

[19] 中华人民共和国国家知识产权局



[12] 发明专利说明书

专利号 ZL 99812091. X

[51] Int. Cl.

C12N 15/12 (2006.01)

C07K 14/72 (2006.01)

G01N 33/50 (2006.01)

G01N 33/566 (2006.01)

[45] 授权公告日 2007 年 8 月 1 日

[11] 授权公告号 CN 1329511C

[22] 申请日 1999.10.12 [21] 申请号 99812091. X

[30] 优先权

[32] 1998.10.13 [33] US [31] 09/170,496

[86] 国际申请 PCT/US1999/023938 1999.10.12

[87] 国际公布 WO2000/022129 英 2000.4.20

[85] 进入国家阶段日期 2001.4.13

[73] 专利权人 阿瑞那制药公司

地址 美国加利福尼亚州

[72] 发明人 多米尼克·P·比汉

德里克·T·查默斯 廖王葵

[56] 参考文献

WO9721731 A 1997.6.19

WO9838217 A 1998.9.3

"CONSTITUTIVE ACTIVATION OF THEALPHA1B - ADRENERGIC RECEPTOR BY ALL AMINOACID SUBSTITUTIONSAT A SINGLE SITE" KJELSBORG M. A. ET AL. :, JOURNAL OF BIOLOGICAL CHEMISTRY,, Vol. 267 No. 3 1992

"CONSTITUTIVELY ACTIVE G PROTEIN - COUPLEDRECEPTORS: POTENTIAL MECHANISMS OFRECEPTOR ACTIVATION" SCHEER A. ET AL. :, JOURNAL OF RECEPTOR AND SIGNAL TRANSDUCTION RESEARCH,, Vol. 17 No. 1/03 1997

审查员 周 洋

[74] 专利代理机构 中原信达知识产权代理有限责任公司

代理人 杨 青 樊卫民

权利要求书 3 页 说明书 284 页 附图 19 页

[54] 发明名称

非内源的被组成型活化的人 G 蛋白偶联的受体

[57] 摘要

在此公开的是内源人 G 蛋白偶联受体 (GPCR) 的组成型活化的非内源形式, 该受体含有 (a) 下列氨基酸区域 (从 C - 末端到 N - 末端走向) 和/或 (b) 横跨 GPCR 的跨膜 - 6 (TM6) 和细胞内环 - 3 (IC3) 区域的下列核酸序列区域 (3' 到 5' 走向), 分别是: (a) $P^1 AA_{15} X$ 和/或 (b) $P^{\text{密码子}} (AA - \text{密码子})_{15} X_{\text{密码子}}$ 。在最优选的实施例中, P^1 和 $P^{\text{密码子}}$ 分别是内源脯氨酸和编码脯氨酸的内源核酸编码区, 它位于非内源 GPCR 的 TM6 内; AA_{15} 和 (AA - 密码子) 15 分别是 15 个内源氨基酸残基和 15 个编码内源氨基酸残基的密码子; X 和 $X_{\text{密码子}}$ 分别是非内源赖氨酸和编码赖氨酸的非内源核酸编码区, 它位于非内

源 GPCR 的 IC3 内。因为最优选包含这些突变的非内源人 GPCR 被引进哺乳动物细胞并被用作筛选候选化合物, 所以包含这些突变的非内源人 GPCR 在本质上不需要被纯化和分离 (即它们被引入到哺乳动物细胞的细胞膜上), 尽管如此纯化与分离的非内源的人 GPCR 是在本公开的范围之内。

1. 一种创造内源人 G 蛋白偶联受体(GPCR)的非内源的、组成型活化形式的方法，其中所说的内源 GPCR 包括一个跨膜 6 区和一个细胞内环 3 区，该方法包括如下步骤：

(a) 选择在跨膜-6 区含有脯氨酸残基的内源人 GPCR；

(b) 在从羧基末端指向氨基末端方向上，识别从步骤(a)所说的脯氨酸残基起算为第 16 位的内源氨基酸残基；

(c) 把步骤(b)识别的氨基酸残基改造为非内源的氨基酸残基以创造内源人 GPCR 的非内源形式；和

(d) 通过测量非内源形式的胞内信号与由内源 GPCR 诱导的信号相比之后的差异，确定步骤(c)的内源人 GPCR 的非内源形式是否是组成型活化的。

2. 权利要求 1 的方法，其中在从羧基末端指向氨基末端方向上距离跨膜 6 区的所说的脯氨酸残基两个残基的氨基酸残基是色氨酸。

3. 权利要求 1 或 2 的方法，其中在从羧基末端指向氨基末端方向上，从所说的脯氨酸残基起算为第 16 位的内源氨基酸残基已改造成赖氨酸。

4. 权利要求 1 或 2 的方法，其中在从羧基末端指向氨基末端方向上，从所说的脯氨酸残基起算为第 16 位的内源氨基酸残基已改造成丙氨酸。

5. 权利要求 1 或 2 的方法，其中在从羧基末端指向氨基末端方向上，从所说的脯氨酸残基起算为第 16 位的内源氨基酸残基已改造成精氨酸。

6. 权利要求 1 或 2 的方法，其中在从羧基末端指向氨基末端方向

上，从所说的脯氨酸残基起算为第 16 位的内源氨基酸残基已改造成组氨酸。

7. 一种直接鉴定选自非内源的组成型活化的人 G 蛋白偶联受体的反激活剂、激活剂和部分激活剂的化合物的方法，其中所说的受体含有一个跨膜-6 区和细胞内环-3 区，该方法包括权利要求 1 的步骤(a)-(d)，并进一步包括如下步骤：

(e) 用步骤(d)鉴定的非内源的、组成型活化的 GPCR 接触候选化合物；和

(f) 通过测量所说的被接触的受体的化合物效应，确定所说的化合物是否是所说受体的反激活剂、激活剂或部分激活剂。

8. 权利要求 7 的方法，其中在从羧基末端指向氨基末端方向上，从所说的脯氨酸残基起算为第 16 位的内源氨基酸残基已改造成赖氨酸。

9. 权利要求 7 的方法，其中在从羧基末端指向氨基末端方向上，从所说的脯氨酸残基起算为第 16 位的内源氨基酸残基已改造成丙氨酸。

10. 权利要求 7 的方法，其中在从羧基末端指向氨基末端方向上，从所说的脯氨酸残基起算为第 16 位的内源氨基酸残基已改造成精氨酸。

11. 权利要求 7 的方法，其中在从羧基末端指向氨基末端方向上，从所说的脯氨酸残基起算为第 16 位的内源氨基酸残基已改造成组氨酸。

12. 权利要求 7 的方法，其中被直接鉴定的化合物是反激活剂。

-
13. 权利要求 7 的方法，其中被直接鉴定的化合物是激活剂。
 14. 权利要求 7 的方法，其中被直接鉴定的化合物是部分激活剂。
 15. 权利要求 7 的方法，其中通过步骤(f)被直接鉴定为反激活剂，
激活剂或部分激活剂的化合物能配制成药物组合物。

非内源的被组成型活化的人 G 蛋白偶联的受体

特此声明，本发明要求下列在先申请的优先权：美国专利申请号 09/170,496(1998 年 10 月 13 日提出申请)、美国专利申请号 08/839,449 (1997 年 4 月 14 日提出申请，现在已放弃)、美国专利申请号 09/060,188 (1998 年 4 月 14 日提出申请)、美国临时申请号 60/090,783(1998 年 6 月 26 日提出申请)和美国临时申请号 60/095,677(1998 年 8 月 7 日提出申请)。前述的每个申请都以全文引入本申请作参考。

本发明的领域

本专利申请文件所公开的发明涉及跨膜受体，具体地讲，涉及被改造的人 G 蛋白偶联受体(GPCR)，它可被组成型活化。在最优选情况下，被改造的人 GPCR 可用于筛选药用化合物。

本发明的背景

尽管在人体内有很多种类的受体，但到目前为止最丰富和最与治疗有关的是 G 蛋白偶联受体(GPCR)。据估计，在人类基因组内有大约 100,000 个基因，它们中的大约 2%即 2,000 个基因被估计用来编码 GPCR。已识别出与其中的约 100 个 GPCR 结合的内源配体。由于在发现内源 GPCR 和发现它的内源配体之间有显著的时间延迟，因此可预测，其余的 1900 种 GPCR 将在它们的内源配体被识别之前的很长时间内被识别和鉴定。其实，人类基因组计划正在快速测序人类的 100,000 个基因，这表明在今后几年内，其余的人 GPCR 将被完全测序。然而，尽管付出了对人类基因组测序的努力，但仍不清楚科学家如何能够快速、有力和有效率地利用这样的信息来提高和增强人类的健康状态。本发明正是指向这个重要目的。

包括 GPCR 在内，其内源配体已被认识的受体被称为“已知”受

体，内源配体尚不知晓的受体被称为“孤儿”受体。这种区别并不仅是语义上的，尤其是对 GPCR 来说。GPCR 代表着药物产品开发的一个重要领域：60%的处方药物开发自 100 个已知 GPCR 中的大约 20 个。因此，孤儿 GPCR 将是推进药物工业增长、扩张、增强和发展的机会，就像是金子对于 19 世纪晚期的加利福尼亚一样。然而，孤儿受体在涉及新药物发现时有一个严重缺陷。这是因为，发现和开发药物的传统途径既需要利用受体又需要利用它的内源配体。因此，迄今为止，孤儿 GPCR 带给本领域的只是用于发现新药的具有诱惑力且未开发的资源。

在探索潜在治疗药物的传统途径下，一般是受体先被识别。在探索药物的努力开始之前，一般会启动精细、费时和昂贵的程序以识别、分离和产生受体的内源配体---对于每个受体，此过程可花费 3 到 10 年，成本大约为每个受体 500 万美元。在探索药物的传统工作可开始之前，必须先消耗这些时间和资金。这是因为，传统的药物探索技术依赖于所谓的“竞争性结合检测”，其中，假定的治疗剂是通过受体被“筛选”的，其目的是找到这样的化合物，它或者能阻止内源配体与受体结合(拮抗剂)、或者能促进或模仿与受体结合的配体的作用(激活剂)。其总体目标是要识别出这样的化合物，它们在配体与受体结合时能阻止细胞活化(拮抗剂)，或者当配体与受体适当地结合时能促进或增强细胞活性(激活剂)。由定义可知，孤儿 GPCR 的内源配体尚未被识别，因此不可能应用传统药物发现技术去发现针对这些受体的独特的新治疗药物。正如下面将揭示的，本发明能够克服传统药物发现技术所导致的这些和其他严格限制。

GPCR 都具有一个相同的基元(motif)。所有这些受体具有七个由 22 到 24 个疏水氨基酸组成的序列，它们组成七个 α 螺旋，每个 α 螺旋都跨过膜(每个跨度都以数字表示，例如，跨膜-1(TM-1)、跨膜-2(TM-2)等)。跨膜螺旋通过氨基酸链连接，在细胞膜的外部即“细胞外”一边的氨基酸链分别在跨膜-2 和跨膜-3、跨膜-4 和跨膜-5、跨膜

-6 和跨膜-7 之间(这些分别被称为“细胞外”区 1、2 和 3 (EC-1、EC-2 和 EC-3))。在细胞膜内部即“细胞内”一边,跨膜螺旋也通过氨基酸链进行连接,这些氨基酸链分别在跨膜-1 和跨膜-2、跨膜-3 和跨膜-4、跨膜-5 和跨膜-6 之间(这些分别被称为“细胞内”区 1、2 和 3 (IC-1、IC-2 和 IC-3))。受体的“羧基”(“C”)端是在细胞内的区域,受体的“氨基”(“N”)端在细胞外的区域。图 1 描绘了与 G 蛋白偶联的受体的一般结构。

一般来说,当内源配体与受体结合时(经常被称为受体的“活化”),细胞内区域的构象发生变化,以容许细胞内区域和细胞内“G-蛋白”进行偶联。尽管存在其他 G 蛋白,但当前已被识别的 G-蛋白是 Gq、Gs、Gi 和 Go。内源配体活化的 GPCR 与 G-蛋白的偶联引发一个信号级联过程(被称为“信号传导”)。在通常情形下,信号传导最终导致细胞活化或细胞抑制。据认为,受体的 IC-3 环与羧基端都和 G 蛋白相互作用。本发明的一个重要焦点就涉及 GPCR 的跨膜-6 (TM6) 区域和细胞内-3 (IC3) 区域。

在生理条件下,GPCR 存在于细胞膜上,并在“非活化”状态和“活化”状态这两种不同构象之间保持平衡。如在图 2 中所图示的那样,在非活性状态下的受体不能与细胞内信号传导途径相偶联以产生生物学反应。受体构象向活性状态的转变就使它与传导途径相偶联(通过 G-蛋白)并产生生物学反应。

受体可被内源配体或药物等化合物稳定在活性状态。近来的发现提供了除内源配体或药物之外能够促进和稳定受体到活性状态构象的方法,这包括但不限于对受体的氨基酸序列的修饰。这些方法通过模仿与受体结合的内源配体的作用来有效地稳定活性状态的受体。通过如此的配体非依赖性方法形成的稳定被称为“组成型受体活化”。

如上所述,将孤儿受体用于筛选目的是不可能的。这是因为涉及

筛选化合物的传统“信条”需要受体的配体已知。由定义可知，这种方法对于孤儿受体来说是没有用的。因此，在坚持用这个教条的方法去发现药物时实质上在教导本领域去废弃孤儿受体的应用，除非和直到受体的内源配体被发现。考虑到大约有 2000 个 G 蛋白偶联受体且其中的大多数是孤儿受体，如此信条就会与具有创造性的唯一且独特的药物探索方法相对立。

关于不同种类 GPCR 的核酸和/或氨基酸序列的信息总结于下表 A。因为本发明在此公开的一个重要焦点直接指向孤儿 GPCR，所以很多下面所引述的参考都是涉及孤儿 GPCR 的。然而，本列表并非要在法律或在其他意义上进行以下暗示或解释，即在此公开的本发明仅可应用于 GPCR 或者下面所特别列举的 GPCR。此外，一些被分离的受体本身并非是本申请的主题；例如，参考国际互联网上列举 GPCR 的 G 蛋白偶联受体数据库(署名的发明人和受让人都与这个站点没有任何关系)。其他的 GPCR 是由本发明受让人所有的发明申请的主题，它们并没有在下面列出(包括 GPR3、GPR6 和 GPR12；参见美国临时申请号 60/094879)：

表 A

受体名称	参考文献
GPR1	23 基因组学(Genomics) 609 (1994)
GPR4	14 DNA 和 细胞生物学(Cell Biology) 25 (1995)
GPR5	14 DNA 和细胞生物学(Cell Biology) 25 (1995)
GPR7	28 基因组学(Genomics) 84 (1995)
GPR8	28 基因组学(Genomics) 84 (1995)
GPR9	184 实验医学杂志(J.Exp.Med.) 963 (1996)
GPR10	29 基因组学(Genomics) 335 (1995)
GPR15	32 基因组学(Genomics) 462 (1996)
GPR17	70 神经化学杂志(J Neurochem.) 1357 (1998)
GPR18	42 基因组学(Genomics) 462 (1997)
GPR20	187 基因(Gene) 75 (1997)
GPR21	187 基因(Gene) 75 (1997)

GPR22	187 基因(Gene) 75 (1997)
GPR24	398 FEBS Lett. 253 (1996)
GPR30	45 基因组学(Genomics) 607 (1997)
GPR31	42 基因组学(Genomics) 519 (1997)
GPR32	50 基因组学(Genomics) 281 (1997)
GPR40	239 生化生物物理研究通讯(Biochem. Biophys. Res. Commun.) 543 (1997)
GPR41	239 生化生物物理研究通讯(Biochem. Biophys. Res. Commun.) 543 (1997)
GPR43	239 生化生物物理研究通讯(Biochem. Biophys. Res. Commun.) 543 (1997)
APJ	136 基因(Gene) 355 (1993)
BLR1	22 欧洲免疫学杂志(Eur. J. Immunol.) 2759 (1992)
CEPR	231 生化生物物理研究通讯(Biochem. Biophys. Res. Commun.) 651 (1997)
EBI1	23 基因组学(Genomics) 643 (1994)
EBI2	67 病毒学杂志(J. Virol.) 2209 (1993)
ETBR-LP2	424 FEBS Lett. 193 (1998)
GPCR-CNS	54 Brain Res. Mol. Brain Res.152 (1998); 45 基因组学(Genomics) 68 (1997)
GPR-NGA	394 FEBS Lett. 325 (1996)
H9	386 FEBS Lett. 219 (1996)
HBA954	1261 生化生物物理学报(Biochim. Biophys. Acta) 121 (1995)

HG38	247 生化生物物理研究通讯(Biochem. Biophys. Res. Commun.) 266 (1998)
HM74	5 国际免疫学(Int. Immunol.) 1239 (1993)
OGR1	35 基因组学(Genomics) 397 (1996)
V28	163 基因(Gene) 295 (1995)

正如下面详细公开的那样,应用突变盒修饰人 GPCR 内源序列会导致人 GPCR 的组成型活化。这些非内源性的可组成型活化的人 GPCR 尤其可用于筛选候选化合物,以直接识别例如与药物相关的化合物。

本发明的概述

在此公开的是非内源的人 G 蛋白偶联受体，它包含：作为(a)最优选的氨基酸序列区域(从 C-末端到 N-末端走向)和/或(b)最优选核酸序列区域(3'到 5'走向)的横跨 GPCR 的跨膜-6 (TM6) 和细胞内环-3 (IC3) 区域：

(a) $P^1 AA_{15} X$

其中：

(1) P^1 是位于 GPCR 的 TM6 区域内的一个氨基酸残基，其中 P^1 是从(i)内源的 GPCR 的脯氨酸残基和(ii)除脯氨酸之外的非内源的氨基酸残基中选择而来；

(2) AA_{15} 是 15 个氨基酸，它们是从(a)内源的 GPCR 的氨基酸、(b)非内源的氨基酸残基和(c)内源的 GPCR 氨基酸和非内源氨基酸的组合中选择而来，除非位于 GPCR 的 TM6 区域内的 15 个氨基酸残基都不是脯氨酸；和

(3) X 是位于所述 GPCR 的 IC3 区域内的非内源氨基酸残基，优选是从赖氨酸、组氨酸和精氨酸中选择而来，最优选地是赖氨酸，但是如果在 X 位置的内源氨基酸是赖氨酸，那么此时 X 是非赖氨酸的氨基酸，优选是丙氨酸；

和/或

(b) $P^{\text{密码子}} (AA\text{-密码子})_{15} X^{\text{密码子}}$

其中：

(1) $P^{\text{密码子}}$ 是位于 GPCR 的 TM6 区域内的一个核苷酸序列，其中 $P^{\text{密码子}}$ 编码从(i)内源的 GPCR 的脯氨酸残基和(ii)除脯氨酸之外的非内源的氨基酸残基中选择出来的氨基酸；

(2) $(AA\text{-密码子})_{15}$ 是编码 15 个氨基酸的 15 个密码子，这些氨基酸是从(a)内源的 GPCR 的氨基酸、(b)非内源的氨基酸残基和(c)内源 GPCR 的氨基酸和非内源氨基酸的组合中选择而来，除非在位于 GPCR 的 TM6 区域内的 15 个内源密码子都不编码脯氨酸；

(3) $X^{\text{密码子}}$ 是编码位于所述 GPCR 的 IC3 区域内的氨基酸残基的

核苷酸，其中 X_{密码子} 编码非内源氨基酸，它们优选是从由赖氨酸、组氨酸和精氨酸中选择而来，最优选地是赖氨酸，但是当在 X_{密码子} 位置的内源编码区域编码赖氨酸，那么此时 X_{密码子} 编码除赖氨酸以外的氨基酸，优选是丙氨酸。

针对这些序列盒所使用的内源和非内源术语是相对于内源 GPCR 而言。例如，一旦内源脯氨酸残基位于特定 GPCR 的 TM6 区域之内且从它算起的第 16 个氨基酸被识别以用于发生突变来组成型活化该受体，那么也有可能突变内源脯氨酸残基(即，一旦标记物被定位且将发生突变的第 16 个氨基酸被确定，就可能突变标记物本身)，尽管最优选脯氨酸残基不被突变。相似地是，尽管最优选 AA₁₅ 保持它们的内源形式，但这些氨基酸也可被突变。在人 GPCR 的非内源形式中唯一必须被突变的氨基酸是 X，即，由从 P¹ 开始的第 16 个残基构成的内源氨基酸不能保持其内源形式而必须被突变，这正如在此进一步公开的那样。重述一遍，尽管优选在人 GPCR 的非内源形式中，P¹ 和 AA₁₅ 保持它们的内源形式(即，与它们的野生型形式一样)，但一旦 X 被识别和被突变，任何和/或所有的 P¹ 和 AA₁₅ 都可以被突变。这对核苷酸序列也同样适用。那么，当在位置 X 的内源氨基酸是赖氨酸时，在此 GPCR 的非内源形式中的 X 是非赖氨酸的氨基酸，优选是丙氨酸。

因此，作为假设的情形，如果内源 GPCR 在上述位置具有如下的内源氨基酸序列：

P-AACCTTGGRRRDDDE-Q

那么下面的任何一个情形和假设中的序列盒都将落在公开的范围之内(非内源氨基酸以粗体字表示)：

P-AACCTTGGRRRDDDE-K

P-AACCTTHIGPRDDDE-K

P-ADEETTGGPRRDDDE-A
P-LLKFMSTWZLVAAPQ-K
A-LLKFMSTWZLVAAPQ-K

也可能在 AA₁₅ 内加入氨基酸残基，但此方法并没有特别的进展。其实，在最优选的实施方案中，在非内源的人 GPCR 与内源的 GPCR 中唯一的氨基酸差别就是在 X 位置的氨基酸；此氨基酸自身的突变导致该受体的组成型活化。

因此，在特别优选的实施方案中，P¹ 和 P^{密码子} 分别是内源脯氨酸和编码脯氨酸的内源核苷酸编码区；X 和 X^{密码子} 分别是非内源赖氨酸或丙氨酸和编码赖氨酸或丙氨酸的非内源核苷酸编码区，其中最优选是赖氨酸。因为最优选带有这些突变的非内源人 GPCR 包含在哺乳动物细胞内并用于筛选候选化合物，因此，尽管分离和纯化的非内源人 GPCR 是在本发明公开的范围之内，但带有突变的非内源的人 GPCR 本身并不需要纯化和分离(即它们被包含在哺乳动物细胞的细胞膜内)。涉及非内源人 GPCR 的基因-靶向和转基因非人哺乳动物(优选是大鼠和小鼠)也在本发明范围之内；特别是，基因靶向的哺乳动物是最优选的，这是因为可把人 GPCR 的非内源形式导入这些动物，以代替非人哺乳动物中的内源 GPCR 编码区(产生这样的非人哺乳动物的技术是周知的，利用人编码区替代这些非人哺乳动物的蛋白质编码区；例如，参见美国专利号 5,777,194)。

已经发现内源人 GPCR 的这些变化可使 GPCR 组成型活化，以致于非内源的被组成型活化的人 GPCR 尤其可被用来直接筛选候选化合物，而不需要内源配体，这正如在此将要进一步揭示的那样。因此，应用这些材料的方法和用这些方法识别的产品也在如下的公开范围之内。

图示的简短说明

图 1 表示与 G 蛋白偶联的受体的大体结构，标出的数字表示跨膜的螺旋、细胞内环和细胞外环。

图 2 示意典型的 G 蛋白偶联受体的活化和非活化两种状态，和活性状态与第二信使传导途径的偶联。

图 3 是优选的载体 pCMV 的序列图示，其中包括限制酶切位点的位置。

图 4 是进行如下比较所测得的信号图示：pCMV、组成型活化的非内源 GPR30 对于 GPR6 介导的 CRE-Luc 报告基因活化的抑制作用、内源 GPR30 对于 GPR6 介导的 CRE-Luc 报告基因活化的抑制作用。

图 5 是进行如下比较所测得的信号图示：pCMV、组成型活化的非内源 GPR17 对于 GPR3 介导的 CRE-Luc 报告基因活化抑制作用，内源 GPR17 对于 GPR3 介导的 CRE-Luc 报告基因活化的抑制作用。

图 6 是比较 pCMV 对照、内源 APJ 和非内源 APJ 所测得的信号图示。

图 7 提供了与其内源型进行比较的非内源人 5-HT_{2A} 受体产生 IP₃ 的图解说明。

图 8 是 GPR1(8A)、GPR30(8B)和 APJ(8C)的点印迹结果。

详细描述

本科学文献涉及受体并采用一些术语来描述对受体具有不同作用的配体。为了清楚和前后一致，在本发明文献中将由始至终使用下列定义。在这些定义与这些词语的其他定义冲突时，选择下列定义：

激活剂 意味着激活细胞内反应的化合物，此时它们结合受体或促进 GTP 与膜结合。

在此应用的氨基酸缩写列于下表：

丙氨酸	ALA	A
精氨酸	ARG	R
天冬酰胺	ASN	N
天冬氨酸	ASP	D
半胱氨酸	CYS	C
谷氨酸	GLU	E
谷氨酰胺	GLN	Q
甘氨酸	GLY	G
组氨酸	HIS	H
异亮氨酸	ILE	I
亮氨酸	LEU	L
赖氨酸	LYS	K
蛋氨酸	MET	M
苯丙氨酸	PHE	F
脯氨酸	PRO	P
丝氨酸	SER	S
苏氨酸	THR	T
色氨酸	TRP	W
酪氨酸	TYR	Y
缬氨酸	VAL	V

部分激活剂 意味着这样的化合物，它们与受体结合时，激活细胞内反应或者促进 GTP 与膜结合的程度低于激活剂。

拮抗剂 意味着这样的化合物，它和激活剂在同一位点与受体竞争性地结合，但不激活由受体的活性形式引起的细胞内反应，并可因此抑制由激活剂或部分激活剂促进的细胞内反应。拮抗剂在没有激活剂或部分激活剂的情形下并不削弱基本细胞内反应。

候选化合物 意味着一个将经受筛选技术检验的分子(例如但不限于化学化合物)。优选的“候选化合物”并不包括对公众来说已知选自受体的反激活剂、激活剂或拮抗剂的化合物，它们以前已通过非直接的识别方法被确定(“非直接识别的化合物”)；更优选不包括先前已经确定至少在一种哺乳动物中具有治疗效果的已被非直接识别的化合物；并且，最优选不包括先前已经确定的在人体中具有治疗用途的已被非直接识别的化合物。

密码子 意味着一组三个核苷酸(或者核苷酸的等价物)，它们一般包括一个与磷酸基团偶联的核苷(腺苷(A)、鸟苷(G)、胞苷(C)、尿苷(U)和胸苷(T))，当被翻译时，它们编码一个氨基酸。

化合物效应 意味着一个化合物抑制或者刺激受体功能的能力的量度，它与受体结合亲和力相对。一个优选的测定化合物效应的方法是通过测量 $[^{35}\text{S}]\text{GTP}\gamma\text{S}$ 的结合，这一点将在本发明的实施例部分中进一步公开。

被组成型活化的受体 意味着易受组成型受体活化的受体。与本发明在此公开的相一致，一个非内源的被组成型活化的人 G 蛋白偶联受体是突变后包括氨基酸盒 $\text{P}^1\text{AA}_{15}\text{X}$ 的受体，这正如在下面详细描述的那样。

组成型受体活化 意味着不利用它的内源配体或其化学等价物与受体结合的方法而使在活性状态下的受体稳定。优选地是，被本发明的组成型受体活化的 G 蛋白偶联受体与 GPCR 的内源形式相比，对组成型活化所测得的信号的反应有至少 10% 的差异(增高或者降低，如具体情况可能的那样)，更优选在如此比较的反应中有大约 25% 的差异，最优选在如此比较的反应中有大约 50% 的差异。当用于直接识别候选化合物的目的时，最优选信号差异至少为 50%，以使在内源信号和非内源信号之间有足够的差异，从而在被选择的候选化合物之间产生区

别。在最多的情形下，“差异”将是信号的增加；然而，正如下面详细将要叙述的那样，对 Gs-偶联的 GPCR 来说，测量的“差异”优选是降低。

接触 意味着把至少两部分放在一起，无论是在体外系统还是在体内系统中。

直接识别或被直接识别，与术语“候选化合物”相联系，意味着筛选针对组成型活化的 G 蛋白偶联受体的候选化合物。本术语在任何情形下都不应被解释或被理解为被包括或包括术语“非直接地识别”或“非直接地被识别”。

内源 意味着由物种的基因组天然产生的物质。关于内源的 GPCR，意味着由人体、昆虫、植物、细菌或病毒天然产生的物质，这些只作为例证但却不是限制。与之相对比，术语“非内源”在本文中 意味着不是由物种的基因组天然产生的物质。例如在其内源形式下并非组成型活化的受体，当应用在此公开的盒使之突变并因而使之组成型活化时，此受体被最优选地指称为“非内源的被组成型活化的受体”，这只作为例证而不是限制。两个用语都可被用来描述“体内”和“体外”系统。在筛选过程中，内源的或非内源的受体可被用于体外筛选系统，其中受体在哺乳动物的细胞表面表达，这也只作为例证而不是限制。作为进一步的例子而不是限制，当操作哺乳动物的基因组以包括非内源组成型活化受体时，可以通过体内系统筛选候选化合物。

宿主细胞 意味着能在其中插入质粒和/或载体的细胞。在原核宿主细胞情形下，当宿主细胞复制时质粒典型地以自主分子方式复制(在一般情况下，质粒在复制后被分离出来以被引入真核宿主细胞中)；在真核宿主细胞情形下，质粒被整合进宿主细胞的细胞 DNA 中，因而，当真核细胞复制时，质粒复制。为在此公开的本发明的目的，宿

主细胞优选是真核细胞，更优选是哺乳动物细胞，最优选地是从 293、293T 和 COS-7 细胞中选择出来的细胞。

非直接地识别或非直接地被识别 意味着发现药物的传统方法，该方法涉及对内源受体特异的内源配体的识别、筛选针对受体的候选化合物、确定那些干扰或竞争配体-受体相互反应的化合物、测量化合物对至少一个与活化受体相关的第二信使途径影响的效率。

抑制，与用语“反应”相联系，意味着在一个化合物存在时一个反应被降低或阻止，这正好与该化合物不存在时相反。

反激活剂 意味着这样的化合物，它们与内源受体或受体的组成型活化形式结合，并且将由受体的活性形式引发的基本细胞内反应抑制到正常基础水平以下，该活性水平是在没有激活剂或部分激活剂的情况下观察的，或者它们降低 GTP 与膜的结合。与在没有反激活剂情况下的基本反应相比，基本细胞内反应在反激活剂的存在下优选被抑制至少 30%、更优选至少 50%、最优选至少 75%。。

已知受体 意味着其特异的内源配体已被识别的内源受体。

配体 意味着对内源的天然产生的受体特异的内源的天然产生的分子。

关于内源受体的核苷酸和/或氨基酸序列的突变 意味着这些内源序列的特定改造，从而使内源的非组成型活化受体的突变型能造成受体的组成型活化。对于特定序列的等价物，人受体的后续突变型被认为是人受体的首次突变的等价物，如果(a)后续突变型受体的组成型活化水平与受体的首次突变所表明的在本质上一样；和(b)在后续突变型受体和受体的首次突变之间的序列同源性的百分数是至少 80%，更优选地是至少 90%，最优选地是至少 95%。在理想的情况下，考虑到

在此公开的用于进行组成型活化的最优选的盒包括在内源和非内源型 GPCR 之间发生变化的单一氨基酸和/或密码子(即 X 或 X_{密码子})，序列同源性的百分数应是至少 98%。

孤儿受体 意味着这样的内源受体，其特异的内源配体尚未被识别或尚未知。

药物组合物 意味着包括至少一种活性成分的组合物，借助此活性成分可以研究该组合物可在哺乳动物(例如但不限于人体)中特定的效果。本领域的那些普通技术人员将能够理解和正确评价那些适于确定活性成分是否具有基于技术人员需要的预期效果的技术。

质粒 意味着载体和 cDNA 的结合体。一般，为 cDNA 复制和/或表达蛋白质的目的将质粒引进宿主细胞。

刺激，与术语“反应”相联系，意味着当一种化合物存在时比当它不存在时反应增强。

横跨，涉及被定义的核苷酸序列或者被定义的氨基酸序列，意味着该序列位于至少两个不同且明确限定的区域之内。例如，在一个长度为 10 个氨基酸的氨基酸序列中，其中 10 个中的 3 个是在 GPCR 的 TM6 区，其余的 7 个是在 GPCR 的 IC3 区，这 10 个氨基酸就可被描述为横跨 GPCR 的 TM6 和 IC3 区。

针对 cDNA 的载体 意味着能够将至少一个 cDNA 掺入其中且能导入到宿主细胞中的环形 DNA。

下面部分的顺序安排是为了表达效果，而不能被解释为对下面的公开或权利要求的限制。

A.引言

受体的传统研究一直是基于这样的前置假定(基于历史),即内源配体必须首先被识别,然后才能发现可以作用于受体的拮抗剂和其他分子。甚至在拮抗剂被首先发现的情况下,搜索的目光也立即延伸到查找内源配体上去。即使在发现组成型活化受体之后,这种思维模式也一直在受体研究中持续。在此之前没有被认识到的是,是受体的活性状态对发现受体的激活剂、部分激活剂和反激活剂是最有用的。对于那些因为受体的过度活化和不够活化而导致的疾病来说,希望得到的治疗药物是能分别用来减少受体的活性状态或增强受体活性的化合物,而并不需要是对抗内源配体的拮抗剂。这是因为,一个降低或增强活化态受体活性的化合物并不需要结合在和内源配体一样的位点上。因而,正如本发明的一个方法所说的那样,对治疗性化合物的任何搜索可通过筛选针对配体非依赖性活性态的化合物而开始。

筛选针对非内源的组成型活化的 GPCR 的候选化合物,这可直接识别与这些细胞表面受体作用的候选化合物,而根本无需先了解或使用此受体的内源配体。通过确定表达和/或过度表达这些 GPCR 的内源形式的身体内部区域,有可能确定与这些受体的表达和/或过度表达相关的疾病/紊乱状态;这种方法在本发明中得到公开。

B.疾病/紊乱识别和/或选择

最优选采用本发明的材料识别针对非内源的组成型活化的 GPCR 的反激活剂。如此的反激活剂是治疗与这些受体有关的疾病的药物探索中先导化合物的理想候选者。因为可直接识别针对这些受体的反激活剂、部分激活剂或激活剂,因此有可能开发和搜索针对与这些受体有关的疾病和紊乱的药物组合物。例如,检查患病和正常组织样品中这些受体的存在,现在不仅仅是学术研究的问题,也是在孤儿受体的情形下通过识别来寻找内源配体的研究道路上所致力解决的问题。可在健康和患病组织的宽广范围内进行组织检查。如此的组织检查提供了把特异受体与疾病/紊乱相联系的优选第一步骤。

优选内源 GPCR 的 DNA 序列被用来制作探针，用于在组织样品中 GPCR 表达的放射标记 cDNA 或 RT-PCR 识别。在疾病组织中受体的存在，或者与正常组织相比在疾病组织中受体的浓度提高或降低，可被优选地用来识别与那种疾病的关联。用这种方法也可很好地把受体定位于器官的区域。基于受体被定位于其中的特定组织的已知功能，受体假想的功能性角色可被推导出来。

C. “人 GPCR 脯氨酸标记物” 算法规则和非内源的组成型活化的人 GPCR 的形成

在生物技术领域所面临的许多挑战中，包括从一个物种搜集遗传信息并把该信息和其他物种的信息相联系的不可预测性---在本领域中，没有比编码核酸和蛋白质的遗传序列这个问题更能困扰人的。因此，为了一致性并考虑到本领域的高不可预测性，下述发明用的哺乳动物术语局限于人 GPCR---把本发明应用于其他哺乳动物物种，尽管具有潜在可能，但并不仅仅只是生搬硬套式的应用。

一般来说，当企图把从一种相关的蛋白质序列或物种中得到的普遍“规则”应用于其他序列或物种时，本领域一般是求助于序列的对齐比较，即把序列线性化并希望在两个或更多的序列之间发现相同的区域。尽管很有用，但此方法并不经常能产生有意义的信息。在 GPCR 情形下，尽管所有 GPCR 的一般结构基元是相同的，但 TM、EC 和 IC 在长度上的不同导致这样的对齐方法在从一种 GPCR 到另一种时变得很困难。因此，尽管可以期望应用一种普遍方法，例如从一种 GPCR 到另一种的组成型活化方法，但由于从一种到另一种 GPCR 在序列长度、一致性等上存在很大的不同，一种可普遍适用的和实际上成功的突变对齐方法在本质上是是不可能的。作个类比，如此的一种方法与这样一种景况相似：让一位旅行者从 A 点开始旅程，给他很多指向 B 点的不同的地图，但在任何一个地图上却没有任何比例尺或距离的标记物，然后让旅行者仅利用这些地图找出通向目的地 B 的最短和最有效

的路径。在这样的情况下，通过以下手段被简化任务：拥有(a)在每个地图上都有一个共同的“地方标记物”，和(b)测量从每一个地方标记物到目的地 B 的距离的能力，然后，这将容许旅行者选择从开始点 A 通向目的地 B 的最有效的路径。

在本质上，本发明的一个特点是提供在人 GPCR 中的这样一个坐标，它可以容许形成组成型活化的人 GPCR。

正如此技术领域中所评价的那样，细胞的跨膜区域是高度疏水的；因此，应用通常的疏水测绘技术，本领域的技术人员就可确定 GPCR 的 TM 区，特别是 TM6(同样的方法也可用来确定 GPCR 的 EC 区和 IC 区)。已经发现，在人 GPCR 的 TM6 区，一个共同的脯氨酸残基(一般靠近 TM6 的中间)是组成型活化的“标记物”。通过从脯氨酸标记物数过 15 个氨基酸，第 16 个氨基酸(定位在 IC3 环上)在从其内源形式突变为非内源形式时导致受体的组成型活化。为方便起见，我们把这称为“人 GPCR 脯氨酸标记物”算法规则。尽管在此位置的非内源氨基酸可以是任何一种氨基酸，但最优选的非内源氨基酸还是赖氨酸。尽管并不希望被任何理论所束缚，我们还是相信该位置的本身是独特的，并且，本位置上的突变会影响受体发生组成性的活化。

我们注意到，例如，当在第 16 位置上的内源氨基酸已经是赖氨酸时(如在 GPR4 和 GPR32 中的那样)，那么为了让 X 是一个非内源氨基酸，它必不是赖氨酸；因此，在内源 GPCR 的第 16 位位置上是内源赖氨酸残基的情形下，非内源 GPCR 在该位置上优选不是赖氨酸的氨基酸，优选是丙氨酸、组氨酸和精氨酸。进一步注意到，确定了 GPR4 看起来被与 Gs 相联并在其内源形式下活化(数据没有列出)。

因为仅有 20 种天然氨基酸(尽管可以利用非天然形成的氨基酸)，为此第 16 位置的替代而选择特别的非内源氨基酸是可行的，并且容许有效选择适合研究者需要的非内源氨基酸。然而，正如提示，在第

16 位更优选的非内源氨基酸是赖氨酸、组氨酸、精氨酸和丙氨酸，其中赖氨酸是最优选的。可认为本领域的普通技术人员有能力用熟练的方法改造密码子的序列以产生所期望的突变。

也发现，在偶然而并非是经常的情形下，脯氨酸残基标记物在 TM6 中位于 W2 之后(即，W2P¹AA₁₅X)，其中 W 是色氨酸，2 是任何氨基酸残基。

我们的发现否定了对于本领域常常应用不可预测且复杂的序列对齐方法的需求。其实，尽管在本质上是一个规则，但我们的发现的重要性就在于它可以容易的方法应用于人 GPCR 上，其可被本领域技术人员灵巧地简化，并得到独特和高度有用的终产品，即被组成型活化的人 GPCR。因为需要很多年和很多资金来确定人 GPCR 的内源配体（正由人类基因组计划揭示），本发明不仅降低积极探索这种序列信息所必要的时间，也能显著地节约成本。本方法能够真正证实人类基因组计划的重要性，因为它不仅准许应用基因信息来理解 GPCR 在疾病等中的角色，也能提供提高人类健康状况的可能。

D. 候选化合物的筛选

1. GPCR 的筛选测定技术

当一种 G 蛋白受体变为组成型活化时，它与 G 蛋白(例如，Gq、Gs、Gi、Go)偶联并刺激释放 GTP，其后 GTP 与 G 蛋白结合。接着，借助受体在正常情况下失活，G 蛋白作为 GTP 酶慢慢地把 GTP 水解为 GDP，然而，包括本发明的非内源的组成型活化的人 GPCR 在内，组成型活化的受体继续把 GDP 转化为 GTP。GTP 非可水解的类似物 [³⁵S]GTPγS，可被用来监测与表达组成型活化受体的膜上的 G 蛋白加强了结合。据报道，[³⁵S]GTPγS 可被用来监测在配体存在或不存在的情形下 G 蛋白与膜的偶联。在本领域中著名和可行的其他例证中有此种监测的一个例证，它由 Traynor 和 Nahorski 在 1995 年所报道。本测定系统的一个优选的应用是为了初步筛选候选化合物，因为本系统

对所有蛋白-偶联受体一般可行，而不考虑与受体的细胞内结构域相互作用的那一种特别的 G 蛋白。

B 2. 特定的 GPCR 筛选测定技术

C 一旦应用“一般”G 蛋白偶联的受体测定方法(即筛选是激活剂、部分激活剂或反激活剂的化合物的方法)识别出候选化合物，优选进一步筛选以确认作用在受体位点的化合物。例如，应用“一般”测定方法识别的化合物可以不与受体结合，但也可以仅仅从细胞内结构域与 G 蛋白“解偶联”。

a. Gs 和 Gi

Gs 刺激腺苷酸环化酶。另一方面，Gi(和 Go)抑制该酶。腺苷酸环化酶催化 ATP 向 cAMP 的转化；因此，与 Gs 蛋白偶联的组成型活化的 GPCR 与升高的细胞内 cAMP 水平相关联。在另一方面，与 Gi(或 Go)蛋白偶联的组成型活化的 GPCR 与降低的细胞内 cAMP 水平相关联。一般情况参见“突触传导的非直接机制(Indirect Mechanisms of Synaptic Transmission)”，第 8 章，从神经到大脑(From Neuron To Brain)(第三版)，Nichols, J.G.等编，Sinauer Associates, Inc. (1992)。因此，检测 cAMP 的方法可被用来确定一个竞争性的化合物是否是受体的反激活剂(即这样的化合物将降低 cAMP 的水平)等。本领域已知的测定 cAMP 的不同方法可以被利用；最优选的方法依赖于在基于 ELISA 的方法中应用抗-cAMP 的抗体。可被应用的另一类测定方法是一种全细胞第二信使报告基因系统测定法。基因上的启动子驱动由一个特别的基因所编码的蛋白质的表达。环 AMP 通过以下步骤促进基因的表达，即它响应促进 cAMP 的 DNA 结合蛋白或转录因子(CREB)的结合，转录因子接着在被称为 cAMP 效应元件的特别位点与启动子结合并驱动基因表达。报告基因系统可被构建为具有一个启动子，该启动子在报告基因的前面含有多个 cAMP 效应元件，例如 β -半乳糖苷酶或荧光素酶。因而，一个被组成型活化的连接 Gs 的受体引起 cAMP 的积累，cAMP 接着激活报告蛋白质的基因和表达。 β -半乳

糖苷酶或荧光素酶等报告蛋白质可用标准生化方法检测到(Chen 等, 1995)。对于偶联 Gi(或 Go)的 GPCR 来说, 它能降低 cAMP 水平, 一种筛选反激活剂等的方法是基于应用与 Gs 相连接的受体(并因而降低 cAMP 水平), 该筛选方法在实施例部分被公开, 其针对 GPR17 和 GPR30。

b. Go 和 Gq

Go 和 Gq 与磷脂酶 C 的活化相联系, 磷脂酶随后水解磷酸酯 PIP₂, 并释放两种细胞内信使: 二酰甘油(DAG)和肌醇-1,4,5-三磷酸(IP₃)。积累增加的 IP₃ 与 Gq-和 Go-关联的受体相关联。一般情况参见“突触传导的非直接机制(Indirect Mechanisms of Synaptic Transmission)”, 第 8 章, 从神经到大脑(From Neuron To Brain)(第三版), Nichols, J.G.等编, Sinauer Associates, Inc. (1992)。测定 IP₃ 积累的方法可被用来确定一个候选化合物是否是例如针对 Gq-或 Go-关联受体等的反激活剂(即如此的化合物能降低 IP₃ 的水平)。Gq 关联受体也可用 AP1 报告基因测定方法来检测, 因为 Gq 依赖的磷脂酶 C 引起含有 AP1 元件的基因活化; 因而, 活化的 Gq 关联受体将导致如此基因的表达增高, 而其反激活剂将导致如此表达的降低, 激活剂将导致如此表达的升高。进行如此测定的商业可得的方法是可行的。

E. 药物化学

在一般但并非经常的情况下对候选化合物直接识别与通过组合化学技术产生的化合物联合使用, 其中随机制备几千种化合物用于此分析。如此筛选的结果一般将是具有独特中心结构的化合物; 其后, 这些化合物围绕着一个优选的中心结构而被优选进行额外的化学修饰, 以进一步加强其药用性质。这样的技术在该领域中是已知的, 并不需要在本专利文件中详细描述。

F. 药物组合物

为进一步开发而选择出的候选化合物可应用本领域周知的技术制

剂成药物组合物。适宜的药物可接受的载体在本领域中是可得的；例如，参见 Remington's Pharmaceutical Sciences, 第 16 版, 1980, Mack Publishing Co.(Oslo 等编)。

G. 其他应用

尽管公开的非内源人 GPCR 的一个优选的应用是为了直接识别作为反激活剂、激活剂或部分激活剂(优选地作为药物使用)的候选化合物, 这些受体也可被用于研究之用。例如, 导入这些受体的体外或体内系统可被用来阐释和理解受体在正常和患病的人体状况中的作用, 也可理解当它应用于理解信号级联反应时组成型活化的角色。这些非内源受体的一个价值由于其独特的特点是它们作为研究工具的用途被强化, 公开的受体可被用来理解特殊受体在人体中的作用, 即使在其内源配体被识别之前。公开的受体的其他应用对于本领域的技术人员将是明显的, 特别是当他们阅读了本申请文件之后。

实施例

下面的实施例是为了说明而非限制本发明的目的。根据本发明文件的叙述, 可以在基于与 TM6 中的脯氨酸残基有关的位置的人 GPCR 的 IC3 环中应用突变盒以组成型活化受体, 并且, 尽管在此公开了特定的核苷酸和氨基酸序列, 但可认为本领域中的那些普通技术人员拥有这样的能力, 即通过对这些序列的简单的修饰就可得到与下面所报道的相同或基本相似的结果。基于技术人员特殊需要的特殊的序列突变方法是在技术人员的了解范围之内。

实施例 1

制备人内源 GPCR

不同的 GPCR 应用在如下的实施例中。一些人内源 GPCR 在表达载体中提供(如下面致谢), 其他的人内源 GPCR 是用公众可得的序列信息从头合成的。

1. GPR1(GenBank 入藏登记号: U13666)

GPR1 的人 cDNA 序列由 Brian O'Dowd(多伦多大学)在 pRcCMV 中提供。GPR1 cDNA(1.4kb 片段)作为一个 NdeI-XbaI 片段被从 pRcCMV 载体中切下, 并被亚克隆进 pCMV 载体的 NdeI-XbaI 位点(参见图 3)。人 GPR1 的核苷酸(SEQ ID NO: 1)和氨基酸(SEQ ID NO: 2)序列其后被确定和证实。

