図 3 G】



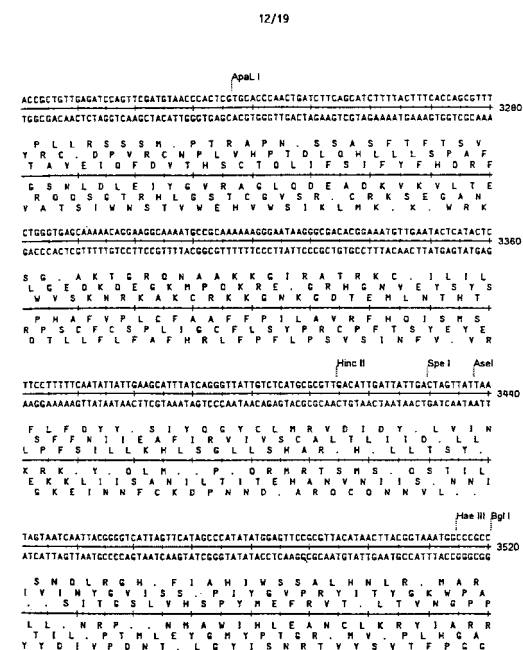
【図 3 H】



【図 3 I】



【図 3 J】



【図 3 K】

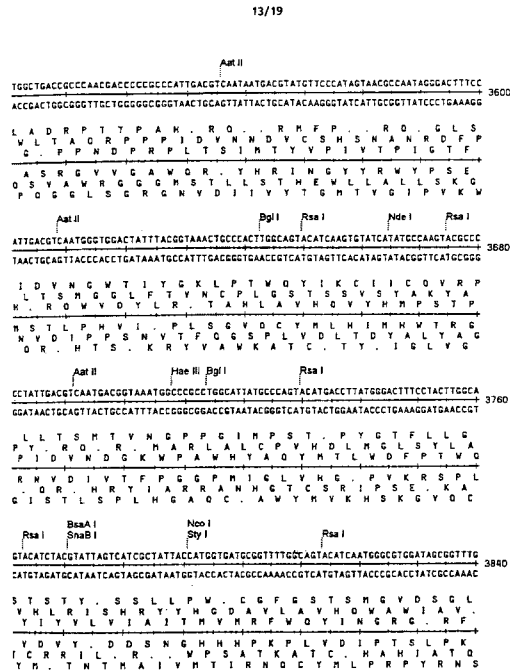


FIGURE 3K

【図 3 L】

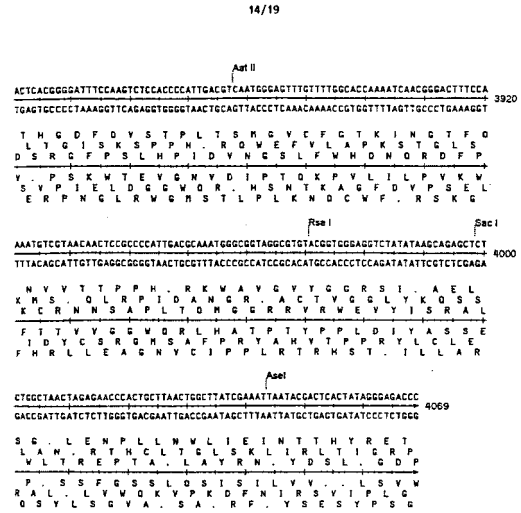


FIGURE 3L

【図 4】

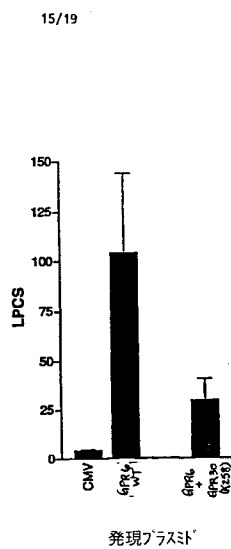


FIGURE 4

【図 5】

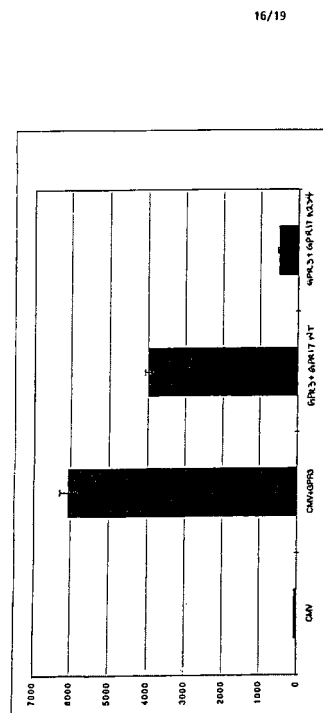


FIGURE 5

【図 6】

17/19

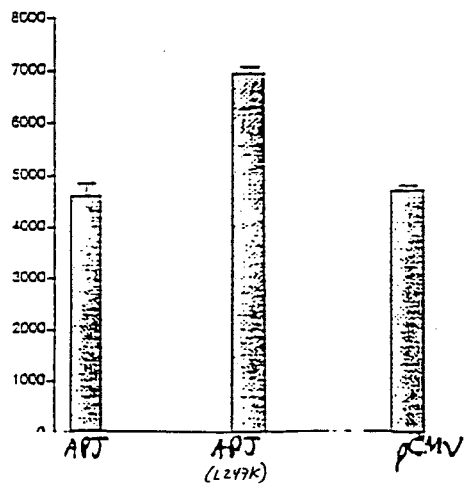


FIGURE 6

【図 7】

18/19

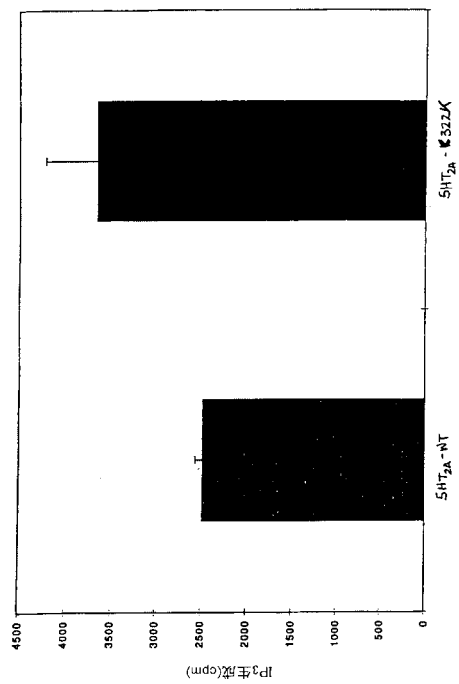


FIGURE 7

【図 8 A - C】

19/19

FIGURE 8A

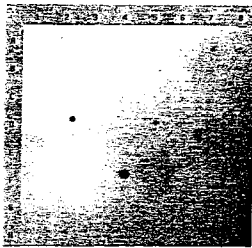


FIGURE 8B

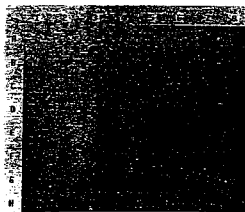
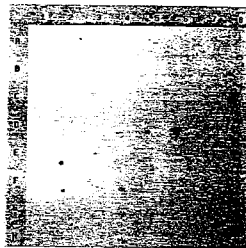


FIGURE 8C

フロントページの続き

(51)Int.Cl.		F I