2. GPR4(GenBank 入藏登记号: L36148, U35399, U21051)

GPR4 的人 cDNA 序列由 Brian O'Dowd(多伦多大学)在 pRcCMV 中提供。GPR1 cDNA(1.4kb 片段)作为一个 ApaI(平端)-XbaI 片段被从 pRcCMV 载体中切下, 并被亚克隆进(把大部分 5'未翻译区除去)pCMV 载体的 HindIII(平端)-XbaI 位点。人 GPR4 的核苷酸序列(SEQ ID NO: 3)和氨基酸(SEQ ID NO: 4)序列其后被确定和证实。

3. GPR5(GenBank 入藏登记号: L36149)

按如下步骤产生人 GPR5 的 cDNA 并将其克隆进 pCMV 表达载体: 以基因组 DNA 为模板, 用 rTth 聚合酶(Perkin Elmer)和制造商提供的缓冲系统进行 PCR, 其中用每个引物 0.25 μ M、4 种核苷酸中每种 0.2mM。循环条件是进行 30 个循环: 94 $^{\circ}$ C 1 分钟; 64 $^{\circ}$ C 1 分钟; 72 $^{\circ}$ C 1.5 分钟。5'PCR 引物包括一个 EcoR I 位点, 其序列为:

5'-TATGAATTCAGATGCTCTAAACGTCCTGC-3' (SEQ ID NO:5)

3'引物包括 BamH I 位点, 其序列为:

5'-TCCGGATCCACCTGCACCTGCGCCTGCACC-3' (SEQ ID NO:6)

1.1kb PCR 片段被用 EcoR I 和 BamH I 消化并被克隆进 pCMV 表达载体的 EcoR I -BamH I 位点。人 GPR5 的核酸(SEQ ID NO:7)和氨基酸(SEQ ID NO:8)序列其后被确定和证实。

4. GPR7(GenBank 入藏登记号: U22491)

按如下步骤产生人 GPR7 的 cDNA 并把其克隆进 pCMV 表达载体: PCR 条件-以基因组 DNA 为模板, 用 rTth 聚合酶(Perkin Elmer)和制造商提供的缓冲系统进行 PCR, 其中使用每个引物 0.25 μ M、4 种核苷酸中每种 0.2mM。循环条件是进行 30 个循环: 94 $^{\circ}$ C 1 分钟; 62 $^{\circ}$ C 1 分钟; 72 $^{\circ}$ C 1 分钟 20 秒。5'PCR 引物包括一个 Hind III 位点, 且其序列为:

5'-GCAAGCTTGGGGGACGCCAGGTCGCCGGCT-3' (SEQ ID NO:9)

3'引物包括 BamH I 位点, 其序列为:

5'-GCGGATCCGGACGCTGGGGGAGTCAGGCTGC-3' (SEQ ID NO:10)

1.1kb PCR 片段被用 Hind III 和 BamH I 消化并被克隆进 pCMV 表达载体的 Hind III-BamH I 位点。人 GPR7 的核酸(SEQ ID NO:11)和氨基酸(SEQ ID NO:12)序列其后被确定和证实。

5. GPR8(GenBank 入藏登记号: U22492)

按如下步骤产生人 GPR8 的 cDNA 并把其克隆进 pCMV 表达载体: 以基因组 DNA 为模板, 用 rTth 聚合酶(Perkin Elmer)和制造商提供的缓冲系统进行 PCR, 其中使用每个引物 0.25 μ M、4 种核苷酸中每种 0.2mM。循环条件是进行 30 个循环: 94 $^{\circ}$ C 1 分钟; 62 $^{\circ}$ C 1 分钟; 72 $^{\circ}$ C 1 分钟 20 秒。5'PCR 引物包括一个 EcoR I 位点, 其序列为:

5'-CGGAATTCGTCAACGGTCCCAGCTACAATG-3' (SEQ ID NO:13)

3'引物包括 BamH I 位点, 其序列为:

5'-ATGGATCCCAGGCCCTTCAGCACCGCAATAT-3' (SEQ ID NO:14)

1.1kb PCR 片段被用 EcoR I 和 BamH I 消化并被克隆进 pCMV 表

达载体的 EcoR I -BamH I 位点。所有测序的 4 个 cDNA 克隆包含可能的多态现象，它涉及第 206 个氨基酸从 Arg 转变为 Gln。暂且不论这个差别，人 GPR8 的核酸(SEQ ID NO:15)和氨基酸(SEQ ID NO:16)序列其后被确定和证实。

6. GPR9(GenBank 入藏登记号: X95876)

按如下步骤产生人 GPR9 的 cDNA 并把其克隆进 pCMV 表达载体：以克隆为模板(由 Brian O'Dowd 提供)，用 pfu 聚合酶(Stratagene)和制造商提供的加有 10%DMSO 的缓冲系统进行 PCR，其中使用每种引物 0.25 μ M、4 种核苷酸每种 0.5mM。循环条件是进行 25 个循环：94 $^{\circ}$ C 1 分钟；56 $^{\circ}$ C 1 分钟；72 $^{\circ}$ C 2.5 分钟。5'PCR 引物包括一个 EcoR I 位点，其序列为：

5'-ACGAATTCAGCCATGGTCCTTGAGGTGAGTGACCACCAA
GTGCTAAAT-3' (SEQ ID NO:17)

3'引物包括 BamH I 位点，其序列为：

5'-GAGGATCCTGGAATGCGGGGAAGTCAG-3' (SEQ ID NO:18)

1.2kb PCR 片段被用 EcoR I 消化并被克隆进 pCMV 表达载体的 EcoR I -Sam I 位点。人 GPR9 的核酸(SEQ ID NO:19)和氨基酸(SEQ ID NO:20)序列其后被确定和证实。

7. GPR9-6(GenBank 入藏登记号: U45982)

按如下步骤产生人 GPR9-6 的 cDNA 并把其克隆进 pCMV 表达载体：以基因组 DNA 为模板，用 rTth 聚合酶(Perkin Elmer)和制造商提供的缓冲系统进行 PCR，其中使用每个引物 0.25 μ M、4 种核苷酸中每种 0.2mM。循环条件是进行 30 个循环：94 $^{\circ}$ C 1 分钟；62 $^{\circ}$ C 1 分钟；72 $^{\circ}$ C 1 分钟 20 秒。5'PCR 引物被用如下序列激酶化：

5'-TTAAGCTTGACCTAATGCCATCTTGTGTCC-3' (SEQ ID
NO:21)

3'引物包括 BamH I 位点，其序列为：

5'-TTGGATCCAAAAGAACCATGCACCTCAGAG-3' (SEQ ID NO:22)

1.2kb PCR 片段被用 BamH I 消化并被克隆进 pCMV 表达载体的 EcoRV-BamH I 位点。人 GPR9-6 的核酸(SEQ ID NO:23)和氨基酸(SEQ ID NO:24)序列其后被确定和证实。

8. GPR10(GenBank 入藏登记号: U32672)

GPR10 的人 cDNA 序列由 Brian O'Dowd(多伦多大学)在 pRcCMV 中提供。GPR10 cDNA(1.3kb 片段)作为一个 EcoRI-XbaI 片段被从 pRcCMV 载体中切下, 并被亚克隆进 pCMV 载体的 EcoRI-XbaI 位点。人 GPR10 的核酸(SEQ ID NO: 25)和氨基酸(SEQ ID NO: 26)序列其后被确定和证实。

9. GPR15 (GenBank 入藏登记号: U34806)

GPR15 的人 cDNA 序列由 Brian O'Dowd(多伦多大学)在 pCDNA3 中提供。GPR15 cDNA(1.5kb 片段)作为一个 HindIII-Bam 片段被从 pCDNA3 载体中切下, 并被亚克隆进 pCMV 载体的 HindIII-Bam 位点。人 GPR15 的核酸(SEQ ID NO: 27)和氨基酸(SEQ ID NO: 28)序列其后被确定和证实。

10. GPR17(GenBank 入藏登记号: Z94154)

按如下步骤产生人 GPR17 的 cDNA 并把其克隆进 pCMV 表达载体: 以基因组 DNA 为模板, 用 rTth 聚合酶(Perkin Elmer)和制造商提供的缓冲系统进行 PCR, 其中使用每个引物 0.25 μ M、4 种核苷酸每种 0.2mM。循环条件是进行 30 个循环: 94 $^{\circ}$ C 1 分钟; 56 $^{\circ}$ C 1 分钟; 72 $^{\circ}$ C 1 分钟 20 秒。5'PCR 引物包括一个 EcoR I 位点, 其序列为:

5'-CTAGAATTCTGACTCCAGCCAAAGCATGAAT-3' (SEQ ID NO:29)

3'引物包括 BamH I 位点, 其序列为:

5'-GCTGGATCCTAAACAGTCTGCGCTCGGCCT-3' (SEQ ID NO:30)

1.1kb PCR 片段被用 EcoR I 和 BamH I 消化并被克隆进 pCMV 表达载体的 EcoR I -BamH I 位点。人 GPR17 的核酸(SEQ ID NO:31)和氨基酸(SEQ ID NO:32)序列其后被确定和证实。

11. GPR18(GenBank 入藏登记号: L42324)

按如下步骤产生人 GPR18 的 cDNA 并把其克隆进 pCMV 表达载体: 以基因组 DNA 为模板, 用 rTth 聚合酶(Perkin Elmer)和制造商提供的缓冲系统进行 PCR, 其中使用每个引物 0.25 μ M、4 种核苷酸中每种 0.2mM。循环条件是进行 30 个循环: 94 $^{\circ}$ C 1 分钟; 54 $^{\circ}$ C 1 分钟; 72 $^{\circ}$ C 1 分钟 20 秒。5'PCR 引物被用如下序列激酶化:

5'-ATAAGATGATCACCCTGAACAATCAAGAT-3' (SEQ ID NO:33)

3'引物包括 EcoR I 位点, 其序列为:

5'-TCCGAATTCATAACATTTCACTGTTTATATTGC-3' (SEQ ID NO:34)

1.0kb PCR 片段被用 EcoR I 消化并被克隆进 pCMV 表达载体的平端-EcoR I 位点。所有 8 个被测序的 cDNA 克隆含有 4 种可能的多态, 其中涉及第 12 位的氨基酸从 Thr 变为 Pro, 第 86 位的氨基酸从 Ala 变为 Glu, 第 97 位的氨基酸从 Ile 变为 Leu, 第 310 位的氨基酸从 Leu 变为 Met。暂且不论这些改变, 人 GPR18 的核酸(SEQ ID NO:35)和氨基酸(SEQ ID NO:36)序列其后被确定和证实。

12. GPR20(GenBank 入藏登记号: U66579)

按如下步骤产生人 GPR20 的 cDNA 并把其克隆进 pCMV 表达载体: 以基因组 DNA 为模板, 用 rTth 聚合酶(Perkin Elmer)和制造商提供的缓冲系统进行 PCR, 其中使用每个引物 0.25 μ M、4 种核苷酸中每

体：以基因组 DNA 为模板，用 rTth 聚合酶(Perkin Elmer)和制造商提供的缓冲系统进行 PCR，其中使用每个引物 0.25 μ M、4 种核苷酸中每种 0.2mM。循环条件是进行 30 个循环：94 $^{\circ}$ C 1 分钟；50 $^{\circ}$ C 1 分钟；72 $^{\circ}$ C 1.5 分钟。5'PCR 引物被用如下序列激酶化：

5'-TCCCCCGGGAAAAAACAACCTGCTCCAAA-3' (SEQ ID NO:45)

3'引物包括 BamH I 位点，其序列为：

5'-TAGGATCCATTTGAATGTGGATTGTTGTTGAAA-3' (SEQ ID NO:46)

1.38kb PCR 片段被用 BamH I 消化并被克隆进 pCMV 表达载体的 EcoRV-BamH I 位点。人 GPR22 的核酸(SEQ ID NO:47)和氨基酸(SEQ ID NO:48)序列其后被确定和证实。

15. GPR24(GenBank 入藏登记号：U71092)

按如下步骤产生人 GPR24 的 cDNA 并把其克隆进 pCMV 表达载体：以基因组 DNA 为模板，用 rTth 聚合酶(Perkin Elmer)和制造商提供的缓冲系统进行 PCR，其中使用每个引物 0.25 μ M、4 种核苷酸中每种 0.2mM。循环条件是进行 30 个循环：94 $^{\circ}$ C 1 分钟；56 $^{\circ}$ C 1 分钟；72 $^{\circ}$ C 1 分钟 20 秒。5'PCR 引物含有具有如下序列的 Hind III 位点：

5'-GTGAAGCTTGCCTCTGGTGCCTGCAGGAGG-3' (SEQ ID NO:49)

3'引物包括 EcoR I 位点，其序列为：

5'-GCAGAATTCCCGGTGGCGTGTGTTGTGGTGCCC-3' (SEQ ID NO:50)

1.3kb PCR 片段被用 Hind III 和 EcoR I 消化并被克隆进 pCMV 表达载体的 Hind III 和 EcoR I 位点。人 GPR24 的核酸(SEQ ID NO:51)和氨基酸(SEQ ID NO:52)序列其后被确定和证实。

16. GPR30(GenBank 入藏登记号: U63917)

按如下步骤产生人 GPR30 的 cDNA 并克隆化: 从基因组 DNA 中扩增 GPR30(1128bp 的长度)的编码序列, 并应用下面的引物:

5'-GGCGGATCCATGGATGTGACTTCCCAA-3' (SEQ ID NO:53)

和

5'-GGCGGATCCCTACACGGCACTGCTGAA-3' (SEQ ID NO:54)。

然后用“TOPO-TA 克隆试剂盒”(Invitrogen, #K4500-01)跟随制造商的指令把扩增的产品克隆进商业可得的载体 PCR2.1(Invitrogen)中。用 BamH I 消化来释放全长 GPR30 插入子, 用琼脂糖凝胶电泳使之从载体中分离, 用 Sephaglas Bandprep™ 试剂盒(Pharmacia, #27-9285-01)按照制造商的指令纯化。人 GPR30 的核酸(SEQ ID NO:55)和氨基酸(SEQ ID NO:56)序列其后被确定和证实。

17. GPR31(GenBank 入藏登记号: U65402)

按如下步骤产生人 GPR31 的 cDNA 并把其克隆进 pCMV 表达载体: 以基因组 DNA 为模板, 用 rTth 聚合酶(Perkin Elmer)和制造商提供的缓冲系统进行 PCR, 其中使用每个引物 0.25 μ M、4 种核苷酸中每种 0.2mM。循环条件是进行 30 个循环: 94 $^{\circ}$ C 1 分钟; 58 $^{\circ}$ C 1 分钟; 72 $^{\circ}$ C 2 分钟。5'PCR 引物含有具有如下序列的 EcoR I 位点:

5'-AAGGAATTCACGGCCGGGTGATGCCATTCCC-3' (SEQ ID NO:57)

3'引物包括 BamH I 位点, 其序列为:

5'-GGTGGATCCATAAACACGGGCGTTGAGGAC-3' (SEQ ID NO:58)

1.0kb PCR 片段被用 EcoR I 和 BamH I 消化并被克隆进 pCMV 表达载体的 EcoR I - BamH I 位点。人 GPR31 的核酸(SEQ ID NO:59)和氨基酸(SEQ ID NO:60)序列其后被确定和证实。

18. GPR32(GenBank 入藏登记号: AF045764)

按如下步骤产生人 GPR32 的 cDNA 并把其克隆进 pCMV 表达载体: 以基因组 DNA 为模板, 用 rTth 聚合酶(Perkin Elmer)和制造商提供的缓冲系统进行 PCR, 其中使用每个引物 0.25 μ M、4 种核苷酸中每种 0.2mM。循环条件是进行 30 个循环: 94 $^{\circ}$ C 1 分钟; 56 $^{\circ}$ C 1 分钟; 72 $^{\circ}$ C 1 分钟 20 秒。5'PCR 引物含有具有如下序列的 EcoR I 位点:

5'-TAAGAATTCCATAAAAATTATGGAATGG-3' (SEQ ID NO:243)

3'引物包括 BamH I 位点, 其序列为:

5'-CCAGGATCCAGCTGAAGTCTTCCATCATTC-3' (SEQ ID NO:244)

1.1kb PCR 片段被用 EcoR I 和 BamH I 消化并被克隆进 pCMV 表达载体的 EcoR I - BamH I 位点。人 GPR32 的核酸(SEQ ID NO:245)和氨基酸(SEQ ID NO:246)序列其后被确定和证实。

19. GPR40(GenBank 入藏登记号: AF024687)

按如下步骤产生人 GPR40 的 cDNA 并把其克隆进 pCMV 表达载体: 以基因组 DNA 为模板, 用 rTth 聚合酶(Perkin Elmer)和制造商提供的缓冲系统进行 PCR, 其中使用每个引物 0.25 μ M、4 种核苷酸中每种 0.2mM。循环条件是进行 30 个循环: 94 $^{\circ}$ C 1 分钟; 65 $^{\circ}$ C 1 分钟; 72 $^{\circ}$ C 1 分钟 10 秒。5'PCR 引物含有具有如下序列的 EcoR I 位点:

5'-GCAGAATTCGGCGGCCCCATGGACCTGCCCCC-3' (SEQ ID NO:247)

3'引物包括 BamH I 位点, 其序列为:

5'-GCTGGATCCCCCGAGCAGTGGCGTTACTTC-3' (SEQ ID NO:248)

1kb PCR 片段被用 EcoR I 和 BamH I 消化并被克隆进 pCMV 表

达载体的 EcoR I - BamH I 位点。人 GPR40 的核酸(SEQ ID NO:249)和氨基酸(SEQ ID NO:250)序列其后被确定和证实。

20. GPR41(GenBank 入藏登记号: AF024688)

按如下步骤产生人 GPR41 的 cDNA 并把其克隆进 pCMV 表达载体: 以基因组 DNA 为模板, 用 rTth 聚合酶(Perkin Elmer)和制造商提供的缓冲系统进行 PCR, 其中使用每个引物 0.25 μ M、4 种核苷酸中每种 0.2mM。循环条件是进行 30 个循环: 94 $^{\circ}$ C 1 分钟; 65 $^{\circ}$ C 1 分钟; 72 $^{\circ}$ C 1 分钟 10 秒。5'PCR 引物含有具有如下序列的 Hind III 位点:

5'-CTCAAGCTTACTCTCTCTCACCAGTGGCCAC-3' (SEQ ID NO:251)

3'引物被用下列序列激酶化:

5'-CCCTCCTCCCCCGGAGGACCTAGC-3' (SEQ ID NO:252)

1kb PCR 片段被用 Hind III 消化并被克隆进 pCMV 表达载体的 Hind III-平端位点。人 GPR41 的核酸(SEQ ID NO:253)和氨基酸(SEQ ID NO:254)序列其后被确定和证实。

21. GPR43(GenBank 入藏登记号: AF024690)

按如下步骤产生人 GPR43 的 cDNA 并把其克隆进 pCMV 表达载体: 以基因组 DNA 为模板, 用 rTth 聚合酶(Perkin Elmer)和制造商提供的缓冲系统进行 PCR, 其中使用每个引物 0.25 μ M、4 种核苷酸中每种 0.2mM。循环条件是进行 30 个循环: 94 $^{\circ}$ C 1 分钟; 65 $^{\circ}$ C 1 分钟; 72 $^{\circ}$ C 1 分钟 10 秒。5'PCR 引物含有具有如下序列的 Hind III 位点:

5'-TTTAAGCTTCCCCTCCAGGATGCTGCCGGAC-3' (SEQ ID NO:255)

3' 引物包括 EcoR I 位点, 其序列为:

5'-GGCGAATTCTGAAGGTCCAGGGAACTGCTA-3' (SEQ ID

NO:256)

1kb PCR 片段被用 Hind III和 EcoR I 消化并被克隆进 pCMV 表达载体的 Hind III- EcoR I 位点。人 GPR43 的核酸(SEQ ID NO:257)和氨基酸(SEQ ID NO:258)序列其后被确定和证实。

22. APJ(GenBank 入藏登记号: U03642)

人 APJ 的 cDNA(在 pRcCMV 载体中)由 Brian O'Dowd(多伦多大学)提供。人 APJ 的 cDNA 作为一个 EcoR I -XbaI(平端)片段被从 pRcCMV 载体中切下, 并被亚克隆进 pCMV 载体的 EcoR I -SmaI 位点。人 APJ 的核苷酸(SEQ ID NO: 61)和氨基酸(SEQ ID NO: 62)序列其后被确定和证实。

23. BLR1(GenBank 入藏登记号: X68149)

按如下步骤产生人 BLR1 的 cDNA 并把其克隆进 pCMV 表达载体: 以胸腺 cDNA 为模板, 用 rTth 聚合酶(Perkin Elmer)和制造商提供的缓冲系统进行 PCR, 其中使用每个引物 0.25 μ M、4 种核苷酸每种 0.2mM。循环条件是进行 30 个循环: 94 $^{\circ}$ C 1 分钟; 62 $^{\circ}$ C 1 分钟; 72 $^{\circ}$ C 1 分钟 20 秒。5'PCR 引物含有具有如下序列的 EcoR I 位点:

5'-TGAGAATTCTGGTGACTCACAGCCGGCACAG-3' (SEQ ID NO:63)

3' 引物包括 BamH I 位点, 其序列为:

5'-GCCGGATCCAAGGAAAAGCAGCAATAAAAGG-3' (SEQ ID NO:64)

1.2kb PCR 片段被用 EcoR I 和 BamH I 消化并被克隆进 pCMV 表达载体的 EcoR I -BamH I 位点。人 BLR1 的核酸(SEQ ID NO:65)和氨基酸(SEQ ID NO:66)序列其后被确定和证实。

24. CEPR(GenBank 入藏登记号: U77827)

按如下步骤产生人 CEPR 的 cDNA 并把其克隆进 pCMV 表达载体: 以基因组 cDNA 为模板, 用 rTth 聚合酶(Perkin Elmer)和制造商提供的缓冲系统进行 PCR, 其中使用每个引物 0.25 μ M、4 种核苷酸中每种 0.2mM。循环条件是进行 30 个循环: 94 $^{\circ}$ C 1 分钟; 65 $^{\circ}$ C 1 分钟; 72 $^{\circ}$ C 1 分钟 20 秒。5'PCR 引物被如下序列激酶化:

5'-CAAAGCTTGAAAGCTGCACGGTGCAGAGAC-3' (SEQ ID NO:67)

3' 引物包括 BamH I 位点, 其序列为:

5'-GCGGATCCCGAGTCACACCCTGGCTGGGCC-3' (SEQ ID NO:68)

1.2kb PCR 片段被用 BamH I 消化并被克隆进 pCMV 表达载体的 EcoRV-BamH I 位点。人 CEPR 的核酸(SEQ ID NO:69)和氨基酸(SEQ ID NO:70)序列其后被确定和证实。

25. EBI1(GenBank 入藏登记号: L31581)

按如下步骤产生人 EBI1 的 cDNA 并把其克隆进 pCMV 表达载体: 以胸腺 cDNA 为模板, 用 rTth 聚合酶(Perkin Elmer)和制造商提供的缓冲系统进行 PCR, 其中使用每个引物 0.25 μ M、4 种核苷酸中每种 0.2mM。循环条件是进行 30 个循环: 94 $^{\circ}$ C 1 分钟; 62 $^{\circ}$ C 1 分钟; 72 $^{\circ}$ C 1 分钟 20 秒。5'PCR 引物包括 EcoR I 位点, 其序列为:

5'-ACAGAATTCCTGTGTGGTTTTACCGCCCAG-3' (SEQ ID NO:71)

3' 引物包括 BamH I 位点, 其序列为:

5'-CTCGGATCCAGGCAGAAGAGTCGCCTATGG-3' (SEQ ID NO:72)

1.2kb PCR 片段被用 EcoR I 和 BamH I 消化并被克隆进 pCMV 表达载体的 EcoR I -BamH I 位点。人 EBI1 的核酸(SEQ ID NO:73)和氨基酸(SEQ ID NO:74)序列其后被确定和证实。

26. EBI2(GenBank 入藏登记号: L08177)

按如下步骤产生人 EBI2 的 cDNA 并把其克隆进 pCMV 表达载体: 以 cDNA 克隆为模板(由 Kevin Lynch 提供, University of Virginia Health Sciences Center; 应用的载体未被本资料识别), 用 pfu 聚合酶(Stratagene)和制造商提供的缓冲系统并辅以 10%DMSO 进行 PCR, 其中使用每个引物 0.25 μ M、4 种核苷酸中每种 0.5mM。循环条件是进行 30 个循环: 94 $^{\circ}$ C 1 分钟; 60 $^{\circ}$ C 1 分钟; 72 $^{\circ}$ C 1 分钟 20 秒。5'PCR 引物包括 EcoR I 位点, 其序列为:

5'-CTGGAATTCACCTGGACCACCAATGGATA-3' (SEQ ID NO:75)

3' 引物包括 BamH I 位点, 其序列为:

5'-CTCGGATCCTGCAAAGTTTGTTCATACAGTT-3' (SEQ ID NO:76)

1.2kb PCR 片段被用 EcoR I 和 BamH I 消化并被克隆进 pCMV 表达载体的 EcoR I -BamH I 位点。人 EBI2 的核酸(SEQ ID NO:77)和氨基酸(SEQ ID NO:78)序列其后被确定和证实。

27. ETBR-LP2(GenBank 入藏登记号: D38449)

按如下步骤产生人 ETBR-LP2 的 cDNA 并把其克隆进 pCMV 表达载体: 以脑 cDNA 为模板, 用 rTth 聚合酶(Perkin Elmer)和制造商提供的缓冲系统进行 PCR, 其中使用每个引物 0.25 μ M、4 种核苷酸中每种 0.2mM。循环条件是进行 30 个循环: 94 $^{\circ}$ C 1 分钟; 65 $^{\circ}$ C 1 分钟; 72 $^{\circ}$ C 1.5 分钟。5'PCR 引物包括 EcoR I 位点, 其序列为:

5'-CTGGAATTCTCCTGCTCATCCAGCCATGCGG-3' (SEQ ID

NO:79)

3' 引物包括 BamH I 位点, 其序列为:

5'-CCTGGATCCCCACCCCTACTGGGGCCTCAG-3' (SEQ ID NO:80)

1.5kb PCR 片段被用 EcoR I 和 BamH I 消化并被克隆进 pCMV 表达载体的 EcoR I -BamH I 位点。人 ETBR-LP2 的核酸(SEQ ID NO:81)和氨基酸(SEQ ID NO:82)序列其后被确定和证实。

28. GHSR(GenBank 入藏登记号: U60179)

按如下步骤产生人 GHSR 的 cDNA 并把其克隆进 pCMV 表达载体: 以海马 cDNA 为模板, 用 TaqPlus Precision 聚合酶(Stratagene)和制造商提供的缓冲系统进行 PCR, 其中使用每个引物 0.25 μ M、4 种核苷酸中每种 0.2mM。循环条件是进行 30 个循环: 94 $^{\circ}$ C 1 分钟; 68 $^{\circ}$ C 1 分钟; 72 $^{\circ}$ C 1 分钟 10 秒。对于第一轮 PCR, 5'PCR 引物序列为:

5'-ATGTGGAACGCGACGCCAGCG-3' (SEQ ID NO:83)

3' 引物序列为:

5'-TCATGTATTAATACTAGATTCT-3' (SEQ ID NO:84)

2 毫升第一轮 PCR 产物被用作第二轮 PCR 模板, 其中 5'引物被如下序列所激酶化:

5'-TACCATGTGGAACGCGACGCCAGCGAAGAGCCGGGGT-3' (SEQ ID NO:85)

3'PCR 引物包括 EcoR I 位点, 其序列为:

5'-CGGAATTCATGTATTAATACTAGATTCTGTCCAGGCCCG-3' (SEQ ID NO:86)

1.1kb PCR 片段被用 EcoR I 消化并被克隆进 pCMV 表达载体的

平端-EcoR I 位点。人 GHSR 的核酸(SEQ ID NO:87)和氨基酸(SEQ ID NO:88)序列其后被确定和证实。

29. GPCR-CNS(GenBank 入藏登记号: AFO17262)

按如下步骤产生人 GPCR-CNS 的 cDNA 并把其克隆进 pCMV 表达载体: 以脑 cDNA 为模板, 用 rTth 聚合酶(Perkin Elmer)和制造商提供的缓冲系统进行 PCR, 其中使用每个引物 0.25 μ M、4 种核苷酸中每种 0.2mM。循环条件是进行 30 个循环: 94 $^{\circ}$ C 1 分钟; 65 $^{\circ}$ C 1 分钟; 72 $^{\circ}$ C 2 分钟。5'PCR 引物包括 Hind III 位点, 其序列为:

5'-GCAAGCTTGTGCCCTCACCAAGCCATGCGAGCC-3' (SEQ ID NO:89)

3' 引物包括 EcoR I 位点, 其序列为:

5'-CGGAATTCAGCAATGAGTTCCGACAGAAGC-3' (SEQ ID NO:90)

1.9kb PCR 片段被用 Hind III 和 EcoR I 消化并被克隆进 pCMV 表达载体的 Hind III-EcoR I 位点。所有 9 个被测序的克隆包含涉及一个 S284C 变化的潜在多态现象。暂且不论这个差别, 人 GPCR-CNS 的核酸(SEQ ID NO:91)和氨基酸(SEQ ID NO:92)序列其后被确定和证实。

30. GPR-NGA(GenBank 入藏登记号: U55312)

按如下步骤产生人 GPR-NGA 的 cDNA 并把其克隆进 pCMV 表达载体: 以基因组 cDNA 为模板, 用 rTth 聚合酶(Perkin Elmer)和制造商提供的缓冲系统进行 PCR, 其中使用每个引物 0.25 μ M、4 种核苷酸中每种 0.2mM。循环条件是进行 30 个循环: 94 $^{\circ}$ C 1 分钟; 56 $^{\circ}$ C 1 分钟; 72 $^{\circ}$ C 1.5 分钟。5'PCR 引物包括 EcoR I 位点, 其序列为:

5'-CAGAATTCAGAGAAAAAAGTGAATATGGTTTTT-3' (SEQ ID NO:93)

3' 引物包括 BamH I 位点, 其序列为:

5'-TTGGATCCCTGGTGCATAACAATTGAAAGAAT-3' (SEQ ID NO:94)

1.3kb PCR 片段被用 EcoR I 和 BamH I 消化并被克隆进 pCMV 表达载体的 EcoR I - BamH I 位点。人 GPR-NGA 的核酸(SEQ ID NO:95)和氨基酸(SEQ ID NO:96)序列其后被确定和证实。

31. H9(GenBank 入藏登记号: U52219)

按如下步骤产生人 HB954 的 cDNA 并把其克隆进 pCMV 表达载体: 以垂体 cDNA 为模板, 用 rTth 聚合酶(Perkin Elmer)和制造商提供的缓冲系统进行 PCR, 其中使用每个引物 0.25 μ M、4 种核苷酸中每种 0.2mM。循环条件是进行 30 个循环: 94 $^{\circ}$ C 1 分钟; 62 $^{\circ}$ C 1 分钟; 72 $^{\circ}$ C 2 分钟。5'PCR 引物包括 Hind III位点, 其序列为:

5'-GGAAAGCTTAACGATCCCCAGGAGCAACAT-3' (SEQ ID NO:97)

3' 引物包括 BamH I 位点, 其序列为:

5'-CTGGGATCCTACGAGAGCATTTTTTCACACAG-3' (SEQ ID NO:98)

1.9kb PCR 片段被用 Hind III和 BamH I 消化并被克隆进 pCMV 表达载体的 Hind III-BamH I 位点。与公开的序列相比较, 还识别出一个不同的同种型并称其为“H9b”, 其在细胞质尾部有一个 12 个 bp 的框插入。两个同种型都含有涉及氨基酸 P320S 和 G448A 改变的潜在多态现象。同种型 H9a 含有另一个涉及氨基酸 S493N 改变的潜在多态现象, 而同种型 H9b 含有另两个额外的涉及氨基酸 I502T 和 A532T(相当于同种型 H9a 的氨基酸 528)改变的潜在多态现象。人 H9 的核酸(SEQ ID NO:99)和氨基酸(SEQ ID NO:100)序列其后被确定和证

实(在下面的部分, 两个同种型都依据人 GPCR 脯氨酸标记物规则进行突变)。

32. HB954(GenBank 入藏登记号: D38449)

按如下步骤产生人 HB954 的 cDNA 并把其克隆进 pCMV 表达载体: 以脑 cDNA 为模板, 用 rTth 聚合酶(Perkin Elmer)和制造商提供的缓冲系统进行 PCR, 其中使用每个引物 0.25 μ M、4 种核苷酸每种 0.2mM。循环条件是进行 30 个循环: 94 $^{\circ}$ C 1 分钟; 58 $^{\circ}$ C 1 分钟; 72 $^{\circ}$ C 2 分钟。5'PCR 引物包括 Hind III 位点, 其序列为:

5'-TCCAAGCTTCGCCATGGGACATAACGGGAGCT-3' (SEQ ID NO:101)

3' 引物包括 EcoR I 位点, 其序列为:

5'-CGTGAATTCCAAGAATTTACAATCCTTGCT-3' (SEQ ID NO:102)

1.6kb PCR 片段被用 Hind III 和 EcoR I 消化并被克隆进 pCMV 表达载体的 Hind III- EcoR I 位点。人 HB954 的核酸(SEQ ID NO:103)和氨基酸(SEQ ID NO:104)序列其后被确定和证实。

33. HG38(GenBank 入藏登记号: AF062006)

按如下步骤产生人 HB38 的 cDNA 并把其克隆进 pCMV 表达载体: 以脑 cDNA 为模板, 用 rTth 聚合酶(Perkin Elmer)和制造商提供的缓冲系统进行 PCR, 其中使用每个引物 0.25 μ M、4 种核苷酸中每种 0.2mM。循环条件是进行 30 个循环: 94 $^{\circ}$ C 1 分钟; 56 $^{\circ}$ C 1 分钟; 72 $^{\circ}$ C 1 分钟 30 秒。进行两次 PCR 反应以分别获得 5'和 3'片段。对于 5'片段, 5'PCR 引物包括 Hind III 位点, 其序列为:

5'-CCCAAGCTTCGGGCACCATGGACACCTCCC-3' (SEQ ID NO:259)

3' 引物包括 BamH I 位点，其序列为：

5'-ACAGGATCCAAATGCACAGCACTGGTAAGC-3' (SEQ ID NO:260)

这个 5' 的 1.5kb PCR 片段被用 Hind III 和 BamH I 消化并被克隆进 pCMV 表达载体的 Hind III-BamH I 位点。对于 3' 引物，5' PCR 引物被如下序列激酶化：

5'-CTATAACTGGGTTACATGGTTTAAC-3' (SEQ ID NO:261)

3' 引物包括 EcoR I 位点，其序列为：

5'-TTTGAATTCACATATTAATTAGAGACATGG-3' (SEQ ID NO:262)

1.4kb 的 3' PCR 片段被用 EcoR I 消化并被亚克隆进 pCMV 表达载体的平端-EcoR I 位点。接着 5' 片段和 3' 片段通过一个共同的 EcoRV 位点被连接在一起，得到全长 cDNA 克隆。人 HG38 的核酸(SEQ ID NO:263)和氨基酸(SEQ ID NO:264)序列其后被确定和证实。

34. HM74(GenBank 入藏登记号：D10923)

按如下步骤产生人 HM74 的 cDNA 并把其克隆进 pCMV 表达载体：以基因组 cDNA 或者胸腺 cDNA 为模板，用 rTth 聚合酶(Perkin Elmer)和制造商提供的缓冲系统进行 PCR，其中使用每个引物 0.25 μ M、4 种核苷酸每种 0.2mM。循环条件是进行 30 个循环：94 $^{\circ}$ C 1 分钟；65 $^{\circ}$ C 1 分钟；72 $^{\circ}$ C 1 分钟 20 秒。5' PCR 引物包括 EcoR I 位点，其序列为：

5'-GGAGAATTCAGTAGGCGAGGCGCTCCATC-3' (SEQ ID NO:105)

3' 引物被如下序列激酶化：

5'-GGAGGATCCAGGAAACCTTAGGCCGAGTCC-3' (SEQ ID

NO:106)

1.3kb PCR 片段被用 EcoR I 消化并被克隆进 pCMV 表达载体的 EcoR I -Sma I 位点。测序的克隆揭示了涉及一个 N94K 改造的潜在多态现象。暂且不论这个差别，人 HM74 的核酸(SEQ ID NO:107)和氨基酸(SEQ ID NO:108)序列其后被确定和证实。

35. MIG(GenBank 入藏登记号: AFO44600 和 AFO44601)

按如下步骤产生人 MIG 的 cDNA 并把其克隆进 pCMV 表达载体: 以基因组 cDNA 为模板, 用 TaqPlus Precision 聚合酶(Stratagene)(第一轮 PCR)或 pfu 聚合酶(Stratagene)(第二轮 PCR)和制造商提供的缓冲系统进行 PCR, 其中使用每个引物 0.25 μ M、4 种核苷酸每种 0.2mM (TaqPlus Precision)或 0.5mM (pfu)。当用 pfu 时, 在缓冲液中包括 10% 的 DMSO。循环条件是进行 30 个循环: 94 $^{\circ}$ C 1 分钟; 65 $^{\circ}$ C 1 分钟; 在 72 $^{\circ}$ C (a)第一轮 PCR 1 分钟和(b)第二轮 PCR 2 分钟。因为在编码区有一个内含子, 分别应用两套引物以产生重叠的 5'和 3'片段。5'片段 PCR 引物是:

5'-ACCATGGCTTGCAATGGCAGTGCGGCCAGGGGGCACT-3'
(外部有义)(SEQ ID NO:109)

和

5'-CGACCAGGACAAACAGCATCTTGGTCACTTGTCTCCGGC-
3'(内部反义)(SEQ ID NO:110)。

3' 片段 PCR 引物为:

5'-GACCAAGATGCTGTTTGTCTGGTCGTGGTGTCTTGGCAT-
3' (内部有义) (SEQ ID NO:111)和

5'-CGGAATTCAGGATGGATCGGTCTCTTGCTGCGCCT-3' (具
有 EcoRI 位点的外部反义) (SEQ ID NO:112)。

通过下列方法把 5'和 3'片段连接在一起: 应用第一轮 PCR 做模

板,并用激酶化的外部有义引物和外部反义引物进行第二轮 PCR。1.2kb PCR 片段被用 EcoR I 消化并被克隆进 pCMV 表达载体的平端-EcoR I 位点。人 MIG 的核酸(SEQ ID NO:113)和氨基酸(SEQ ID NO:114)序列其后被确定和证实。

36.OGR1(GenBank 入藏登记号: U48405)

按如下步骤产生人 OGR1 的 cDNA 并把其克隆进 pCMV 表达载体:以基因组 cDNA 为模板,用 rTth 聚合酶(Perkin Elmer)和制造商提供的缓冲系统进行 PCR,其中使用每个引物 0.25 μ M、4 种核苷酸中每种 0.2mM。循环条件是进行 30 个循环:94 $^{\circ}$ C 1 分钟;65 $^{\circ}$ C 1 分钟;72 $^{\circ}$ C 1 分钟 20 秒。5'PCR 引物被如下序列激酶化:

5'-GGAAGCTTCAGGCCCAAAGATGGGGAACAT-3' (SEQ ID NO:115)

3' 引物包括 BamH I 位点,其序列为:

5'-GTGGATCCACCCGCGGAGGACCCAGGCTAG-3' (SEQ ID NO:116)

1.1kb PCR 片段被用 BamH I 消化并被克隆进 pCMV 表达载体的 EcoRV- BamH I 位点。人 OGR1 的核酸(SEQ ID NO:117)和氨基酸(SEQ ID NO:118)序列其后被确定和证实。

37. 5-羟色胺 5HT_{2A}

编码人内源 5HT_{2A} 受体的 cDNA 通过 RT-PCR 而得到:应用人脑 poly-A⁺ RNA;来自 5'未翻译区的 5'引物,其具有如下 Xho I 限制位点:

5'-GACCTCGAGTCCTTCTACACCTCATC-3'(SEQ ID NO:119)

来自 3'未翻译区具有如下 Xba I 位点的 3'引物:

5'-TGCTCTAGATTCCAGATAGGTGAAAACCTTG-3'(SEQ ID NO:120)

用 TaqPlus™ Precision 聚合酶(Stratagene)或 rTth™ 聚合酶(Perkin Elmer)和制造商提供的缓冲系统进行 PCR，其中使用每个引物 0.25μM、4 种核苷酸每种 0.2mM。循环条件是进行 30 个循环：94℃ 1 分钟；57℃ 1 分钟；72℃ 2 分钟。1.5kb 的 PCR 片段被用 Xba I 消化并被亚克隆进 pBluescript 的 EcoRV-Xba I 位点。得到的 cDNA 克隆被完全测序，发现它编码被公布序列的两种氨基酸的变化。第一个是在 N-末端细胞外结构域的一个 T25N 突变；第二个是一个 H452Y 突变。因为使用两个不同商业来源(从 Stratagene 得到 TaqPlus™ 和从 Perkin Elmer 得到的 rTth™)的 Taq 多聚酶通过两个独立的 PCR 反应得到 cDNA 克隆且 cDNA 克隆含有相同的两个突变，所以这些突变好象是代表序列多态现象而不是 PCR 错误。除了这些例外，人 5HT_{2A} 的核酸(SEQ ID NO:121)和氨基酸(SEQ ID NO:122)序列其后被确定和证实。

38. 5-羟色胺 5HT_{2C}

编码人内源 5HT_{2C} 受体的 cDNA 通过 RT-PCR 而得到：其中应用人脑 poly-A⁺ RNA。从 5'和 3'未翻译区得到 5'和 3'引物，其序列为：

5'-GACCTCGAGGTTGCTTAAGACTGAAGC-3'(SEQ ID NO:123)

5'-ATTTCTAGACATATGTAGCTTGTACCG-3'(SEQ ID NO:124)

人 5HT_{2C} 的核酸(SEQ ID NO:125)和氨基酸(SEQ ID NO:126)序列其后被确定和证实。

39. V28(GenBank 入藏登记号：U20350)

按如下步骤产生人 V28 的 cDNA 并把其克隆进 pCMV 表达载体：以脑 cDNA 为模板，用 rTth 聚合酶(Perkin Elmer)和制造商提供的缓冲系统进行 PCR，其中使用每个引物 0.25μM、4 种核苷酸每种 0.2mM。循环条件是进行 30 个循环：94℃ 1 分钟；65℃ 1 分钟；72℃ 1 分钟 20 秒。5'PCR 引物引物包括 Hind III位点，其序列为：

5'-GGTAAGCTTGGCAGTCCACGCCAGGCCTTC-3' (SEQ ID

NO:127)

3' 引物包括 EcoR I 位点，其序列为：

5'-TCCGAATTCTCTGTAGACACAAGGCTTTGG-3' (SEQ ID NO:128)

1.1kb PCR 片段被用 Hind III和 EcoR I 消化并被克隆进 pCMV 表达载体的 Hind III-EcoR I 位点。人 V28 的核酸(SEQ ID NO:129)和氨基酸(SEQ ID NO:130)序列其后被确定和证实。

实施例 2

制备非内源的人 GPCR

1.定点诱变

把建立在此处公开的人 GPCR 脯氨酸标记物方法基础上的诱变使用于前述的内源人 GPCR，应用 Transformer Site-Directed Mutagenesis 试剂盒(Clontech)并按照生产商的说明进行。对于此诱变方法，使用一个突变探针和一个选择标记物探针(除非另外说明，SEQ ID NO: 132 的探针始终是同样的)，用于特定序列的这些探针序列列于下表 B(括号中的数字是 SEQ ID NO)。为方便起见，引入人 GPCR 的密码子突变也被以标准的形式标出：

表 B

受体鉴定者 (密码子突变)	突变探针序列(5'-3') (SEQ ID NO)	选择标记物探针序列 (5'-3') (SEQ ID NO)
GPR1 (F245K)	GATCTCCAGTAGGCATA <u>AAGT</u> GGACAATTCTGG (131)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAG (132)
GPR4 (K223A)	AGAAGGCCAAGATCG <u>CGCGG</u> CTGGCCCTCA (133)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT
GPR5 (V224K)	CGGCGCCACCGCACG <u>AAAAA</u> GCTCATCTTC	CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT

GPR7 (T250K)	(134) GCCAAGAAGCGGGTGAAGTT CCTGGTGGTGGCA (135)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
GPR8 (T259K)	CAGGCGGAAGGTGAAAGTCC TGGTCCTCGT (136)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
GPR9 (M254K)	CGGCGCTGCGGGCCAGCG GCTGGTGGTGGTG (137)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
GPR9-6 (L241K)	CCAAGCACAAAGCCAGAGAA GTGACCATCAC (138)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
GPR10 (F276K)	GCGCCGGCGCACCAAAATGCT TGCTGGTGGT (139)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
GPR15 (I240K)	CAAAAAGCTGAAGAAATCTA AGAAGATCATCTTTATTGTCG (140)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
GPR17 (V234K)	CAAGACCAAGGCAAAACGCA TGATCGCCAT (141)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
GPR18 (I231K)	GTCAAGGAGAAGTCCAAAAG GATCATCATC (142)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
GPR20 (M240K)	CGCCGCGTGCGGGCCAGCA GCTCCTGCTC (143)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
GPR21 (A251K)	CCTGATAAGCGCTATAAAAT GGTCCTGTTTCGA (144)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
GPR22 (F312K)	GAAAGACAAAAGAGAGTCA AGAGGATGTCCTTTATTG (145)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
GPR24 (T304K)	CGGAGAAAGAGGGTGA AAC GCACAGCCATCGCC (146)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
GPR30 (L258K)	其它途径：见下文	其它途径：见下文
GPR31 (Q221K)	AAGCTTCAGCGGGCCAGGC ACTGGTCACC (147)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
GPR32 (K255A)	CATGCCAACCGGCCGCGAG GCTGCTGCTGGT (279)	ACCAGCAGCAGCCTCGCGGG CCGGTTGGCATG (280)
GPR40 (A223K)	CGGAAGCTGCGGGCCAAATG GGTGGCCGGC (265)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
GPR41	CAGAGGAGGGTGAAGGGGCT GTTGGCG	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT

(A223K)	(266)	
GPR43 (V221K)	GGCGGCGCCGAGCCAAGGGG CTGGCTGTGG (267)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT
APJ (L247K)	其它途径; 见下文	其它途径; 见下文
BLR1 (V258K)	CAGCGGCAGAAAGGCCAAAA GGGTGGCCATC (148)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT
CEPR (L258K)	CGGCAGAAAGCGAAGCGCAT GATCCTCGCG (149)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT

EBI1 (I262K)	GAGCGCAACAAGGCCAAAA AGGTGATCATC (150)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT
EBI2 (L243K)	GGTGTAAACAAAAAGGCTAA AAACACAATTATTCTTATT (151)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT
ETBR-LP2 (N358K)	GAGAGCCAGCTCAAGAGCAC CGTGGTG (152)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT
GHSR (V262K)	CCACAAGCAAACCAAGAAAA TGCTGGCTGT (153)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT
GPCR-CNS (N491K)	CTAGAGAGTCAGATGAAGTG TACAGTAGTGGCAC (155)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT
GPR-NGA (I275K)	CGGACAAAAGTGAAAACTAA AAAGATGTTCTTCATT (156)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT
H9a and H9b (F236K)	GCTGAGGTTGCAATAAACT AACCATGTTTGTG (157)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT
HB954 (H265K)	GGGAGGCCGAGCTGAAAGCC ACCCTGCTC (158)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT
HG38 (V765K)	GGGACTGCTCTATGAAAAAA CACATTGCCCTG (268)	CATCAAGTGTATCATGTGCC AAGTACGCCC (154)

HM74 (I230K)	CAAGATCAAGAGAGCCA CCTTCATCATG (159)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT
MIG (T273K)	CCGGAGACAAGTGAAGAAG ATGCTGTTGTG (160)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT
OGR1 (Q227K)	GCAAGGACCAGATCAAGCGG CTGGTGCTCA (161)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAGT
5-羟色胺 5HT _{2A} (C322K)	其它途径; 见下文	其它途径; 见下文
5-羟色胺 5HT _{2C} (S310K)	其它途径; 见下文	其它途径; 见下文

V28 (I230K)	CAAGAAAGCCAAAGCCAAG AAACTGATCCTTCTG (162)	CTCCTTCGGTCCTCCTATCGT TGTCAGAAAGT
----------------	---	--------------------------------------

然后，非内源的人 GPCR 被测序，得到且被证实的核酸和氨基酸序列被列于本发明所附的“序列列表”，如下的表 C 为其摘要：

表 C

突变的 GPCR	核酸序列列表	氨基酸序列列表
GPR1 (F245K)	SEQ.ID.NO.: 163	SEQ.ID.NO.: 164
GPR4 (K223A)	SEQ.ID.NO.: 165	SEQ.ID.NO.: 166
GPR5 (V224K)	SEQ.ID.NO.: 167	SEQ.ID.NO.: 168
GPR7 (T250K)	SEQ.ID.NO.: 169	SEQ.ID.NO.: 170
GPR8 (T259K)	SEQ.ID.NO.: 171	SEQ.ID.NO.: 172
GPR9 (M254K)	SEQ.ID.NO.: 173	SEQ.ID.NO.: 174
GPR9-6 (L241K)	SEQ.ID.NO.: 175	SEQ.ID.NO.: 176
GPR10 (F276K)	SEQ.ID.NO.: 177	SEQ.ID.NO.: 178
GPR15 (I240K)	SEQ.ID.NO.: 179	SEQ.ID.NO.: 180
GPR17 (V234K)	SEQ.ID.NO.: 181	SEQ.ID.NO.: 182
GPR18 (I231K)	SEQ.ID.NO.: 183	SEQ.ID.NO.: 184
GPR20 (M240K)	SEQ.ID.NO.: 185	SEQ.ID.NO.: 186
GPR21 (A251K)	SEQ.ID.NO.: 187	SEQ.ID.NO.: 188
GPR22 (F312K)	SEQ.ID.NO.: 189	SEQ.ID.NO.: 190
GPR24 (T304K))	SEQ.ID.NO.: 191	SEQ.ID.NO.: 192
GPR30	SEQ.ID.NO.: 193	SEQ.ID.NO.: 194

(L258K)		
GPR31 (Q221K)	SEQ.ID.NO.: 195	SEQ.ID.NO.: 196
GPR32 (K255A)	SEQ.ID.NO.: 269	SEQ.ID.NO.: 270
GPR40 (A223K)	SEQ.ID.NO.: 271	SEQ.ID.NO.: 272
GPR41 (A223K)	SEQ.ID.NO.: 273	SEQ.ID.NO.: 274
GPR43 (V221K)	SEQ.ID.NO.: 275	SEQ.ID.NO.: 276
APJ (L247K)	SEQ.ID.NO.: 197	SEQ.ID.NO.: 198
BLR1 (V258K)	SEQ.ID.NO.: 199	SEQ.ID.NO.: 200
CEPR (L258K)	SEQ.ID.NO.: 201	SEQ.ID.NO.: 202
EBI1 (I262K)	SEQ.ID.NO.: 203	SEQ.ID.NO.: 204
EBI2 (L243K)	SEQ.ID.NO.: 205	SEQ.ID.NO.: 206
ETBR-LP2 (N358K)	SEQ.ID.NO.: 207	SEQ.ID.NO.: 208
GHSR (V262K)	SEQ.ID.NO.: 209	SEQ.ID.NO.: 210
GPCR-CNS (N491K)	SEQ.ID.NO.: 211	SEQ.ID.NO.: 212
GPR-NGA (I275K)	SEQ.ID.NO.: 213	SEQ.ID.NO.: 214
H9a (F236K)	SEQ.ID.NO.: 215	SEQ.ID.NO.: 216
H9b (F236K)	SEQ.ID.NO.: 217	SEQ.ID.NO.: 218
HB954 (H265K)	SEQ.ID.NO.: 219	SEQ.ID.NO.: 220
HG38 (V765K)	SEQ.ID.NO.: 277	SEQ.ID.NO.: 278
HM74 (I230K)	SEQ.ID.NO.: 221	SEQ.ID.NO.: 222
MIG (T273K)	SEQ.ID.NO.: 223	SEQ.ID.NO.: 224
OGR1 (Q227K)	SEQ.ID.NO.: 225	SEQ.ID.NO.: 226
5-羟色胺 5HT _{2A} (C322K)	SEQ.ID.NO.: 227	SEQ.ID.NO.: 228
5-羟色胺 5HT _{2C} (S310K)	SEQ.ID.NO.: 229	SEQ.ID.NO.: 230
V28 (I230K)	SEQ.ID.NO.: 231	SEQ.ID.NO.: 232

2. 利用脯氨酸标记物算法规则的其他突变方法：APJ、5-羟色胺 5HT_{2A}、5-羟色胺 5HT_{2C} 和 GPR30

尽管上述定点诱变方法是特别优选地，但其他方法也可以用来创造如此的突变；那些熟悉本领域的人员知道用所选择的方法突变 GPCR 以适合技术人员的特别需要。

a. APJ

制备非内源的人 APJ 受体通过突变 L247K 而完成。合成两个含有此突变的寡核苷酸：

5'-GGCTTAAGAGCATCATCGTGGTGCTGGTG-3' (SEQ ID NO: 233)

5'-GTCACCACCAGCACACGATGATGCTCTTAAGCC-3' (SEQ ID NO: 234)

两个寡核苷酸被退火，并被用来取代人的内源 APJ 的 NaeI-BstE II 片段，以产生非内源的人 APJ。

b. 5-羟色胺 5HT_{2A}

包含点突变 C322K 的 cDNA 通过利用包括氨基酸 322 的限制酶位点 Sph I 构建。含有 C322K 突变的引物：

5'-CAAAGAAAGTACTGGGCATCGTCTTCTTCCT-3' (SEQ ID NO: 235)

与从受体的 3' 未翻译区得到的引物一起应用：

5'-TGCTCTAGATTCCAGATAGGTGAAAACCTTG-3' (SEQ ID NO: 236)

以进行 PCR(在上述条件下)。得到的 PCR 片段然后被用来经 T4 多聚酶补平的 Sph I 位点去替代内源 5HT_{2A} cDNA 的 3' 末端。

c. 5-羟色胺 5HT_{2C}

包含 S310K 突变的 cDNA 通过用编码目的突变的合成双链寡核苷酸去替代包括氨基酸 310 的 Sty I 限制片段而构建。应用的意义链具有如下序列：

5'-CTAGGGGCACCATGCAGGCTATCAACAATGAAAGAAAAG
CTAAGAAAGTC-3'(SEQ ID NO: 237)

应用的反义链具有如下序列：

5'-CAAGGACTTTCTTAGCTTTTCTTTCATTGTTGATAGCCTGC
ATGGTGCCC-3'(SEQ ID NO: 238)

d. GPR30

在产生非内源 GPR30 之前，几个独立的 pCR2.1/GPR30 分离物被整体测序以识别不发生经 PCR 产生的突变的克隆。没有突变的克隆用 EcoR I 消化，并通过用 EcoR I 消化 pCI-Neo 并把从 pCR2.1/GPR30 得到的 EcoR I -释放的 GPR30 片段亚克隆，将内源 GPR30 cDNA 片段转移进由 CMV 驱动表达的质粒 pCI-Neo(Promega)，以产生 pCI/GPR30。其后，按照制造商的说明，用 Quick-Change™ 定点诱变试剂盒(Stratagene, #200518)把位于密码子 258 上的亮氨酸突变为赖氨酸，引物如下：

5'-CGGCGGCAGAAGGCGAAACGCATGATCCTCGCGGT-
3'(SEQ ID NO: 239)和

5'-ACCGCGAGGATCATGCGTTTCGCCTTCTGCCGCCG-3'(SEQ
ID NO:240)

实施例 3

(内源和突变的)受体表达

尽管在本领域中有多种细胞可用于蛋白质的表达，但最优选应用的是哺乳动物细胞。据预测，其基本原因是实用性，即例如表达 GPCR

的酵母细胞的应用，有可能把一种非哺乳动物细胞引入到程序中，此细胞可能不(其实，对于酵母来说，是不)包括偶联受体、遗传机制和分泌途径，而这些是经过进化用于哺乳动物系统的。因此，在非哺乳动物细胞中得到的结果，尽管是可能有用的，但并不如从哺乳动物细胞中得到的结果优选。在哺乳动物细胞中，COS-7、293 和 293T 细胞是特别优选的，尽管应用的特定哺乳动物细胞可按技术人员的特别需要而被判定。

除非在此说明，应用如下的程序表达内源和非内源人 GPCR。表 D 列举用于表达 GPCR 的哺乳动物细胞和数量。

表 D

受体名称 (内源或非内源的)	哺乳动物细胞 (应用的数量)
GPR17	293(2×10^4)
GPR30	293(4×10^4)
APJ	COS-7(5×10^6)
ETBR-LP2	293(1×10^7)
	293T(1×10^7)
GHSR	293(1×10^7)
	293T(1×10^7)
MIG	293(1×10^7)
5-羟色胺 5HT _{2A}	293T(1×10^7)
5-羟色胺 5HT _{2C}	293T(1×10^7)

第一天，哺乳动物细胞被接种到板上。第二天，准备两支试管(比例是每板用于一支试管)：通过混合 20 μ g DNA(例如 pCMV 载体、带有内源受体 cDNA 的 pCMV 载体、带有非内源受体 cDNA 的 pCMV 载体)在 1.2 ml 无血清的 DMEM(Irvine Scientific, Irvine, CA)来制备试管 A；通过混合 120 μ l lipofectamine (Gibco BRL)在 1.2ml 无血清 DMEM

中制备试管 B。把试管 A 和 B 互倾混合(几次),然后在室温下温育 30-45 分钟。组合物被称为“转染组合物”。植出的细胞用 1XPBS 洗涤,然后加入 10ml 无血清的 DMEM。把 2.4ml 转染组合物加入到细胞中去,然后在 37°C/5% CO₂ 下温育 4 小时。接着通过抽吸移去转染组合物,然后加入 25ml 的 DMEM/10%胎牛血清。接着细胞在 37°C/5% CO₂ 温育。72 小时后收获细胞并用来进行分析。

1. Gi 偶联受体: 与 Gs 偶联受体共转染

对于 GPR320 来说,已经确定本受体与 G 蛋白 Gi 偶联。已知 Gi 抑制腺苷酸环化酶,这对 ATP 向 cAMP 的转化催化是必须的。因此,非内源的、被组成型活化的 GPR30 将被期望与 cAMP 水平的下降相关。尽管可通过测定下降的 cAMP 水平对非内源的、可被组成型活化的 GPR30 进行测定证实,但它可通过联合应用 Gs 偶联受体而优选地被测定。例如,Gs 偶联的受体将刺激腺苷酸环化酶,并因而与 cAMP 的升高相关联。本专利申请的受让人已经发现孤儿受体 GPR6 是一个内源的、被组成型活化的 GPCR。GPR6 与 Gs 偶联。因此,当被共转染时,可以很容易证实由一个假想的 GPR30 突变导致其组成型活化:即当与内源的、被组成型活化的 GPR6/非内源的、被组成型活化的 GPR30(这导致一个相对更低水平的 cAMP)相比时,内源的、被组成型活化的 GPR6/内源的非组成型活化的 GPR30 细胞将导致 cAMP 水平的升高。探查 cAMP 的测定方法可用来确定一个候选化合物是否是例如针对 Gs 关联受体的反激活剂(即这样的化合物将降低 cAMP 的水平)或是 Gi 关联受体(或 Go 关联受体)的反向激活剂(即此候选化合物将增加 cAMP 的水平)。在本领域中已知有多种测量 cAMP 的方法可被应用;一个优选的方法依赖于抗-cAMP 抗体的应用。另一个最优选的方法是利用全细胞第二信使报告基因系统试验。基因上的启动子引发特定基因编码的蛋白质表达。环 AMP 促进基因表达的过程,是它先促进响应 cAMP 的 DNA 结合蛋白质或转录因子(CREB)的结合,其后它们在被称为是 cAMP 效应元件的特殊位点与启动子结合并促进基因的表达。报告基因系统可被构建为含有一个启动子,该启动

子在报告基因如 β -半乳糖苷酶或荧光素酶前含有多个 cAMP 效应元件。因而，一个被活化的受体例如 GPR6 就引起 cAMP 的积聚，这然后又活化基因和促进报告蛋白质的表达。最优选 293 细胞与 GPR6(或另外一个 Gs 偶联受体)和 GPR30(或另外一个 Gi 偶联受体)质粒共转染，优选比例为 1: 1，最优选比例为 1: 4。因为 GPR6 是一个内源的、被组成型活化的受体并可刺激 cAMP 的产生，所以 GPR6 可强烈地激活报告基因和它的表达。报告蛋白质例如 β -半乳糖苷酶或荧光素酶可接着被用标准生化测定方法检测到(Chen 等，1995)。内源的、组成型活化的 GPR6 与内源的、非组成型活化的 GPR30 共转染可导致荧光素酶报告蛋白的增加；相反地，内源的、组成型活化的 GPR6 与非内源的、组成型活化的 GPR30 共转染可导致荧光素酶表达的急剧下降。在本领域中已知几个报告基因质粒并可用于测量第二信使试验。据认为，对于熟练的技术人员来说，很容易基于技术人员的特别需要为一个特别的基因表达确定合适的报告质粒。尽管可将多种可得的细胞用于表达，但哺乳动物细胞是最优选的，在这些类型中，293 细胞是最优选的。293 细胞被报告基因质粒 pCRE-Luc/GPR6 和非内源的、组成型活化的 GPR30 转染，其中使用 Mammalian Transfection™试剂盒(Stratagene, #200285) CaPO₄ 沉淀程序，并依照制造商的指令(关于公开的内源 GPR6 序列，参见 28 Genomics 347(1995))。沉淀含有 400ng 报告基因、80ng CMV-表达质粒(具有 1: 4 的 GPR6 对内源 GPR30 或非内源 GPR30 的比例)和 20ng CMV-SEAP(编码分泌的碱性磷酸酶的转染控制质粒)。把 50%的沉淀分到 96 孔组织培养板(包含 4×10^4 细胞/孔)的 3 个孔中；其余的 50%丢弃。第二天上午换培养液。转染开始 48 小时后，细胞被裂解，并按每个制造商的指令用 Luclite™试剂盒(Packard, Cat.#6016911)、Trilux 1450 Microbeta™液体闪烁和发光计数器(Wallac)检测荧光素酶的活性。用 GraphPad Prism 2.0a(GraphPad Software Inc.)分析数据。

对于已被确定是偶联 Gi 的 GPR17 来说，可基于在本质上是应用另一个偶联 Gs 的内源受体 GPR3，来对前述方法加以改进并使用。(参

见 23 Genomics 609(1994)和 24 Genomics 391(1994))。最优选应用 293 细胞。这些细胞被植到 96 孔培养板上,密度为每孔 2×10^4 细胞,在第二天按照制造商的指令用 Lipofectamine Reagent (BRL)转染。按如下程序为每 6 孔的转染制备一个 DNA/脂的组合物:在 100 μ l DMEM 中的 260ng 质粒 DNA 与在 100 μ l DMEM 中的脂轻轻混合(260ng 质粒 DNA 含有 200ng 8 $\times$ CRE-Luc 报告质粒(参见下面),50ng pCMV 含有内源受体或非内源受体或仅有 pCMV,10ng GPRS 表达质粒(GPRS 在 pcDNA3(Invitrogen)中)。8 $\times$ CRE-Luc 报告基因质粒按如下方法制备:通过在 p β gal-Basic 载体(Clontech)中的 BglV-Hind III位点克隆大鼠生长激素释放抑制因子启动子(-71/+51),得到载体 SRIF- β -gal。用腺病毒模板 AdpCF126CCRE8(参见 7 Human Gene Therapy 1883(1996))通过 PCR 得到 cAMP 效应元件的 8 个拷贝,把它克隆进 SRIF- β -gal 载体的 Kpn-BglV 位点,得到 8 $\times$ CRE- β -gal 报告基因载体。8 $\times$ CRE-Luc 报告基因质粒是通过在 HindIII-BamH I 位点用从 pGL3-基本载体(Promega)中得到的荧光素酶基因置换 8 $\times$ CRE- β -gal 报告基因载体中的 β -半乳糖苷酶基因而得到。在室温温育 30 分钟之后,DNA/脂组合物用 400 μ l DMEM 稀释,把 100 μ l 得到的被稀释的组合物加到每个孔中。在细胞培养温育箱中温育 4 小时之后,在每个孔中加入带有 10% FCS 的 100 μ l DMEM。第二天上午,被转化的细胞被换成每孔使用带有 10% FCS 的 200 μ l DMEM。8 小时之后,在一次 PBS 洗涤之后,每孔被换成 100 μ l 无酚红的 DMEM。下一天用 LucLite™ 报告基因测定试剂盒(Packard),按照制造商的指令测定荧光素酶活性,在 1450 Microbeta™ 闪烁发光计数器(Wallac)上记数。

图 4 表明,组成型活化的 GPR30 抑制 GPR6 介导的 CRE-Luc 报告基因在 293 细胞中的活化。在表达载体 pCMV 中,测定荧光素酶为大约 4.1 相对光单位。内源 GPR30 表达的荧光素酶是在大约 8.5 相对光单位,而非内源的、组成型活化的 GPR30(L258K)分别表达大约 3.8 和 3.1 相对光单位的荧光素酶。用内源 GPR30 与内源 GPR6 以 1:4 的比例共转染,可显著地增加荧光素酶的表达达到大约 104.1 相对光单

位。用非内源 GPR30(L258K)与内源 GPR6 以相同的比例共转染，可显著地降低表达，它们分别在大约 18.2 和 29.5 相对光单位。对于 GPR17，当与 GPR3 共转染时，也可观察到相似的结果，如在图 5 中所示的那样。

实施例 3

确定非内源 GPCR 的组成性活性的测定方法

A. 细胞膜结合试验

1. [³⁵S]GTPγS 试验

当 G 蛋白偶联受体在其活性状态，并作为配体结合或者作为组成型活化的结果时，受体与 G 蛋白偶联并刺激 GDP 的释放和其后 GTP 与 G 蛋白的结合。G 蛋白-受体复合物的 α 亚基作为 GTP 酶并慢慢地水解 GTP 为 GDP，在此点受体通常发生失活。组成型活化受体继续把 GDP 转化为 GTP。不可水解的 GTP 类似物 [³⁵S]GTPγS，可被用来展示 [³⁵S]GTPγS 与表达组成型活化受体的膜的增强的结合。应用 [³⁵S]GTPγS 结合测定组成型活化的优点是：(a)它对所有 G 蛋白偶联受体是普遍适用的；(b)它邻近细胞膜表面，在此处较少可能拣到遇到影响细胞内级联反应的分子。

此试验利用 G 蛋白偶联受体的刺激 [³⁵S]GTPγS 与表达相关受体的细胞膜结合的能力。因此本测定可用于直接识别法去筛选针对已知、孤儿和组成型活化 G 蛋白偶联受体的候选化合物。本测定是普遍的并可用于针对所有 G 蛋白偶联受体的药物发现。

[³⁵S]GTPγS 试验：在 20mM HEPES、1 至大约 20mM 的 MgCl₂(尽管 20mM 是优选的，但这个剂量可针对结果的最优化进行调整)、pH 7.4、含有在 0.3 和 1.2nM 之间的 [³⁵S]GTPγS(尽管 1.2 是优选的，但这个剂量可针对结果的最优进行调整)、12.5 到 75μg 膜蛋白(例如，COS-7 细胞表达受体，本剂量可为最优化进行调整，尽管 75μg 是优选的)和 1μM GDP(这个剂量可针对结果的最优化进行改造)的结合缓冲液中温

育一小时。接着加入麦胚凝集素小珠(25 μ l, Amersham), 组合物在室温下再温育 30 分钟, 然后试管在 1500 $\times$ g、室温下离心 5 分钟, 并在闪烁计数器上记数。

另一个花费更少但同样可适用的方法已被识别, 它也可满足大规模筛选的需要。Flash platesTM 和 WallacTM 闪烁带可被用来格式化高处理量的 [³⁵S]GTP γ S 结合测定。进一步, 利用此技术, 本方法可用于已知的 GPCR, 它在通过 [³⁵S]GTP γ S 的结合监测化合物效应的同时, 同时监测与受体结合的由氚标记的配体。这之所以可能的是因为 Wallac β 计数器可以把能量窗口切换成监测氚和 ³⁵S 标记的探针。本方法也可用于侦查导致受体活化的其他类型的膜活化事件。例如, 本方法可用于监测许多受体的 ³²P 磷酸化(针对 G 蛋白偶联受体和酪氨酸激酶受体)。当膜被离心到孔的底部时, 结合的 [³⁵S]GTP γ S 或 ³²P 磷酸化的受体将要活化包被在孔上的闪烁剂。Scinti® 带(Wallac)已被用来展示这个原理。另外, 本方法也可通过应用放射标记的配体用来测量与受体结合的配体。以相似的方式, 当放射标记的结合的配体被离心到孔底时, 闪烁带标记在位于标记的配体附近, 这导致活化并被检测到。

基于前述的程序, 比较空白对照(pCMV)、内源 APJ 和非内源 APJ 的代表性结果被图示于图 6。

2. 腺苷酸环化酶

设计用来进行基于细胞的测定的 Flash PlateTM 腺苷酸环化酶试剂盒(New England Nuclear; 目录号 SMP004A)被改进以应用于未加工的胞浆膜。闪烁板的孔含有闪烁剂包被层, 其中含有识别 cAMP 的特异抗体。在孔中产生的 cAMP 通过直接和放射性 cAMP 示踪物竞争与 cAMP 抗体结合而被定量。下面是对测量在表达受体的膜上 cAMP 水平变化程序的简短描述。

在转染后大约 3 天收获转染细胞。通过在含有 20mM pH 7.4 的

HEPES 和 10mM MgCl_2 的缓冲液中均质化悬浮的细胞来制备细胞膜。均质化是在冰上用 Brinkman Polytron™ 进行大约 10 秒钟。得到的均质化物在 4℃、49,000×g 离心 15 分钟。得到的沉淀物接着在含有 20mM pH 7.4 的 HEPES 和 0.1mM EDTA 缓冲液中悬浮，均质化 10 秒钟，然后在 4℃、49,000×g 离心 15 分钟。得到的沉淀可被贮藏在 -4℃ 备用。在测量的当天，膜沉淀物在室温下被缓慢解冻，在含有 20mM pH 7.4 的 HEPES 和 10mM MgCl_2 的缓冲液中重新悬浮(这些数量可被优化，尽管在此列举的数值是优选的)，得到最终蛋白质浓度为 0.60mg/ml(重新悬浮的膜放置在冰上备用)。

按照制造商的指令制备和维持 cAMP 标准品和检测缓冲液(含有 2μCi 示踪物 [^{125}I cAMP(100μl)] 的 11ml 检测缓冲液)。为筛选用的试验缓冲液被新鲜制备，它含有 20mM pH 7.4 的 HEPES、10mM MgCl_2 、20mM (Sigma)、0.1 单位/ml 肌酸磷酸激酶(Sigma)、50μM GTP (Sigma) 和 0.2mM ATP(Sigma)；试验缓冲液可在冰上贮存备用。首先加入 50μl 试验缓冲液、接着加入 50μl 膜悬浮物到 NEN Flash Plate，以开始试验。得到的测定组合物在室温下温育 60 分钟，然后加入 100μl 检测缓冲液。培养板接着再温育 2-4 小时，然后用 Wallac MicroBeta 液闪计数器计数。cAMP/孔的数值从标准 cAMP 曲线外推，该曲线包括在每个测定板之内。应用前述的分析 MIG 的测定方法。

B. 基于报告基因的测定

1. CREB 报告基因测定(Gs 偶联受体)

检测 Gs 刺激的方法依赖于转化因子 CREB 的已知性质，它是以 cAMP 依赖的方式被活化的。应用 PathDetect CREB trans-Reporting System(Stratagene, Catalogue #219010)来检测在 293 和 293T 细胞中 Gs 偶联的活性。用上述系统的质粒成分和编码内源的或突变的受体的指明的表达质粒转染细胞，其使用哺乳动物细胞转染试剂盒(Stratagene, Catalogue #200285)并按照制造商的指令。简短而言，400ng pFR-Luc (荧光素酶报告基因质粒含有 Gal4 识别序列)、40ng pFA2-CREB (Gal4-

CREB 融合蛋白含有 Gal4 DNA 结合域)、80ng CMV-受体表达质粒(包括受体)和 20ng CMV-SEAP(分泌的碱性磷酸酶表达质粒;碱性磷酸酶活性在转染细胞的培养基中测量,以控制在样品间转染效率的变化)在磷酸钙沉淀中按照试剂盒的指令进行混合。把沉淀的一半等量地分布在 96 孔培养板的 3 个孔中,保持细胞过夜,第二天上午置换新鲜培养基。转染后 48 小时,如上述 GPR30 系统所说的方法处理细胞并测定荧光素酶活性。此测定用于 GHSR。

2. AP1 报告基因测定(Gq 偶联的受体)

测定 Gq 刺激依赖的方法依赖于 Gq 依赖的磷脂酶 C 已知的特性,即它可引起在其启动子含有 AP1 元件的基因活化。按照上述 CREB 报告基因测定所说的程序,使用 Pathdetect AP-1 cis-Reporting System (Stratagene, Catalogue #219073),其中只是将磷酸钙沉淀的组分改为 410ng pAP1-Luc、80ng 受体表达质粒和 20ng CMV-SEAP。本测定用于 ETBR-LP2。

C. 细胞内 IP3 累积的测定

在第一天,含有 5-羟色胺受体(内源的和突变的)的细胞被接种于 24 孔培养板上,一般是 1×10^5 细胞/孔。在第二天转染细胞,首先混合在 50 μ l/孔无血清 DMEM 中的 0.25 μ g DNA 和在 50 μ l/孔无血清 DMEM 中的 2 μ l lipofectamine。轻轻地混合溶液并在室温下温育 15-30 分钟。用 0.5ml PBS 洗涤细胞,把 400 μ l 无血清培养基与转染培养基混合并加到细胞中。然后在 37 $^{\circ}$ C/5% CO₂ 下温育细胞 3-4 小时,再移去转染培养基,替换为 1ml/孔常规培养基。在第三天,用 ³H-肌醇标记细胞。简短地说,移去培养基,细胞用 0.5ml PBS 洗涤,接着加入 0.5ml/孔无肌醇/无血清培养基(GIBCO BRL)和 0.25 μ Ci/孔 ³H-肌醇,在 37 $^{\circ}$ C/5% CO₂ 下温育细胞 16-18 小时。在第四天,用 0.5ml PBS 洗涤细胞,加入 0.45ml 试验培养基,其中含有无肌醇/无血清培养基 10 μ M 巴吉林 10mM 氯化锂或 0.4ml 试验培养基和 50 μ l 10 $\times$ ketaserin (ket) 以得到 10 μ M 的终浓度。然后在 37 $^{\circ}$ C 温育细胞 30 分钟。用 0.5ml PBS

洗涤细胞，加入 200 μ l/孔新鲜的/冰冷的终止液(1M KOH、18mM 硼酸钠、3.8mM EDTA)。溶液在冰上放置 5-10 分钟或直到细胞被溶解，然后用 200 μ l 新鲜的/冰冷的中和液(7.5% HCl)中和。然后把裂解物转移到 1.5ml 离心管中，加入 1ml/管氯仿/甲醇(1: 2)。然后使溶液涡旋 15 秒钟，把上层上样至 Biorad AG1-X8 阴离子交换树脂(100-200 目)。首先，树脂以 1:1.25 W/V 的比例用水洗涤，向柱中加载 0.9ml 的上层溶液。用 10ml 5mM 肌醇和 10ml 5mM 的硼酸钠/60mM 甲酸钠洗涤柱子。肌醇三磷酸酯被洗提入液闪管中，其中含有 10ml 液闪鸡尾，它有 2ml 0.1M 甲酸/1M 甲酸铵。通过用 10ml 0.1M 甲酸/3M 甲酸铵洗涤和用 ddH_2O 洗涤两次来再生交换柱，柱子贮存在 4 $^{\circ}\text{C}$ 的水中。

图 7 图示了从包括 C322K 突变的人 5-HT_{2A} 受体产生 IP₃。尽管这些结果表明脯氨酸突变算法规则可组成型活化本受体，但为了应用这样一个受体而筛选识别可能的治疗物，优选更大的差别。然而，因为活化的受体可被用来理解和阐释组成型活化的角色和用来识别可被进一步检查的化合物，我们相信这个差别本身在区分人 5-HT_{2A} 受体的内源和非内源形式时是有用的。

D. 结果概要

检测到的 GPCR 结果列于表 E，其中百分数的增加表示，观察到的非内源 GPCR 结果与内源 GPCR 相比较时的百分数差异；这些结果后面的括号里面标明的是应用的检测方法。进一步，应用的测定系统在括号中列出(并且，在应用不同的宿主细胞时，两者都列出)。这些结果表明，可应用多种方法确定人 GPCR 的非内源形式的组成型活化。相信本领域的熟练技术人员在基于前述并参考本领域中的信息后可具有选择和/或最佳化适于研究者的特别需要的特别测定方法的能力。

表 E

受体识别者 (密码子突变)	差异的百分数
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)
5-羟色胺 5-HT _{2A} (C322K)	33.2(IP ₃)
5-羟色胺 5-HT _{2C} (S310K)	39.1(IP ₃)

实施例 6

内源孤儿 GPCR 的组织分布

应用商业可得的人组织斑点印迹系统，探查内源孤儿 GPCR 以确定这些受体被定位的区域。除非在下面指明，完整受体 cDNA(放射标记的)被用作探针：按照制造商的指令，应用 Prime-It II™ 随机引物标记试剂盒(Stratagene, #300385)，用完整受体 cDNA(从载体中切下)产生放射标记的探针。把人 RNA Master Blot™ (Clontech, #7770-1)与 GPCR 放射标记的探针杂交，并按照制造商的指令在严格的条件下洗涤。印迹曝光于 Kodak BioMax 放射自显影底片，在-80℃过夜。

代表性的斑点印迹结果在表 8 中以 GPR1(8A)、GPR30(8B)和 APJ(8C)列出，针对所有受体的结果摘要列于表 F。

表 F

GPCR	组织分布(相对于斑点印迹中其他组织的最高水平)
GPR1	胎盘、卵巢、肾上腺
GPR4	广泛；在心脏、肺、肾上腺、甲状腺、脊髓中最高
GPR5	胎盘、胸腺、胎儿胸腺、在脾、胎儿脾中水平较低
GPR7	肝脏、脾脏、脊髓、胎盘
GPR8	没有检测到表达
GPR9-6	胸腺、胎儿胸腺，在小肠中水平较低
GPR18	脾脏、淋巴结、胎儿脾、睾丸
GPR20	广泛
GPR21	广泛；非常低的量
GPR22	心脏、胎儿心脏，在大脑中水平较低
GPR30	胃
GPR31	广泛
BLR1	脾脏
CEPR	胃、肝脏、甲状腺、壳
EBI1	胰脏，在淋巴样组织中水平较低
EBI2	淋巴样组织、主动脉、肺、脊髓
ETBR-LP2	广泛；脑组织
GPCR-CNS	脑，在睾丸、胎盘中水平较低
GPR-NGA	垂体，在脑中水平较低
H9	垂体
HB954	主动脉、小脑，在其他大部分组织中水平较低
HM74	脾脏、白细胞、骨髓、乳腺、肺、气管
MIG	在肾脏、肝脏、胰脏、肺和脾脏中水平较低
ORG1	垂体、胃、胎盘
V28	脑、脾脏、外周白细胞

基于前述的信息，注意到可评定人 GPCR 在患病组织中的分布；然后，在“正常”和患病组织中的对比性评定可被用来确定在疾病状态下一个特别受体的过度表达和不足表达的可能性。当希望利用人 GPCR 的非内源形式进行筛选以直接识别可能与治疗相关的候选化合物时，注意到反激活剂可用于治疗由特定人 GPCR 过度表达引起的疾病和紊乱，而激活剂或部分激活剂对于治疗特定人 GPCR 不足表达引起的疾病和紊乱是有用的。

正如被期望的，利用本领域技术人员周知的技术(例如，原位杂交)，更详细的受体细胞定位可被用于识别特殊细胞，其中感兴趣的受体在特殊细胞所在的组织中表达。

预期本发明文件提到的每一个专利、申请和印刷出版物，都以全文作为参考引入。

正如本领域的技术人员将认知的那样，可以对本发明的优选实施方案做多种变化和修饰而不背离本发明的精神。预期所有这些变化都落在本发明的范围之内。

虽然本领域的普通技术人员可以得到许多不同的载体为内源和非内源 GPCR 的目的使用，但最好是用 pCMV 载体。按照国际承认用于专利程序的微生物保存布达佩斯条约，该载体于 1998 年 10 月 13 日保存在美国典型培养物保藏中心 (American Type Culture Collection) (ATCC) (10801 University Blvd, Manassas, VA20110-2209 USA7)。该载体由 ATCC 在 1998 年____月____日进行了检验并在 1998 年____年____日测定了其存活性。ATCC 为 pCMV 给出了下列保藏号---

序列表

(1) 一般资料:

(i) 申请人: 多米尼克·P·比汉; 德里克·T·查默斯; 廖王蓁

(ii) 发明名称: 非内源的被组成型活化的人 G 蛋白偶联的受体

(iii) 序列数: 280

(iv) 通讯地址:

(A) 收信人: 阿瑞那制药公司

(B) Nancy Ridge 大道 6166 号

(C) 城市: 圣地亚哥

(D) 州: 加利福尼亚州

(E) 国家: 美国

(F) 邮编: 92121

(v) 计算机可读形式:

(A) 介质类型: 软盘

(B) 计算机: IBM 个人兼容机

(C) 操作系统: PC-DOS/MS-DOS

(D) 软件: PatentIn Release #1.0, #1.30 版

(vi) 本申请资料:

(A) 申请号:

(B) 申请日:

(C) 分类号

(viii) 代理人信息

(A) 姓名: Burgoon, Richard P.

(B) 登记号: 34787

(ix) 电讯信息:

(A) 电话: (858)453-7200

(B) 电传: (858)453-7210

(2) SEQ ID NO:1 的资料:

(i) 序列特征:

(A) 长度: 1068 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 1 的序列描述:

```

ATGGAAGATT TGGAGGAAAC ATTATTTGAA GAATTTGAAA ACTATTCCTA TGACCTAGAC 60
TATTACTCTC TGGAGTCTGA TTTGGAGGAG AAAGTCCAGC TGGGAGTTGT TCACTGGGTC 120
TCCCTGGTGT TATATTGTTT GGCTTTTGTG CTGGGAATTC CAGGAAATGC CATCGTCATT 180
TGGTTCACGG GGCTCAAGTG GAAGAAGACA GTCACCACTC TGTGGTTCCT CAATCTAGCC 240
ATTGCGGATT TCATTTTCTT TCTCTTTCTG 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 ACTCCGGGA CACTGTGGAG 540
TTCAATAATC ATACTCTTTG CTATAACAAT TTTCAGAAGC ATGATCCTGA CCTACTTTG 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 ACCATTCAAC ACAATAGCTA TTCCACCAT 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 CTGGAACAG CTCAATAA 1068

```

(3) SEQ ID NO:2 的资料:

(i) 序列特征:

(A) 长度: 355 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:2 的序列描述:

```

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			
	340	345	350
Thr Ala Gln			

355

(4) SEQ ID NO: 3 的资料:

(i) 序列特征:

(A) 长度: 1089 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 3 的序列描述:

```

ATGGGCAACC ACACGTGGGA GGGCTGCCAC GTGGA CTGCG CCGTGGACCA 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 GCTTCGCCCCG CCTGCGCCGC GTCAAGACCG CCGTGGCCGT GAGTCCGTG 420
GTCTGGGCCA CGGAGCTGGG CGCCAACTCG GCGCCCCTGT TCCATGACGA GCTCTCCGA 480
GACCGCTACA ACCACACCTT CTGCTTTGAG AAGTTCCCA TGGAAGGCTG GGTGGCCTGG 540
ATGAACCTCT ATCGGGTGTT CGTGGGCTTC CTCTTCCCGT GGGCGCTCAT GCTGCTGTCG 600
TACCGGGGCA TCCTGCGGGC CGTGCGGGGC AGCGTGTCCT CCGAGCGCCA GGAGAAGGCC 660
AAGATCAAGC GGCTGGCCCT CAGCCTCATC GCCATCGTGC TGGTCTGCTT TCGGCCCTAT 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
CTCACCTGG AGACCCCACT CACCTCCAAG AGGAACAGCA CAGCCAAAGC CATGACTGGC 1020
AGCTGGGCGG CCACTCCGCC TTCCCAGGGG GACCAGGTGC AGCTGAAGAT GCTGCCGCCA 1080
GCACAATGA

```

(5) SEQ ID NO:4 的资料:

(i) 序列特征:

(A) 长度: 362 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:4 的序列描述:

```

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
 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

(6) SEQ ID NO:5 的资料:

(i) 序列特征:

- (A) 长度: 30 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:5 的序列描述:

TATGAATTCA GATGCTCTAA ACGTCCCTGC 30

(7) SEQ ID NO:6 的资料:

(i) 序列特征:

- (A) 长度: 30 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:6 的序列描述:

TCCGGATCCA CCTGCACCTG CGCCTGCACC 30

(8) SEQ ID NO: 7 的资料:

(i) 序列特征:

- (A) 长度: 1002 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 7 的序列描述:

```

ATGGAGTCCT CAGGCAACCC AGAGAGCACC ACCTTTTTTT ACTATGACCT TCAGAGCCAG 60
CCGTGTGAGA ACCAGGCCTG GGTCTTTGCT ACCCTCGCCA CCACTGTCCT GTACTGCCCTG 120
GTGTTTCTCC TCAGCCTAGT GGGCAACAGC CTGGTCCTGT GGGTCCTGGT GAAGTATGAG 180
AGCCTGGAGT CCCTACCAA CATCTTCATC CTCAACCTGT GCCTCTCAGA CCTGGTGTTT 240
GCCTGCTTGT TGCCTGTGTG GATCTCCCCA TACCACTGGG GCTGGGTGCT GGGAGACTTC 300
CTCTGCAAAC TCCTCAATAT GATCTTCTCC ATCAGCCTCT ACAGCAGCAT CTTCTTCCTG 360
ACCATCATGA CCATCCACCG CTACCTGTCG GTAGTGAGCC CCCTCTCCAC CCTGCGCGTC 420
CCCACCCTCC GCTGCCGGGT GCTGGTGACC ATGGCTGTGT GGGTAGCCAG CATCCTGTCC 480
TCCATCCTCG ACACCATCTT CCACAAGGTG CTTTCTTCGG GCTGTGATTA TTCCGAACTC 540
ACGTGGTACC TCACCTCCGT CTACCAGCAC AACCTCTTCT TCCTGCTGTC CCTGGGGATT 600
ATCCTGTTCT GCTACGTGGA GATCCTCAGG ACCCTGTTCC GCTCAGCTC CAAGCGGCGC 660
CACCGCACGG TCAAGTCAT CTTCGCCATC GTGGTGGCCT ACTTCCTCAG CTGGGGTCCC 720
TACAACTTCA CCCTGTTTCT GCAGACGCTG TTTCCGACCC AGATCATCCG GAGCTGCGAG 780
GCCAAACAGC AGCTAGAATA CGCCTGCTC ATCTGCCGCA ACCTCGCCTT CTCCCACTGC 840
TGCTTTAACC CGGTGCTCTA TGTCTTCGTG GGGGTCAAGT TCCGCACACA CCTGAAACAT 900
GTTCTCCGGC AGTTCTGGTT CTGCCGGCTG CAGGCACCCA GCCCAGCCTC GATCCCCAC 960
TCCCTGGTG CCTTCGCCTA TGAGGGCGCC TCCTTCTACT GA 1002

```

(9) SEQ ID NO:8 的资料:

(i) 序列特征:

(A) 长度: 333 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:8 的序列描述:

```

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

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) SEQ ID NO:9 的资料:

(i) 序列特征:

(A) 长度: 30 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 9 的序列描述:

GCAAGCTTGG GGGACGCCAG GTCGCCGGCT

30

(11) SEQ ID NO:10 的资料:

(i) 序列特征:

(A) 长度: 31 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:10 的序列描述:

gcggatccgg acgctggggg agtcaggctg c

31

(12) SEQ ID NO: 11 的资料:

(i) 序列特征:

(A) 长度: 987 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 11 的序列描述:

```

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
GCCGTGTGGG GGATCGTCAC ACTCGTCGTG CTGCCCTTCG CAGTCTTCGC CCGGCTAGAC 540
GACGAGCAGG GCCGGCGCCA GTGCGTGCTA GTCTTTCCGC AGCCCGAGGC CTTCTGGTGG 600
CGCGCGAGCC GCCTCTACAC GTCGTGCTG GGCTTCGCCA TCCCCGTGTC CACCATCTGT 660
GTCCTCTATA CCACCTGCT GTGCCGGCTG CATGCCATGC GGCTGGACAG CCACGCCAAG 720
GCCCTGGAGC GCGCCAAGAA GCGGGTGACC TTCCTGGTGG TGGCAATCCT GGCGGTGTGC 780
CTCCTCTGCT GGACGCCTA CCACCTGAGC ACCGTGGTGG CGCTCACCAC CGACCTCCCG 840
CAGACGCCGC TGGTCATCGC TATCTCCTAC TTCATACCA GCCTGACGTA CGCCAACAGC 900
TGCTCAACC CCTTCCTCTA CGCCTTCCTG GACGCCAGCT TCCGCAGGAA CCTCCGCCAG 960
CTGATAACTT GCCGCGCGGC AGCCTGA 987

```

(13) SEQ ID NO:12 的资料:

(i) 序列特征:

(A) 长度: 328 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:12 的序列描述:

```

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
          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		

(14) SEQ ID NO:13 的资料:

(i) 序列特征:

- (A) 长度: 30 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:13 的序列描述:

CGGAATTCGT CAACGGTCCC AGCTACAATG 30

(15) SEQ ID NO:14 的资料:

(i) 序列特征:

- (A) 长度: 31 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:14 的序列描述:

ATGGATCCCA GGCCCTTCAG CACCGCAATA T 31

(16) SEQ ID NO: 15 的资料:

(i) 序列特征:

(A) 长度: 1002 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 15 的序列描述:

```

ATGCAGGCCG CTGGGCACCC AGAGCCCCTT GACAGCAGGG GCTCCTTCTC CCTCCCCACG 60
ATGGGTGCCA ACGTCTCTCA GGACAATGGC ACTGGCCACA ATGCCACCTT CTCCGAGCCA 120
CTGCCGTTCC TCTATGTGCT CCTGCCC GCC GTGTACTCCG GGATCTGTGC TGTGGGGCTG 180
ACTGGCAACA CGGCCGTCAT CCTTGTAATC CTAAGGGCGC CCAAGATGAA GACGGTGACC 240
AACGTGTTCA TCCTGAACCT GGCCGTCGCC GACGGGCTCT TCACGCTGGT ACTGCCCGTC 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 TCACGGTCCT GGTCTGCCCC 540
TTCTTCTCTT TCGCTGGCGT CTACAGCAAC GAGCTGCAGG TCCAAGCTG TGGGCTGAGC 600
TTCCCGTGGC CCGAGCGGGT CTGGTTCAAG GCCAGCCGTG TCTACACTTT GGTCTGGGC 660
TTCGTGCTGC CCGTGTGCAC CATCTGTGTG CTCTACACAG ACCTCCTGCG CAGGCTGCGG 720
GCCGTGCGGC TCCGCTCTGG AGCCAAGGCT CTAGGCAAGG CCAGGCGGAA GGTGACCGTC 780
CTGGTCTCTG TCGTGCTGGC CGTGTGCCTC CTCTGCTGGA CGCCCTTCCA CCTGGCCTCT 840
GTCGTGGCCC TGACCACGGA CCTGCCCCAG ACCCACTGG TCATCAGTAT GTCCTACGTC 900
ATCACCAGCC TCACGTACGC CAACTCGTGC CTGAACCCCT TCCTCTACGC CTTTCTAGAT 960
GACAACTTCC GGAAGAACTT CCGCAGCATA TTGCGGTGCT GA 1002

```

(17) SEQ ID NO:16 的资料:

(i) 序列特征:

(A) 长度: 333 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 16 的序列描述:

```

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
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 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
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
Asp Asn Phe Arg Lys Asn Phe Arg Ser Ile Leu Arg Cys	325	330	

(18) SEQ ID NO:17 的资料:

(i) 序列特征:

(A) 长度: 48 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 17 的序列描述:

ACGAATTCAG CCATGGTCCT TGAGGTGAGT GACCACCAAG TGCTAAAT

48

(19) SEQ ID NO:18 的资料:

(i) 序列特征:

(A) 长度: 27 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:18 的序列描述:

GAGGATCCTG GAATGCGGGG AAGTCAG

27

(20) SEQ ID NO: 19 的资料:

(i) 序列特征:

(A) 长度: 1107 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 19 的序列描述:

ATGGTCCTTG AGGTGAGTGA CCACCAAGTG CTAAATGACG CCGAGGTTGC CGCCCTCCTG 60
 GAGAACTTCA GCTCTTCCTA TGACTATGGA 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 CCCC GGCCCG CGTGACCCCTC ACCTGCCTGG CTGTCTGGGG GCTCTGCCTG 540
 CTTTTCGCCC TCCCAGACTT CATCTTCCTG TCGGCCACC ACGACGAGCG CCTCAACGCC 600
 ACCCACTGCC AATACAACTT CCCACAGGTG GGCCGCACGG CTCTGCGGGT GCTGCAGCTG 660
 GTGGCTGGCT TTCTGCTGCC CCTGCTGGTC ATGGCCTACT GCTATGCCCA CATCCTGGCC 720
 GTGCTGCTGG TTCCAGGGG CCAGCGGCGC CTGCGGGCCA TGCGGCTGGT GGTGGTGGTC 780
 GTGGTGGCCT TTGCCCTCTG CTGGACCCCC TATCACCTGG TGGTGGCTGGT 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

```

(21) SEQ ID NO:20 的资料:

(i) 序列特征:

(A) 长度: 368 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:20 的序列描述:

```

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 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					
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					
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					

(22) SEQ ID NO:21 的资料:

(i) 序列特征:

- (A) 长度: 30 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:21 的序列描述:

TTAAGCTTGA CCTAATGCCA TCTTGTGTCC

30

(23) SEQ ID NO:22 的资料:

(i) 序列特征:

- (A) 长度: 30 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组)

(xi) SEQ ID NO:22 的序列描述:

TTGGATCCAA AAGAACCATG CACCTCAGAG

30

(24) SEQ ID NO: 23 的资料:

(i) 序列特征:

- (A) 长度: 1074 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 23 的序列描述:

ATGGCTGATG ACTATGGCTC TGAATCCACA TCTTCCATGG AAGACTACGT TAACTTCAAC 60
TTCCTGACT TCTACTGTGA GAAAAACAAT GTCAGGCAGT TTGCGAGCCA TTCCTCCCA 120
CCCTGTACT 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
CTAAAAGTGA CCATCACTGT CCTGACCGTC TTTGTCTTGT CTCAGTTTCC CTACAAGTGC 780
ATTTTGTGGG TGCAGACCAT TGACGCCTAT GCCATGTTCA TCTCCAAGTGC TGCCGTTTCC 840
ACCAACATTG ACATCTGCTT CCAGGTCACC CAGACCATCG CCTTCTTCCA CAGTTGCCTG 900
AACCCTGTTT TCTATGTTTT TGTGGGTGAG AGATTCCGCC GGGATCTCGT GAAAACCTG 960
AAGAACTTGG GTTGCATCAG CCAGGCCAG TGGGTTTCAT TTACAAGGAG AGAGGGAAGC 1020
TTGAAGCTGT CGTCTATGTT GCTGGAGACA ACCTCAGGAG CACTCTCCCT CTGA 1074

(25) SEQ ID NO:24 的资料:

(i) 序列特征:

- (A) 长度: 357 个氨基酸
- (B) 类型: 氨基酸
- (C) 链型:

(D) 拓扑学：不相关

(ii) 分子类型：蛋白质

(xi) SEQ ID NO: 24 的序列描述：

```

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

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

(26) SEQ ID NO:25 的资料:

(i) 序列特征:

- (A) 长度: 1110 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 25 的序列描述:

A

1

(27) SEQ ID NO:26 的资料:

(i) 序列特征:

- (A) 长度: 369 个氨基酸
- (B) 类型: 氨基酸
- (C) 链型:
- (D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:26 的序列描述:

Met Ala Ser Ser 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
 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 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

(28) SEQ ID NO: 27 的资料:

(i) 序列特征:

(A) 长度: 1083 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 27 的序列描述:

ATGGACCCAG AAGAACTTC AGTTATTTG GATTATTACT ATGCTACGAG CCCAACTCT 60
 GACATCAGGG AGACCCACTC CCATGTTCTT TACACCTCTG TCTTCCTTCC AGTCTTTTAC 120
 ACAGCTGTGT TCCTGACTGG AGTGCTGGGG AACCTTGTTT TCATGGGAGC GTTGCATTTC 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 ATAAACTCA TATGGTCCCT GGTGGCCTTA 600
 ATTTTCACCT TTTTGTCCC TTTGTGAGC ATTGTGACCT GCTACTGTTG CATTGCAAGG 660
 AAGCTGTGTG CCCATTACCA GCAATCAGGA AAGCACAACA AAAAGCTGAA GAAATCTATA 720
 AAGATCATCT TTATTGTCGT GGCAGCCTTT CTTGTCTCCT GGCTGCCCTT CAATACTTTC 780
 AAGTTCCTGG CCATTGTCTC TGGGTGCGG 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

(29) SEQ ID NO:28 的资料:

(i) 序列特征:

(A) 长度: 360 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:28 的序列描述:

```

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
340 345 350

Arg Arg Lys Arg Ser Val Ser Leu
355 360

(30) SEQ ID NO:29 的资料:

(i) 序列特征:

- (A) 长度: 31 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:29 的序列描述:

CTAGAATTCT GACTCCAGCC AAAGCATGAA T

31

(31) SEQ ID NO:30 的资料:

(i) 序列特征:

- (A) 长度: 30 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:30 的序列描述:

GCTGGATCCT AACAGTCTG CGCTCGGCCT

30

(32) SEQ ID NO: 31 的资料:

(i) 序列特征:

(A) 长度: 1020 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 31 的序列描述:

```

ATGAATGGCC TTGAAGTGGC TCCCCAGGT CTGATCACCA ACTTCTCCCT GGCCACGGCA 60
GAGCAATGTG GCCAGGAGAC GCCACTGGAG AACATGCTGT TCGCCTCCTT CTACCTTCTG 120
GATTTTATCC TGGCTTTAGT TGGCAATACC CTGGCTCTGT GGCTTTTCAT CCGAGACCAC 180
AAGTCCGGGA CCCCAGCCAA CGTGTTCCTG ATGCATCTGG CCGTGGCCGA CTGTGCTGTC 240
GTGCTGGTCC TGCCCACCCG CCTGGTCTAC CACTTCTCTG GGAACCACTG GCCATTTGGG 300
GAAATCGCAT GCCGTCTCAC CGGCTTCCTC TTCTACCTCA ACATGTACGC CAGCATCTAC 360
TTCCTCACCT GCATCAGCGC CGACCGTTTC CTGGCCATTG TGCACCCGGT CAAGTCCCTC 420
AAGTCCGCGA GGCCCCTCTA CGCACACCTG GCCTGTGCCT TCCTGTGGGT GGTGGTGGCT 480
GTGGCCATGG CCCCAGTCTG GGTGAGCCCA CAGACCGTGC AGACCAACCA CACGGTGGTC 540
TGCCTGCAGC TGTACCGGGA GAAGGCCTCC CACCATGCCC TGGTGTCCCT GGCAGTGGCC 600
TTCACCTTCC CGTTCATCAC CACGGTCACC TGCTACCTGC TGATCATCCG CAGCCTGCGG 660
CAGGGCCTGC GTGTGGAGAA GCGCCTCAAG ACCAAGGCAG TCGCATGAT CGCCATAGTG 720
CTGGCCATCT TCCTGGTCTG CTTCGTGCCC TACCACGTCA ACCGCTCCGT CTACGTGCTG 780
CACTACCGCA GCCATGGGGC CTCCTGCGCC ACCCAGCGCA TCCTGGCCCT GGCAAACCGC 840
ATCACCTCCT GCCTCACCAG CCTCAACGGG GCACTCGACC CCATCATGTA TTTCTTCGTG 900
GCTGAGAAGT TCGCCACGC CCTGTGCAAC TTGCTCTGTG GCAAAGGCT CAAGGGCCCC 960
CCCCCAGCT TCGAAGGGAA AACCAACGAG AGCTCGCTGA GTGCCAAGTC AGAGCTGTGA 1020

```

(33) SEQ ID NO:32 的资料:

(i) 序列特征:

(A) 长度: 339 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 32 的序列描述:

```

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
 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

(34) SEQ ID NO:33 的资料:

(i) 序列特征:

- (A) 长度: 29 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 33 的序列描述:

ATAAGATGAT CACCCTGAAC AATCAAGAT 29

(35) SEQ ID NO:34 的资料:

(i) 序列特征:

- (A) 长度: 33 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:34 的序列描述:

TCCGAATTCA TAACATTCA CTGTTTATAT TGC 33

(36) SEQ ID NO: 35 的资料:

(i) 序列特征:

- (A) 长度: 996 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 35 的序列描述:

ATGATCACCC TGAACAATCA AGATCAACCT GTCACCTTTA ACAGCTCACA TCCAGATGAA 60
TACAAAATTG CAGCCCTTGT CTTCTATAGC TGTATCTTCA TAATTGGATT ATTTGTAAAC 120
ATCACTGCAT TATGGGTTTT CAGTTGTACC ACCAAGAAGA GAACCACGGT AACCATCTAT 180

```

ATGATGAATG TGGCATTAGT GGA CT TGATA TTTATAATGA CTTTACCCTT TCGAATGTTT 240
TATTATGCAA AAGATGCATG GCCATTGGA 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 AGATTTCTGA CATCATCTAT 540
CTAAAAGCTG TGAACGTGCT GAACCTCACT CGACTGACAT TTTTTTCTT GATTCCTTTG 600
TTCATCATGA TTGGGTGCTA CTTGGTCATT ATTCATAATC TCCTTCACGG CAGGACGTCT 660
AAGCTGAAAC CCAAAGTCAA GGAGAAGTCC ATAAGGATCA TCATCACGCT GCTGGTGCAG 720
GTGCTCGTCT GCTTTATGCC CTTCCACATC TGTTTCGCTT TCCTGATGCT GGAACGGGG 780
GAGAACAGTT ACAATCCCTG GGGAGCCTT ACCACCTTC TCATGAACCT CAGCACGTGT 840
CTGGATGTGA TTCTCTACTA CATCGTTTCA AAACAATTTC AGGCTCGAGT CATTAGTGTC 900
ATGCTATACC GTAATTACCT TCGAAGCCTG CGCAGAAAAA GTTCCGATC TGGTAGTCTA 960
AGGTCATAA GCAATATAA CAGTGAAATG TTATGA 996

```

(37) SEQ ID NO:36 的资料:

(i) 序列特征:

(A) 长度: 331 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:36 的序列描述:

```

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 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 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		
		330

(38) SEQ ID NO:37 的资料:

(i) 序列特征:

- (A) 长度: 28 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:37 的序列描述:

CCAAGCTTCC AGGCCTGGGG TGTGCTGG

(39) SEQ ID NO:38 的资料:

(i) 序列特征:

(A) 长度: 29 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:38 的序列描述:

ATGGATCCTG ACCTTCGGCC CCTGGCAGA

29

(40) SEQ ID NO: 39 的资料:

(i) 序列特征:

(A) 长度: 1077 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 39 的序列描述:

ATGCCCTCTG TGTCTCCAGC GGGGCCCTCG GCCGGGGCAG TCCCCAATGC CACCGCAGTG 60
 ACAACAGTGC GGACCAATGC CAGCGGGCTG GAGGTGCCCC TGTTCCACCT GTTTGCCCGG 120
 CTGGACGAGG AGCTGCATGG CACCTTCCCA GGCTGTGCG TGGCGTGAT GGCGGTGCAC 180
 GGAGCCATCT TCCTGGCAGG GCTGGTGCTC AACGGGCTGG CGCTGTACGT CTTCTGCTGC 240
 CGCACCCGGG CCAAGACACC CTCAGTCATC TACACCATCA ACCTGGTGGT GACCGATCTA 300
 CTGGTAGGGC TGTCCTGCC CACGCGCTTC GCTGTGTA CTACGGCGCCAG GGGCTGCCTG 360
 CGCTGTGCCT TCCCGCACGT CCTCGGTAC TTCCTCAACA TGCACTGCTC CATCCTCTTC 420
 CTCACCTGCA TCTGCGTGGA CCGCTACCTG GCCATCGTGC GGCCCGAAGG CTCCCGCCGC 480
 TGCCGCCAGC CTGCCTGTGC CAGGGCCGTG TCGCCTTCG TGTGGCTGGC CGCCGGTGCC 540
 GTCACCCTGT CGGTGCTGGG CGTGACAGGC AGCCGGCCCT GCTGCCGTGT CTTGCGCTG 600
 ACTGTCTGG AGTTCCTGCT GCCCCTGCTG GTCATCAGCG TGTTTACCGG CCGCATCATG 660
 TGTGCACTGT CGCGGCCGGG TCTGCTCCAC CAGGGTCGCC AGCGCCGCGT GCGGGCCATG 720
 CAGTCTCTGC TCACGGTGCT CATCATCTTT CTCGTCTGCT TCACGCCCTT CCACGCCCGC 780
 CAAGTGCCCG TGGCGCTGTG GCCCGACATG CCACACCACA CGAGCCTCGT GGTCTACCAC 840
 GTGGCCGTGA CCCTCAGCAG CCTCAACAGC TGATGGACC CCATCGTCTA CTGCTTCGTC 900
 ACCAGTGGCT TCCAGGCCAC CGTCCGAGGC CTCTTCGGCC AGCACGGAGA GCGTGAGCCC 960
 AGCAGCGGTG ACGTGGTCAG CATGCACAGG AGTCCAAGG GCTCAGGCCG TCATCACATC 1020
 CTCAGTGCCG GCCCTCACGC CCTCACCCAG GCCCTGGCTA ATGGGCCCGA GGCTTAG 1077

(41) SEQ ID NO:40 的资料:

(i) 序列特征:

(A) 长度: 358 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学：不相关

(ii) 分子类型：蛋白质

(xi) SEQ ID NO: 40 的序列描述:

```

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

```

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

(42) SEQ ID NO:41 的资料:

(i) 序列特征:

- (A) 长度: 30 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 41 的序列描述:

GAGAATTCAC TCCTGAGCTC AAGATGAACT 30

(43) SEQ ID NO:42 的资料:

(i) 序列特征:

- (A) 长度: 30 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 42 的序列描述:

CGGGATCCCC GTAAC TGAGC CACTTCAGAT 30

(44) SEQ ID NO:43 的资料:

(i) 序列特征:

(A) 长度: 1050 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 43 的序列描述:

```

ATGAACTCCA CCTTGGATGG TAATCAGAGC AGCCACCCTT TTTGCCTCTT GGCATTTGGC 60
TATTTGGAAA CTGTCAATTT TTGCCTTTTG GAAGTATTGA TTATTGTCTT TCTAACTGTA 120
TTGATTATTT CTGGCAACAT CATTGTGATT TTTGTATTTC ACTGTGCACC TTTGTTGAAC 180
CATCACACTA CAAGTTATTT TATCCAGACT ATGGCATATG CTGACCTTTT TGTTGGGGTG 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
TTCTGCCTT 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 CCTGTCCTGA TAAGCGCTAT GCCATGGTCC TGTTTCGAAT CACTAGTGTA 780
TTTTACATCC TCTGGTTGCC ATATATCATC TACTTCTGT TGGAAAGCTC CACTGGCCAC 840
AGCAACCGCT TCGCATCCTT CTTGACCACC TGGCTTGCTA TTAGTAACAG TTTCTGCAAC 900
TGTGTAATTT ATAGTCTCTC CAACAGTGTA TTCCAAAGAG GACTAAAGCG CCTCTCAGGG 960
GCTATGTGTA CTTCTGTGTC AAGTCAGACT ACAGCCAACG ACCCTTACAC AGTTAGAAGC 1020
AAAGGCCCTC TTAATGGATG TCATATCTGA 1050

```

(45) SEQ ID NO:44 的资料:

(i) 序列特征:

(A) 长度: 349 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 44 的序列描述:

```

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

```

98

325330335

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

345

(46) SEQ ID NO: 45 的资料:

- (i) 序列特征:
- (A) 长度: 30 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形
- (ii) 分子类型: DNA (基因组的)
- (xi) SEQ ID NO: 45 的序列描述:
- TCCCCCGGGA AAAAAACCAA CTGCTCCAAA
- 30

(47) SEQ ID NO:46 的资料:

- (i) 序列特征:
- (A) 长度: 31 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形
- (ii) 分子类型: DNA (基因组的)
- (xi) SEQ ID NO:46 的序列描述:
- TAGGATCCAT TTGAATGTGG ATTTGGTGAA A
- 31

(48) SEQ ID NO:47 的资料:

- (i) 序列特征:
- (A) 长度: 1302 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形
- (ii) 分子类型: DNA (基因组的)
- (xi) SEQ ID NO:47 的序列描述:
- WCHTRVHGNR AWCANMWRWM MWDNMAHMGs BDRVWNNYMK DDWBHHRTAN GKGBBKDWMH 60
- ATMWNWNNN WMHASRTHHT MSNWRMANRG ARTMSNWRMA NRGARABACK SATGTGTTTT 120
- TCTCCCATTC TGGAATCAA CATGCAGTCT GAATCTAACA TTACAGTGCG AGATGACATT 180
- GATGACATCA ACACCAATAT GTACCAACCA CTATCATATC CGTTAAGCTT TCAAGTGTCT 240
- CTCACCGGAT TTCTTATGTT AGAAATTGTG TTGGGACTTG GCAGCAACCT CACTGTATTG 300
- GTACTTTACT GCATGAAATC CAACTTAATC AACTCTGTCA GTAACATTAT TACAATGAAT 360
- CTTCATGTAC TTGATGTAAT AATTTGTGTG GGATGTATTC CTCTAACTAT AGTTATCCTT 420

```

CTGCTTTCAC TGGAGAGTAA CACTGCTCTC ATTTGCTGTT TCCATGAGGC TTGTGTATCT 480
TTTGCAAGTG TCTCAACAGC AATCAACGTT TTTGCTATCA CTTTGGACAG ATATGACATC 540
TCTGTAAAAC CTGCAAACCG AATTCTGACA ATGGGCAGAG CTGTAATGTT AATGATATCC 600
ATTTGGATTT TTTCTTTTTT CTCTTTCCTG ATTCCTTTTA TTGAGGTAAA TTTTTCAGT 660
CTTCAAAGTG GAAATACCTG GGAAAACAAG ACACCTTTAT GTGTCAGTAC AAATGAATAC 720
TACACTGAAC TGGGAATGTA TTATCACCTG TTAGTACAGA TCCCAATATT CTTTTTCACT 780
GTTGTAGTAA TGTTAATCAC ATACACCAAA ATACTTCAGG CTCTTAATAT TCGAATAGGC 840
ACAAGATTTT CAACAGGGCA GAAGAAGAAA GCAAGAAAGA AAAAGACAAT TTCTCTAACC 900
ACACAACATG AGGCTACAGA CATGTCACAA AGCAGTGGTG GGAGAAATGT AGTCTTTGGT 960
GTAAGAACTT CAGTTTCTGT AATAATTGCC CTCGGCGAG CTGTGAAACG ACACCGTGAA 1020
CGACGAGAAA GACAAAAGAG AGTCTTCAGG ATGTCTTTAT TGATTATTTT TACATTTCTT 1080
CTCTGCTGGA CACCAATTC TGTTTTAAAT ACCACCATT TATGTTTAGG CCCAAGTGAC 1140
CTTTTAGTAA AATTAAGATT GTGTTTTTTA GTCATGGCTT ATGGAACAAC TATATTTTAC 1200
CCTCTATTAT ATGCATTAC TAGACAAAAA TTTCAAAGG TCTTGAAAAG TAAAATGAAA 1260
AAGCGAGTTG TTTCTATAGT AGAAGCTGAT CCCCTGCCTA ATAATGCTGT AATACACAAC 1320
TCTTGATAG ATCCCAAAG AAACAAAAA ATTACCTTTG AAGATAGTGA AATAAGAGAA 1380
AAACGTTTAG TGCCTCAGGT TGTACAGAC TAG 1413

```

(49) SEQ ID NO: 48 的资料:

(i) 序列特征:

(A) 长度: 433 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 48 的序列描述:

```

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 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
 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

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

(50) SEQ ID NO: 49 的资料:

(i) 序列特征:

(A) 长度: 30 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 49 的序列描述:

GTGAAGCTTG CCTCTGGTGC CTGCAGGAGG

30

(51) SEQ ID NO:50 的资料:

(i) 序列特征:

(A) 长度: 31 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 50 的序列描述:

GCAGAATTCC CGGTGGCGTG TTGTGGTGCC C

31

(52) SEQ ID NO:51 的资料:

(i) 序列特征:

(A) 长度: 1209 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 51 的序列描述:

```

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 TGTCCCAGG 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
GGGAGACCA 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 GCCTCCAGC GCAGCATCCG GCTGCGGACA 900
AAGAGGGTGA CCCGCACAGC CATCGCCATC TGTCTGGTCT TCTTTGTGTG CTGGGCACCC 960
TACTATGTGC TACAGTGAC 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) SEQ ID NO:52 的资料:

(i) 序列特征:

(A) 长度: 402 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 52 的序列描述:

```

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 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
 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

(54) SEQ ID NO:53 的资料:

- (i) 序列特征:
 - (A) 长度: 27 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑学: 线形
- (ii) 分子类型: DNA (基因组的)
- (xi) SEQ ID NO: 53 的序列描述:
GGCGGATCCA TGGATGTGAC TTCCCAA

27

(55) SEQ ID NO: 54 的资料:

- (i) 序列特征:
 - (A) 长度: 27 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑学: 线形
- (ii) 分子类型: DNA (基因组的)
- (xi) SEQ ID NO: 54 的序列描述:
GGCGGATCCC TACACGGCAC TGCTGAA

27

(56) SEQ ID NO:55 的资料:

- (i) 序列特征:
 - (A) 长度: 1128 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑学: 线形
- (ii) 分子类型: DNA (基因组的)
- (xi) SEQ ID NO:55 的序列描述:
ATGGATGTGA CTTCCTAAGC CCGGGGCGTG GGCCTGGAGA TGTACCCAGG CACCGCGCAC 60
GCTGCGGCCC CCAACACCAC CTCCCCGAG CTCAACCTGT CCCACCCGCT CCTGGGCACC 120

```

GCCCTGGCCA ATGGGACAGG TGAGCTCTCG GAGCACCAGC AGTACGTGAT CGGCCTGTTC 180
CTCTCGTGCC TCTACACCAT CTTCTCTTC 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 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 CGCACCGGCA CCGTGGGCTG CGGCCCCGGC GGCAGAAGGC GTCGCGCATG 780
ATCCTCGCGG TGGTGTGTTT CTTCTTCGTC TGCTGGCTGC CGGAGAACGT CTTTCATCAGC 840
GTGCACCTCC TGCAGCGGAC GCAGCCTGGG GCCGCTCCCT GCAAGCAGTC TTTCCGCCAT 900
GCCCCCCCC TCACGGGCCA CATTGTCAAC CTCGCCGCCT TCTCCAACAG CTGCCTAAAC 960
CCCCTCATCT ACAGCTTTCT CGGGGAGACC TTCAGGGACA AGCTGAGGCT GTACATTGAG 1020
CAGAAAACAA ATTTGCCGGC CCTGAACCGC TTCTGTCACG CTGCCCTGAA GGCCGTCATT 1080
CCAGACAGCA CCGAGCAGTC GGATGTGAGG TTCAGCAGTG CCGTGTGA 1128

```

(57) SEQ ID NO:56 的资料:

(i) 序列特征:

(A) 长度: 375 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:56 的序列描述:

```

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
 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
 305 310 315 320

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

(58) SEQ ID NO: 57 的资料:

- (i) 序列特征:
 - (A) 长度: 31 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑学: 线形
- (ii) 分子类型: DNA (基因组的)
- (xi) SEQ ID NO: 57 的序列描述:
AAGGAATTCA CGGCCGGTG ATGCCATTCC C

31

(59) SEQ ID NO: 58 的资料:

- (i) 序列特征:
 - (A) 长度: 30 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑学: 线形
- (ii) 分子类型: DNA (基因组的)
- (xi) SEQ ID NO: 58 的序列描述:
GGTGGATCCA TAAACACGGG CGTTGAGGAC

30

(60) SEQ ID NO:59 的资料:

- (i) 序列特征:
 - (A) 长度: 960 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑学: 线形
- (ii) 分子类型: DNA (基因组的)
- (xi) SEQ ID NO: 59 的序列描述:

atgccattcc caaactgctc agcccccagc actgtggtgg ccacagctgt ggggtgtcttg 60
ctggggctgg agtgtgggt gggctgtctg ggcaacgagg tggcgtgtg gaccttcctg 120
ttccgggtca ggggtgtgaa gccgtacgt gtctacctgc tcaacctggc cctggctgac 180
ctgtgttgg ctgcgtgcct gcctttcctg gccgccttct acctgagcct ccaggcttgg 240
catctgggcc gtgtgggctg ctgggccctg cgcttcctgc tggacctcag ccgcagcgtg 300
gggatggcct tccctggccgc cgtggccttg gaccggtagc tccgtgtggt ccacctcgg 360
cttaaggtea acctgctgtc tctcaggcg gccctggggg tctcgggcct cgtctggctc 420
ctgatggtcg ccctcacctg ccgggcttg ctcatctctg aggccgccca gaactccacc 480
aggtgccaca gtttctactc cagggcagac ggctccttca gcatcatctg gcaggaagca 540
ctctcctgcc ttcagtttgt cctccccctt ggctcctatg tgttctgcaa tgcaggcatc 600
atcagggtc tccagaaaag actccgggag cctgagaaac agcccaagct tcagcgggcc 660

```

caggcactgg tcaccttggg ggtgggtgtg ttgtctctgt gctttctgcc ctgcttcctg 720
gccagagtcc tgatgcacat cttccagaat ctggggagct gcagggccct ttgtgcagtg 780
gctcatacct cggatgtcac gggcagcctc acctacctgc acagtgtcgt caaccccggtg 840
gtatactgct tctccagccc caccctcagg agctcctatc ggaggggtctt ccacaccctc 900
cgaggcaaag ggcaggcagc agagcccca gatttcaacc ccagagactc ctattcctga 960

```

(61) SEQ ID NO:60 的资料:

(i) 序列特征:

(A) 长度: 319 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 60 的序列描述:

```

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 Gln Ala Leu Val
 210 215 220
 Thr Leu Val Val Val Leu Phe Ala Leu Cys Phe Leu Pro Cys Phe Leu
 225 230 235 240
 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
 Gln Ala Ala Glu Pro Pro Asp Phe Asn Pro Arg Asp Ser Tyr Ser
 305 310 315

(62) SEQ ID NO:61 的资料:

(i) 序列特征:

(A) 长度: 1143 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 61 的序列描述:

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 CTTCTGGGTG 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
TGATGCCCT ACCACCTGGT GAAGACGCTG TACATGCTGG GCAGCCTGCT GCACTGGCCC 840
TGTAAGTTG ACCTCTTCCT CATGAACATC TTCCCCTACT GCACCTGCAT CAGCTACGTC 900
AACAGCTGCC TCAACCCCTT CCTCTATGCC TTTTCGACC CCCGCTTCG CCAGGCCTGC 960
ACCTCCATGC TCTGCTGTGG CCAGAGCAGG TGCGCAGGCA CCTCCCACAG CAGCAGTGGG 1020
GAGAAGTCAG CCAGCTACTC TTCGGGGCAC AGCCAGGGGC CCGGCCCCAA CATGGGCAAG 1080
GGTGAGAAC AGATGCACGA GAAATCCATC CCCTACAGCC AGGAGACCCT TGTGGTTGAC 1140
TAG 1143

```

(63) SEQ ID NO:62 的资料:

(i) 序列特征:

(A) 长度: 380 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 62 的序列描述:

```

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
          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

```

(64) SEQ ID NO: 63 的资料:

(i) 序列特征:

- (A) 长度: 31 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 63 的序列描述:

TGAGAATTCT GGTGACTCAC AGCCGGCACA G

31

(65) SEQ ID NO:64 的资料:

(i) 序列特征:

(A) 长度: 31 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:64 的序列描述:

GCCGGATCCA AGGAAAAGCA GCAATAAAAG G

31

(66) SEQ ID NO:65 的资料:

(i) 序列特征:

(A) 长度: 1119 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:65 的序列描述:

ATGAACTACC CGCTAACGCT GGAAATGGAC CTCGAGAACC TGGAGGACCT GTTCTGGGAA 60
CTGGACAGAT TGGACAACCTA TAACGACACC TCCCTGGTGG AAAATCATCT CTGCCCTGCC 120
ACAGAGGGTC CCCTCATGGC CTCCTTCAAG GCCGTGTTCG 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 ATTGTCCAG CCGTCCATGC CTACCGCCAC 480
CGCCGCCTCC TCTCCATCCA CATCACCTGT GGGACCATCT GGCTGGTGGG CTTCTCCTT 540
GCCTTGCCAG AGATTCTCTT CGCCAAAGTC AGCCAAGGCC ATCACAACAA CTCCCTGCCA 600
CGTTGCACCT TCTCCCAAGA GAACCAAGCA GAAACGCATG CCTGGTTCAC CTCCCGATT 660
CTCTACCATG TGGCGGGATT CCTGCTGCCC ATGCTGGTGA TGGGCTGGTG CTACGTGGGG 720
GTAGTGCACA GGTTCGCCA GGCCAGCGG CGCCCTCAGC GGCAGAAGGC AGTCAGGGTG 780
GCCATCCTGG TGACAAGCAT CTTCTTCCTC TGCTGGTCAC CCTACCACAT CGTCATCTTC 840
CTGGACACCC TGGCGAGGCT GAAGGCCGTG GACAATACCT GCAAGCTGAA TGGCTCTCTC 900
CCCGTGGCCA TCACCATGTG TGAGTTCCTG GGCCTGGCCC ACTGCTGCCT CAACCCCATG 960
CTCTACACTT TCGCCGGCGT GAAGTTCGC AGTGACCTGT CGCGGCTCCT GACCAAGCTG 1020
GGCTGTACCG GCCCTGCCTC CCTGTGCCAG CTCTTCCCTA GCTGGCGCAG GAGCAGTCTC 1080
TCTGAGTCAG AGAATGCCAC CTCTCTCACC ACGTTCTAG 1119

(67) SEQ ID NO:66 的资料:

(i) 序列特征:

(A) 长度: 372 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:66 的序列描述:

```

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

(68) SEQ ID NO: 67 的资料:

(i) 序列特征:

(A) 长度: 30 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 67 的序列描述:

CAAAGCTTGA AAGCTGCACG GTGCAGAGAC

30

(69) SEQ ID NO:68 的资料:

(i) 序列特征:

(A) 长度: 30 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组)

(xi) SEQ ID NO:68 的序列描述:

GCGGATCCCG AGTCACACCC TGGCTGGGCC

30

(70) SEQ ID NO:69 的资料:

(i) 序列特征:

(A) 长度: 1128 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:69 的序列描述:

```

ATGGATGTGA CTTCCCAAGC CCGGGGCGTG GGCCTGGAGA TGTACCCAGG CACCGCGCAG 60
CCTGCGGCCC CCAACACCAC CTCCCCGAG CTCAACCTGT CCCACCCGCT CCTGGGCACC 120
GCCCTGGCCA ATGGGACAGG TGAGCTCTCG GAGCACCAGC AGTACGTGAT CGGCCTGTTC 180
CTCTCGTGCC TCTACACCAT CTTCTCTTC CCCATCGGCT TTGTGGGCAA CATCCTGATC 240
CTGGTGGTGA ACATCAGCTT CCGCGAGAAG ATGACCATCC CCGACCTGTA CTTATCAAC 300
CTGGCGGTGG CGGACCTCAT CCTGGTGGCC GACTCCCTCA TTGAGGTGTT CAACCTGCAC 360
GAGCGGTACT ACGACATCGC CGTCTGTGC 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 CGAGGTCACG 660
CTGGGCTTCA TCGTGCCCTT CGCCATCATC GGCTGTGCT ACTCCCTCAT TGTCCGGGTG 720
CTGGTCAGGG CGCACCAGCA CCGTGGGCTG CGGCCCCGGC GGCAGAAGGC GCTCCGCATG 780
ATCCTCGCGG TGGTGCTGGT CTTCTTCGTC TGCTGGCTGC CGGAGAACGT CTTATCAGC 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 TTCTGTCACG CTGCCCTGAA GGCCGTCATT 1080
CCAGACAGCA CCGAGCAGTC GGATGTGAGG TTCAGCAGTG CCGTGTAG 1128

```

(71) SEQ ID NO:70 的资料:

(i) 序列特征:

(A) 长度: 375 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:70 的序列描述:

```

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

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

275280285

Pro Gly Ala Ala Pro Cys Lys Gln Ser Phe Arg His Ala His Pro Leu

290295300

Thr Gly His Ile Val Asn Leu Thr Ala Phe Ser Asn Ser Cys Leu Asn

305310315320

Pro Leu Ile Tyr Ser Phe Leu Gly Glu Thr Phe Arg Asp Lys Leu Arg

325330335

Leu Tyr Ile Glu Gln Lys Thr Asn Leu Pro Ala Leu Asn Arg Phe Cys

340345350

His Ala Ala Leu Lys Ala Val Ile Pro Asp Ser Thr Glu Gln Ser Asp

355360365

Val Arg Phe Ser Ser Ala Val

370375

(72) SEQ ID NO: 71 的资料:

(i) 序列特征:

- (A) 长度: 30 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 71 的序列描述:
ACAGAATTCC TGTGTGGTTT TACCGCCCAG

30

(73) SEQ ID NO:72 的资料:

(i) 序列特征:

- (A) 长度: 30 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 72 的序列描述:
CTCGGATCCA GGCAGAAGAG TCGCCTATGG

30

(74) SEQ ID NO:73 的资料:

(i) 序列特征:

(A) 长度: 1137 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 73 的序列描述:

```

ATGGACCTGG GGAACCAAT GAAAAGCGTG CTGGTGGTGG CTCTCCTTGT CATTTTCCAG 60
GTATGCCTGT GTCAAGATGA GGTCACGGAC GATTACATCG GAGACAACAC CACAGTGGAC 120
TACACTTGT TCGAGTCTTT GTGCTCCAAG AAGGACGTGC GGAAC TTAA AGCCTGGTTC 180
CTCCCTATCA TGTACTCCAT CATTTGTTTC GTGGGCCTAC TGGGCAATGG GCTGGTCGTG 240
TTGACCTATA TCTATTTCAA GAGGCTCAAG ACCATGACCG ATACCTACCT GCTCAACCTG 300
GCGGTGGCAG ACATCCTCTT CCTCCTGACC CTTCCCTTCT GGGCCTACAG CGCGGCCAAG 360
TCCTGGGTCT TCGGTGTCCA CTTTGTGCAAG CTCATCTTTG CCATCTACAA GATGAGCTTC 420
TTCAGTGGCA TGTCCTACT 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 TCCCCTGCT GGCCATGAGC 720
TTCTGTTACC TTGTCATCAT CCGCACCTG CTCCAGGCAC GCAACTTTGA GCGCAACAAG 780
GCCATCAAGG TGATCATCGC TGTGGTCGTG GTCTTCATAG TCTTCCAGCT GCCCTACAAT 840
GGGGTGGTCC TGGCCAGAC GGTGGCCAAC TTCAACATCA CCAGTAGCAC CTGTGAGCTC 900
AGTAAGCAAC TCAACATCGC CTACGACGTC ACCTACAGCC TGGCCTGCGT CCGCTGCTGC 960
GTCAACCCTT TCTTGACGC CTTTCATCGC GTCAAGTTCC GCAACGATCT CTTCAAGCTC 1020
TTCAAGGACC TGGGTGCTT CAGCCAGGAG CAGTCCGGC AGTGGTCTTC CTGTCGGCAC 1080
ATCCGGCGCT CCTCCATGAG TGTGGAGGCC GAGACCACCA CCACCTTCTC CCCATAG 1137

```

(75) SEQ ID NO:74 的资料:

(i) 序列特征:

(A) 长度: 378 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 74 的序列描述:

```

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			
	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
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																	

(76) SEQ ID NO:75 的资料:

(i) 序列特征:

- (A) 长度: 32 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 75 的序列描述:

CTGGAATTCA CCTGGACCAC CACCAATGGA TA 32

(77) SEQ ID NO: 76 的资料:

(i) 序列特征:

- (A) 长度: 30 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 76 的序列描述:

CTCGGATCCT GCAAAGTTTG TCATACAGTT 30

(78) SEQ ID NO:77 的资料:

(i) 序列特征:

- (A) 长度: 1085 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:77 的序列描述:

```

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 TGATTCTGA TATACTTTTT 240
ACCACGGCTT TGCCTACACG AATAGCCTAC TATGCAATGG GCTTTGACTG GAGAATCGGA 300
GATGCCTTGT GTAGGATAAC TGCCTAGTGT TTTTACATCA ACACATATGC AGGTGTGAAC 360
TTTATGACCT GCCTGAGTAT TGACCGCTTC ATTGCTGTGG TGCACCCTCT ACGCTACAAC 420
AAGATAAAAA GGATTGAACA TGCAAAGGC 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
TGCAAACCTC TCAGAACTGC CAAACAAAAC CCACTCACTG AGAAATCTGG TGTAAACAAA 720
AAGGCTCTCA 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 AAAGGTTATG 960
AGGATGCTGA AACGGCAAGT CAGTGTATCG ATTTCTAGTG CTGTGAAGTC AGCCCTGAA 1020
GAAAATTCAC GTGAAATGAC AGAAACGCAG ATGATGATAC ATTCCAAGTC TTCAAATGGA 1080
AAGTGA 1086

```

(79) SEQ ID NO:78 的资料:

(i) 序列特征:

(A) 长度: 361 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:78 的序列描述:

```

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					
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) SEQ ID NO: 79 的资料:	
(i) 序列特征:	
(A) 长度: 31 个碱基对	
(B) 类型: 核酸	
(C) 链型: 单链	
(D) 拓扑学: 线形	
(ii) 分子类型: DNA (基因组的)	
(xi) SEQ ID NO: 79 的序列描述:	
CTGGAATTCT CCTGCTCATC CAGCCATGCG G	31
(81) SEQ ID NO: 80 的资料:	
(i) 序列特征:	
(A) 长度: 30 个碱基对	
(B) 类型: 核酸	
(C) 链型: 单链	
(D) 拓扑学: 线形	
(ii) 分子类型: DNA (基因组的)	
(xi) SEQ ID NO: 80 的序列描述:	
CCTGGATCCC CACCCCTACT GGGGCCTCAG	30
(82) SEQ ID NO:81 的资料:	
(i) 序列特征:	
(A) 长度: 1446 个碱基对	
(B) 类型: 核酸	
(C) 链型: 单链	
(D) 拓扑学: 线形	
(ii) 分子类型: DNA (基因组的)	
(xi) SEQ ID NO: 81 的序列描述:	
ATGCGGTGGC TGTGGCCCT GGCTGTCTCT CTTGCTGTGA TTTTGGCTGT GGGGCTAAGC 60	
AGGGTCTCTG GGGGTGCCCC CCTGCACCTG GGCAGGCACA GAGCCGAGAC CCAGGAGCAG 120	
CAGAGCCGAT CCAAGAGGGG CACCGAGGAT GAGGAGGCCA AGGGCGTGCA GCAGTATGTG 180	
CCTGAGGAGT GGGCGGAGTA CCCCCGCCCC ATTCACCCTG CTGGCCTGCA GCCAACCAAG 240	
CCCTTGGTGG CCACCAGCCC TAACCCCGAC AAGGATGGGG GCACCCAGAG CAGTGGGCAG 300	
GAACTGAGGG GCAATCTGAC AGGGGCACCA GGGCAGAGGC TACAGATCCA GAACCCCTG 360	
TATCCGGTGA CCGAGAGCTC CTACAGTGCC TATGCCATCA TGCTTCTGGC GCTGGTGGTG 420	
TTTTCGGTGG GCATTGTGGG CAACCTGTGC GTCATGTGCA TCGTGTGGCA CAGCTACTAC 480	
CTGAAGAGCG CCTGGAATC CATCCTTGCC AGCCTGGCCC TCTGGGATTT TCTGGTCCTC 540	

```

TTTTTCTGCC TCCCTATTGT CATCTTCAAC GAGATCACCA AGCAGAGGCT ACTGGGTGAC 600
GTTTCTTGTC 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 CCCCACCATG 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 TGCACCCTCC 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 TCACCCAC TCCTGCCCT GGGCACACCT 1440
TGCTGA 1446

```

(83) SEQ ID NO:82 的资料:

(i) 序列特征:

(A) 长度: 481 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 82 的序列描述:

```

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
 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

(84) SEQ ID NO:83 的资料:

(i) 序列特征:

- (A) 长度: 22 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 83 的序列描述:

ATGTGGAACG CGACGCCAG CG 22

(85) SEQ ID NO:84 的资料:

(i) 序列特征:

- (A) 长度: 22 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 84 的序列描述:

TCATGTATTA ATACTAGATT CT 22

(86) SEQ ID NO: 85 的资料:

(i) 序列特征:

- (A) 长度: 38 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 85 的序列描述:

TACCATGTGG AACGCGACGC CCAGCGAAGA GCCGGGGT

38

(87) SEQ ID NO:86 的资料:

(i) 序列特征:

(A) 长度: 39 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:86 的序列描述:

CGGAATTCAT GTATTAATAC TAGATTCTGT CCAGGCCCG

39

(88) SEQ ID NO:87 的资料:

(i) 序列特征:

(A) 长度: 1101 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:87 的序列描述:

ATGTGGAACG CGACGCCAG CGAAGAGCCG GGGTTCAACC TCACACTGGC CGACCTGGAC 60
TGGGATGCTT CCCCCGCAA CGACTCGCTG GGCGACGAGC TGCTGCAGCT CTTCCCCGCG 120
CCGCTGCTGG CGGGCGTCAC AGCCACCTGC GTGGCACTCT TCGTGGTGGG TATCGCTGGC 180
AACCTGCTCA CCATGCTGGT GGTGTCGCGC TTCCGCGAGC TGCGCACCAC CACCAACCTC 240
TACCTGTCCA GCATGGCCTT CTCCGATCTG CTCATCTTCC TCTGCATGCC CCTGGACCTC 300
GTTTCGCCTCT GGCAGTACCG GCCCTGGAAC TTCGGCGACC TCCTCTGCAA ACTCTTCCAA 360
TTCGTCAGTG AGAGCTGCAC CTACGCCACG GTGCTACCA TCACAGCGCT GAGCGTCGAG 420
CGCTACTTCG CCATCTGCTT CCCACTCCGG GCCAAGGTGG TGGTCACCAA GGGGCGGGTG 480
AAGCTGGTCA TCTTCGTCAT CTGGGCCGTG GCCTTCTGCA GCGCCGGGCC CATCTTCGTG 540
CTAGTCGGGG TGGAGCACGA GAACGGCACC GACCTTGGG ACACCAACGA GTGCCGCCCC 600
ACCGAGTTTG CGGTGCGCTC TGGACTGCTC ACGGTCATGG TGTGGGTGTC CAGCATCTTC 660
TTCTTCCTTC CTGTCTTCTG TCTCACGGTC CTCTACAGTC TCATCGGCAG GAAGCTGTGG 720
CGGAGGAGGC GCGGCGATGC TGTCTGGGT GCCTCGCTCA GGGACCAGAA CCACAAGCAA 780
ACCGTAAAA 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
 GAACCCCTTCT CCCAGAGAAA GCTCTCCACT CTGAAAGATG AAAGTTCTCG GGCCTGGACA 1080
 GAATCTAGTA TTAATACATG A 1101

(89) SEQ ID NO: 88 的资料:

(i) 序列特征:

(A) 长度: 366 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 88 的序列描述:

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	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
      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) SEQ ID NO: 89 的资料:

(i) 序列特征:

(A) 长度: 33 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 89 的序列描述:

GCAAGCTTGT GCCCTACCA AGCCATGCGA GCC

33

(91) SEQ ID NO:90 的资料:

(i) 序列特征:

(A) 长度: 30 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 90 的序列描述:

CGGAATTCAG CAATGAGTTC CGACAGAAGC

30

(92) SEQ ID NO:91 的资料:

(i) 序列特征:

(A) 长度: 1842 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 91 的序列描述:

ATGCGAGCCC CGGGCGCGCT TCTCGCCCGC ATGTCGCGGC TACTGCTTCT GCTACTGCTC 60
 AAGGTGTCTG CCTCTTCTGC CCTCGGGGTC GCCCCTGCGT CCAGAAACGA AACTTGCTCTG 120
 GGGGAGAGCT GTGCACCTAC AGTGATCCAG CGCCGCGGCA GGGACGCCTG GGGACCGGGA 180
 AATTCTGCAA GAGACGTTCT GCGAGCCCGA GCACCCAGGG AGGAGCAGGG GGCAGCGTTT 240
 CTTGCGGGAC CCTCCTGGGA CCTGCCGGCG GCCCCGGGCC GTGACCCGGC TGCAGGCAGA 300
 GGGGCGGAGG CGTCGGCAGC CGGACCCCGG GGACCTCCAA CCAGGCCACC TGGCCCCCTG 360
 AGGTGGAAG GTGCTCGGGG TCAGGAGCCT TCTGAACTT TGGGGAGAGG GAACCCACG 420
 GCCCTCCAGC TCTTCCTTCA GATCTCAGAG GAGGAAGAGA AGGTCCCAG AGGCGCTGGC 480
 ATTTCCGGGC GTAGCCAGGA GCAGAGTGTG AAGACAGTCC CCGGAGCCAG CGATCTTTT 540
 TACTGGCCAA GGAGAGCCGG GAACTCCAG GGTTCACC ACCAAGCCCCT GTCCAAGACG 600
 GCCAATGGAC TGGCGGGGCA CGAAGGGTGG ACAATTGCAC TCCCGGGCCG GGCCTGGCC 660
 CAGAATGGAT CCTTGGGTGA AGGAATCCAT GAGCCTGGGG GTCCCCGCGG GGGAAACAGC 720
 ACGAACCGGC GTGTGAGACT GAAGAACCC 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
 GAGGTGCTT CTCTGGGAGT CACCACTTTC ACCTTATGTG CTCTGTGCAT AGACCGCTTC 1080
 CGTGTGCCA CCAACGTACA GATGTACTAC GAAATGATCG AAAACTGTTC CTCAACAAC 1140
 GCCAACTTG CTGTTATATG GGTGGGAGCT CTATTGTTAG CACTTCCAGA AGTTGTTCTC 1200
 CGCCAGCTGA GCAAGGAGGA TTTGGGGTTT AGTGGCCGAG CTCCGGCAGA AAGGTGCATT 1260
 ATTAAGATCT CTCCTGATTT ACCAGACACC ATCTATGTT TAGCCCTCAC CTACGACAGT 1320
 GCGAGACTGT GGTGGTATTT TGGCTGTTAC TTTTGTGTC CCACGCTTT CACCATCACC 1380
 TGCTCTCTAG TGAAGCGAG GAAAATCCGC AAAGCAGAGA AAGCCTGTAC CCGAGGGAAT 1440
 AAACGGCAGA TTCAACTAGA GAGTCAGATG AACTGTACAG TAGTGGCACT GACCATTTTA 1500
 TATGGATTT GCATTATTCC TGAAAATATC TGCAACATT TTAAGCCTA CATGGCTACA 1560
 GGGGTTTCAC AGCAGACAAT GGACCTCCTT AATATCATCA GCCAGTTCCT TTTGTTCTTT 1620
 AAGTCCTGTG TCACCCAGT CCTCCTTTTC TGTCTCTGCA AACCTTCAG TCGGCCTTC 1680

ATGGAGTGCT GCTGCTGTTG CTGTGAGGAA TGCATTGAGA AGTCTTCAAC GGTGACCAGT 1740
 GATGACAATG ACAACGAGTA CACCACGGAA CTCGAACTCT CGCCTTTCAG TACCATACGC 1800
 CGTGAAATGT CCACTTTTGC TTCTGTCGGA ACTCATTGCT GA 1842

(93) SEQ ID NO:92 的资料:

(i) 序列特征:

(A) 长度: 613 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 92 的序列描述:

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 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
 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

[illegible]

(94) SEQ ID NO:93 的资料:

(i) 序列特征:

(A) 长度: 34 个碱基对

(B) 类型: 核酸

(C) 链型：单链

(D) 拓扑学：线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 93 的序列描述:

CAGAATTCAG AGAAAAAAG TGAATATGGT TTTT

34

(95) SEQ ID NO: 94 的资料:

(i) 序列特征:

(A) 长度: 32 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 94 的序列描述:

TTGGATCCCT GGTGCATAAC AATTGAAAGA AT

32

(96) SEQ ID NO:95 的资料:

(i) 序列特征:

(A) 长度: 1248 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:95 的序列描述:

```

ATGGTTTTTG CTCACAGAAT GGATAACAGC AAGCCACATT TGATTATTCC TACACTTCTG 60
GTGCCCCTCC 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
TTCGTCCTGC 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 AAGTTCCTT GTTTTCACAG CTATCACATG GATATCCTTT 960
AGTTCCTCAG CCTCTAAACC TACTCTGTAT TCAATTTATA ATGCCAATTT TCGGAGAGGG 1020
ATGAAAGAGA CTTTTTGCAT GTCCTCTATG AAATGTTACC GAAGCAATGC CTATACTATC 1080
ACAACAAGTT CAAGGATGGC CAAAAA AAC TACGTTGGCA TTTCAGAAAT CCCTTCCATG 1140
GCCAAACTA TTACCAAAGA CTCGATCTAT GACTCATTG ACAGAGAAGC CAAGGAAAAA 1200
AAGCTTGCTT GGCCATTAA CTCAAATCCA CCAATACTT TTGTCTAA 1248