G 0 1 N 33/566 (2006.01)		G 0 1 N 33/566
C 1 2 N 1/15 (2006.01)		C 1 2 N 1/15
C 1 2 N 1/19 (2006.01)		C 1 2 N 1/19
C 1 2 N 1/21 (2006.01)		C 1 2 N 1/21
C 1 2 N 5/10 (2006.01)		C 1 2 N 5/00 1 0 1

(72)発明者 ベハン, ドミニク・ピー
 アメリカ合衆国カリフォルニア州 9 2 1 3 1 サンディエゴ・ロックスボロコート 1 1 4 7 2

(72)発明者 シヤルマース, デレク・テイ
 アメリカ合衆国カリフォルニア州 9 2 0 7 5 ソラナビーチ・ロングデンレイン 3 4 7

(72)発明者 リオウ, チエン・ダブリュー
 アメリカ合衆国カリフォルニア州 9 2 1 2 9 サンディエゴ・サリックスプレイス 7 6 6 8

審査官 濱田 光浩

(56)参考文献 国際公開第 9 8 / 0 3 8 2 1 7 (WO, A 1)
 国際公開第 9 7 / 0 2 1 7 3 1 (WO, A 1)
 J. Biol. Chem., 1 9 9 2 年, vol. 267, no. 3, p. 1430-1433
 J. Recept. Signal Transduction Res., 1 9 9 7 年, vol. 17, no. 1-3, p. 57-73

(58)調査した分野(Int.Cl., D B名)
 C12N 15/00
 CA/BIOSIS/MEDLINE/WPIDS(STN)