```

(97) SEQ ID NO:96 的资料:

(i) 序列特征:

(A) 长度: 415 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:96 的序列描述:

```

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	Ser
275					280					285				
Trp	Leu	Pro	Phe	His	Val	Ala	Gln	Leu	Trp	His	Pro	His	Glu	Gln
290					295					300				
Tyr	Lys	Lys	Ser	Ser	Leu	Val	Phe	Thr	Ala	Ile	Thr	Trp	Ile	Ser
305					310					315				
Ser	Ser	Ser	Ala	Ser	Lys	Pro	Thr	Leu	Tyr	Ser	Ile	Tyr	Asn	Ala
325					330					335				
Phe	Arg	Arg	Gly	Met	Lys	Glu	Thr	Phe	Cys	Met	Ser	Ser	Met	Lys
340					345					350				
Tyr	Arg	Ser	Asn	Ala	Tyr	Thr	Ile	Thr	Thr	Ser	Ser	Arg	Met	Ala
355					360					365				
Lys	Asn	Tyr	Val	Gly	Ile	Ser	Glu	Ile	Pro	Ser	Met	Ala	Lys	Thr
370					375					380				
Thr	Lys	Asp	Ser	Ile	Tyr	Asp	Ser	Phe	Asp	Arg	Glu	Ala	Lys	Glu
385					390					395				
Lys	Leu	Ala	Trp	Pro	Ile	Asn	Ser	Asn	Pro	Pro	Asn	Thr	Phe	Val
405					410					415				

(98) SEQ ID NO:97 的资料:

(i) 序列特征:

- (A) 长度: 30 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:97 的序列描述:

GGAAAGCTTA ACGATCCCCA GGAGCAACAT

30

(99) SEQ ID NO: 98 的资料:

(i) 序列特征:

- (A) 长度: 31 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 98 的序列描述:

CTGGGATCCT ACGAGAGCAT TTTTCACACA G

31

(100) SEQ ID NO: 99 的资料:

(i) 序列特征:

(A) 长度: 1842 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 99 的序列描述:

```

ATGGGGCCCA CCCTAGCGGT TCCCACCCCC TATGGCTGTA TTGGCTGTAA GCTACCCAG 60
CCAGAATACC CACCGGCTCT AATCATCTTT ATGTTCTGCG CGATGGTTAT CACCATCGTT 120
GTAGACCTAA TCGGCAACTC CATGGTCATT TTGGCTGTGA CGAAGAACAA GAAGTCCGG 180
AATTCTGGCA ACATCTTCGT GGTCACTCTC TCTGTGGCCG ATATGCTGGT GGCCATCTAC 240
CCATACCCTT 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
CTGCCCCAACA 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 TCCTCTTTGC 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 GAGGCCCGTA CCCTGGCCCG CGCCCGTGCC 1020
CATGCTCGCG ACCAAGCTCG TGAACAAGAC CGTGCCCATG CCTGTCCTGC TGTGGAGGAA 1080
ACCCCGATGA ATGTCCGGA TGTTCATTA CCTGGTGATG CTGCAGCTGG CCACCCGAC 1140
CGTGCCTCTG GCCACCCTAA GCCCCATTCC AGATCCTCCT CTGCCTATCG CAAATCTGCC 1200
TCTACCCACC ACAAGTCTGT CTTAGCCAC TCCAAGGCTG CCTCTGGTCA CCTCAAGCCT 1260
GTCTCTGGCC ACTCCAAGCC TGCCTCTGGT CACCCCAAGT CTGCCACTGT CTACCCTAAG 1320
CCTGCCTCTG TCCATTCAA GGGTGACTCT GTCCATTCA AGGGTGACTC TGTCCATTTC 1380
AAGCCTGACT CTGTTCAATT 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
CCCGAGATCC CTGCCATTGC CCACCCTGTG TCTGACGACA GTGACCTCCC TGAGTCGGCC 1680
TCTAGCCCTG CCGCTGGGCC CACCAAGCCT GCTGCCAGCC AGCTGGAGTC TGACACCATC 1740
GCTGACCTTC CTGACCCTAC TGTAAGTCACT ACCAGTACCA ATGATTACCA TGATGTCGTG 1800
GTTGTTGATG TTGAAGATGA TCCTGATGAA ATGGCTGTGT GA 1842

```

(101) SEQ ID NO:100 的资料:

(i) 序列特征:

(A) 长度: 613 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 100 的序列描述:

```

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 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			
	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

500505510

Thr Ala Asp Tyr Pro Lys Pro Ala Thr Thr Ser His Pro Lys Pro Ala

515520525

Ala Ala Asp Asn Pro Glu Leu Ser Ala Ser His Cys Pro Glu Ile Pro

530535540

Ala Ile Ala His Pro Val Ser Asp Asp Ser Asp Leu Pro Glu Ser Ala

545550555560

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

565570575

Ser Asp Thr Ile Ala Asp Leu Pro Asp Pro Thr Val Val Thr Thr Ser

580585590

Thr Asn Asp Tyr His Asp Val Val Val Asp Val Glu Asp Asp Pro

595600605

Asp Glu Met Ala Val

610

(102) SEQ ID NO:101 的资料:

(i) 序列特征:

- (A) 长度: 32 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 101 的序列描述:

TCCAAGCTTC GCCATGGGAC ATAACGGGAG CT

32

(103) SEQ ID NO:102 的资料:

(i) 序列特征:

- (A) 长度: 30 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 102 的序列描述:

CGTGAATTCC AAGAATTTAC AATCCTTGCT

30

(104) SEQ ID NO:103 的资料:

(i) 序列特征:

(A) 长度: 1548 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 103 的序列描述:

```

ATGGGACATA ACGGGAGCTG GATCTCTCCA AATGCCAGCG AGCCGCACAA CGCGTCCGGC 60
GCCGAGGCTG CGGGTGTGAA CCGCAGCGCG CTCGGGGAGT TCGGCGAGGC GCAGCTGTAC 120
CGCCAGTTCA CCACCACCGT GCAGGTCGTC ATCTTCATAG GCTCGCTGCT CGGAAACTTC 180
ATGGTGTTAT GGTCAACTTG CCGCACAACC GTGTTCAAAT CTGTCACCAA CAGGTTTCATT 240
AAAAACCTGG CCTGCTCGGG GATTTGTGCC AGCCTGGTCT GTGTGCCCTT CGACATCATC 300
CTCAGCACCA GTCCTCACTG TTGCTGGTGG ATCTACACCA TGCTCTTCTG CAAGGTCGTC 360
AAATTTTTCG 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
GTGTTCTCTT 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 TTCCGCTCTT 900
TTGCTGCTCA CTGCTGTTTG GCTGCCCAAA GTCTCCCTGC TGGCAAACCC TGTTCCTCTT 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 TGGCACCAGC AGCCCTGTG 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

```

(105) SEQ ID NO: 104 的资料:

(i) 序列特征:

(A) 长度: 515 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 104 的序列描述:

```

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

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
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

(106) SEQ ID NO:105 的资料:

(i) 序列特征:

(A) 长度: 29 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:105 的序列描述:

GGAGAATTCA CTAGGCGAGG CGCTCCATC

29

(107) SEQ ID NO:106 的资料:

(i) 序列特征:

(A) 长度: 30 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:106 的序列描述:

GGAGGATCCA GGAAACCTTA GGCCGAGTCC

30

(108) SEQ ID NO: 107 的资料:

(i) 序列特征:

(A) 长度: 1164 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 107 的序列描述:

ATGAATCGGC ACCATCTGCA GGATCACTTT CTGGAAATAG ACAAGAAGAA CTGCTGTGTG 60
 TTCCGAGATG ACTTCATTGC CAAGGTGTTG CCGCCGGTGT TGGGGCTGGA GTTTATCTTT 120
 GGGCTTCTGG GCAATGGCCT TGCCCTGTGG ATTTTCTGTT TCCACCTCAA GTCCTGGAAA 180
 TCCAGCCGGA TTTTCTGTG CAACCTGGCA GTAGCTGACT TTCTACTGAT CATCTGCCTG 240
 CCGTTCGTGA TGGACTACTA TGTGCGGCGT TCAGACTGGA ACTTTGGGGA CATCCCTTGC 300
 CGGCTGGTGC TCTTCATGTT TGCCATGAAC CGCCAGGGCA GCATCATCTT CCTCACGGTG 360
 GTGGCGGTAG ACAGGTATTT CCGGGTGGTC CATCCCCACC ACGCCCTGAA CAAGATCTCC 420
 AATTGGACAG CAGCCATCAT CTCTTGCCTT CTGTGGGGCA TCACTGTTGG CCTAACAGTC 480
 CACCTCCTGA AGAAGAAGTT GCTGATCCAG AATGGCCCTG CAAATGTGTG CATCAGCTTC 540
 AGCATCTGCC ATACCTTCGG GTGGCAGGAA 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 TATCACTCTC 840

```

AGCTTCACCT ACATGAACAG CATGCTGGAC CCCGTGGTGT ACTACTTCTC CAGCCCATCC 900
TTTCCCAACT TCTTCTCCAC TTTGATCAAC 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) SEQ ID NO: 108 的资料:

(i) 序列特征:

(A) 长度: 387 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 108 的序列描述:

```

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 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

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

```

(110) SEQ ID NO:109 的资料:

(i) 序列特征:

(A) 长度: 37 碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(iv) 反义: 否

(xi) SEQ ID NO: 109 的序列描述:

ACCATGGCTT GCAATGGCAG TCGGCCAGG GGGCACT 37

(111) SEQ ID NO:110 的资料:

(i) 序列特征:

(A) 长度: 39 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(iv) 反义: 是

(xi) SEQ ID NO: 110 的序列描述:

CGACCAGGAC AAACAGCATC TTGGTCACTT GTCTCCGGC 39

(112) SEQ ID NO:111 的资料:

(i) 序列特征:

(A) 长度: 39 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(iv) 反义: 否

(xi) SEQ ID NO: 111 的序列描述:

GACCAAGATG CTGTTTGTCC TGGTCGTGGT GTTTGGCAT 39

(113) SEQ ID NO:112 的资料:

(i) 序列特征:

(A) 长度: 35 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(iv) 反义: 是

(xi) SEQ ID NO: 112 的序列描述:

CGGAATTCAG GATGGATCGG TCTCTTGCTG CGCCT 35

(114) SEQ ID NO: 113 的资料:

(i) 序列特征:

(A) 长度: 1212 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 113 的序列描述:

```

ATGGCTTGCA ATGGCAGTGC GGCCAGGGGG CACTTTGACC CTGAGGACTT GAACCTGACT 60
GACGAGGCAC TGAGACTCAA GTACCTGGGG CCCCAGCAGA CAGAGCTGTT CATGCCCATC 120
TGTGCCACAT ACCTGCTGAT CTTCGTGGTG GGCGCTGTGG GCAATGGGCT GACCTGTCTG 180
GTCATCCTGC GCCACAAGGC CATGCGCACG CCTACCAACT ACTACCTCTT CAGCCTGGCC 240
GTGTCGGACC TGCTGGTGCT GCTGGTGGGC CTGCCCTGG AGCTCTATGA GATGTGGCAC 300
AACTACCCCT TCCTGCTGGG CGTTGGTGGC TGCTATTTC GCACGCTACT GTTTGAGATG 360
GTCTGCCTGG CTCAGTGCT CAACGTCAC GCCCTGAGCG TGGAACGCTA TGTGGCCGTG 420
GTGCACCCAC TCCAGGCCAG GTCCATGGTG ACGCGGGCCC ATGTGCGCCG AGTGCTTGGG 480
GCCGTCTGGG GTCTTGCCAT GCTCTGCTCC CTGCCAACA CCAGCCTGCA CGGCATCCGG 540
CAGCTGCACG TGCCCTGCCG GGGCCCAGTG CCAGACTCAG CTGTTTGCAT GCTGGTCCGC 600
CCACGGGGCC TCTACAACAT GGTAGTGCAG ACCACCGCG TGCTCTTCTT CTGCCTGCCC 660
ATGGCCATCA TGAGCGTGCT CTACCTGCTC ATTGGGCTGC GACTGCGGCG GGAGAGGCTG 720
CTGCTCATGC AGGAGGCCAA GGGCAGGGGC TCTGCAGCAG CCAGGTCCAG ATACACCTGC 780
AGGCTCCAGC AGCACGATCG GGGCCGAGA CAAGTGACCA AGATGCTGTT TGTCTGGTGC 840
GTGGTGTTTG GCATCTGCTG GGCCCCGTT 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 GGTCCACCC CCTGGCTGGG AACGATGGCC CAGAGGCGCA GCAAGAGACC 1200
GATCCATCCT GA 1212

```

(115) SEQ ID NO:114 的资料:

(i) 序列特征:

(A) 长度: 403 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:114 的序列描述:

```

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		
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		
385	390	395 400
Asp Pro Ser		

(116) SEQ ID NO:115 的资料:

(i) 序列特征:

- (A) 长度: 30 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:115 的序列描述:

GGAAGCTTCA GGCCCAAAGA TGGGGAACAT 30

(117) SEQ ID NO: 116 的资料:

(i) 序列特征:

- (A) 长度: 30 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 116 的序列描述:

GTGGATCCAC CCGCGGAGGA CCCAGGCTAG 30

(118) SEQ ID NO:117 的资料:

(i) 序列特征:

(A) 长度: 1098 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组)

(xi) SEQ ID NO:117 的序列描述:

```

ATGGGGAACA TCACTGCAGA CAACTCCTCG ATGAGCTGTA CCATCGACCA TACCATCCAC 60
CAGACGCTGG CCCCGGTGGT CTATGTTACC GTGCTGGTGG TGGGCTTCCC GGCCAACTGC 120
CTGTCCCTCT ACTTCGGCTA CCTGCAGATC AAGGCCGGA ACGAGCTGGG CGTGTACCTG 180
TGCAACCTGA CGGTGGCGA CCTCTTCTAC ATCTGCTCGC TGCCCTTCTG GCTGCAGTAC 240
GTGCTGCAGC ACGACAACCTG GTCTCACGGC GACCTGTCCT GCCAGGTGTG CGGCATCCTC 300
CTGTACGAGA ACATCTACAT CAGCGTGGGC TTCCTCTGCT GCATCTCCGT GGACCGCTAC 360
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 GCGTGCGCC 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 CTTGCTCAGC GAGACCACCC ACCGGGACCT GGCCCGCCTC 900
CGCGGGGCCT GCCTGGCCTT CCTCACCTGC TCCAGGACCG GCCGGGCCAG GGAGGCCTAC 960
CCGCTGGGTG CCCCCGAGGC CTCGGGAAA AGCGGGGCC AGGGTGAGGA GCCCGAGCTG 1020
TTGACCAAGC TCCACCCGGC CTTCCAGACC CTAACCTCGC CAGGGTCGGG CGGGTTCCTC 1080
ACGGGCAGGT TGGCCTAG

```

(119) SEQ ID NO:118 的资料:

(i) 序列特征:

(A) 长度: 365 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:118 的序列描述:

```

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

(120) SEQ ID NO:119 的资料:

(i) 序列特征:

- (A) 长度: 26 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组)

(xi) SEQ ID NO:119 的序列描述:

GACCTCGAGT CCTTCTACAC CTCATC 26

(121) SEQ ID NO: 120 的资料:

(i) 序列特征:

- (A) 长度: 30 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 120 的序列描述:

TGCTCTAGAT TCCAGATAGG TGAAAACCTG 30

(122) SEQ ID NO:121 的资料:

(i) 序列特征:

- (A) 长度: 1416 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 121 的序列描述:

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 TATTTCTCTG TGTCACTGTC 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 AAGGCATTTT TGAATATCAT TGCTGTTTGG 600
ACCATATCAG TAGGTATATC CATGCCAATA CCACTCTTTG GGCTACAGGA CGATTCGAAG 660
GTCTTTAAGG AGGGGAGTTG CTTACTCGCC GATGATAACT TTGTCCTGAT CGGCTCTTTT 720
GTGTCATTTT 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 CATCTGCAAA GAGTCCTGCA ATGAGGATGT CATTGGGGCC 1080
CTGCTCAATG TGTTTGTGTG GATCGGTTAT CTCTCTTCAG CAGTCAACCC ACTAGTCTAC 1140
ACACTGTTC AACAAGACCTA TAGGTCAGCC TTTTCACGGT ATATTCAGTG TCAGTACAAG 1200
GAAAACAAAA AACCATGCA GTTAATTTTA GTGAACACAA TACCGGCTTT GGCCTACAAG 1260
TCTAGCCAAC TTCAAATGGG ACAAAAAAAG AATTCAAAGC AAGATGCCAA GACAACAGAT 1320
AATGACTGCT CAATGGTTGC TCTAGGAAAG CAGTATTCTG AAGAGGCTTC TAAAGACAAT 1380
AGCGACGGAG TGAATGAAAA GGTGAGCTGT GTGTGA 1416

```

(123) SEQ ID NO:122 的资料:

(i) 序列特征:

(A) 长度: 471 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 122 的序列描述:

```

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
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
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
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	
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											

(124) SEQ ID NO:123 的资料:

(i) 序列特征:

- (A) 长度: 27 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 123 的序列描述:
GACCTCGAGG TTGCTTAAGA CTGAAGC

27

(125) SEQ ID NO: 124 的资料:

(i) 序列特征:

- (A) 长度: 27 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 124 的序列描述:
ATTCTAGAC ATATGTAGCT TGTACCG

27

(126) SEQ ID NO:125 的资料:

(i) 序列特征:

(A) 长度: 1377 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:125 的序列描述:

```

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 CCCCGTCTGG ATTTCTTTAG ATGTTTTATT TTCAACAGCG 420
TCCATCATGC ACCTCTGCGC TATATCGCTG GATCGGTATG TAGCAATACG TAATCCTATT 480
GAGCATAGCC GTTTC AATTC GCGGACTAAG GCCATCATGA AGATTGCTAT TGTGTGGCA 540
ATTTCTATAG GTGTATCAGT TCCTATCCCT GTGATTGGAC TGAGGGACGA AGAAAAGGTG 600
TTCGTGAACA ACACGACGTG CGTGCTCAAC GACCCAAATT TCGTTCTTAT TGGGTCCTTC 660
GTAGCTTTCT TCATACCGCT GACGATTATG GTGATTACGT ATTGCCTGAC CATCTACGTT 720
CTGCGCCGAC AAGCTTTGAT GTTACTGCAC GGCCACACCG AGGAACCGCC TGGACTAAGT 780
CTGGATTTC TGAAGTGCTG CAAGAGGAAT ACGGCCGAGG AAGAGAACTC TGCAAACCT 840
AACCAAGACC AGAACGCACG CCGAAGAAAG AAGAAGGAGA GACGTCCTAG GGGCACCATG 900
CAGGCTATCA ACAATGAAAG AAAAGCTTCG AAAGTCCTTG GGATTGTTTT CTTTGTGTTT 960
CTGATCATGT GGTGCCCATT TTTCATTACC AATATTCTGT CTGTTCTTTG TGAGAAGTCC 1020
TGTAACCAAA AGCTCATGGA AAAGCTTCTG AATGTGTTT TTTGGATTGG CTATGTTTGT 1080
TCAGGAATCA ATCCTCTGGT GTATACTCTG TTCAACAAA 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) SEQ ID NO:126 的资料:

(i) 序列特征:

(A) 长度: 458 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 126 的序列描述:

```

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
 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

```

(128) SEQ ID NO: 127 的资料:

(i) 序列特征:

(A) 长度: 30 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 127 的序列描述:

GGTAAGCTTG GCAGTCCACG CCAGGCCTTC

30

(129) SEQ ID NO:128 的资料:

(i) 序列特征:

(A) 长度: 30 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 128 的序列描述:

TCCGAATTCT CTGTAGACAC AAGGCTTTGG

30

(130) SEQ ID NO:129 的资料:

(i) 序列特征:

(A) 长度: 1068 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 129 的序列描述:

```

ATGGATCAGT TCCCTGAATC AGTGACAGAA AACTTTGAGT ACGATGATTT GGCTGAGGCC 60
TGTTATATTG GGGACATCGT GGTCTTTGGG ACTGTGTTCC TGTCCATATT CTA CTCCGTC 120
ATCTTTGCCA TTGGCCTGGT GGGAAATTG TTGGTAGTGT TTGCCCTCAC CAACAGCAAG 180
AAGCCCAAGA GTGTCACCGA CATTTACCTC CTGAACCTGG CCTTGTCTGA TCTGCTGTTT 240
GTAGCCACTT TGCCCTTCTG GACTCACTAT TTGATAAATG AAAAGGGCCT CCACAATGCC 300
ATGTGCAAAT TCACTACCGC CTTCTTCTTC ATCGGCTTTT TTGGAAGCAT ATTCTTCATC 360
ACCGTCATCA GCATTGATAG GTACCTGGCC ATCGTCCTGG CCGCCAATC CATGAACAAC 420
CGGACCGTGC AGCATGGCGT CACCATCAGC CTAGGCGTCT GGGCAGCAGC CATTTTGGTG 480
GCAGCACCCC AGTTCATGTT CACAAAGCAG AAAGAAAATG AATGCCTTGG TGA CTACCCC 540
GAGGTCTCC AGGAAATCTG GCCCGTGCTC CGCAATGTGG AAACAAATTT TCTTGCTTC 600
CTACTCCCC TGCTCATTAT GAGTTATTGC TACTTCAGAA TCATCCAGAC GCTGTTTTCC 660
TGCAAGAACC ACAAGAAAGC CAAAGCCATT AAAGTATCC TTCTGGTGGT CATCGTGTTT 720
TTCCTCTTCT GGACACCCTA CAACGTTATG ATTTTCCTGG AGACGCTTAA GCTCTATGAC 780
TTCTTTCCCA GTTGTGACAT GAGGAAGGAT CTGAGGCTGG CCCTCAGTGT GACTGAGACG 840
GTTGCATTTA GCCATTGTTG CCTGAATCCT CTCATCTATG CATTTGCTGG GGAGAAGTTC 900
AGAAGATACC TTTACCACCT GTATGGGAAA TGCCTGGCTG TCCTGTGTGG GCGCTCAGTC 960
CACGTTGATT TCTCCTCATC TGAATCACA AGGAGCAGGC ATGGAAGTGT TCTGAGCAGC 1020
AATTTTACTT ACCACACGAG TGATGGAGAT GCATTGCTCC TTCTCTGA 1068

```

(131) SEQ ID NO:130 的资料:

(i) 序列特征:

(A) 长度: 355 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:130 的序列描述:

```

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

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

(132) SEQ ID NO:131 的资料:

(i) 序列特征:

- (A) 长度: 32 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 131 的序列描述:

GATCTCCAGT AGGCATAAGT GGACAATTCT GG 32

(133) SEQ ID NO:132 的资料:

(i) 序列特征:

- (A) 长度: 30 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:132 的序列描述:

CTCCTTCGGT CCTCCTATCG TTGTCAGAAG 30

(134) SEQ ID NO: 133 的资料:

- (i) 序列特征:
 - (A) 长度: 30 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑学: 线形
- (ii) 分子类型: DNA (基因组的)
- (xi) SEQ ID NO: 133 的序列描述:
AGAAGGCCAA GATCGCGCGG CTGGCCCTCA

30

(135) SEQ ID NO:134 的资料:

- (i) 序列特征:
 - (A) 长度: 30 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑学: 线形
- (ii) 分子类型: DNA (基因组的)
- (xi) SEQ ID NO:134 的序列描述:
CGGCGCCACC GCACGAAAAA GCTCATCTTC

30

(136) SEQ ID NO:135 的资料:

- (i) 序列特征:
 - (A) 长度: 33 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑学: 线形
- (ii) 分子类型: DNA (基因组的)
- (xi) SEQ ID NO:135 的序列描述:
GCCAAGAAGC GGGTGAAGTT CCTGGTGGTG GCA

33

(137) SEQ ID NO:136 的资料:

- (i) 序列特征:
 - (A) 长度: 30 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑学: 线形

- (ii) 分子类型: DNA (基因组的)
- (xi) SEQ ID NO:136 的序列描述:
CAGGCGGAAG GTGAAAGTCC TGGTCCTCGT 30

(138) SEQ ID NO:137 的资料:

- (i) 序列特征:
- (A) 长度: 33 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑学: 线形
- (ii) 分子类型: DNA (基因组的)
- (xi) SEQ ID NO:137 的序列描述:
CGGCGCCTGC GGGCCAAGCG GCTGGTGGTG GTG 33

(139) SEQ ID NO: 138 的资料:

- (i) 序列特征:
- (A) 长度: 31 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑学: 线形
- (ii) 分子类型: DNA (基因组的)
- (xi) SEQ ID NO: 138 的序列描述:
CCAAGCACAA AGCCAAGAAA GTGACCATCA C 31

(140) SEQ ID NO: 139 的资料:

- (i) 序列特征:
- (A) 长度: 30 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑学: 线形
- (ii) 分子类型: DNA (基因组的)
- (xi) SEQ ID NO: 139 的序列描述:
GCGCCGGCGC ACCAAATGCT TGCTGGTGGT 30

(141) SEQ ID NO:140 的资料:

- (i) 序列特征:

(A) 长度: 41 个碱基对	
(B) 类型: 核酸	
(C) 链型: 单链	
(D) 拓扑学: 线形	
(ii) 分子类型: DNA (基因组的)	
(xi) SEQ ID NO: 140 的序列描述:	
CAAAAAGCTG AAGAAATCTA AGAAGATCAT CTTTATTGTC G	41
(142) SEQ ID NO:141 的资料:	
(i) 序列特征:	
(A) 长度: 30 个碱基对	
(B) 类型: 核酸	
(C) 链型: 单链	
(D) 拓扑学: 线形	
(ii) 分子类型: DNA (基因组的)	
(xi) SEQ ID NO: 141 的序列描述:	
CAAGACCAAG GCAAAACGCA TGATCGCCAT	30
(143) SEQ ID NO:142 的资料:	
(i) 序列特征:	
(A) 长度: 30 个碱基对	
(B) 类型: 核酸	
(C) 链型: 单链	
(D) 拓扑学: 线形	
(ii) 分子类型: DNA (基因组的)	
(xi) SEQ ID NO:142 的序列描述:	
GTCAAGGAGA AGTCCAAAAG GATCATCATC	30
(144) SEQ ID NO:143 的资料:	
(i) 序列特征:	
(A) 长度: 30 个碱基对	
(B) 类型: 核酸	
(C) 链型: 单链	
(D) 拓扑学: 线形	
(ii) 分子类型: DNA (基因组的)	
(xi) SEQ ID NO: 143 的序列描述:	

CGCCGCGTGC GGGCCAAGCA GCTCCTGCTC 30

(145) SEQ ID NO: 144 的资料:

(i) 序列特征:

- (A) 长度: 33 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:144 的序列描述:
CCTGATAAGC GCTATAAAAT GGTCTGTTC CGA 33

(146) SEQ ID NO:145 的资料:

(i) 序列特征:

- (A) 长度: 36 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:145 的序列描述:
GAAAGACAAA AGAGAGTCAA GAGGATGTCT TTATTG 36

(147) SEQ ID NO:146 的资料:

(i) 序列特征:

- (A) 长度: 33 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:146 的序列描述:
CGGAGAAAGA GGGTGAAACG CACAGCCATC GCC 33

(148) SEQ ID NO:147 的资料:

(i) 序列特征:

- (A) 长度: 30 个碱基对
- (B) 类型: 核酸

- (C) 链型: 单链
- (D) 拓扑学: 线形
- (ii) 分子类型: DNA (基因组的)
- (xi) SEQ ID NO:147 的序列描述:
AAGCTTCAGC GGGCCAAGGC ACTGGTCACC 30
- (149) SEQ ID NO:148 的资料:
- (i) 序列特征:
- (A) 长度: 30 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形
- (ii) 分子类型: DNA (基因组的)
- (xi) SEQ ID NO:148 的序列描述:
CAGCGGCAGA AGGCAAAAAG GGTGGCCATC 30
- (150) SEQ ID NO:149 的资料:
- (i) 序列特征:
- (A) 长度: 30 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形
- (ii) 分子类型: DNA (基因组的)
- (xi) SEQ ID NO:149 的序列描述:
CGGCAGAAGG CGAAGCGCAT GATCCTCGCG 30
- (151) SEQ ID NO:150 的资料:
- (i) 序列特征:
- (A) 长度: 30 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形
- (ii) 分子类型: DNA (基因组的)
- (xi) SEQ ID NO:150 的序列描述:
GAGCGCAACA AGGCCAAAAA GGTGATCATC 30

(152) SEQ ID NO:151 的资料:

(i) 序列特征:

(A) 长度: 39 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:151 的序列描述:

GGTGTAACA AAAAGGCTAA AAACACAATT ATTCTTATT

39

(153) SEQ ID NO:152 的资料:

(i) 序列特征:

(A) 长度: 27 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:152 的序列描述:

GAGAGCCAGC TCAAGAGCAC CGTGCTG

27

(154) SEQ ID NO:153 的资料:

(i) 序列特征:

(A) 长度: 30 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:153 的序列描述:

CCACAAGCAA ACCAAGAAAA TGCTGGCTGT

30

(155) SEQ ID NO:154 的资料:

(i) 序列特征:

(A) 长度: 30 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)
(xi) SEQ ID NO:154 的序列描述:
CATCAAGTGT ATCATGTGCC AAGTACGCCC 30

(156) SEQ ID NO:155 的资料:
(i) 序列特征:
(A) 长度: 34 个碱基对
(B) 类型: 核酸
(C) 链型: 单链
(D) 拓扑学: 线形
(ii) 分子类型: DNA (基因组的)
(xi) SEQ ID NO:155 的序列描述:
CTAGAGAGTC AGATGAAGTG TACAGTAGTG GCAC 34

(157) SEQ ID NO:156 的资料:
(i) 序列特征:
(A) 长度: 36 个碱基对
(B) 类型: 核酸
(C) 链型: 单链
(D) 拓扑学: 线形
(ii) 分子类型: DNA (基因组的)
(xi) SEQ ID NO:156 的序列描述:
CTAGAGAGTC AGATGAAGTG TACAGTAGTG GCAC 34

(158) SEQ ID NO: 157 的资料:
(i) 序列特征:
(A) 长度: 33 个碱基对
(B) 类型: 核酸
(C) 链型: 单链
(D) 拓扑学: 线形
(ii) 分子类型: DNA (基因组的)
(xi) SEQ ID NO: 157 的序列描述:
GCTGAGGTTC GCAATAAACT AACCATGTTT GTG 33

(159) SEQ ID NO: 158 的资料:
(i) 序列特征:

(A) 长度: 29 个碱基对	
(B) 类型: 核酸	
(C) 链型: 单链	
(D) 拓扑学: 线形	
(ii) 分子类型: DNA (基因组的)	
(xi) SEQ ID NO: 158 的序列描述:	
GGGAGGCCGA GCTGAAAGCC ACCCTGCTC	29

(160) SEQ ID NO:159 的资料:	
(i) 序列特征:	
(A) 长度: 31 个碱基对	
(B) 类型: 核酸	
(C) 链型: 单链	
(D) 拓扑学: 线形	
(ii) 分子类型: DNA (基因组的)	
(xi) SEQ ID NO: 159 的序列描述:	
CAAGATCAAG AGAGCCAAAA CCTTCATCAT G	31

(161) SEQ ID NO:160 的资料:	
(i) 序列特征:	
(A) 长度: 31 个碱基对	
(B) 类型: 核酸	
(C) 链型: 单链	
(D) 拓扑学: 线形	
(ii) 分子类型: DNA (基因组的)	
(xi) SEQ ID NO: 160 的序列描述:	
CCGGAGACAA GTGAAGAAGA TGCTGTTTGT C	31

(162) SEQ ID NO:161 的资料:	
(i) 序列特征:	
(A) 长度: 30 个碱基对	
(B) 类型: 核酸	
(C) 链型: 单链	
(D) 拓扑学: 线形	
(ii) 分子类型: DNA (基因组的)	
(xi) SEQ ID NO: 161 的序列描述:	

GCAAGGACCA GATCAAGCGG CTGGTGCTCA

30

(163) SEQ ID NO:162 的资料:

(i) 序列特征:

(A) 长度: 34 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 162 的序列描述:

CAAGAAAGCC AAAGCCAAGA AACTGATCCT TCTG

34

(164) SEQ ID NO: 163 的资料:

(i) 序列特征:

(A) 长度: 1068 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 163 的序列描述:

ATGGAAGATT TGGAGGAAAC ATTATTTGAA GAATTTGAAA ACTATTCCTA TGACCTAGAC 60
 TATTACTCTC TGGAGTCTGA TTTGGAGGAG AAAGTCCAGC TGGGAGTTGT TCACTGGGTC 120
 TCCCTGGTGT TATATTGTTT GGCTTTTGTG CTGGGAATTC CAGGAAATGC CATCGTCATT 180
 TGGTTCACGG GGCTCAAGTG GAAGAAGACA GTCACCACTC TGTGGTTCCT CAATCTAGCC 240
 ATTGCGGATT TCATTTTCTT TCTCTTTCTG 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 ACTCCGGGA CACTGTGGAG 540
 TTCAATAATC ATACTCTTTG CTATAACAAT TTTCAGAAGC ATGATCCTGA CCTCACTTTG 600
 ATCAGGCACC ATGTTCTGAC TTGGGTGAAA TTTATCATTG GCTATCTCTT CCCTTTGCTA 660
 ACAATGAGTA TTTGCTACTT GTGTCTCATC TTCAAGGTGA AGAAGCGAAC AGTCCTGATC 720
 TCCAGTAGGC ATAAGTGGAC AATTCTGGTT GTGGTTGTGG CCTTTGTGGT TTGCTGGACT 780
 CCTTATCACC TGTTTAGCAT TTGGGAGCTC ACCATTCAAC ACAATAGCTA TTCCACCAT 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) SEQ ID NO:164 的资料:

(i) 序列特征:

(A) 长度: 355 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:164 的序列描述:

```

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			

(166) SEQ ID NO:165 的资料:

(i) 序列特征:

- (A) 长度: 1089 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:165 的序列描述:

```

ATGGGCAACC ACACGTGGGA GGGCTGCCAC GTGGACTCGC GCGTGGACCA CCTCTTCCG 60
CCATCCCTCT ACATCTTTGT CATCGGCGTG GGGCTGCCCA CCAACTGCCT GGCTCTGTGG 120
GCGGCTTACC GCCAGGTGCA ACAGCGCAAC GAGCTGGGCG TCTACCTGAT GAACCTCAGC 180
ATCGCCGACC TGCTGTACAT CTGCACGCTG CCGCTGTGGG TGGACTACTT CCGTGCACCAC 240
GACAACTGGA TCCACGGCCC CGGGTCCTGC AAGCTCTTTG GGTTCATCTT CTACACCAAT 300
ATCTACATCA GCATCGCCTT CCTGTGCTGC ATCTCGGTGG ACCGCTACCT GGCTGTGGCC 360
CACCACTCC GCTTCGCCCG CCTGCGCCGC GTCAAGACCG CCGTGGCCGT GAGCTCCGTG 420
GTCTGGGCCA CGGAGCTGGG CGCCAACCTG GCGCCCCTGT TCCATGACGA GCTCTCCGA 480
GACCGCTACA ACCACACCTT CTGCTTTGAG AAGTTCCCA TGGAAGGCTG GGTGGCCTGG 540
ATGAACCTCT ATCGGGTGTT CGTGGGCTTC CTCTCCCGT GGGCGCTCAT GCTGCTGTGG 600
TACCGGGGCA TCCTGCGGGC CGTGCGGGGC AGCGTGTTCA CCGAGCGCCA GGAGAAGGCC 660
AAGATCGCGC GGCTGGCCCT CAGCCTCATC GCCATCGTGC TGGTCTGCTT TGCGCCCTAT 720
CACGTGCTCT TGCTGTCCCG CAGCGCCATC TACCTGGGCC GCCCCTGGGA CTGCGGCTTC 780
GAGGAGCGCG TCTTTTCTGC ATACCACAGC TACTGGCTT TCACCAGCCT CAACTGTGTG 840
GCGGACCCCA TCCTCTACTG CCTGGTCAAC GAGGGCGCCC GCAGCGATGT GGCCAAGGCC 900

```

```

CTGCACAACC TGCTCCGCTT TCTGGCCAGC GACAAGCCCC AGGAGATGGC CAATGCCTCG 960
CTCACCTGG AGACCCCACT CACCTCCAAG AGGAACAGCA CAGCCAAAGC CATGACTGGC 1020
AGCTGGGCGG CCACTCCGCC TTCCAGGGG GACCAGGTGC AGCTGAAGAT GCTGCCGCCA 1080
GCACAATGA 1089

```

(167) SEQ ID NO:166 的资料:

(i) 序列特征:

(A) 长度: 362 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:166 的序列描述:

```

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		
	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	

(168) SEQ ID NO: 167 的资料:

(i) 序列特征:

- (A) 长度: 1002 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 167 的序列描述:

```

ATGGAGTCCT CAGGCAACCC AGAGAGCACC ACCTTTTTTT ACTATGACCT TCAGAGCCAG 60
CCGTGTGAGA ACCAGGCCTG GGTCTTTGCT ACCCTCGCCA CCACTGTCCT GTACTGCCTG 120
GTGTTTCTCC TCAGCCTAGT GGGCAACAGC CTGGTCCTGT GGGTCCTGGT GAAGTATGAG 180
AGCCTGGAGT CCCTACCAA CATCTTCATC CTCAACCTGT GCCTCTCAGA CCTGGTGTTT 240

```

```

GCCTGCTTGT TGCCTGTGTG GATCTCCCCA TACCACTGGG GCTGGGTGCT GGGAGACTTC 300
CTCTGCAAAC TCCTCAATAT GATCTTCTCC ATCAGCCTCT ACAGCAGCAT CTTCTTCTCG 360
ACCATCATGA CCATCCACCG CTACCTGTCG GTAGTGAGCC CCCTCTCCAC CCTGCGCGTC 420
CCCACCCTCC GCTGCCGGGT 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
CACCGCACGA AAAAGCTCAT CTTCGCCATC GTGGTGGCCT ACTTCCTCAG CTGGGGTCCC 720
TACAACTTCA CCCTGTTTCT GCAGACGCTG TTTCGGACCC AGATCATCCG GAGCTGCGAG 780
GCCAAACAGC AGCTAGAATA CGCCTGCTC ATCTGCCGCA ACCTCGCCTT CTCCCACTGC 840
TGCTTTAACC CGGTGCTCTA TGTCTTCGTG GGGGTCAAGT TCCGCACACA CCTGAAACAT 900
GTTCTCCGGC AGTTCTGGTT CTGCCGGCTG CAGGCACCCA GCCCAGCCTC GATCCCCCAC 960
TCCCCTGGTG CCTTCGCCTA TGAGGGCGCC TCCTTCTACT GA 1002

```

(169) SEQ ID NO:168 的资料:

(i) 序列特征:

(A) 长度: 333 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:168 的序列描述:

```

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 Lys			
	210	215	220
Lys Leu Ile Phe Ala Ile Val Val Ala Tyr Phe Leu Ser Trp Gly Pro			
	225	230	235
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
Ser Pro Gly Ala Phe Ala Tyr Glu Gly Ala Ser Phe Tyr			
	325	330	

(170) SEQ ID NO:169 的资料:

(i) 序列特征:

- (A) 长度: 987 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:169 的序列描述:

ATGGACAACG CCTCGTTCTC GGAGCCCTGG CCCGCCAACG CATCGGGCCC GGACCCGGCG 60

```

CTGAGCTGCT CCAACGCGTC GACTCTGGCG CCGCTGCCGG CGCCGCTGGC GGTGGCTGTA 120
CCAGTTGTCT ACGCGGTGAT CTGCGCCGTG GGTCTGGCGG GCAACTCCGC CGTGCTGTAC 180
GTGTTGCTGC GGGCGCCCG 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
GCCGTGTGGG GGATCGTCAC ACTCGTCGTG CTGCCCTTCG CAGTCTTCGC CCGGCTAGAC 540
GACGAGCAGG GCCGGCGCCA GTGCGTGCTA GTCTTTCCGC AGCCCGAGGC CTTCTGGTGG 600
CGCGCGAGCC GCCTCTACAC GTCGTGCTG GGCTTCGCCA TCCCGTGTG CACCATCTGT 660
GTCCTCTATA CCACCTGCT GTGCCGGCTG CATGCCATGC GGCTGGACAG CCACGCCAAG 720
GCCCTGGAGC GCGCCAAGAA GCGGGTGAAG TTCCTGGTGG TGGCAATCCT GCGGTTGTGC 780
CTCCTCTGCT GGACGCCCTA CCACCTGAGC ACCGTGGTGG CGCTCACCAC CGACCTCCCG 840
CAGACGCCGC TGGTCATCGC TATCTCCTAC TTCATCACCA GCCTGACGTA CGCCAACAGC 900
TGCTCAACC CCTTCCTCTA CGCTTCCTG GACGCCAGCT TCCGCAGGAA CCTCCGCCAG 960
CTGATAACTT GCCGCGCGGC AGCCTGA 987

```

(171) SEQ ID NO:170 的资料:

(i) 序列特征:

(A) 长度: 328 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:170 的序列描述:

```

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		
130	135	140
Arg Arg Val Ala Gly Arg Thr Tyr Ser Ala Ala Arg Ala Val Ser Leu		
145	150	155
Ala Val Trp Gly Ile Val Thr Leu Val Val Leu Pro Phe Ala Val Phe		
	165	170
Ala Arg Leu Asp Asp Glu Gln Gly Arg Arg Gln Cys Val Leu Val Phe		
	180	185
Pro Gln Pro Glu Ala Phe Trp Trp Arg Ala Ser Arg Leu Tyr Thr Leu		
	195	200
Val Leu Gly Phe Ala Ile Pro Val Ser Thr Ile Cys Val Leu Tyr Thr		
	210	215
Thr Leu Leu Cys Arg Leu His Ala Met Arg Leu Asp Ser His Ala Lys		
	225	230
Ala Leu Glu Arg Ala Lys Lys Arg Val Lys Phe Leu Val Val Ala Ile		
	245	250
Leu Ala Val Cys Leu Leu Cys Trp Thr Pro Tyr His Leu Ser Thr Val		
	260	265
Val Ala Leu Thr Thr Asp Leu Pro Gln Thr Pro Leu Val Ile Ala Ile		
	275	280
Ser Tyr Phe Ile Thr Ser Leu Thr Tyr Ala Asn Ser Cys Leu Asn Pro		
	290	295
Phe Leu Tyr Ala Phe Leu Asp Ala Ser Phe Arg Arg Asn Leu Arg Gln		
	305	310
Leu Ile Thr Cys Arg Ala Ala Ala		
	325	

(172) SEQ ID NO: 171 的资料:

(i) 序列特征:

(A) 长度: 1002 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 171 的序列描述:

```

ATGCAGGCCG CTGGGCACCC AGAGCCCTT GACAGCAGGG GCTCCTTCTC CCTCCCCACG 60
ATGGGTGCCA ACGTCTCTCA GGACAATGGC ACTGGCCACA ATGCCACCTT CTCCGAGCCA 120
CTGCCGTTCC TCTATGTGCT CCTGCCCGCC GTGTACTCCG GGATCTGTGC TGTGGGGCTG 180
ACTGGCAACA CGGCCGTCAT CCTTGTAATC 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 TCACGGTCCT GGTTCGCCC 540
TTCTTCTCTT TCGCTGGCGT CTACAGCAAC GAGCTGCAGG TCCAAGCTG TGGGCTGAGC 600
TTCCCGTGGC CCGAGCAGGT CTGGTTCAAG GCCAGCCGTG TCTACACGTT GGTCTGGGC 660
TTCGTGCTGC CCGTGTGCAC CATCTGTGTG CTCTACACAG ACCTCCTGCG CAGGCTGCGG 720
GCCGTGCGGC TCCGCTCTGG AGCCAAGGCT CTAGGCAAGG CCAGGCGGAA GGTGAAAGTC 780
CTGGTCTCTG TCGTGCTGGC 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) SEQ ID NO:172 的资料:

(i) 序列特征:

(A) 长度: 333 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 172 的序列描述:

```

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

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

```

(174) SEQ ID NO:173 的资料:

(i) 序列特征:

(A) 长度: 1107 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学：线形

(ii) 分子类型：DNA（基因组的）

(xi) SEQ ID NO: 173 的序列描述：

```

ATGGTCCTTG AGGTGAGTGA CCACCAAGTG CTAAATGACG CCGAGGTTGC CGCCCTCCTG 60
GAGAACTTCA GCTCTTCCTA TGA CTATGGA 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 CCCCGGCCCG CGTGACCCTC ACCTGCCTGG CTGTCTGGGG GCTCTGCCTG 540
CTTTTCGCCC TCCAGACTT CATCTTCCTG TCGGCCACC ACGACGAGCG CCTCAACGCC 600
ACCCACTGCC AATACAATT CCCACAGGTG GGCCGCACGG CTCTGCGGGT GCTGCAGCTG 660
GTGGCTGGCT TTCTGCTGCC CCTGCTGGTC ATGGCCTACT GCTATGCCCA CATCCTGGCC 720
GTGCTGCTGG TTCCAGGGG 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 ATGTGCTCT TGCGCCTGGG CTGCCCCAAC 1020
CAGAGAGGGC TCCAGAGGCA GCCATCGTCT TCCCGCGGG ATTATCCTG GTCTGAGACC 1080
TCAGAGGCCT CCTACTCGGG CTTGTGA 1107

```

(175) SEQ ID NO:174 的资料：

(i) 序列特征：

(A) 长度：368 个氨基酸

(B) 类型：氨基酸

(C) 链型：

(D) 拓扑学：不相关

(ii) 分子类型：蛋白质

(xi) SEQ ID NO:174 的序列描述：

```

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
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
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	290	295	300
Gly Leu Gly Tyr Met His Cys Cys Leu Asn Pro Leu Leu Tyr Ala Phe	305	310	315
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

(176) SEQ ID NO:175 的资料:

(i) 序列特征:

(A) 长度: 1074 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 175 的序列描述:

```

ATGGCTGATG ACTATGGCTC TGAATCCACA TCTTCCATGG AAGACTACGT TAACTTCAAC 60
TTCCTGACT TCTACTGTGA GAAAAACAAT GTCAGGCAGT TTGCGAGCCA TTCCTCCCA 120
CCCTTGTA CTGCTCGTGT CATCGTGGGT GCCTTGGGCA ACAGTCTGT 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 CCTTCGTTGG CATGGCTTGC 660
TGCTATACCA TCATCATTCA CACCCTGATA CAAGCCAAGA AGTCTTCCAA GCACAAAGCC 720
AAGAAAGTGA CCATCACTGT CCTGACCGTC TTTGTCTTGT CTCAGTTTCC CTACAACATG 780
ATTTTGTGG TGCAGACCAT TGACGCCTAT GCCATGTTCA TCTCCAACATG TGCCGTTTCC 840
ACCAACATTG ACATCTGCTT CCAGGTCACC CAGACCATCG CCTTCTTCCA CAGTTGCCTG 900
AACCTGTTC TCTATGTTTT TGTGGGTGAG AGATTCCGCC GGGATCTCGT GAAAACCTG 960
AAGAACTTGG GTTGATCAG CCAGGCCAG TGGGTTTCAT TTACAAGGAG AGAGGGAAGC 1020
TTGAAGCTGT CGTCTATGTT GCTGGAGACA ACCTCAGGAG CACTCTCCCT CTGA 1074

```

(177) SEQ ID NO: 176 的资料:

(i) 序列特征:

(A) 长度: 357 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 176 的序列描述:

```

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

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

(178) SEQ ID NO:177 的资料:

(i) 序列特征:

- (A) 长度: 1110 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:177 的序列描述:

ATGGCCTCAT CGACCACTCG GGGCCCCAGG GTTCTGACT TATTTTCTGG GCTGCCGCCG 60
 GCGGTCACAA CTCCCGCCAA CCAGAGCGCA GAGGCCTCGG CGGGCAACGG GTCGGTGGCT 120
 GCGCGGACG CTCCAGCCGT CACGCCCTTC CAGAGCCTGC AGCTGGTGCA TCAGCTGAAG 180
 GGGCTGATCG TGCTGCTCTA CAGCGTCGTG GTGGTCGTGG GGCTGGTGGG CAACTGCCTG 240
 CTGGTGCTGG TGATCGCGCG GGTGCCGCGG CTGCACAACG TGACGAACTT CCTCATCGGC 300
 AACCTGGCCT TGTCCGACGT GCTCATGTGC ACCGCCTGCG TGCCGCTCAC GCTGGCCTAT 360
 GCCTTCGAGC CACGCGGCTG GGTGTTCCGC GCGCGCCTGT GCCACCTGGT CTCTTCCTG 420
 CAGCCGGTCA CCGTCTATGT GTCGGTGTTT ACGCTCACCA CCATCGCAGT GGACCGCTAC 480
 GTCGTGCTGG TGACCCGCT 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 CTCCGCAACC GCGTGGTGCC GGGCTGCGTG 780
 ACCCAGAGCC AGGCCGACTG GGACCGCGCT CGGCGCCGGC GCACCAAATG CTTGCTGGTG 840
 GTGGTCGTGG TGGTGTTCCG CGTCTGCTGG CTGCCGCTGC ACGTCTTCAA CCTGCTGCGG 900
 GACCTCGACC CCCACGCCAT CGACCCCTAC GCCTTTGGGC TGGTGCAGCT GCTCTGCCAC 960
 TGGCTCGCCA TGAGTTCGGC CTGCTACAAC CCCTTCATCT ACGCCTGGCT GCACGACAGC 1020
 TTCCGCGAGG AGCTGCGCAA ACTGTTGGTC GCTTGGCCCC GCAAGATAGC CCCCATGGC 1080
 CAGAATATGA CCGTCAGCGT GGTTCATCTGA 1110

(179) SEQ ID NO:178 的资料:

(i) 序列特征:

- (A) 长度: 369 个氨基酸
- (B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:178 的序列描述:

```

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

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 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
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

(180) SEQ ID NO: 179 的资料:

(i) 序列特征:

(A) 长度: 1083 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 179 的序列描述:

```

ATGGACCCAG AAGAACTTC AGTTTATTTG GATTATTACT ATGCTACGAG CCCAAACTCT 60
GACATCAGGG AGACCACTC CCATGTTCTT TACACCTCTG TCTTCCTTCC AGTCTTTTAC 120
ACAGCTGTGT TCCTGACTGG AGTGCTGGGG AACCTTGTTT TCATGGGAGC GTTGCATTTC 180
AAACCCGGCA GCCGAAGACT GATCGACATC TTTATCATCA ATCTGGCTGC CTCTGACTTC 240
ATTTTCTTGT TCACATTGCC TCTCTGGGTG GATAAAGAAG CATCTCTAGG ACTGTGGAGG 300
ACGGGCTCCT TCCTGTGCAA AGGGAGCTCC TACATGATCT CCGTCAATAT GCACTGCAGT 360
GTCCTCCTGC TCACTTGCAT 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 GAAATCTAAG 720
AAGATCATCT TTATTGTCGT GGCAGCCTTT CTTGTCTCCT GGCTGCCCTT CAATACTTTC 780
AAGTTCCTGG CCATTGTCTC TGGGTTGCGG CAAGAACACT ATTTACCCTC AGCTATTCTT 840
CAGCTTGATA TGGAGGTGAG TGGACCCTTG GCATTGCCA ACAGCTGTGT CAACCCTTTC 900
ATTTACTATA TCTTCGACAG CTACATCCGC CGGGCCATTG TCCACTGCTT GTGCCCTTGC 960

```

```

CTGAAAACT ATGACTTTGG GAGTAGCACT GAGACATCAG ATAGTCACCT CACTAAGGCT 1020
CTCTCCACCT TCATTCATGC AGAAGATTTT GCCAGGAGGA GGAAGAGGTC TGTGTCACTC 1080
TAA 1083

```

(181) SEQ ID NO: 180 的资料:

(i) 序列特征:

(A) 长度: 360 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 180 的序列描述:

```

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
 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

(182) SEQ ID NO:181 的资料:

(i) 序列特征:

- (A) 长度: 1020 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 181 的序列描述:

ATGAATGGCC TTGAAGTGGC TCCCCAGGT CTGATCACCA ACTTCTCCCT GGCCACGGCA 60
 GAGCAATGTG GCCAGGAGAC GCCACTGGAG AACATGCTGT TCGCCTCCTT CTACCTTCTG 120
 GATTTTATCC TGGCTTTAGT TGGCAATACC CTGGCTCTGT GGCTTTTCAT CCGAGACCAC 180
 AAGTCCGGGA CCCC GGCCAA CGTGTCCTG ATGCATCTGG CCGTGGCCGA CTTGTCTGTC 240
 GTGCTGGTCC TGCCACCCG CCTGGTCTAC CACTTCTCTG GGAACCACTG GCCATTTGGG 300
 GAAATCGCAT GCCGTCTCAC CGGCTTCCTC TTCTACCTCA ACATGTACGC CAGCATCTAC 360

TTCCTCACCT GCATCAGCGC CGACCGTTTC CTGGCCATTG TGCACCCGGT CAAGTCCCTC 420
 AAGCTCCGCA GGCCCTCTA CGCACACCTG GCCTGTGCCT TCCTGTGGGT GGTGGTGGCT 480
 GTGGCCATGG CCCCCTGCT GGTGAGCCCA CAGACCGTGC AGACCAACCA CACGGTGGTC 540
 TGCCTGCAGC TGTACCGGA GAAGGCCTCC CACCATGCCC TGGTGTCCCT GGCAGTGGCC 600
 TTCACCTTCC CGTTCATCAC CACGGTCACC TGCTACCTGC TGATCATCCG CAGCCTGCGG 660
 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
 CTTGAGAAAT TCCGCCCGC CTTGTGCAC TTGCTCTGTG GCAAAAGGCT CAAGGGCCCG 960
 CCCCCAGCT TCGAAGGGAA AACCAACGAG AGCTCGCTGA GTGCCAAGTC AGAGCTGTGA 1020

(183) SEQ ID NO:182 的资料:

(i) 序列特征:

(A) 长度: 339 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 182 的序列描述:

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			
	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	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			

(184) SEQ ID NO:183 的资料:

(i) 序列特征:

(A) 长度: 996 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 183 的序列描述:

ATGATCACCC TGAACAATCA AGATCAACCT GTCCCTTTTA ACAGCTCACA TCCAGATGAA 60
TACAAAATTG CAGCCCTTGT CTTCTATAGC TGTATCTTCA TAATTGGATT ATTTGTAAAC 120

```

ATCACTGCAT TATGGGTTTT CAGTTGTACC ACCAAGAAGA GAACACGGT AACCATCTAT 180
ATGATGAATG TGGCATTAGT GGAATTGATA TTTATAATGA CTTTACCCTT TCGAATGTTT 240
TATTATGCAA AAGATGAATG GCCATTTGGA GAGTACTTCT GCCAGATTCT TGGAGCTCTC 300
ACAGTGTITT ACCCAAGCAT TGCTTTATGG CTTCTTGCCT TTATTAGTGC TGACAGATAC 360
ATGGCCATTG TACAGCCGAA GTACGCCAAA GAACTTAAAA ACACGTGCAA AGCCGTGCTG 420
GCGTGTGTGG GAGTCTGGAT AATGACCCTG ACCACGACCA CCCCTCTGCT ACTGCTCTAT 480
AAAGACCCAG ATAAAGACTC CACTCCCGCC ACCTGCCTCA AGATTTCTGA CATCATCTAT 540
CTAAAAGCTG TGAACGTGCT GAACCTCACT CGACTGACAT TTTTCTTCTT GATTCCTTTG 600
TTCATCATGA TTGGGTGCTA CTTGGTCATT ATTCATAATC TCCTTCACGG CAGGACGTCT 660
AAGCTGAAAC CCAAAGTCAA GGAGAAGTCC AAAAGGATCA TCATCACGCT GCTGGTGCAG 720
GTGCTCGTCT GCTTTATGCC CTTCCACATC TGTTTCGCTT TCCTGATGCT GGAACGGGG 780
GAGAATAGTT ACAATCCCTG GGGAGCCTTT ACCACCTTCC TCATGAACCT CAGCACGTGT 840
CTGATGTGA TTCTCTACTA CATCGTTTCA AAACAATTTC AGGCTCGAGT CATTAGTGTC 900
ATGTATACC GTAATTACCT TCGAAGCATG CGCAGAAAAA GTTCCGATC TGGTAGTCTA 960
AGGTCACTAA GCAATATAAA CAGTGAAATG TTATGA 996

```

(185) SEQ ID NO:184 的资料:

(i) 序列特征:

(A) 长度: 331 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 184 的序列描述:

```

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

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

(186) SEQ ID NO: 185 的资料:

(i) 序列特征:

- (A) 长度: 1077 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 185 的序列描述:

ATGCCCTCTG TGTCTCCAGC GGGGCCCTCG GCCGGGCGAG TCCCAATGC CACCGCAGTG 60

ACAACAGTGC GGACCAATGC CAGCGGGCTG GAGGTGCCCC GTTCCACCT GTTTGCCCGG 120
 CTGGACGAGG AGCTGCATGG CACCTTCCCA GGCCTGTGCG TGGCGCTGAT GCGGTGCAC 180
 GGAGCCATCT TCCTGGCAGG GCTGGTGCTC AACGGGCTGG CGCTGTACGT CTTCTGCTGC 240
 CGCACCCGGG CCAAGACACC CTCAGTCATC TACACCATCA ACCTGGTGGT GACCGATCTA 300
 CTGGTAGGGC TGTCCTGCC CACGCGCTTC GCTGTGTACT ACGGCGCCAG GGGCTGCCTG 360
 CGCTGTGCCT TCCCGCACGT CCTCGGTAC TTCTCAACA TGA CTGCTC CATCTCTTC 420
 CTCACCTGCA TCTGCGTGGA CCGCTACCTG GCCATCGTGC GGGCCGAAGG CTCCCGCCGC 480
 TGCCGCCAGC CTGCCTGTGC CAGGGCCGTG TGCGCCTTCG TGTGGCTGGC CGCCGGTGCC 540
 GTCACCTGT CGGTGCTGGG CGTGACAGGC AGCGGGCCCT GCTGCCGTGT CTTTGCCTG 600
 ACTGTCTGG AGTTCCTGCT GCCCCTGCTG GTCATCAGCG TGTTTACCGG CCGCATCATG 660
 TGTGCACTGT CGCGGCCGGG TCTGCTCCAC CAGGGTCGCC AGCGCCGCGT GCGGGCCAAG 720
 CAGCTCTGCT TCACGGTGCT CATCATCTTT CTCGTCTGCT TCACGCCCTT CCACGCCCGC 780
 CAAGTGCCG TGGCGCTGTG GCGGACATG CCACACCACA CGAGCCTCGT GGTCTACCAC 840
 GTGGCCGTGA CCCTCAGCAG CCTCAACAGC TGATGGACC CCATCGTCTA CTGCTTCGTC 900
 ACCAGTGGCT TCCAGGCCAC CGTCCGAGGC CTCTTCGGCC AGCACGGAGA GCGTGAGCCC 960
 AGCAGCGGTG ACGTGGTCAG CATGCACAGG AGTCCAAGG GCTCAGGCCG TCATCACATC 1020
 CTCAGTGCCG GCCCTCACGC CCTCACCCAG GCCCTGGCTA ATGGGCCGA GGCTTAG 1077

(187) SEQ ID NO:186 的资料:

(i) 序列特征:

(A) 长度: 358 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:186 的序列描述:

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 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 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) SEQ ID NO:187 的资料:

(i) 序列特征:

(A) 长度: 1050 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:187 的序列描述:

```

ATGAACTCCA CCTTGGATGG TAATCAGAGC AGCCACCCTT TTTGCCTCTT GGCATTTGGC 60
TATTTGAAA CTGTCAATTT TTGCCTTTTG GAAGTATTGA TTATTGTCTT TCTAACTGTA 120
TTGATTATTT CTGGCAACAT CATTGTGATT TTTGTATTTC ACTGTGCACC TTTGTTGAAC 180
CATCACACTA CAAGTATTTT TATCCAGACT ATGGCATATG CTGACCTTTT TGTTGGGGTG 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 CCTTTTCCA CTGGGGCAA 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 TGTTTCGAAT CACTAGTGTA 780
TTTTACATCC TCTGGTTGCC ATATATCATC TACTTCTTGT TGGAAAGCTC CACTGGCCAC 840
AGCAACCGCT TCGCATCCTT CTTGACCACC TGGCTTGCTA TTAGTAACAG TTTCTGCAAC 900
TGTGTAATTT ATAGTCTCTC CAACAGTGTA TTCCAAAGAG GACTAAAGCG CCTCTCAGGG 960
GCTATGTGTA CTTCTGTGTC AAGTCAGACT ACAGCCAACG ACCCTTACAC AGTTAGAAGC 1020
AAAGGCCCTC TTAATGGATG TCATATCTGA 1050

```

(189) SEQ ID NO: 188 的资料:

(i) 序列特征:

(A) 长度: 349 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 188 的序列描述:

```

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
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
Ile Ser Glu Arg Gln Ala Arg Phe Ser Ser Gln Ser Gly Glu Thr Gly	225	230	235
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
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

(190) SEQ ID NO: 189 的资料:

(i) 序列特征:

(A) 长度: 1302 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 189 的序列描述:

```

ATGTGTTTT CTCCATTCT GGAAATCAAC ATGCAGTCTG AATCTAACAT TACAGTCCGA 60
GATGACATTG ATGACATCAA CACCAATATG TACCAACCAC TATCATATCC GTTAAGCTTT 120
CAAGTGCTCT TCACCGGATT TCTTATGTGA 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 TGTAATGTGA 480
ATGATATCCA TTTGGATTTT TTCTTTTTTC TCTTTCCTGA TTCCTTTTAT TGAGGTAAAT 540
TTTTTCAGTC TTCAAAGTGG AAATACCTGG GAAAAACAAGA CACTTTTATG TGTCAGTACA 600
AATGAATACT AACTGAACT GGAATGTAT TATCACCTGT TAGTACAGAT CCCAATATT 660
TTTTTCACTG 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 GTCAAGAGGA TGTCTTTATT GATTATTTCT 960
ACATTTCTTC TCTGCTGGAC ACCAATTCTT GTTTTAAATA CCACCATTTT ATGTTTAGGC 1020
CCAAGTGACC TTTTAGTAAA ATTAAGATTG TGTTTTTTAG TCGTGGCTTA TGGAACAAC 1080
ATATTTCAAC CTCTATTATA TGCATTCAC AGACAAAAAT TTCAAAGGT CTTGAAAAGT 1140
AAAATGAAAA AGCGAGTTGT TTCTATAGTA GAAGCTGATC CCCTGCCTAA TAATGCTGTA 1200
ATACACAACT CTGGATAGA TCCCAAAAGA AACAAAAAAA TTACCTTTGA AGATAGTGAA 1260
ATAAGAGAAA AACGTTTAGT GCCTCAGGTT GTCACAGACT AG 1302

```

(191) SEQ ID NO:190 的资料:

(i) 序列特征:

(A) 长度: 433 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 190 的序列描述:

```

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
 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) SEQ ID NO:191 的资料:

(i) 序列特征:

- (A) 长度: 1209 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 191 的序列描述:

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 CCGTGTTCGG 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 TGACGTCCTC 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
ATCGTGCTCT GTGAGACGTT CCGCAAACGC TTGGTCCTGT CGGTGAAGCC TGCAGCCCAG 1140
GGGCAGCTTC GCGCTGTCAG CAACGCTCAG ACGGCTGACG AGGAGAGGAC AGAAAGCAAA 1200
GGCACCTGA                                     1209

```

(193) SEQ ID NO:192 的资料:

(i) 序列特征:

(A) 长度: 402 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 192 的序列描述:

```

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
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			

(194) SEQ ID NO:193 的资料:

(i) 序列特征:

(A) 长度: 1128 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 193 的序列描述:

```

ATGGATGTGA CTTCCTAAGC CCGGGGCGTG GGCCTGGAGA TGTACCCAGG CACCGCGCAC 60
GCTGCGGCCC CCAACACCAC CTCCCCGAG CTCAACCTGT CCCACCGCT 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 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 GAAACGCATG 780
ATCCTCGCGG TGGTGTGGT CTTCTTCGTC TGCTGGCTGC CGGAGAACGT CTTTCATCAGC 840
GTGCACCTCC TGCAGCGGAC GCAGCCTGGG GCCGCTCCCT GCAAGCAGTC TTTCCGCCAT 900
GCCACCCCC TCACGGGCCA CATTGTCAAC CTCGCCGCT TCTCCAACAG CTGCCTAAAC 960
CCCCTCATCT ACAGCTTCTT CGGGGAGACC TTCAGGACA AGCTGAGGCT GTACATTGAG 1020
CAGAAAACAA ATTTGCCGGC CCTGAACCGC TTCTGTACG CTGCCCTGAA GGCCGTCATT 1080
CCAGACAGCA CCGAGCAGTC GGATGTGAGG TTCAGCAGTG CCGTGTGA 1128

```

(195) SEQ ID NO: 194 的资料:

(i) 序列特征:

(A) 长度: 375 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 194 的序列描述:

```

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
 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
 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) SEQ ID NO:195 的资料:

(i) 序列特征:

(A) 长度: 960 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:195 的序列描述:

```

ATGCCATTCC CAAACTGCTC AGCCCCAGC ACTGTGGTGG CCACAGCTGT GGGTGTCTTG 60
CTGGGGCTGG AGTGTGGGCT GGGTCTGCTG GGCAACGCGG TGGCGCTGTG GACCTTCCTG 120
TTCCGGGTCA GGGTGTGGAA GCCGTACGCT GTCTACCTGC TCAACCTGGC CCTGGCTGAC 180
CTGCTGTTGG CTGCGTGCCT GCCTTTCCTG GCCGCCTTCT ACCTGAGCCT CCAGGCTTGG 240
CATCTGGGCC GTGTGGGCTG CTGGGCCCTG CGCTTCCTGC TGGACCTCAG CCGCAGCGTG 300
GGGATGGCCT TCCTGGCCGC CGTGGCTTTG GACCGGTACC TCCGTGTGGT CCACCCTCGG 360
CTTAAGGTCA ACCTGTGTC 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
AAGGCACTGG TCACCTTGGT GGTGGTGCTG TTTGCTCTGT GCTTTCTGCC CTGCTTCCTG 720
GCCAGAGTCC TGATGCACAT CTCCAGAAT CTGGGGAGCT GCAGGGCCCT TTGTGCAGTG 780
GCTCATACCT CGGATGTAC GGGCAGCCTC ACCTACCTGC ACAGTGTCGT CAACCCCGTG 840
GTATACTGCT TCTCCAGCCC CACCTTCAGG AGCTCCTATC GGAGGGTCTT CCACACCCTC 900
CGAGGCAAAG GGCAGGCAGC AGAGCCCCCA GATTTCACC CCAGAGACTC CTATTCCTGA 960

```

(197) SEQ ID NO:196 的资料:

(i) 序列特征:

(A) 长度: 319 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:196 的序列描述:

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

(198) SEQ ID NO:197 的资料:

(i) 序列特征:

(A) 长度: 1143 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:197 的序列描述:

```

ATGGAGGAAG GTGGTGATTT TGACAACTAC TATGGGGCAG ACAACCAGTC TGAGTGTGAG 60
TACACAGACT GGAAATCCTC GGGGGCCCTC ATCCCTGCCA TCTACATGTT GGTCTTCCTC 120
CTGGGCACCA CGGGAACGG TCTGGTGCTC TGGACCGTGT TTCGGAGCAG CCGGGAGAAG 180
AGGCGCTCAG CTGATATCTT CATTGCTAGC CTGGCGGTGG CTGACCTGAC CTTCGTGGTG 240
ACGCTGCCCC TGTGGGCTAC CTACACGTAC CGGGACTATG ACTGGCCCTT TGGGACCTTC 300
TTCTGCAAGC TCAGCAGCTA CCTCATCTTC GTCAACATGT ACGCCAGCGT CTTCTGCCTC 360
ACCGGCCTCA GCTTCGACCG CTACCTGGCC ATCGTGAGGC CAGTGGCCAA TGCTCGGCTG 420
AGGTGCGGG TCAGCGGGGC CGTGGCCACG GCAGTTCTTT GGGTGCTGGC CGCCCTCCTG 480
GCCATGCCTG TCATGGTGTT ACGCACCACC GGGGACTTGG AGAACACCAC TAAGGTGCAG 540
TGCTACATGG ACTACTCCAT GGTGGCCACT GTGAGCTCAG AGTGGGCCTG GGAGGTGGGC 600
CTTGGGTCT 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
AACAGTGCC TCAACCCCTT CCTCTATGCC TTTTTCGACC CCCGCTCCG CCAGGCCTGC 960
ACCTCCATGC TCTGCTGTGG CCAGAGCAGG TGCGCAGGCA CCTCCCACAG CAGCAGTGGG 1020
GAGAAGTCAG CCAGCTACTC TTCGGGGCAC AGCCAGGGGC CCGGCCCAA CATGGGCAAG 1080
GGTGGAGAAC AGATGCACGA GAAATCCATC CCCTACAGCC AGGAGACCCT TGTGGTTGAC 1140
TAG 1143

```

(199) SEQ ID NO: 198 的资料:

(i) 序列特征:

(A) 长度: 380 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 198 的序列描述:

```

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
          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

(200) SEQ ID NO: 199 的资料:

(i) 序列特征:

(A) 长度: 1119 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 199 的序列描述:

ATGAACTACC CGTAACGCT GGAAATGGAC CTCGAGAACC TGGAGGACCT GTTCTGGGAA 60
CTGGACAGAT TGGACAATA TAACGACACC TCCCTGGTGG AAAATCATCT CTGCCCTGCC 120
ACAGAGGGTC CCCTCATGGC CTCCTTCAAG GCCGTGTTCC TGCCCGTGGC CTACAGCCTC 180
ATCTTCCTCC TGGGCGTGAT CGGCAACGTC CTGGTGCTGG TGATCCTGGA GCGGCACCGG 240
CAGACACGCA GTTCCACGGA GACCTTCCTG TTCCACCTGG CCGTGGCCGA CCTCTGCTG 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 CTTCTCCTCT 540
GCCTTGCCAG AGATTCTCTT CGCCAAAGTC AGCCAAGGCC ATCACAACAA CTCCCTGCCA 600
CGTTGCACCT TCTCCCAAGA GAACCAAGCA GAAACGCATG CCTGGTTTAC CTCCCGATTG 660
CTCTACCATG TGGCGGGATT CCTGCTGCCC ATGCTGGTGA TGGGCTGGTG CTACGTGGGG 720
GTAGTGACA GGTTCGCCA GGCCAGCGG CGCCCTCAGC GGCAGAAGGC AAAAAGGGTG 780
GCCATCCTGG TGACAAGCAT CTTCTTCCTC TGCTGGTCAC CCTACCACAT CGTCATCTTC 840
CTGGACACCC TGGCGAGGCT GAAGGCCGTG GACAATACCT GCAAGCTGAA TGGCTCTCTC 900
CCCGTGGCCA TCACCATGTG TGAGTTCCTG GGCCTGGCCC ACTGCTGCCT CAACCCCATG 960
CTTACACTT TCGCCGCGT GAAGTTCGCG AGTGACCTGT CGCGGCTCCT GACCAAGCTG 1020
GGCTGTACCG GCCCTGCCTC CCTGTGCCAG CTCTTCCCTA GCTGGCGCAG GAGCAGTCTC 1080

TCTGAGTCAG AGAATGCCAC CTCTCTCACC ACGTTCTAG

1119

(201) SEQ ID NO:200 的资料:

(i) 序列特征:

(A) 长度: 372 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 200 的序列描述:

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
 245 250 255
 Ala Lys Arg Val Ala Ile Leu Val Thr Ser Ile Phe Phe Leu Cys Trp
 260 265 270
 Ser Pro Tyr His Ile Val Ile Phe Leu Asp Thr Leu Ala Arg Leu Lys
 275 280 285
 Ala Val Asp Asn Thr Cys Lys Leu Asn Gly Ser Leu Pro Val Ala Ile
 290 295 300
 Thr Met Cys Glu Phe Leu Gly Leu Ala His Cys Cys Leu Asn Pro Met
 305 310 315 320
 Leu Tyr Thr Phe Ala Gly Val Lys Phe Arg Ser Asp Leu Ser Arg Leu
 325 330 335
 Leu Thr Lys Leu Gly Cys Thr Gly Pro Ala Ser Leu Cys Gln Leu Phe
 340 345 350
 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) SEQ ID NO:201 的资料:

(i) 序列特征:

(A) 长度: 1128 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 201 的序列描述:

ATGGATGTGA CTCCCAAGC CCGGGGCGTG GGCCTGGAGA TGTACCCAGG CACCGCGCAG 60
 CCTGCGGCCC CCAACACCAC CTCCCCGAG CTCAACCTGT CCCACCGCT CCTGGGCACC 120
 GCCCTGGCCA ATGGGACAGG TGAGCTCTCG GAGCACCAGC AGTACGTGAT CGGCCTGTTC 180
 CTCTCGTGCC TCTACACCAT CTTCTCTTC 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 GTTCGCACAC 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 CGGCCCCGGC GGCAGAAGGC GAAGCGCATG 780
ATCCTCGCGG TGGTGTGGT CTTCTTCGTC TGCTGGCTGC CGGAGAACGT CTTTCATCAGC 840
GTGCACCTCC TGCAGCGGAC GCAGCCTGGG GCCGCTCCCT GCAAGCAGTC TTTCCGCCAT 900
GCCACCCCC TCACGGGCCA CATTGTCAAC CTCACGCCT TCTCCAACAG CTGCCTAAAC 960
CCCCTCATCT ACAGCTTCT CGGGGAGACC TTCAGGACA AGCTGAGGCT GTACATTGAG 1020
CAGAAAACAA ATTTGCCGGC CCTGAACCGC TTCTGTACG CTGCCCTGAA GGCCGTCATT 1080
CCAGACAGCA CCGAGCAGTC GGATGTGAGG TTCAGCAGTG CCGTGTAG 1128

```

(203) SEQ ID NO:202 的资料:

(i) 序列特征:

(A) 长度: 375 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 202 的序列描述:

```

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 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

```

(204) SEQ ID NO:203 的资料:

(i) 序列特征:

(A) 长度: 1137 个碱基对

(B) 类型: 核酸

(C) 链型：单链

(D) 拓扑学：线形

(ii) 分子类型：DNA（基因组的）

(xi) SEQ ID NO: 203 的序列描述:

```

ATGGACCTGG GGAACCAAT GAAAAGCGTG CTGCTGGTGG CTCTCCTTGT CATTTTCCAG 60
GTATGCCTGT GTCAAGATGA GGTCACGGAC GATTACATCG GAGACAACAC CACAGTGGAC 120
TACACTTTGT TCGAGTCTTT GTGCTCCAAG AAGGACGTGC GGAACCTTAA AGCCTGGTTC 180
CTCCCTATCA TGTACTCCAT CATTTGTTTC GTGGGCCTAC TGGGCAATGG GCTGGTCTGT 240
TTGACCTATA TCTATTCAA GAGGCTCAAG ACCATGACCG ATACCTACCT GCTCAACCTG 300
GCGGTGGCAG ACATCCTCTT CCTCCTGACC CTTCCCTTCT GGGCCTACAG CGCGGCCAAG 360
TCCTGGGTCT TCGGTGTCCA CTTTGTCAAG CTCATCTTTG CCATCTACAA GATGAGCTTC 420
TTCAGTGGCA TGCTCCTACT TCTTGCATC 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 TCCCCCTGCT GGCCATGAGC 720
TTCTGTTACC TTGTATCAT CCGCACCCCTG CTCCAGGCAC GCAACTTTGA GCGCAACAAG 780
GCCAAAAGG TGATCATCGC TGTGGTCGTG GTCTTCATAG TCTTCCAGCT GCCCTACAAT 840
GGGGTGGTCC TGGCCAGAC GGTGGCCAAC TTCAACATCA CCAGTAGCAC CTGTGAGCTC 900
AGTAAGCAAC TCAACATCGC CTACGACGTC ACCTACAGCC TGGCCTGCGT CCGCTGCTGC 960
GTCAACCCCT TCTTGTACGC CTTATCGGC GTCAAGTTCC GCAACGATCT CTTCAAGCTC 1020
TTCAAGGACC TGGGCTGCCT CAGCCAGGAG CAGCTCCGGC AGTGGTCTTC CTGTGCGCAC 1080
ATCCGGCGCT CCTCCATGAG TGTGGAGGCC GAGACCACCA CCACCTTCTC CCCATAG 1137

```

(205) SEQ ID NO: 204 的资料:

(i) 序列特征:

(A) 长度: 378 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 204 的序列描述:

```

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
      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 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                    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

(206) SEQ ID NO:205 的资料:

(i) 序列特征:

(A) 长度: 1086 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:205 的序列描述:

```

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

```

(207) SEQ ID NO:206 的资料:

(i) 序列特征:

(A) 长度: 361 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:206 的序列描述:

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
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
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
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
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
Arg Met Leu Lys Arg Gln Val Ser Val Ser Ile Ser Ser Ala Val Lys		
	325	330
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	

(208) SEQ ID NO: 207 的资料:

(i) 序列特征:

- (A) 长度: 1446 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 207 的序列描述:

```

ATGCGGTGGC TGTGGCCCTT GGCTGTCTCT CTTGCTGTGA 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 CAACCTGTGC GTCATGTGCA TCGTGTGGCA CAGCTACTAC 480
CTGAAGAGCG CCTGGAATC CATCCTTGCC AGCCTGGCCC TCTGGGATTT TCTGGTCCCTC 540
TTTTTCTGCC TCCCTATTGT CATCTTCAAC GAGATCACCA AGCAGAGGCT ACTGGGTGAC 600
GTTTCTTGTC 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 CCCCACCATG 840
GGCACCCCTG ACTCATGCAT CATGAAACCC TCAGCCAGCC TGCCCGAGTC CCTGTATTCA 900
CTGGTGATGA CCTACCAGAA CGCCCGCATG TGGTGGTACT TTGGCTGCTA CTTCTGCCTG 960
CCCATCCTCT TCACAGTCAC CTGCCAGCTG GTGACATGGC GGGTGCAGAG 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 GGAGTGCAGC 1320
GGGGCTTCGG AGGCCTCTGC TGCCAATGGG TCGGACAACA AGCTCAAGAC CGAGGTGTCC 1380

```

TCTTCATCT ACTTCCACAA GCCCAGGGAG TCACCCCCAC TCCTGCCCCT GGGCACACCT 1440
TGCTGA 1446

(209) SEQ ID NO: 208 的资料:

(i) 序列特征:

(A) 长度: 481 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 208 的序列描述:

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 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 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
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		

465470475480

Cys

(210) SEQ ID NO:209 的资料:

- (i) 序列特征:

(A) 长度: 1101 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形
- (ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 209 的序列描述:

ATGTGGAACG CGACGCCAG CGAAGAGCCG GGGTTCAACC TCACACTGGC CGACCTGGAC 60

TGGGATGCTT CCCCCGGCAA CGACTCGCTG GCGACGAGC TGCTGCAGCT CTCCCCGCG 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 CTACGCCAGG 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 GCGGCGATGC TGTGCTGGGT GCCTCGCTCA GGGACCAGAA CCACAAGCAA 780

ACCAAGAAAA TGCTGGCTGT AGTGGTGTTT GCCTTCATCC TCTGCTGGCT CCCCTCCAC 840

GTAGGGCGAT ATTTATTTTC CAAATCCTTT GAGCCTGGCT CCTTGGAGAT TGCTCAGATC 900

AGCCAGTACT GCAACCTCGT GTCCTTTGTC CTCTTCTACC TCAGTGCTGC CATCAACCCC 960

ATTCTGTACA ACATCATGTC CAAGAAGTAC CGGGTGGCAG TGTTCAGACT TCTGGGATTC 1020

GAACCTTCTT CCCAGAGAAA GCTCTCCACT CTGAAAGATG AAAGTTCTCG GGCCTGGACA 1080

GAATCTAGTA TTAATACATG A 1101

(211) SEQ ID NO:210 的资料:

- (i) 序列特征:

(A) 长度: 366 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 210 的序列描述:

Met Trp Asn Ala Thr Pro Ser Glu Glu Pro Gly Phe Asn Leu Thr Leu

151015

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					
		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 Lys 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			
	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

(212) SEQ ID NO:211 的资料:

(i) 序列特征:

(A) 长度: 1842 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 211 的序列描述:

```

ATGCGAGCCC CGGGCGCGCT TCTCGCCCGC ATGTCGCGGC TACTGCTTCT GCTACTGCTC 60
AAGGTGTCTG CCTCTTCTGC CCTCGGGGTC GCCCTGCGT CCAGAAACGA AACTTGTCTG 120
GGGAGAGCT GTGCACCTAC AGTGATCCAG CGCCGCGGCA GGGACGCCTG GGGACCGGGA 180
AATTCTGCAA GAGACGTTCT GCGAGCCCGA GCACCCAGGG AGGAGCAGGG GGCAGCGTTT 240
CTTGCGGGAC CCTCCTGGGA CTGCGCGCG GCCCGGGCC GTGACCCGCG TGCAGGCAGA 300
GGGGCGGAGG CGTCGGCAGC CGGACCCCGG GGACCTCAA CCAGGCCACC TGGCCCCTGG 360
AGGTGGAAG GTGCTCGGG TCAGGAGCCT TCTGAACTT TGGGAGAGG GAACCCACG 420
GCCCTCCAGC TCTTCCTTCA GATCTCAGAG GAGGAAGAGA AGGTCCCAG AGGCGCTGGC 480
ATTTCCGGGC GTAGCCAGGA GCAGAGTGTG AAGACAGTCC CCGGAGCCAG CGATCTTTT 540
TACTGGCCAA GGAGAGCCGG GAAACTCCAG GGTTCACAC ACAAGCCCCT GTCCAAGACG 600
GCCAATGGAC TGGCGGGGCA CGAAGGGTGG ACAATTGCAC TCCCGGGCCG GCGCTGGCC 660
CAGAATGGAT CCTTGGGTGA AGGAATCCAT GAGCCTGGG GTCCCGCCG GGGAAACAGC 720
ACGAACCGGC GTGTGAGACT GAAGAACCCC TTCTACCGC TGACCCAGGA GTCCTATGGA 780
GCCTACGCGG TCATGTGTCT GTCCGTGGT ATCTTCGGG CCGGCATCAT TGGCAACCTG 840
GCGGTGATGT GCATCGTGTG CCACAACCTAC TACATGCGG GCATCTCAA CTCCCTCTTG 900
GCCAACCTGG CCTTCTGGGA CTTTCTCATC ATCTTCTTCT GCCTTCGCT GGTCTCTTC 960
CACGAGCTGA CCAAGAAGTG GCTGCTGGAG GACTTCTCCT GCAAGATCGT GCCCTATATA 1020
GAGGTGCGCT CTCTGGGAGT CACCACTTTC ACCTTATGTG CTCTGTGCAT AGACCGCTTC 1080
CGTGTGCCA CCAACGTACA GATGTACTAC GAAATGATCG AAAATTGTTC CTCAACAAC 1140
GCCAACTTG CTGTTATATG GGTGGGAGCT CTATTGTTAG CACTTCCAGA AGTTGTTCTC 1200
CGCCAGCTGA GCAAGGAGGA TTTGGGGTTT AGTGCCGAG CTCCGGCAGA AAGGTGCATT 1260
ATTAAGATCT CTCCTGATTT ACCAGACACC ATCTATGTT TAGCCCTCAC CTACGACAGT 1320
GCGAGACTGT GGTGGTATTT TGGCTGTAC TTTTGTTCG CCACGCTTT CACCATCACC 1380
TGCTCTCTAG TGA CTGCGAG GAAAATCCGC AAAGCAGAGA AAGCCTGTAC CCGAGGGAAT 1440
AAACGGCAGA TTCAACTAGA GAGTCAGATG AAGTGTACAG TAGTGGCACT GACCATTTTA 1500
TATGGATTTT GCATTATTCC TGAAAATATC TGCAACATTG TTAGTGCCTA CATGGCTACA 1560

```

GGGGTTTCAC AGCAGACAAT GGACCTCCTT AATATCATCA GCCAGTTCCT TTTGTTCTTT 1620
 AAGTCCTGTG TCACCCAGT CCTCCTTTTC TGTCTCTGCA AACCCCTTCAG TCGGGCCTTC 1680
 ATGGAGTGCT GCTGCTGTTG CTGTGAGGAA TGCATTCAGA AGTCTTCAAC GGTGACCAGT 1740
 GATGACAATG ACAACGAGTA CACCACGGAA CTCGAACTCT CGCCTTTCAG TACCATACGC 1800
 CGTGAAATGT CCACTTTTGC TTCTGTCGGA ACTCATTGCT GA 1842

(213) SEQ ID NO:212 的资料:

(i) 序列特征:

(A) 长度: 613 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 212 的序列描述:

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 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
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

465470475480

Lys Arg Gln Ile Gln Leu Glu Ser Gln Met Lys Cys Thr Val Val Ala

485490495

Leu Thr Ile Leu Tyr Gly Phe Cys Ile Ile Pro Glu Asn Ile Cys Asn

500505510

Ile Val Thr Ala Tyr Met Ala Thr Gly Val Ser Gln Gln Thr Met Asp

515520525

Leu Leu Asn Ile Ile Ser Gln Phe Leu Leu Phe Phe Lys Ser Cys Val

530535540

Thr Pro Val Leu Leu Phe Cys Leu Cys Lys Pro Phe Ser Arg Ala Phe

545550555560

Met Glu Cys Cys Cys Cys Cys Cys Glu Glu Cys Ile Gln Lys Ser Ser

565570575

Thr Val Thr Ser Asp Asp Asn Asp Asn Glu Tyr Thr Thr Glu Leu Glu

580585590

Leu Ser Pro Phe Ser Thr Ile Arg Arg Glu Met Ser Thr Phe Ala Ser

595600605

Val Gly Thr His Cys

610

(214) SEQ ID NO: 213 的资料:

(i) 序列特征:

- (A) 长度: 1248 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 213 的序列描述:

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

TTCGTCCTGC 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 AACTATTTC TCCCCTCCTC TTGGGAAGGC 660
ACTGCCTACA CTGTCATCCA CTTCTTGGTG GGCTTTGTGA TTCCATCTGT CCTCATAATT 720
TTATTTTACC AAAAGGTCAT AAAATATATT TGGAGAATAG GCACAGATGG CCGAACGGTG 780
AGGAGGACAA TGAACATTGT CCCTCGGACA AAAGTGAAAA CTAAAAAGAT 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
GCCAAACTA TTACCAAAGA CTCGATCTAT GACTCATTG ACAGAGAAGC CAAGGAAAAA 1200
AAGCTTGCTT GGCCCATTA CTCAAATCCA CCAAATACTT TTGTCTAA 1248

```

(215) SEQ ID NO:214 的资料:

(i) 序列特征:

(A) 长度: 415 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:214 的序列描述:

```

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 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 Glv 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

(216) SEQ ID NO:215 的资料:

(i) 序列特征:

(A) 长度: 1842 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:215 的序列描述:

```

ATGGGGCCCA CCCTAGCGGT TCCCACCCCA 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
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 CGTCCTCCCT 600
CTCCTCATCG TGGGTTTCTG CTACGTGAGG ATCTGGACCA AAGTGTGGC GGCCCGTGAC 660
CCTGCAGGGC AGAATCCTGA CAACCAACTT GCTGAGGTTT GCAATAAACT AACCATGTTT 720
GTGATCTTCC TCCTCTTTGC AGTGTGCTGG TGCCCTATCA ACGTGCTCAC TGTCTTGGTG 780
GCTGTCACTC CGAAGGAGAT GGCAGGCAAG ATCCCCAACT GGCTTTATCT TGCAGCCTAC 840
TTCATAGCCT ACTTCAACAG CTGCCTCAAC GCTGTGATCT ACGGGCTCCT CAATGAGAAT 900
TTCCGAAGAG AATACTGGAC CATCTTCCAT GCTATGCGGC ACCCTATCAT ATTCTTCTCT 960
GGCCTCATCA GTGATATTCG TGAGATGCAG GAGGCCCGTA CCCTGGCCCG CGCCCGTGCC 1020
CATGCTCGCG ACCAAGCTCG TGAACAAGAC CGTGCCCATG CCTGTCCTGC TGTGGAGGAA 1080
ACCCGATGA ATGTCCGGAA TGTCCATTA CCTGGTGATG CTGCAGCTGG CCACCCGAC 1140
CGTGCTCTG GCCACCCTAA GCCCCATTCC AGATCCTCCT CTGCCTATCG CAAATCTGCC 1200
TCTACCCACC ACAAGTCTGT CTTAGCCAC TCCAAGGCTG CCTCTGGTCA CCTCAAGCCT 1260
GTCTCTGGCC ACTCCAAGCC TGCTCTGGT CACCCCAAGT CTGCCACTGT CTACCCTAAG 1320
CCTGCTCTG TCCATTTCAA GGCTGACTCT GTCCATTTCA AGGGTGACTC TGTCCATTTC 1380
AAGCCTGACT CTGTTCAATT CAAGCCTGCT TCCAGCAACC CCAAGCCCAT CACTGGCCAC 1440
CATGTCTCTG CTGGCAGCCA CTCCAAGTCT GCCTTCAATG CTGCCACCAG CCACCCTAAA 1500
CCCATCAAGC CAGCTACCAG CCATGCTGAG CCCACCACTG CTGACTATCC CAAGCCTGCC 1560
ACTACCAGCC ACCCTAAGCC CGCTGCTGCT GACAACCCTG AGCTCTCTGC CTCCATTGC 1620
CCCGAGATCC CTGCCATTGC CCACCCTGTG TCTGACGACA GTGACCTCCC TGAGTCGGCC 1680
TCTAGCCCTG CCGCTGGGCC CACCAAGCCT GCTGCCAGCC AGCTGGAGTC TGACACCATC 1740
GCTGACCTTC CTGACCCTAC TGATGCTACT ACCAGTACCA ATGATTACCA TGATGCTGTC 1800
GTTGTTGATG TTGAAGATGA TCCTGATGAA ATGGCTGTGT GA 1842

```

(217) SEQ ID NO: 216 的资料:

(i) 序列特征:

(A) 长度: 613 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 216 的序列描述:

```

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
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
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
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
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
Ser Ser Pro Ala Ala Gly Pro Thr Lys Pro Ala Ala Ser Gln Leu Glu		
	565	570
Ser Asp Thr Ile Ala Asp Leu Pro Asp Pro Thr Val Val Thr Thr Ser		
	580	585
Thr Asn Asp Tyr His Asp Val Val Val Val Asp Val Glu Asp Asp Pro		
	595	600
Asp Glu Met Ala Val		
610		

(218) SEQ ID NO:217 的资料:

(i) 序列特征:

- (A) 长度: 1854 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组)

(xi) SEQ ID NO:217 的序列描述:

```

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
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 CGTCCTCCCT 600
CTCCTCATCG TGGGTTTCTG CTACGTGAGG ATCTGGACCA AAGTGTGGC GGCCCGTGAC 660
CCTGCAGGGC AGAATCCTGA CAACCAACTT GCTGAGGTTC GCAATAAACT AACCATGTTT 720
GTGATCTTCC TCCTCTTTGC AGTGTGCTGG TGCCCTATCA ACGTGCTCAC TGTCTTGGTG 780
GCTGTCACTC CGAAGGAGAT GGCAGGCAAG ATCCCCAACT GGCTTTATCT TGCAGCCTAC 840
TTCATAGCCT ACTTCAACAG CTGCCTCAAC GCTGTGATCT ACGGGCTCCT CAATGAGAAT 900
TTCCGAAGAG AATACTGGAC CATCTCCAT GCTATGCGGC ACCCTATCAT ATTCTTCTCT 960
GGCCTCATCA GTGATATTCG TGAGATGCAG GAGGCCCGTA CCCTGGCCCG CGCCCGTGCC 1020
CATGCTCGCG ACCAAGCTCG TGAACAAGAC CGTGCCCATG CCTGTCCTGC TGTGGAGGAA 1080
ACCCCGATGA ATGTCCGGAA TGTTCCATTA CCTGGTGATG CTGCAGCTGG CCACCCCGAC 1140
CGTGCCTCTG GCCACCCTAA GCCCCATTCC AGATCCTCCT CTGCCTATCG CAAATCTGCC 1200

```

TCTACCCACC ACAAGTCTGT CTTTAGCCAC TCCAAGGCTG CCTCTGGTCA CCTCAAGCCT 1260
GTCTCTGGCC ACTCCAAGCC TGCCTCTGGT CACCCCAAGT CTGCCACTGT CTACCCTAAG 1320
CCTGCCTCTG TCCATTTCOA GGCTGACTCT GTCCATTTCOA AGGGTGACTC TGTCCATTTC 1380
AAGCCTGACT CTGTTCAATTT 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
GCCTCCCATTT GCCCGAGAT CCCTGCCATT GCCCACCTG TGTCTGACGA CAGTGACCTC 1680
CCTGAGTCGG CCTCTAGCCC TGCCGCTGGG CCCACCAAGC CTGCTGCCAG CCAGCTGGAG 1740
TCTGACACCA TCGCTGACCT TCCTGACCCT ACTGTAGTCA CTACCAGTAC CAATGATTAC 1800
CATGATGTCG TGGTTGTTGA TGTGAAGAT GATCCTGATG AAATGGCTGT GTGA 1854

(219) SEQ ID NO:218 的资料:

(i) 序列特征:

- (A) 长度: 617 个氨基酸
- (B) 类型: 氨基酸
- (C) 链型:
- (D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:218 的序列描述:

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
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
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
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
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
545	550	555
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) SEQ ID NO:219 的资料:

(i) 序列特征:

- (A) 长度: 1548 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组)

(xi) SEQ ID NO:219 的序列描述:

```

ATGGGACATA ACGGGAGCTG GATCTCTCCA AATGCCAGCG AGCCGCACAA CGCGTCCGGC 60
GCCGAGGCTG CGGTGTGAA CCGCAGCGCG CTCGGGGAGT TCGGCGAGGC GCAGCTGTAC 120
CGCCAGTTCA CCACCACCGT GCAGGTCGTC ATCTTCATAG GCTCGTGCT CGGAACTTC 180

```

```

ATGGTGTTAT GGTCAACTTG CCGCACAACC GTGTTCAAAT CTGTCACCAA CAGGTTTCATT 240
AAAAACCTGG CCTGCTCGGG GATTGTGCC AGCCTGGTCT GTGTGCCCTT CGACATCATC 300
CTCAGCACCA GTCCTCACTG TTGCTGGTGG ATCTACACCA TGCTCTTCTG CAAGGTCGTC 360
AAATTTTTCG ACAAAGTATT CTGCTCTGTG ACCATCCTCA GCTTCCCTGC TATTGCTTTG 420
GACAGGTA CTCTAGTCCT 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 TGAAAGCCAC CCTGCTCTCC ATGGTGATGG TCTTCATCTT GTGTAGCGTG 840
CCCTATGCCA CCCTGGTCGT CTACCAGACT GTGCTCAATG TCCCTGACAC TTCGTCTTC 900
TTGCTGCTCA CTGCTGTTTG GCTGCCCAAA GTCTCCCTGC TGGCAAACCC TGTCTCTTT 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 AGCCCTGTG 1320
GAACCTGAAA CATTCCTGA TAAGTATTCC CTGCAGTTTG GCTTTGGGCC TTTTGAGTTG 1380
CCTCCTCAGT GGCTCTCAGA GACCCGAAAC AGCAAGAAGC GGCTGCTTCC CCCCTTGGGC 1440
AACACCCCAG AAGAGCTGAT CCAGACAAAG GTGCCCAAGG TAGGCAGGGT GGAGCGGAAG 1500
ATGAGCAGAA ACAATAAAGT GAGCATTTT CCAAAGGTGG ATTCCTAG 1548

```

(221) SEQ ID NO: 220 的资料:

(i) 序列特征:

(A) 长度: 515 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 220 的序列描述:

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
Ser Val Leu Tyr Pro Leu Glu Arg	Lys Ile Ser Asp Ala Lys Ser Arg	
145	150	155
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
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
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
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		
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
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		

(222) SEQ ID NO:221 的资料:

(i) 序列特征:

- (A) 长度: 1164 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 221 的序列描述:

```

ATGAATCGGC ACCATCTGCA GGATCACTTT CTGGAAATAG ACAAGAAGAA CTGCTGTGTG 60
TTCCGAGATG ACTTCATTGC CAAGGTGTTG CCGCCGGTGT TGGGGCTGGA GTTTATCTTT 120
GGGCTTCTGG GCAATGGCCT TGCCCTGTGG ATTTTCTGTT TCCACCTCAA GTCCTGGAAA 180
TCCAGCCGGA TTTTCTGTT CAACCTGGCA GTAGCTGACT TTCTACTGAT CATCTGCCTG 240
CCGTTCTGTA TGGACTACTA TGTGCGGCGT TCAGACTGGA AGTTTGGGGA CATCCCTTGC 300
CGGCTGGTGC TCTTCATGTT TGCCATGAAC CGCCAGGGCA GCATCATCTT CCTCACGGTG 360
GTGGCGGTAG ACAGGTATTT CCGGGTGGTC CATCCCCACC ACGCCCTGAA CAAGATCTCC 420
AATTGGACAG CAGCCATCAT CTCTGCCTT CTGTGGGGCA TCACTGTTGG CCTAACAGTC 480

```

```

CACCTCTGA AGAAGAAGTT GCTGATCCAG AATGGCCCTG CAAATGTGTG CATCAGCTTC 540
AGCATCTGCC ATACCTTCCG GTGGCACGAA GCTATGTTCC TCCTGGAGTT CCTCTGCCC 600
CTGGGCATCA TCCTGTTCTG CTCAGCCAGA ATTATCTGGA GCCTGCGGCA GAGACAAATG 660
GACCGGCATG CCAAGATCAA GAGAGCCAAA ACCTTCATCA TGGTGGTGGC CATCGTCTTT 720
GTCATCTGCT TCCTTCCCAG CGTGGTTGTG CGGATCCGCA TCTTCTGGCT CCTGCACACT 780
TCGGGCACGC AGAATTGTGA AGTGTACCGC TCGGTGGACC TGGCGTTCTT TATCACTCTC 840
AGCTTCACCT ACATGAACAG CATGCTGGAC CCCGTGGTGT ACTACTTCTC CAGCCCATCC 900
TTTCCCAACT TCTTCTCCAC TTTGATCAAC CGCTGCCTCC AGAGGAAGAT GACAGGTGAG 960
CCAGATAATA ACCGCAGCAC GAGCGTCGAG CTCACAGGGG ACCCCAACAA AACCAGAGGC 1020
GCTCCAGAGG CGTTAATGGC CAACTCCGGT GAGCCATGGA GCCCCTCTTA TCTGGGCCCA 1080
ACCTCAAATA ACCATTCCAA GAAGGGACAT TGTACCAAG AACCAGCATC TCTGGAGAAA 1140
CAGTTGGGCT GTGCATCGA GTAA 1164

```

(223) SEQ ID NO:222 的资料:

(i) 序列特征:

(A) 长度: 387 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 222 的序列描述:

```

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

```

```

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

(224) SEQ ID NO:223 的资料:

(i) 序列特征:

(A) 长度: 1212 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 不相关

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 223 的序列描述:

```

ATGGCTTGCA ATGGCAGTGC GGCCAGGGGG CACTTTGACC CTGAGGACTT GAACCTGACT 60
GACGAGGCAC TGAGACTCAA GTACCTGGGG CCCAGCAGA CAGAGCTGTT CATGCCATC 120
TGTGCCACAT ACCTGCTGAT CTTCTGGTGG GCGCTGTGG GCAATGGGCT GACCTGTCTG 180
GTCATCCTGC GCCACAAGGC CATGCGCAGC CCTACCAACT ACTACCTCTT CAGCCTGGCC 240
GTGTGGGACC TGCTGGTGCT GCTGGTGGGC CTGCCCCTGG AGCTCTATGA GATGTGGCAC 300
AACTACCCCT TCCTGCTGGG CGTTGGTGGC TGCTATTTCC GCACGCTACT GTTTGAGATG 360
GTCTGCCTGG CCTCAGTGCT CAACGTCAGT GCCCTGAGCG TGGAACGCTA TGTGGCCGTG 420
GTGCACCCAC TCCAGGCCAG GTCCATGGTG ACGCGGGCCC ATGTGCGCCG AGTGCTTGGG 480
GCCGTCTGGG GTCTTGCCAT GCTCTGCTCC CTGCCAACA CCAGCCTGCA CGGCATCCGG 540
CAGCTGCAGC TGCCCTGCCG GGGCCCAGTG CCAGACTCAG CTGTTTGCA TGTGGTCCGC 600
CCACGGGCCC TCTACAACAT GGTAGTGCAG ACCACCGCGC TGCTCTTCTT CTGCCTGCCC 660
ATGGCCATCA TGAGCGTGCT CTACCTGCTC ATTGGGCTGC GACTGCGGCG GGAGAGGCTG 720
CTGCTCATGC AGGAGGCCAA GGGCAGGGGC TCTGCAGCAG CCAGGTCCAG ATACACCTGC 780
AGGCTCCAGC AGCAGGATCG GGGCCGGAGA CAAGTGAAGA AGATGCTGTT TGTCTGGTGC 840
GTGGTGTGTT GCATCTGCTG GGCCCCGTTT CACGCCGACC GCGTCATGTG GAGCGTCGTG 900
TCACAGTGGA CAGATGGCCT GCACCTGGCC TTCCAGCAGC 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

```

(225) SEQ ID NO: 224 的资料:

(i) 序列特征:

(A) 长度: 403 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 224 的序列描述:

```

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

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

(226) SEQ ID NO:225 的资料:

(i) 序列特征:

(A) 长度: 1098 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学：线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:225 的序列描述:

ATGGGGAACA	TCACTGCAGA	CAACTCCTCG	ATGAGCTGTA	CCATCGACCA	TACCATCCAC	60
CAGACGCTGG	CCCCGGTGGT	CTATGTTACC	GTGCTGGTGG	TGGGCTTCCC	GGCCAACTGC	120
CTGTCCCTCT	ACTTCGGCTA	CCTGCAGATC	AAGGCCCGGA	ACGAGCTGGG	CGTGTACCTG	180
TGCAACCTGA	CGGTGCCGA	CCTCTTCTAC	ATCTGCTCGC	TGCCCTTCTG	GCTGCAGTAC	240
GTGCTGCAGC	ACGACAACTG	GTCTCACGGC	GACCTGTCTT	GCCAGGTGTG	CGGCATCCTC	300
CTGTACGAGA	ACATCTACAT	CAGCGTGGGC	TTCCTCTGCT	GCATCTCCGT	GGACCGCTAC	360
CTGGCTGTGG	CCCATCCCTT	CCGCTTCCAC	CAGTTCGGGA	CCCTGAAGGC	GGCCGTCGGC	420
GTCAGCGTGG	TCATCTGGGC	CAAGGAGCTG	CTGACCAGCA	TCTACTTCCT	GATGCACGAG	480
GAGGTCATCG	AGGACGAGAA	CCAGCACC GC	GTGTGCTTTG	AGCACTACCC	CATCCAGGCA	540
TGGCAGCGCG	CCATCAACTA	CTACC GCTTC	CTGGTGGGCT	TCCTCTTCCC	CATCTGCCTG	600
CTGCTGGCGT	CCTACCAGGG	CATCTGCGC	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	CCCCGAGGC	CTCCGGGAAA	AGCGGGGCC	AGGGTGAGGA	GCCCCGAGCTG	1020
TTGACCAAGC	TCCACCCGGC	CTTCAGACC	CCTAACTCGC	CAGGGTCGGG	CGGGTTCCCC	1080
ACGGGCAGGT	TGGCCTAG					1098

(227) SEO ID NO:226 的资料:

(i) 序列特征:

(A) 长度: 365 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 226 的序列描述:

```

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 Lys 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	

(228) SEQ ID NO: 227 的资料:

(i) 序列特征:

(A) 长度: 1416 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 227 的序列描述:

```

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 TATTTCTTGA TGTCATTGCT CATAGCTGAT 360
ATGCTGCTGG GTTTCCTTGT CATGCCCGTG TCCATGTAA CCATCCTGTA TGGGTACCGG 420
TGGCCTCTGC CGAGCAAGCT TTGTGCAGTC TGGATTACC TGGACGTGCT CTTCTCCACG 480
GCCTCCATCA TGCACCTCTG CGCCATCTCG CTGGACCGCT ACGTCGCCAT CCAGAATCCC 540
ATCCACCACA GCCGCTCAA CTCCAGAACT AAGGCATTTC TGAATAATCAT TGCTGTTTGG 600
ACCATATCAG TAGGTATATC CATGCCAATA CCAGTCTTTG GGCTACAGGA CGATTGGAAG 660
GTCTTTAAGG AGGGGAGTTG CTTACTCGCC GATGATAACT TTGTCCTGAT CGGCTCTTTT 720
GTGTCAATTT 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
GCAAAGAAGG TGCTGGGCAT CGTCTTCTTC CTGTTTGTGG TGATGTGGTG CCCTTTCTTC 1020
ATCACAAACA TCATGGCCGT CATCTGCAAA GAGTCCTGCA ATGAGGATGT CATTGGGGCC 1080
CTGCTCAATG TGTGTGTTG 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 CAATGGTTGC TCTAGGAAAG CAGTATTCTG AAGAGGCTTC TAAAGACAAT 1380
AGCGACGGAG TGAATGAAAA GGTGAGCTGT GTGTGA 1416

```

(229) SEQ ID NO:228 的资料:

(i) 序列特征:

(A) 长度: 470 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 228 的序列描述:

```

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 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

(230) SEQ ID NO:229 的资料:

(i) 序列特征:

(A) 长度: 1377 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 229 的序列描述:

```
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 TACTTGTGAT GCCCCTGTCT CTCCTGGCAA TCCTTTATGA TTATGTCTGG 360
CCACTACCTA GATATTTGTG CCCCGTCTGG ATTTCTTTAG ATGTTTTATT TTCAACAGCG 420
TCCATCATGC ACCTCTGCGC TATATCGCTG GATCGGTATG TAGCAATACG TAATCCTATT 480
GAGCATAGCC GTTTC AATTC GCGGACTAAG GCCATCATGA AGATTGCTAT TGTTTGGGCA 540
ATTTCTATAG GTGTATCAGT TCCTATCCCT GTGATTGGAC TGAGGGACGA AGAAAAGGTG 600
TTCGTGAACA ACACGACGTG CGTGCTCAAC GACCCAAATT TCGTTCTTAT TGGGTCCTTC 660
GTAGCTTTCT TCATACCGCT GACGATTATG GTGATTACGT ATTGCCTGAC CATCTACGTT 720
CTGCGCCGAC AAGCTTTGAT GTTACTGCAC GGCCACACCG AGGAACCGCC TGGACTAAGT 780
CTGGATTTCC TGAAGTGCTG CAAGAGGAAT ACGGCCGAGG AAGAGAACTC TGCAAACCCT 840
AACCAAGACC AGAACGCACG CCGAAGAAAG AAGAAGGAGA GACGTCCTAG GGGCACCATG 900
CAGGCTATCA ACAATGAAAG AAAAGCTAAG AAAGTCCTTG GGATTGTTTT CTTTGTGTTT 960
CTGATCATGT GGTGCCCATT 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 GGATTAGCAG TGTGTGA 1377
```

(231) SEQ ID NO:230 的资料:

(i) 序列特征:

(A) 长度: 458 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:230 的序列描述:

```

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		
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
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
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	

(232) SEQ ID NO:231 的资料:

(i) 序列特征:

- (A) 长度: 1068 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 231 的序列描述:

```

ATGGATCAGT TCCCTGAATC AGTGACAGAA AACTTTGAGT ACGATGATTT GGCTGAGGCC 60
TGTATATATT GGGACATCGT GGTCTTTGGG ACTGTGTTCC TGTCCATATT CTA CTCCGTC 120
ATCTTTGCCA TTGGCCTGGT GGGAAATTTG TTGGTAGTGT TTGCCCTCAC CAACAGCAAG 180
AAGCCCAAGA GTGTCACCGA CATTACCTC CTGAACCTGG CCTTGCTCTGA TCTGCTGTTT 240
GTAGCCACTT TGCCCTTCTG GACTCACTAT TTGATAAATG AAAAGGGCCT CCACAATGCC 300
ATGTGCAAAT TCACTACCGC CTTCTTCTTC ATCGGCTTTT TTGGAAGCAT ATTCTTCATC 360
ACCGTCATCA GCATTGATAG GTACCTGGCC ATCGTCCTGG CCGCCAACTC CATGAACAAC 420
CGGACCGTGC AGCATGGCGT CACCATCAGC CTAGGCGTCT GGGCAGCAGC CATTTTGGTG 480
GCAGACCCCC AGTTCATGTT CACAAAGCAG AAAGAAAATG AATGCCTTGG TGA CTACCCC 540
GAGGTCCTCC AGGAAATCTG GCCCGTGCTC CGCAATGTGG AAACAAATTT TCTTGGCTTC 600
CTACTCCCCC TGCTCATTAT GAGTTATTGC TACTTCAGAA TCATCCAGAC GCTGTTTTCC 660
TGCAAGAACC ACAAGAAAGC CAAAGCCAAG AAAGTATGCC TTCTGGTGGT CATCGTGTTT 720
TTCCTCTTCT GGACACCCTA CAACGTTATG ATTTTCCTGG AGACGCTTAA GCTCTATGAC 780
TTCTTTCCCA GTTGTGACAT GAGGAAGGAT CTGAGGCTGG CCCTCAGTGT GACTGAGACG 840
GTTGCATTTA GCCATTGTTG CCTGAATCCT CTCATCTATG CATTTGCTGG GGAGAAGTTC 900
AGAAGATACC TTTACCACCT GTATGGGAAA TGCCTGGCTG TCCTGTGTGG GCGCTCAGTC 960
CACGTTGATT TCTCCTCATC TGAATCACA AGGAGCAGGC ATGGAAGTGT TCTGAGCAGC 1020
AATTTTACTT ACCACACGAG TGATGGAGAT GCATTGCTCC TTCTCTGA 1068

```

(233) SEQ ID NO:232 的资料:

(i) 序列特征:

(A) 长度: 355 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:232 的序列描述:

```

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
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
Tyr Cys Tyr Phe Arg Ile Ile Gln Thr Leu Phe Ser Cys Lys Asn His		
210	215	220
Lys Lys Ala Lys Ala Lys Lys Leu Ile Leu Leu Val Val Ile Val Phe		
225	230	235
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
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		

(234) SEQ ID NO: 233 的资料:

(i) 序列特征:

- (A) 长度: 29 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(iv) 反义: 否

(xi) SEQ ID NO: 233 的序列描述:

GGCTTAAGAG CATCATCGTG GTGCTGGTG 29

(235) SEQ ID NO:234 的资料:

(i) 序列特征:

- (A) 长度: 34 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(iv) 反义: 是

(xi) SEQ ID NO:234 的序列描述:

GTCACCACCA GCACCACGAT GATGCTCTTA AGCC 34

(236) SEQ ID NO:235 的资料:

(i) 序列特征:

- (A) 长度: 31 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:235 的序列描述:

CAAAGAAAGT ACTGGGCATC GTCTTCTTCC T 31

(237) SEQ ID NO:236 的资料:

(i) 序列特征:

- (A) 长度: 30 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)
(xi) SEQ ID NO:236 的序列描述:
TGCTCTAGAT TCCAGATAGG TGAAAACCTG 30

(238) SEQ ID NO:237 的资料:
(i) 序列特征:
 (A) 长度: 50 个碱基对
 (B) 类型: 核酸
 (C) 链型: 单链
 (D) 拓扑学: 线形
(ii) 分子类型: DNA (基因组的)
(iv) 反义: 否
(xi) SEQ ID NO:237 的序列描述:
CTAGGGGCAC CATGCAGGCT ATCAACAATG AAAGAAAAGC TAAGAAAGTC 50

(239) SEQ ID NO: 238 的资料:
(i) 序列特征:
 (A) 长度: 50 个碱基对
 (B) 类型: 核酸
 (C) 链型: 单链
 (D) 拓扑学: 线形
(ii) 分子类型: DNA (基因组的)
(iv) 反义: 是
(xi) SEQ ID NO: 238 的序列描述:
CAAGGACTTT CTTAGCTTTT CTTTCATTGT TGATAGCCTG CATGGTGCCC 50

(240) SEQ ID NO: 239 的资料:
(i) 序列特征:
 (A) 长度: 35 个碱基对
 (B) 类型: 核酸
 (C) 链型: 单链
 (D) 拓扑学: 线形
(ii) 分子类型: DNA (基因组的)
(xi) SEQ ID NO: 239 的序列描述:
CGGCGGCAGA AGGCGAAACG CATGATCCTC GCGGT 35

(241) SEQ ID NO:240 的资料:

(i) 序列特征:

(A) 长度: 35 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 240 的序列描述:

ACCGCGAGGA TCATGCGTTT CGCCTTCTGC CGCCG

35

(242) SEQ ID NO:241 的资料:

(i) 序列特征:

(A) 长度: 24 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 241 的序列描述:

GAGACATATT ATCTGCCACG GAGG

24

(243) SEQ ID NO:242 的资料:

(i) 序列特征:

(A) 长度: 24 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:242 的序列描述:

TTGGCATAGA AACCGGACCC AAGG

24

(244) SEQ ID NO:243 的资料:

(i) 序列特征:

(A) 长度: 28 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 243 的序列描述:

taagaattcc ataaaaatta tggaatgg

28

(245) SEQ ID NO: 244 的资料:

- (i) 序列特征:
 - (A) 长度: 30 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑学: 线形
- (ii) 分子类型: DNA (基因组的)
- (xi) SEQ ID NO:244 的序列描述:

CCAGGATCCA GCTGAAGTCT TCCATCATTC30

(246) SEQ ID NO:245 的资料:

- (i) 序列特征:
 - (A) 长度: 1071 个碱基对
 - (B) 类型: 核酸
 - (C) 链型: 单链
 - (D) 拓扑学: 线形
- (ii) 分子类型: DNA (基因组的)
- (xi) SEQ ID NO:245 的序列描述:

ATGAATGGGG TCTCGGAGGG GACCAGAGGC TGCAGTGACA 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 CGATTTATG CTCTCACTGT CTCTGCCCAT TGCCATGTAC 300
TATATTGTCT CCAGGCAGTG GCTCCTCGGA GAGTGGGCCT GCAAACCTCTA CATCACCTTT 360
GTGTTCCCTCA 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 GCTTCCTGGG GCCCTTAGCA ATCATAGGCA CCTGCGCCCA CCTCATCCGG 720
GCCAAGCTCT TGCGGGAGGG CTGGGTCCAT GCCAACCGGC CCGCGAGGCT GCTGCTGGTG 780
CTGGTGAGCG CTTTCTTTAT CTTCTGGTCC CCGTTTAACG TGGTGTGTT GGTCCATCTG 840
TGGCGACGGG TGATGCTCAA GGAAATCTAC CACCCCGGA TGCTGCTCAT CCTCCAGGCT 900
AGCTTTGCCT TGGGCTGTGT CAACAGCAGC CTCAACCCT TCCTCTACGT CTTGTTGGC 960
AGAGATTTC AAGAAAAGTT TTTCCAGTCT TTGACTTCTG CCCTGGCGAG GCGGTTTGA 1020
GAGGAGGAGT TTCTGTCATC CTGTCCCGT GGCAACGCCC CCCGGAATG A 1071

(247) SEQ ID NO:246 的资料:

- (i) 序列特征:
 - (A) 长度: 356 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:246 的序列描述:

```

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

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

(248) SEQ ID NO:247 的资料:

(i) 序列特征:

- (A) 长度: 32 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:247 的序列描述:

GCAGAATTCG GCGGCCCCAT GGACCTGCCC CC 32

(249) SEQ ID NO:248 的资料:

(i) 序列特征:

- (A) 长度: 30 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:248 的序列描述:

GCTGGATCCC CCGAGCAGTG GCGTTACTTC 30

(250) SEQ ID NO:249 的资料:

(i) 序列特征:

(A) 长度: 903 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:249 的序列描述:

```

ATGGACCTGC CCCCCGAGCT 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 TATGCCGCGG 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 TCTCTGCTC TTTTCTGTC CCTTGCCAT CACAGCCTTC 600
TGCTACGTGG GCTGCCTCCG GGCCTGGCC CGCTCCGGCC TGACGCACAG GCGGAAGCTG 660
CGGGCCGCCT GGGTGGCCGG CGGGGCCCTC CTCACGCTGC TGCTCTGCGT AGGACCCTAC 720
AACGCCTCCA ACGTGGCCAG CTTCCTGTAC CCCAATCTAG GAGGCTCCTG GCGGAAGCTG 780
GGGCTCATCA CGGGTGCTG GAGTGTGGTG CTTAATCCGC TGGTGACCGG TTAATTGGGA 840
AGGGGTCTTG GCCTGAAGAC AGTGTGTGCG GCAAGAACGC AAGGGGGCAA GTCCAGAAG 900
TAA 903

```

(251) SEQ ID NO:250 的资料:

(i) 序列特征:

(A) 长度: 300 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:250 的序列描述:

```

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			
	275	280	285
Cys Ala Ala Arg Thr Gln Gly Gly Lys Ser Gln Lys			
	290	295	300

(252) SEQ ID NO:251 的资料:

(i) 序列特征:

(A) 长度: 31 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:251 的序列描述:

CTCAAGCTTA CTCTCTCTCA CCAGTGGCCA C

31

(253) SEQ ID NO:252 的资料:

(i) 序列特征:

(A) 长度: 24 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:252 的序列描述:

CCCTCCTCCC CCGGAGGACC TAGC

24

(254) SEQ ID NO:253 的资料:

(i) 序列特征:

(A) 长度: 1041 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:253 的序列描述:

ATGGATACAG GCCCGACCA GTCCTACTTC TCCGGAATC ACTGGTTCGT CTTCTCGGTG 60
TACCTTCTCA CTTTCTGGT GGGGCTCCCC CTCAACCTGC TGGCCCTGGT GGTCTTCGTG 120
GGCAAGCTGC AGCGCCGCC GGTGGCCGTG GACGTGCTCC TGCTCAACCT GACCGCTCG 180
GACCTGCTCC TGCTGTGTT CCTGCCTTTC CGCATGGTGG AGGCAGCCAA TGGCATGCAC 240
TGGCCCCTGC CCTTCATCCT CTGCCCCTC TCTGGATTCA TCTTCTTAC CACCATCTAT 300
CTACCGCCC TCTTCTGGC AGCTGTGAGC ATTGAACGCT TCCTGAGTGT GGCCACCCA 360
CTGTGGTACA AGACCGGCC GAGGCTGGG CAGGCAGGTC TGGTGAGTGT GGCCTGCTGG 420
CTGTTGGCCT CTGCTCACTG CAGCGTGGTC TACGTCATAG AATTCTCAGG GGACATCTCC 480
CACAGCCAGG GCACCAATGG GACCTGCTAC CTGGAGTTCC GGAAGGACCA GCTAGCCATC 540
CTCCTGCCCC TGCGGCTGGA GATGGCTGTG GTCCTCTTGG TGGTCCCGCT GATCATCACC 600
AGCTACTGCT ACAGCCGCTT GGTGTGGATC CTCGGCAGAG GGGGCAGCCA CCGCCGGCAG 660
AGGAGGGTGG CGGGGCTGTT GCGGCCACG CTGCTCAACT TCCTTGTCTG CTTTGGGCCC 720
TACAACGTGT CCCATGTCGT GGGCTATATC TCGGTGAAA GCCCGGCATG GAGGATCTAC 780
GTGACGCTTC TCAGCACCTT GAACTCCTGT GTCGACCCCT TTGTCTACTA CTTCTCCTCC 840
TCCGGGTTC AAGCCGACTT TCATGAGCTG CTGAGGAGGT TGTGTGGGCT CTGGGGCCAG 900
TGGCAGCAGG AGAGCAGCAT GGAGCTGAAG GAGCAGAAGG GAGGGGAGGA GCAGAGAGCG 960
GACCGACCAG CTGAAAGAAA GACCACTGAA CACTCACAGG GCTGTGGAAC TGGTGGCCAG 1020

GTGGCCTGTG CTGAAAGCTA G

1041

(255) SEQ ID NO:254 的资料:

(i) 序列特征:

(A) 长度: 346 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:254 的序列描述:

```

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

(256) SEQ ID NO:255 的资料:

(i) 序列特征:

- (A) 长度: 31 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:255 的序列描述:

TTTAAGCTTC CCCTCCAGGA TGCTGCCGGA C

31

(257) SEQ ID NO:256 的资料:

(i) 序列特征:

- (A) 长度: 31 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 不相关

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:256 的序列描述:

GGCGAATTCT GAAGGTCCAG GGAAACTGCT A

31

(258) SEQ ID NO: 257 的资料:

(i) 序列特征:

(A) 长度: 993 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 257 的序列描述:

```

ATGCTGCCGG ACTGGAAGAG CTCCTTGATC CTCATGGCTT ACATCATCAT CTTCCTCACT 60
GGCCTCCCTG CCAACCTCCT GGCCCTGCGG GCCTTTGTGG GGCGGATCCG 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 CTTACCGAT AACAGTTGG ACGTGGTGCT GCCCGTGCGG 540
CTGGAGCTGT GCCTGGTGCT CTTCTTCATC CCCATGGCAG TCACCATCTT CTGCTACTGG 600
CGTTTTGTGT GGATCATGCT CTCCCAGCCC CTTGTGGGGG CCCAGAGGCG GCGCCGAGCC 660
GTGGGGCTGG CTGTGGTGAC GCTGCTCAAT TTCCTGGTGT GCTTCGGACC TTACAACGTG 720
TCCCACCTGG TGGGGTATCA CCAGAGAAAA AGCCCTGGT 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) SEQ ID NO: 258 的资料:

(i) 序列特征:

(A) 长度: 362 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 258 的序列描述:

```

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 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
 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

(260) SEQ ID NO:259 的资料:

(i) 序列特征:

(A) 长度: 30 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 259 的序列描述:

CCCAAGCTTC GGCACCATG GACACCTCCC

30

(261) SEQ ID NO:260 的资料:

(i) 序列特征:

(A) 长度: 30 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 260 的序列描述:

ACAGGATCCA AATGCACAGC ACTGGTAAGC

30

(262) SEQ ID NO:261 的资料:

(i) 序列特征:

(A) 长度: 25 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 261 的序列描述:

CTATAACTGG GTTACATGGT TTAAC

25

(263) SEQ ID NO:262 的资料:

(i) 序列特征:

(A) 长度: 30 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 262 的序列描述:

TTTGAATTCA CATATTAATT AGAGACATGG

30

(264) SEQ ID NO: 263 的资料:

(i) 序列特征:

(A) 长度: 2724 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 263 的序列描述:

ATGGACACCT CCCGGCTCGG TGTGCTCCTG TCCTTGCCTG 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 TTCCTGGAGG AGTTACGTCT TCGGGGAAAC 300
GCTCTGACAT ACATTCCCAA GGGAGCATTC ACTGGCCTTT ACAGTCTTAA AGTTCCTTATG 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 ATGCCTTTGG AAACCTCTCC AGCTTGGTAG TTCTACATCT CCATAACAAT 660
AGAATCCACT CCCTGGGAAA GAAATGCTTT GATGGGCTCC ACAGCCTAGA GACTTTAGAT 720
TTAAATTACA ATAACCTGA TGAATTCCCC ACTGCAATTA GGACACTCTC CAACCTTAAA 780
GAACTAGGAT TTCATAGCAA CAATATCAGG TCGATACCTG AGAAAGCATT TGTAGGCAAC 840
CCTTCTCTTA TTACAATACA TTTCTATGAC AATCCCATCC AATTGTGTTG GAGATCTGCT 900
TTTCAACATT TACCTGAACT AAGAACACTG ACTCTGAATG GTGCCTCACA AATAACTGAA 960
TTTCCTGATT 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 AFTTGGCTTG GAACAAAATT GCTATTATTC ACCCCAATGC ATTTTCCACT 1260
TTGCCATCCC TAATAAGCT GGACCTATCG TCCAACCTCC TGTCGTCTTT TCCTATAACT 1320
GGGTACATG GTTTAACTCA CTTAAATTA ACAGGAAATC ATGCCTTACA GAGCTTGATA 1380
TCATCTGAAA ACTTTCCAGA ACTCAAGGTT ATAGAAATGC CTTATGCTTA CCAGTGCTGT 1440
GCATTTGGAG TGTGTGAGAA TGCCTATAAG ATTTCTAATC AATGGAATAA AGGTGACAAC 1500
AGCAGTATGG ACGACCTTCA TAAGAAAGAT GCTGGAATGT TTCAGGCTCA AGATGAACGT 1560
GACCTTGAAG ATTTCTGCT TGACTTTGAG GAAGACCTGA AAGCCCTTCA TTCAGTGCAG 1620
TGTTACCTT CCCCAGGCC CTTCAAACCC TGTGAACACC TGCTTGATGG CTGGCTGATC 1680

```

AGAATTGGAG TGTGGACCAT AGCAGTTCTG GCACTTACTT GTAATGCTTT GGTGACTTCA 1740
ACAGTTTTC AATCCCTCT GTACATTTCC CCCATTAAAC TGTTAATTGG GGTGATCGCA 1800
GCAGTGAACA TGCTCACGGG AGTCTCCAGT GCCGTGCTGG CTGGTGTGGA TCGGTTCACT 1860
TTTGGCAGCT TTGCACGACA TGGTGCCTGG TGGGAGAATG GGGTTGGTTG CCATGTCATT 1920
GGTTTTTTGT CCATTTTTC TTCAGAATCA TCTGTTTTCC TGCTTACTCT GGCAGCCCTG 1980
GAGCGTGGGT TCTCTGTGAA ATATTCTGCA AAATTTGAAA CGAAAGCTCC ATTTTCTAGC 2040
CTGAAAGTAA TCATTTTGCT CTGTGCCCTG CTGGCCTTGA CCATGGCCGC AGTTCCCCTG 2100
CTGGGTGGCA GCAAGTATGG CGCCTCCCCT CTCTGCCTGC CTTTGCCTTT TGGGGAGCCC 2160
AGCACCATGG GCTACATGGT CGCTCTCATC TTGCTCAATT CCCTTTGCTT CCTCATGATG 2220
ACCATGCGCT ACACCAAGCT CACTGCAAT TTGGACAAGG GAGACCTGGA GAATATTTGG 2280
GACTGCTCTA TGGTAAAACA CATTGCCCTG TTGCTCTTCA CCAACTGCAT CCTAAACTGC 2340
CCTGTGGCTT TCTGTCTCT CTCCTCTTTA ATAAACCTTA CATTTATCAG TCCTGAAGTA 2400
ATTAAGTTTA TCCTTCTGGT GGTAGTCCCA CTCCTGCAT GTCTCAATCC CCTCTCTAC 2460
ATCTTGTTCA ATCTCACTT TAAGGAGGAT CTGGTGAGCC TGAGAAAGCA AACCTACGTC 2520
TGGACAAGAT CAAACACCC AAGCTTGATG TCAATTAAT CTGATGATGT CGAAAAACAG 2580
TCCTGTGACT CAACTCAAGC CTTGGTAACC TTTACCAGCT CCAGCATCAC TTATGACCTG 2640
CCTCCCAGTT CCGTGCCATC ACCAGCTTAT CCAGTGACTG AGAGCTGCCA TCTTCTCTCT 2700
GTGGCATTG TCCCATGTCT CTAA 2724

```

(265) SEQ ID NO:264 的资料:

(i) 序列特征:

(A) 长度: 907 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:264 的序列描述:

```

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
 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
 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

(266) SEQ ID NO:265 的资料:

(i) 序列特征:

(A) 长度: 30 个碱基对

(B) 类型: 核酸
(C) 链型: 单链
(D) 拓扑学: 线形
(ii) 分子类型: DNA (基因组的)
(xi) SEQ ID NO:265 的序列描述:
CGGAAGCTGC GGGCCAAATG GGTGGCCGGC 30

(267) SEQ ID NO:266 的资料:
(i) 序列特征:
(A) 长度: 27 个碱基对
(B) 类型: 核酸
(C) 链型: 单链
(D) 拓扑学: 线形
(ii) 分子类型: DNA (基因组的)
(xi) SEQ ID NO:266 的序列描述:
CAGAGGAGGG TGAAGGGGCT GTTGCCG 27

(268) SEQ ID NO: 267 的资料:
(i) 序列特征:
(A) 长度: 30 个碱基对
(B) 类型: 核酸
(C) 链型: 单链
(D) 拓扑学: 线形
(ii) 分子类型: DNA (基因组的)
(xi) SEQ ID NO: 267 的序列描述:
GGCGGCGCCG AGCCAAGGGG CTGGCTGTGG 30

(269) SEQ ID NO:268 的资料:
(i) 序列特征:
(A) 长度: 32 个碱基对
(B) 类型: 核酸
(C) 链型: 单链
(D) 拓扑学: 线形
(ii) 分子类型: DNA (基因组的)
(xi) SEQ ID NO:268 的序列描述:
GGGACTGCTC TATGAAAAA CACATTGCCC TG 32

(270) SEQ ID NO:269 的资料:

(i) 序列特征:

(A) 长度: 1071 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:269 的序列描述:

```

ATGAATGGGG TCTCGAGGG GACCAGAGGC TGCAGTGACA GGCAACCTGG GGTCTGACA 60
CGTGATCGCT CTTGTTCCAG GAAGATGAAC TCTCCGGAT GCCTGTCTGA GGAGGTGGGG 120
TCCCTCCGCC CACTGACTGT GGTATCCTG TCTGCGTCCA TTGTCGTCGG AGTGCTGGGC 180
AATGGGCTGG TGCTGTGGAT GACTGTCTTC CGTATGGCAC GCACGGTCTC CACCGTCTGC 240
TTCTTCCACC TGGCCCTTGC CGATTTTCATG CTCTCACTGT CTCTGCCCAT TGCCATGTAC 300
TATATTGTCT CCAGGCAGTG GTCCTCGGA GAGTGGGCCT GCAAACCTCTA CATCACCTTT 360
GTGTTTCCTCA GCTACTTTGC CAGTAACTGC CTCCTGTCT TCATCTCTGT GGACCGTTGC 420
ATCTCTGTCC TCTACCCCGT CTGGGCCCTG AACCACCGCA CTGTGCAGCG GGCGAGCTGG 480
CTGGCCTTTG GGGTGTGGCT CCTGGCCGCC GCCTTGTGCT CTGCGCACCT GAAATTCCGG 540
ACAACCAGAA AATGGAATGG CTGTACGCAC TGCTACTTGG CGTTCAACTC TGACAATGAG 600
ACTGCCCAGA TTTGGATTGA AGGGGTCTGT GAGGGACACA TTATAGGGAC CATTGGCCAC 660
TTCCTGCTGG GCTTCCTGGG GCCCTTAGCA ATCATAGGCA CCTGCGCCCA CCTCATCCGG 720
GCCAAGCTCT TCGGGGAGGG CTGGGTCCAT GCCAACCGGC CCAAGAGGCT GCTGCTGGTG 780
CTGGTGAGCG CTTTCTTTAT CTTCTGGTCC CCGTTTAACG TGGTGCTGTT GGTCCATCTG 840
TGGCGACGGG TGATGCTCAA GGAAATCTAC CACCCCGGA TGCTGCTCAT CCTCCAGGCT 900
AGCTTTGCCT TGGGTGTGT CAACAGCAGC CTCAACCCCT TCCTCTACGT CTTTCGTTGGC 960
AGAGATTTCC AAGAAAAGTT TTTCCAGTCT TTGACTTCTG CCCTGGCGAG GGCGTTTGA 1020
GAGGAGGAGT TTCTGTCATC CTGTCCCGT GGCAACGCC CCCGGAATG A 1071

```

(271) SEQ ID NO:270 的资料:

(i) 序列特征:

(A) 长度: 356 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:270 的序列描述:

```

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
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
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
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			
	340	345	350
Ala Pro Arg Glu			
	355		

(272) SEQ ID NO: 271 的资料:

(i) 序列特征:

(A) 长度: 903 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 271 的序列描述:

```

ATGGACCTGC CCCCGCAGCT CTCCTTCGGC CTCTATGTGG CCGCCTTTGC GCTGGGCTTC 60
CCGCTCAACG TCCTGGCCAT CCGAGGCGCG ACGGCCACG CCCGGCTCCG TCTCACCCT 120
AGCCTGGTCT ACGCCCTGAA CCTGGGCTGC TCCGACCTGC TGCTGACAGT CTCTCTGCCC 180
CTGAAGGCGG TGGAGGCGCT AGCCTCCGGG GCCTGGCCTC TGCCGGCCTC GCTGTGCCCC 240
GTCTTCGCGG TGGCCCACTT CTTCCCACTC TATGCCGCGG GGGGCTTCCT GGCCGCCCTG 300
AGTGCAGGCC GCTACCTGGG AGCAGCCTTC CCCTTGGGCT ACCAAGCCTT CCGGAGGCCG 360
TGCTATTCCT GGGGGGTGTG CGCGGCCATC TGGGCCCTCG TCCTGTGTCA CCTGGGTCTG 420
GTCTTTGGGT TGGAGGCTCC AGGAGGCTGG CTGGACCACA GCAACACCTC CCTGGGCATC 480
AACACACCGG TCAACGGCTC TCCGGTCTGC CTGGAGGCCT GGGACCCGGC CTCTGCCGGC 540
CCGGCCCGCT TCAGCCTCTC TCTCCTGCTC TTTTCTGTC CCTTGCCAT CACAGCCTTC 600
TGCTACGTGG GCTGCCTCCG GGCCTGCGC CGCTCCGGCC TGACGCACAG GCGGAAGCTG 660
CGGGCCAAAT GGGTGGCCGG CGGGGCCCTC CTCACGCTGC TGCTCTGCGT AGGACCCTAC 720
AACGCCTCCA ACGTGGCCAG CTTCTGTAC CCCAATCTAG GAGGCTCCTG GCGGAAGCTG 780
GGGCTCATCA CGGGTGCCTG GAGTGTGGTG CTTAATCCGC TGGTGACCGG TTAATTGGGA 840
AGGGGTCTTG GCCTGAAGAC AGTGTGTGCG GCAAGAACGC AAGGGGGCAA GTCCAGAAG 900
TAA
903

```

(273) SEQ ID NO:272 的资料:

(i) 序列特征:

(A) 长度: 300 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 272 的序列描述:

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
 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) SEQ ID NO:273 的资料:

(i) 序列特征:

(A) 长度: 1041 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 273 的序列描述:

```

ATG GATA CAG GCCCG ACCA GTCCTACTTC TCCGGAATC ACTGGTTCGT CTTCTCGGTG 60
TACCTTCTCA CTTTCCTGGT GGGGCTCCCC CTCAACCTGC TGGCCCTGGT GGTCTTCGTG 120
GGCAAGCTGC AGCGCCGCC GGTGGCCGTG GACGTGCTCC TGCTCAACCT GACCGCTCG 180
GACCTGCTCC TGCTGCTGTT CCTGCCTTTC CGCATGGTGG AGGCAGCCAA TGGCATGCAC 240
TGGCCCCCTGC CCTTCATCCT CTGCCCCATC TCTGGATTCA TCTTCTTCAC CACCATCTAT 300
CTCACCGCCC TCTTCCTGGC AGCTGTGAGC ATTGAACGCT TCCTGAGTGT GGCCACCCA 360
CTGTGGTACA AGACCCGCC GAGGCTGGGG CAGGCAGGTC TGGTGAGTGT GGCCTGCTGG 420
CTGTTGGCCT CTGCTCACTG CAGCGTGGTC TACGTCATAG AATTCTCAGG GGACATCTCC 480
CACAGCCAGG GCACCAATGG GACCTGCTAC CTGGAGTTCC GGAAGGACCA GCTAGCCATC 540
CTCCTGCCCG TCGGGCTGGA GATGGCTGTG GTCCTCTTTG TGGTCCCGCT GATCATCACC 600
AGCTACTGCT ACAGCCGCCT GGTGTGGATC CTCGGCAGAG GGGGCAGCCA CCGCCGGCAG 660
AGGAGGGTGA AGGGGCTGTT GGCGGCCACG CTGCTCAACT TCCTTGTCTG CTTTGGGCCC 720
TACAACGTGT CCCATGTCGT GGGCTATATC TGCGGTGAAA 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

```

(275) SEQ ID NO:274 的资料:

(i) 序列特征:

(A) 长度: 346 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:274 的序列描述:

```

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
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
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		
225	230	235
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) SEQ ID NO:275 的资料:

(i) 序列特征:

- (A) 长度: 993 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO: 275 的序列描述:

```

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 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 CTCCAGCCC CTTGTGGGGG CCCAGAGGCG GCGCCGAGCC 660
AAGGGGCTGG CTGTGGTGAC GCTGCTCAAT TTCCTGGTGT GCTTCGGACC TTACAACGTG 720
TCCCACCTGG TGGGGTATCA CCAGAGAAAA AGCCCCTGGT GCGGTCAAT 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

```

(277) SEQ ID NO: 276 的资料:

(i) 序列特征:

- (A) 长度: 330 个氨基酸
- (B) 类型: 氨基酸
- (C) 链型:
- (D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO: 276 的序列描述:

```

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

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

(278) SEQ ID NO:277 的资料:

(i) 序列特征:

(A) 长度: 2724 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑学: 线形

(ii) 分子类型: DNA (基因组的)

(xi) SEQ ID NO:277 的序列描述:

```

ATGGACACCT CCCGGCTCGG TGTGCTCCTG TCCTTGCCTG TGCTGCTGCA GCTGGCGACC 60
GGGGGCAGCT CTCCCAGGTC TGGTGTGTTG CTGAGGGGCT GCGCCACACA CTGTCATTGC 120
GAGCCCGACG GCAGGATGTT GCTCAGGGTG GACTGCTCCG ACCTGGGGCT CTCGGAGCTG 180
CCTTCCAACC TCAGCGTCTT CACCTCCTAC CTAGACCTCA GTATGAACAA CATCAGTCAG 240
CTGCTCCCGA ATCCCCTGCC CAGTCTCCGC TTCCTGGAGG AGTTACGTCT TCGGGGAAAC 300
GCTCTGACAT ACATTCCCAA GGGAGCATTC ACTGGCCTTT ACAGTCTTAA AGTTCTTATG 360
CTGCAGAATA ATCAGCTAAG ACACGTACCC ACAGAAGCTC TGCAGAAATT GCGAAGCCTT 420
CAATCCCTGC GTCTGGATGC TAACCACATC AGCTATGTGC CCCCAGCTG TTTCAGTGGC 480
CTGCATTCCC TGAGGCACCT GTGGCTGGAT GACAATGCGT TAACAGAAAT CCCCGTCCAG 540
GCTTTTAGAA GTTTATCGGC ATTGCAAGCC ATGACCTTGG CCCTGAACAA AATACACCAC 600
ATACCAGACT ATGCCTTTGG 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 TACCTGAAC T AAGAACACTG ACTCTGAATG GTGCCTCACA AATAACTGAA 960
TTTCCTGATT 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 ACCCAATGC ATTTTCCACT 1260
TTGCCATCCC TAATAAAGCT GGACCTATCG TCCAACCTCC TGTCGTCTTT TCCTATAACT 1320
GGGTACATG GTTAACTCA CTTAAATTA ACAGGAAATC ATGCCTTACA GAGCTTGATA 1380
TCATCTGAAA ACTTTCCAGA ACTCAAGGTT ATAGAAATGC CTTATGCTTA CCAGTGCTGT 1440
GCATTTGGAG TGTGTGAGAA TGCCTATAAG ATTTCTAATC AATGGAATAA AGGTGACAAC 1500
AGCAGTATGG ACGACCTTCA TAAGAAAGAT GCTGGAATGT TTCAGGCTCA AGATGAACGT 1560
GACCTTGAAG ATTTCTGCT T GACTTTGAG GAAGACCTGA AAGCCCTTCA TTCAGTGCAG 1620

```

```

TGTTACCTT CCCCAGGCC CTTCAAACCC TGTGAACACC TGCTTGATGG CTGGCTGATC 1680
AGAATTGGAG TGTGGACCAT AGCAGTTCTG GCACTTACTT GTAATGCTTT GGTGACTTCA 1740
ACAGTTTTCA GATCCCCTCT GTACATTTCC CCCATTAAAC TGTTAATTGG GGTTCATCGCA 1800
GCAGTGAACA TGCTCACGGG AGTCTCCAGT GCCGTGCTGG CTGGTGTGGA TCGTTTCACT 1860
TTTGGCAGCT TTGCACGACA TGGTGCCTGG TGGGAGAATG GGGTTGGTTG CCATGTCATT 1920
GGTTTTTTGT CCATTTTTGC TTCAGAATCA TCTGTTTTCC TGCTTACTCT GGCAGCCCTG 1980
GAGCGTGGGT TCTCTGTGAA ATATTCTGCA AAATTTGAAA CGAAAGCTCC ATTTTCTAGC 2040
CTGAAAGTAA TCATTTTGCT CTGTGCCCTG CTGGCCTTGA CCATGGCCGC AGTCCCCTG 2100
CTGGGTGGCA GCAAGTATGG CGCCTCCCCT CTCTGCCTGC CTTTGCCTTT TGGGGAGCCC 2160
AGCACCATGG GCTACATGGT CGCTCTCATC TTGCTCAATT CCCTTTGCTT CCTCATGATG 2220
ACCATTGCCT ACACCAAGCT CTACTGCAAT TTGGACAAGG GAGACCTGGA GAATATTTGG 2280
GACTGCTCTA TGA AAAAACA CATTGCCCTG TTGCTCTTCA CCAACTGCAT CCTAAACTGC 2340
CCTGTGGCTT TCTGTCTT CTCCTCTTATAATAACCTTA CATTTATCAG TCCTGAAGTA 2400
ATTAAGTTTA TCCTTCTGGT GGTAGTCCCA CTCCTGCAT GTCTCAATCC CCTTCTCTAC 2460
ATCTTGTTC ATCTCACTT TAAGGAGGAT CTGGTGAGCC TGAGAAAGCA AACCTACGTC 2520
TGGACAAGAT CAAAACACCC AAGCTTGATG TCAATTA ACTCTGATGATGT CGAAAAACAG 2580
TCCTGTGACT CAACTCAAGC CTTGGTAACC TTTACCAGCT CCAGCATCAC TTATGACCTG 2640
CCTCCAGTT CCGTGCCATC ACCAGCTTAT CCAGTGACTG AGAGCTGCCA TCTTCTCTCT 2700
GTGGCATTG TCCATGTCT CTAA 2724

```

(279) SEQ ID NO:278 的资料:

(i) 序列特征:

(A) 长度: 907 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑学: 不相关

(ii) 分子类型: 蛋白质

(xi) SEQ ID NO:278 的序列描述:

```

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
 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
      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

```

(280) SEQ ID NO: 279 的资料:

(i) 序列特征:

-
- (A) 长度: 32 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形
- (ii) 分子类型: DNA (基因组的)
- (xi) SEQ ID NO: 279 的序列描述:
- CATGCCAACC GGCCCGCGAG GCTGCTGCTG GT 32
- (281) SEQ ID NO: 280 的资料:
- (i) 序列特征:
- (A) 长度: 32 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑学: 线形
- (ii) 分子类型: DNA (基因组的)
- (xi) SEQ ID NO: 280 的序列描述:
- ACCAGCAGCA GCCTCGCGGG CCGGTTGGCA TG 32

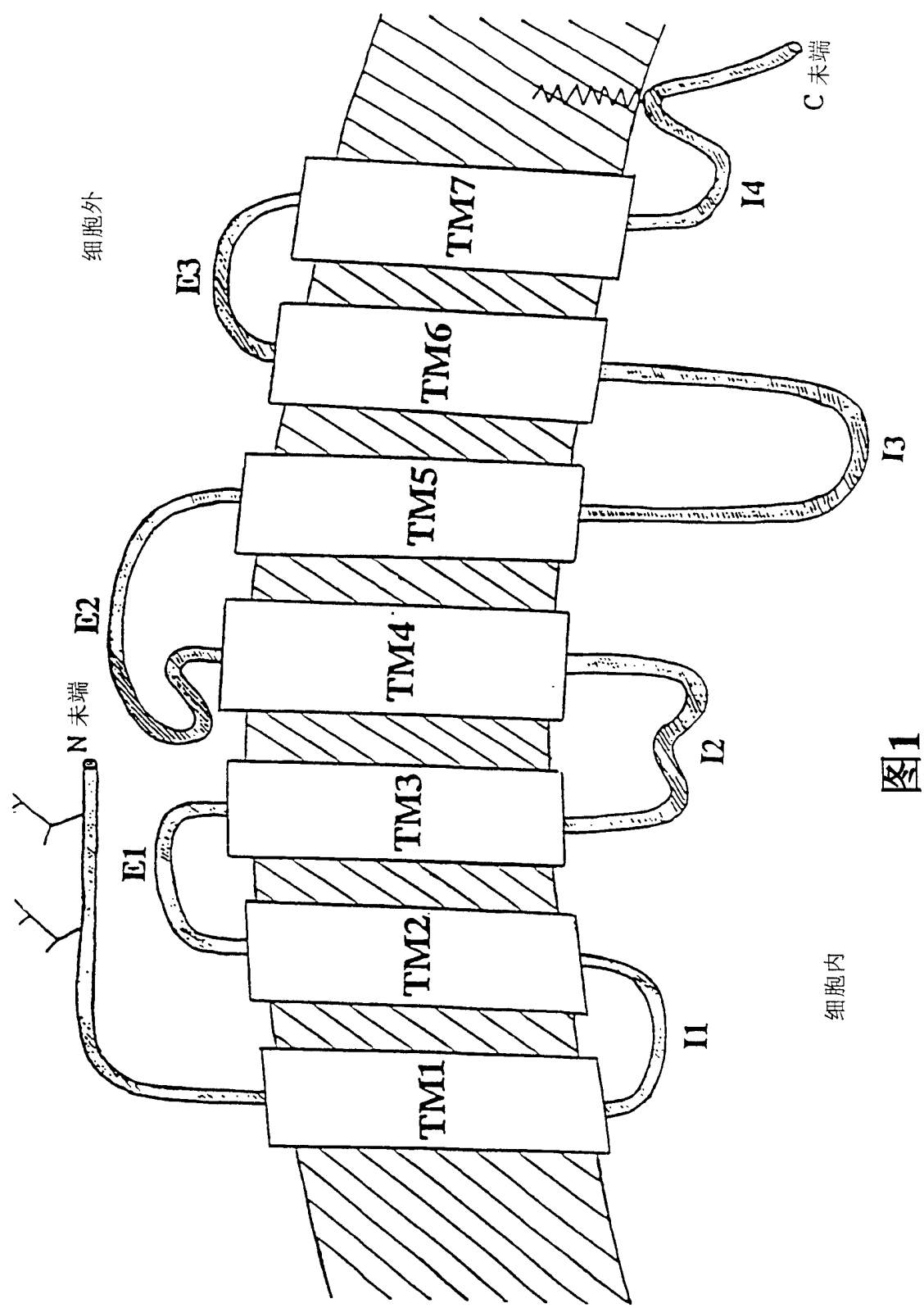


图1

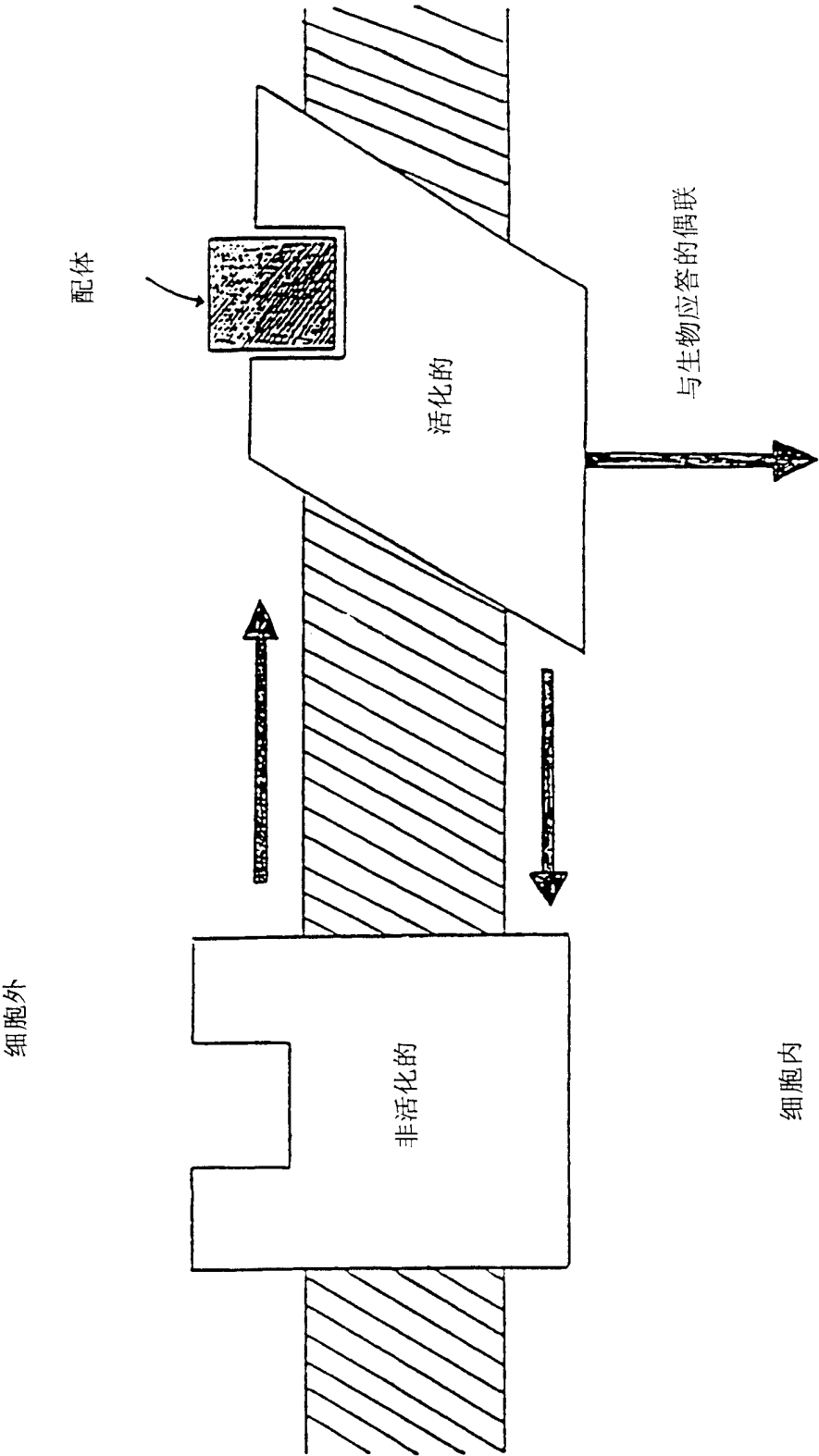


图2

pCMV序列和限制位点

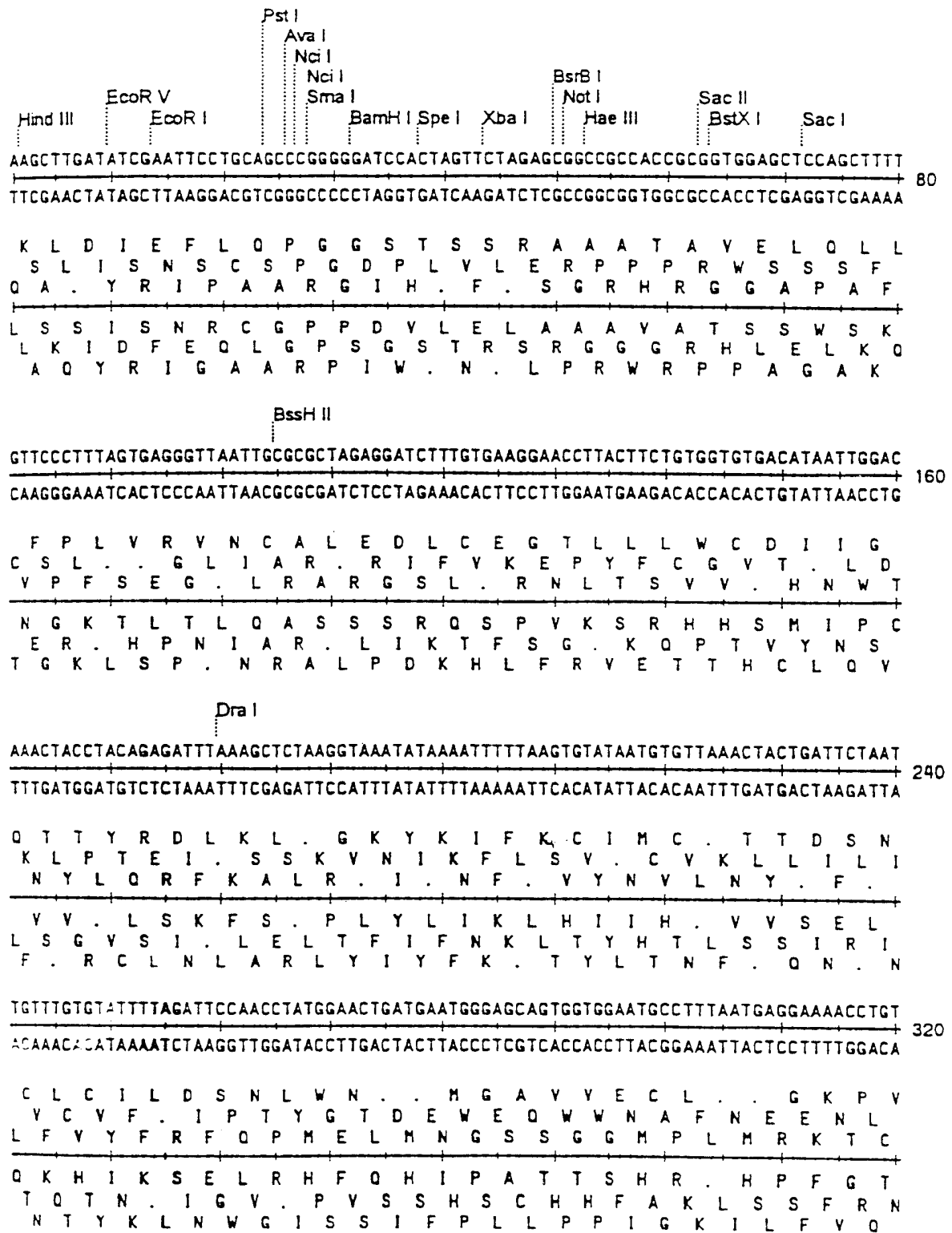


图3A

TTTGCTCAGAAGAAATGCCATCTAGTGATGATGAGGCTACTGCTGACTCTCAACATTCTACTCTCCAAAAAAGAAGAGA 400
 AAACGAGTCTTCTTTACGGTAGATCACTACTACTCCGATGACGACTGAGAGTTGTAAGATGAGGAGGTTTTTCTTCTCT

L L R R N A I G Y C . L S T F Y S S K K E E
 F C S E E M P S S D D E A T A D S O H S T P P K K K R
 F A Q K K C H L V M M R L L L T L N I L L L Q K R R E

K S L L F A M . H H H P . Q O S E V N . E E L F S S F
 Q E S S I G D L S S S A V A S E . C E V G G F F F L
 K A . F F H W R T I I L S S S V R L M R S R W F L L S

Sty I
 AAGGTAGAAGACCCCAAGGACTTTCCTTCAGAATTGCTAAGTTTTTGGAGTCATGCTGTGTTTAGTAATAGAACTCTTGC 480
 TTCCATCTTCTGGGGTTCCTGAAAGGAAGTCTTAACGATTCAAAAACTCAGTACGACACAAATCATTATCTTGAGAACG

K G R R P Q G L S F R I A K F F E S C C V . . . N S C
 K V E D P K D F P S E L L S F L S H A V F S N R T L A
 R . K T P R T F L Q N C . V F . V M L C L V I E L L

P L L G W P S E K L I A L N K S D H Q T . Y Y F E Q
 F T S S G L S K G E S N S L K K L . A T N L L L V R A
 L Y F V G L V K R . F Q . T K Q T M S H K T I S S K S

TTGCTTTGCTATTACACCACAAAGGAAAAAGCTGCACTGCTATACAAGAAAATTATGGAAAAATATTCTGTAACCTTTA 560
 AACGAAACGATAAATGTGGTGTTTCTTTTTTCGACGTGACGATATGTTCTTTTAATACCTTTTTATAAGACATTGGAAT

L L C Y L H H K G K S C T A I Q E N Y G K I F C N L Y
 C F A I Y T T K E K A A L L Y K K I M E K Y S V T F
 L A L L F T P Q R K K L H C Y T R K L W K N I L . P L

K S Q . K C W L P F L Q V A I C S F . P F I N Q L R .
 Q K A I . V V F S F A A S S Y L F I I S F Y E T V K I
 A K S N V G C L F F S C Q . V L F N H F F I R Y G K

AseI
 TAAGTAGGCATAACAGTTATAATCATAACATACTGTTTTTCTTACTCCACACAGGCATAGAGTGTCTGCTATTAATAAC 640
 ATTCATCCGTATTGTCAATATTAGTATTGTATGACAAAAAGAATGAGGTGTGTCGTATCTCACAGACGATAATTATTG

K . A . Q L . S . H T V F S Y S T Q A . S V C Y . .
 I S R H N S Y N H N I L F F L T P H R H R V S A I N N
 . V G I T V I I I T Y C F F L L H T G I E C L L L I T

L Y A Y C N Y D Y C V T K E . E V C A Y L T Q . . Y S
 L L C L L . L . L M S N K R V G C L C L T D A I L L
 Y T P M V T I I M V Y O K K K S W V P M S H R S N I V

Rsa I
 TATGCTCAAAAATTGTGTACCTTTAGCTTTTTAATTTGTAAAGGGGTTAATAAGGAATATTTGATGTATAGTGCCTTGAC 720
 ATACGAGTTTTTAACACATGGAAATCGAAAAATTAACATTTCCCAATTATTCCTTATAAACTACATATCACGGAAGT

L C S K I V Y L . L F N L . R G . . G I F D V . C L D
 Y A Q K L C T F S F L I C K G V N K E Y L M Y S A L T
 M L K N C V P L A F . F V K G L I R N I . C I V P .

H E F I T Y R . S K L K Y L P . Y P I N S T Y H R S
 . A . F N H V K L K K I Q L P T L L S Y K I Y L A K V
 I S L F O T G K A K . N T F P N I L F I Q H I T G Q S

图3B

图3C

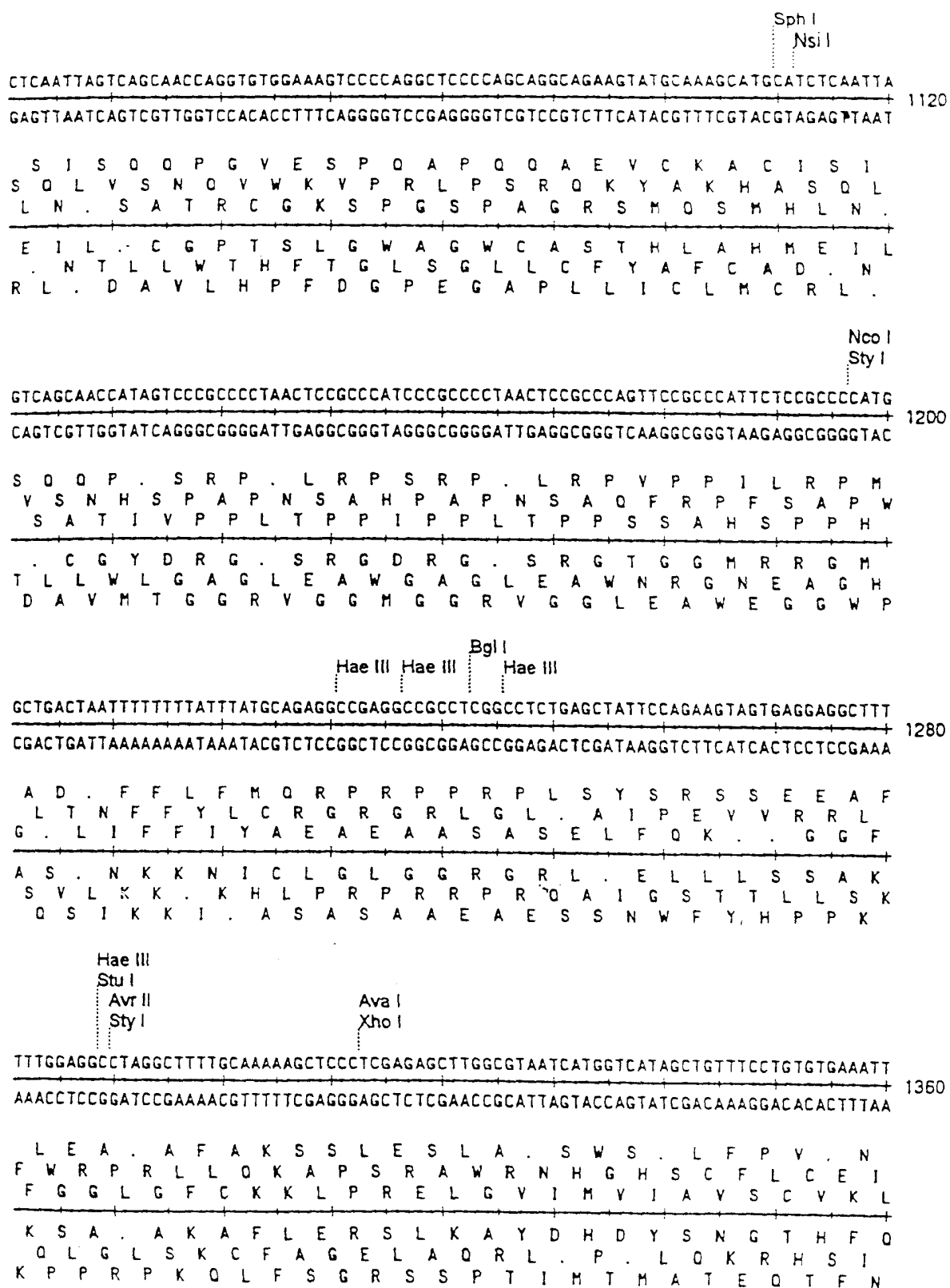


图3D

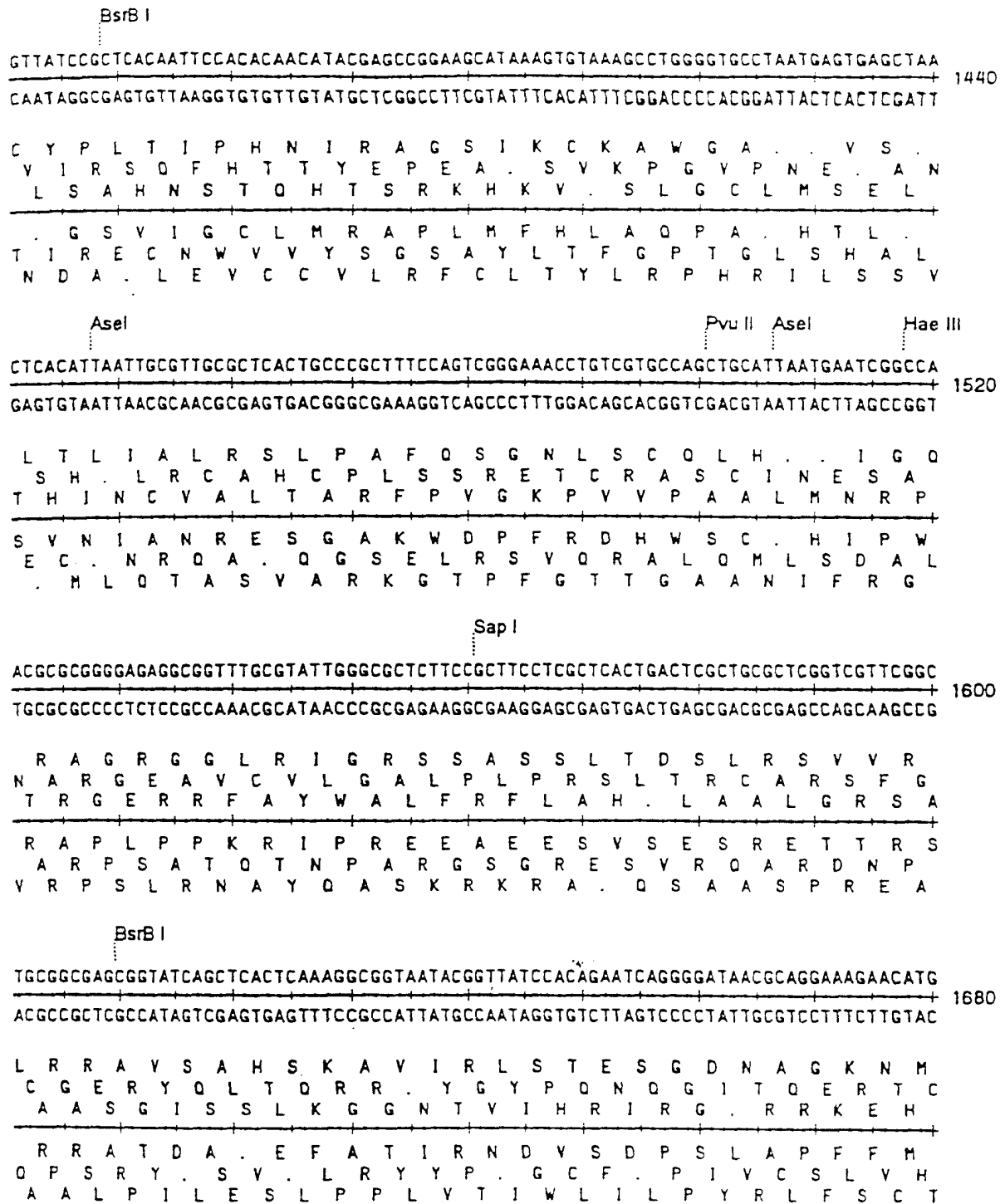


图3E

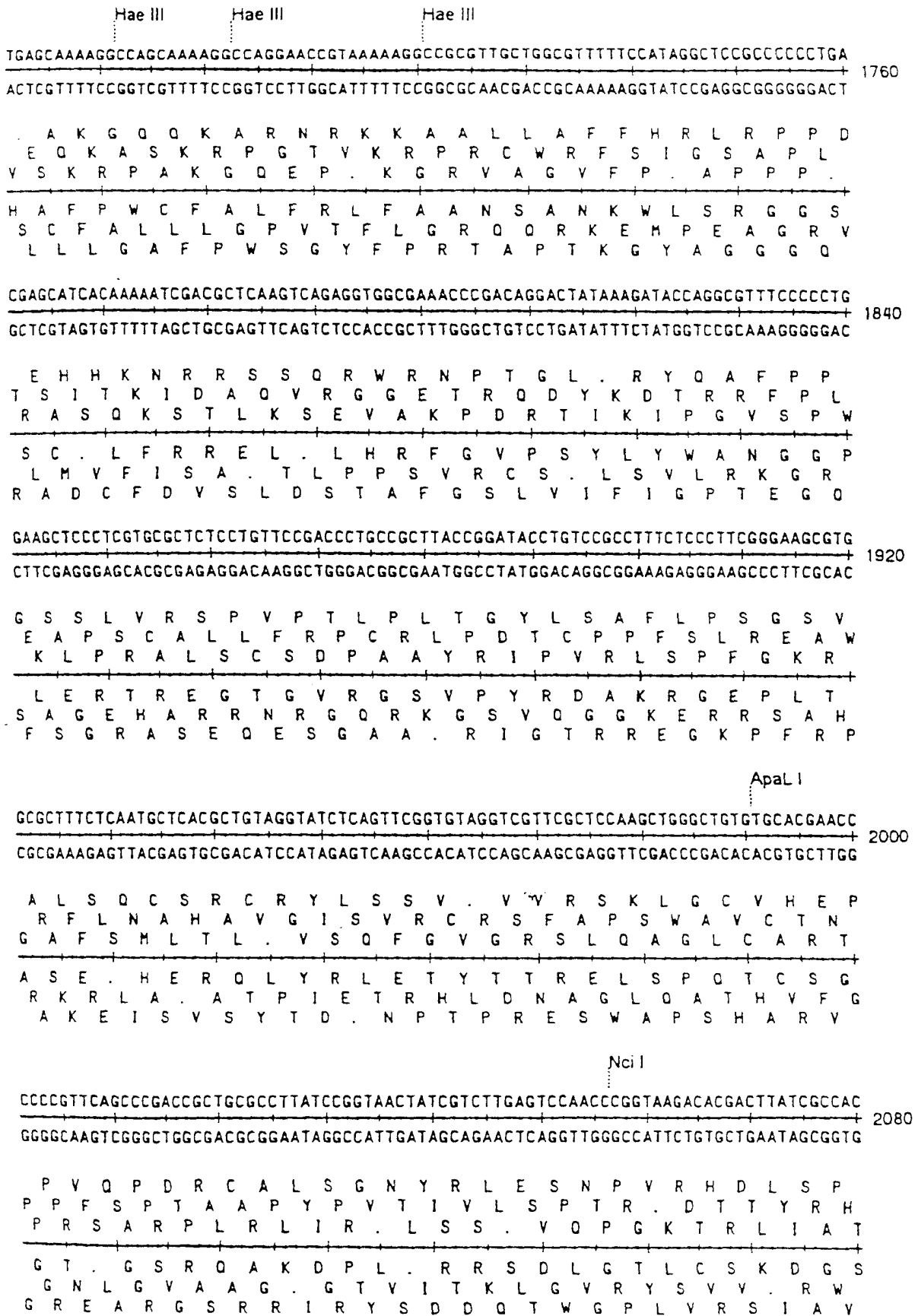


图3F

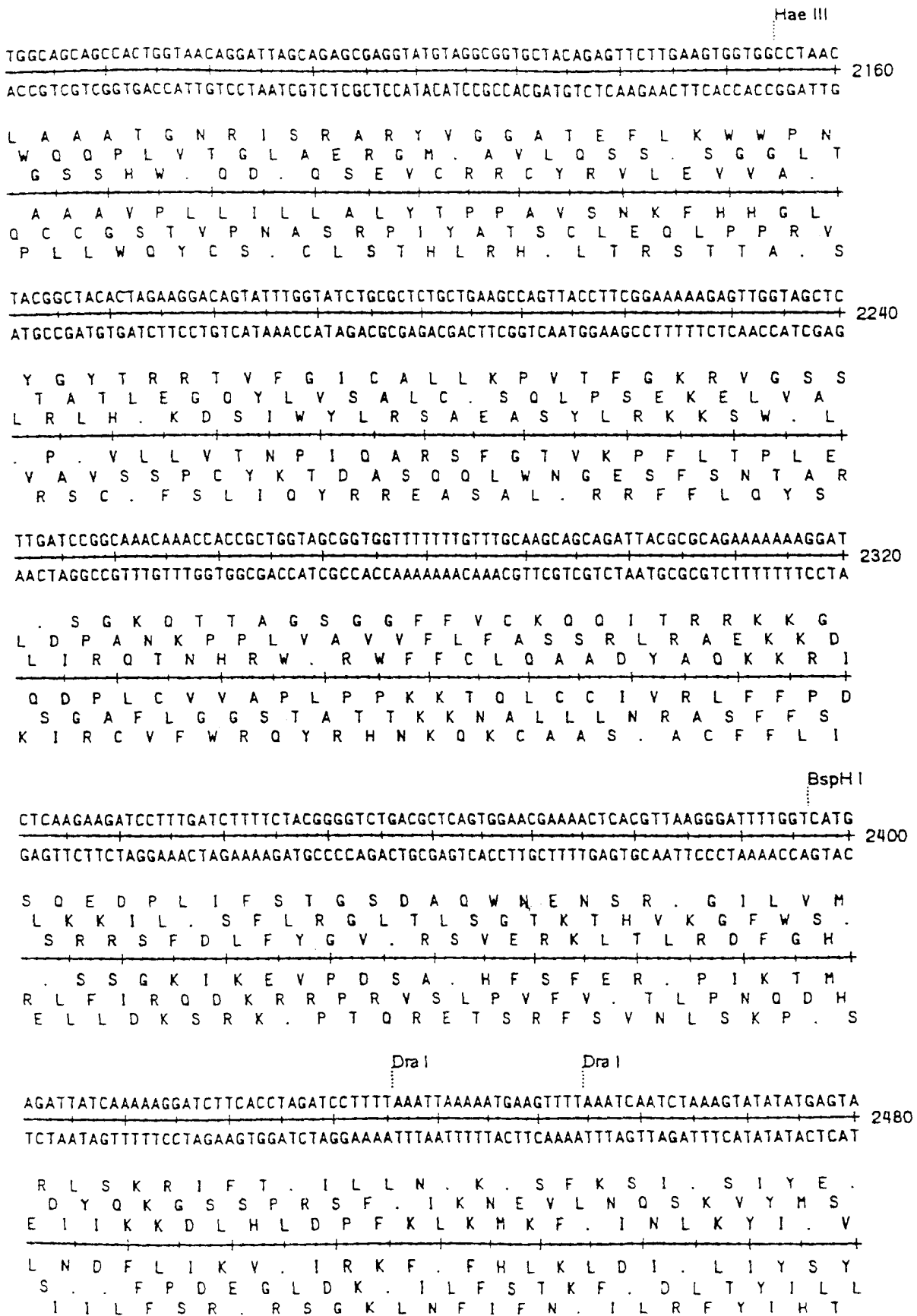


图3G

AACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTG
 TTGAACCAGACTGTCAATGGTTACGAATTAGTCACTCCGTGGATAGAGTCGCTAGACAGATAAAGCAAGTAGGTATCAAC 2560

T W S D S Y Q C L I S E A P I S A I C L F R S S I V
 K L G L T V T N A . S V R H L S Q R S V Y F V H P . L
 N L V . Q L P M L N Q . G T Y L S D L S I S F I H S C
 V Q D S L . W H K I L S A G I E A I Q R N R E D M T A
 S P R V T V L A . D T L C R D . R D T . K T . G Y N
 F K T Q C N G I S L . H P V . R L S R D I E N M W L Q

Hae III

CCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAC
 GGACTGAGGGGCAGCACATCTATTGATGCTATGCCCTCCGAATGGTAGACCGGGTCACGACGTTACTATGGCGCTCTG 2640

A . L P V V . I T T I R E G L P S G P S A A M I P R D
 P D S P S C R . L R Y G R A Y H L A P V L Q . Y R E T
 L T P R R V D N Y D T G G L T I W P Q C C N D T A R
 Q S G T T Y I V V I R S P K G D P G L A A I I G R S
 G S E G D H L Y S R Y P L A . W R A G T S C H Y R S V
 R V G R R T S L . S V P P S V M Q G W H Q L S V A L G

Bgl I

Hae III

Ava II

CCACGCTCACC GGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCTGCAACTTT
 GGTGCGAGTGGCCGAGGTCTAAATAGTCGTTATTTGGTCGGTCGGCCTTCCCGGCTCGCGTCTTACCAGGACGTTGAAA 2720

P R S P A P D L S A I N Q P A G R A E R R S G P A T L
 H A H R L Q I Y Q Q . T S Q P E G P S A E V V L Q L
 P T L T G S R F I S N K P A S R K G R A O K W S C N F
 G R E G A G S K D A I F W G A P L A S R L L P G A V K
 W A . R S W I . . C Y V L W G S P G L A S T T R C S .
 V S V P E L N I L L L G A L R F P R A C F H D Q L K

Ase I

Nci I

Fsp I

ATCCGCTCCATCCAGTCTATTAATTGTTGCCGGAAGCTAGAGTAAGTAGTTCCGCCAGTTAATAGTTTGCACAACGTTG
 TAGCGGGAGGTAGGTCAGATAATTAACAACGGCCCTTCGATCTCATTCATCAAGCGGTCAATTATCAAACGCGTTGCAAC 2800

S A S I Q S I N C C R E A R V S S S P V N S L R N V
 Y P P P S S L L I V A G K L E . V V R Q L I V C A T L
 R L H P V Y . L L P G S . S K . F A S . . F A Q R C
 D A E M W D I L Q Q R S A L T L L E G T L L K R L T T
 G G G D L R N I T A P F S S Y T T R W N I T O A V N
 I R R W G T . . N N G P L . L L Y N A L . Y N A C R Q

TGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTACAGCTCCGGTTCCCAACGATCAAGG
 ACGGTAACGATGTCCGTAGCACCACAGTGCGAGCAGCAAACCATACCGAAGTAAGTCGAGGCCAAGGGTTGCTAGTTCC 2880

A I A T G I V V S R S S F G M A S F S S G S Q R S R
 L P L L Q A S W C H A R R L V W L H S A P V P N D Q G
 C H C Y R H R G V T L V V W Y G F I Q L R F P T I K
 A M A V P M T T D R E D N P I A E N L E P E W R D L
 G N S C A D H H . A R R K T H S . E A G T G L S . P
 W Q . L C P P T V S T T O . P K T . . R N G V I L A

图3H

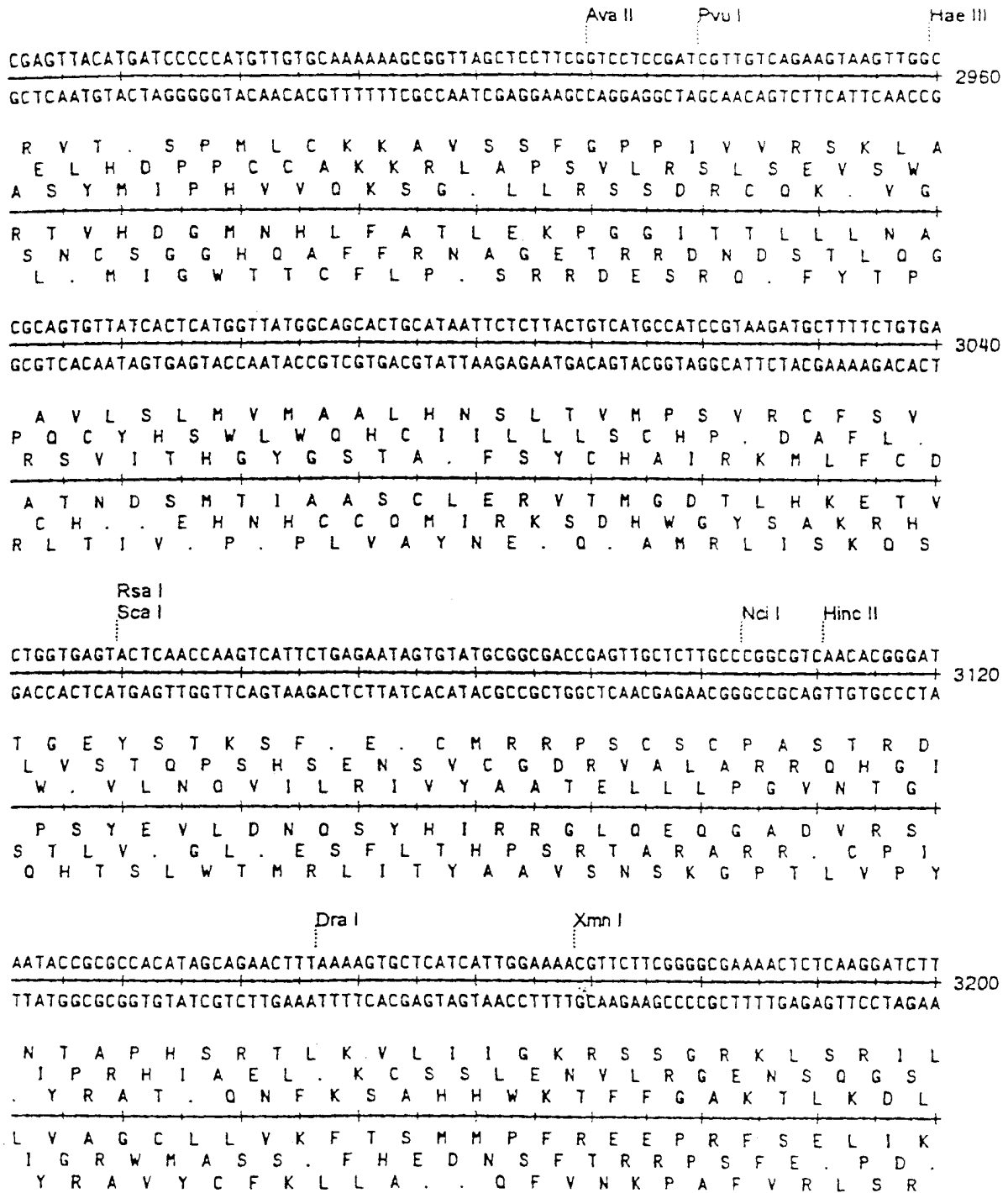


图31

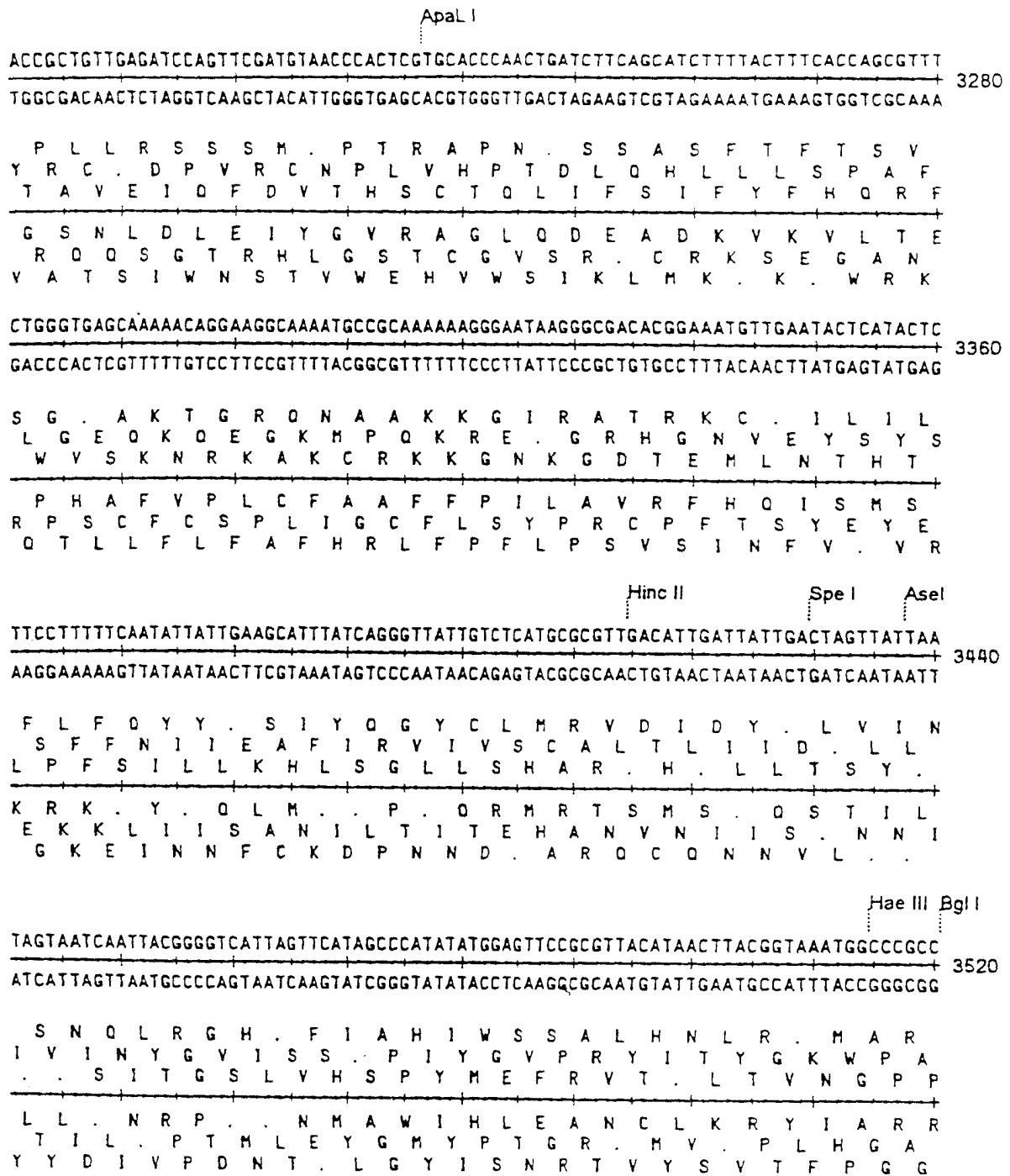


图3J

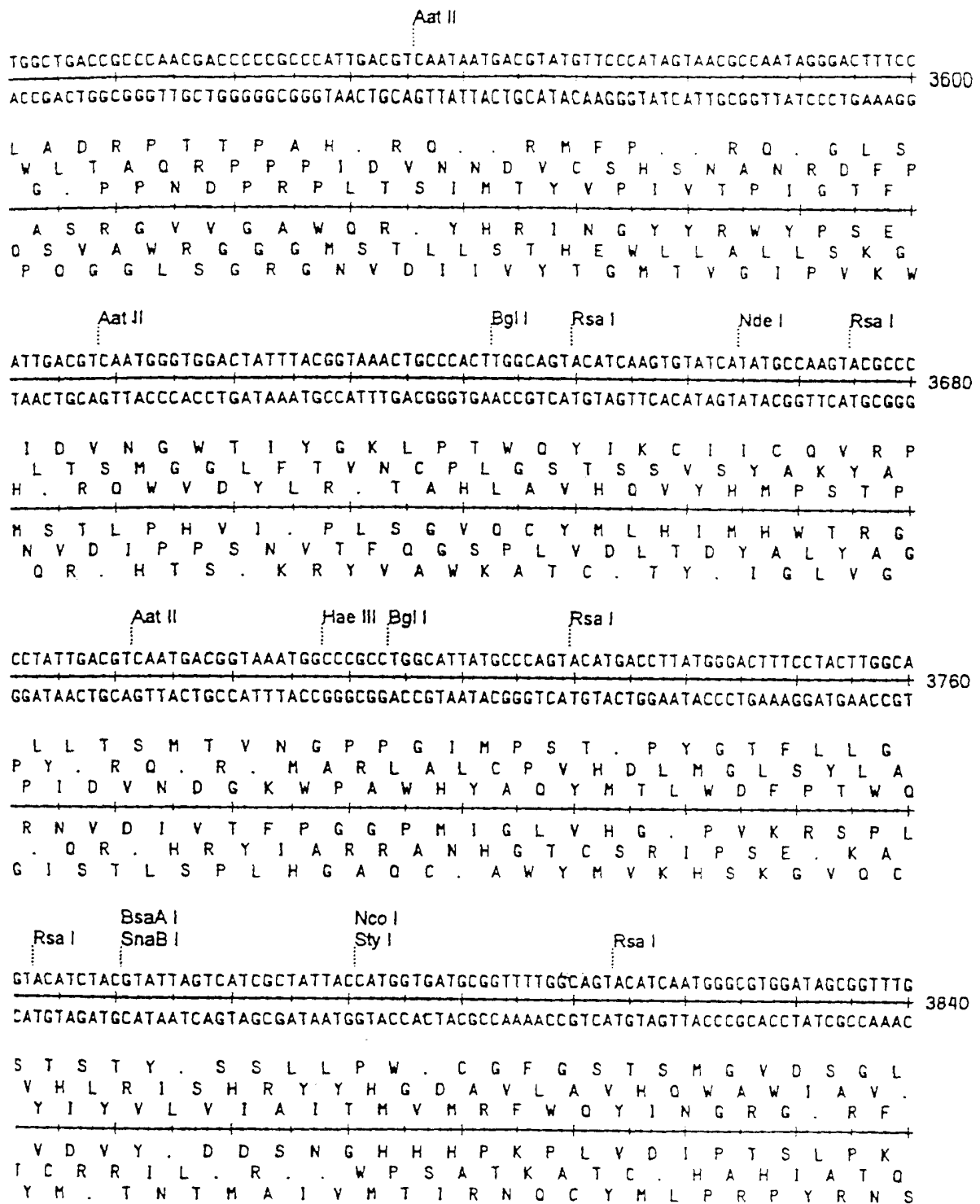


图3K

图3L

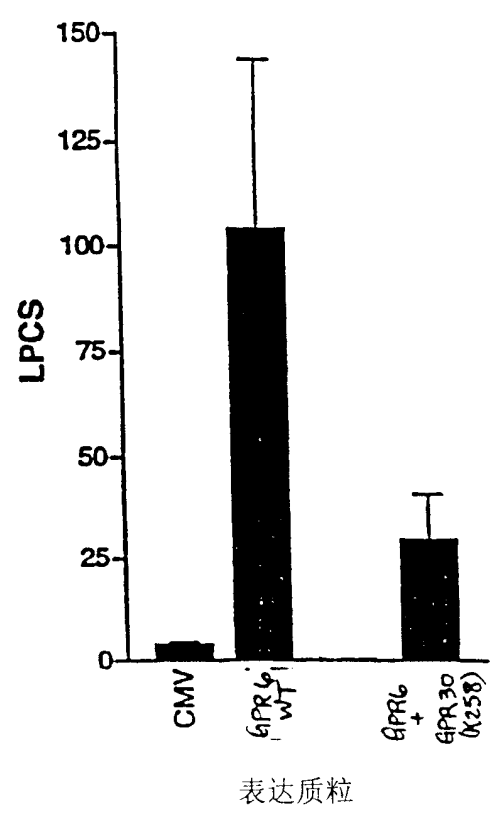


图4

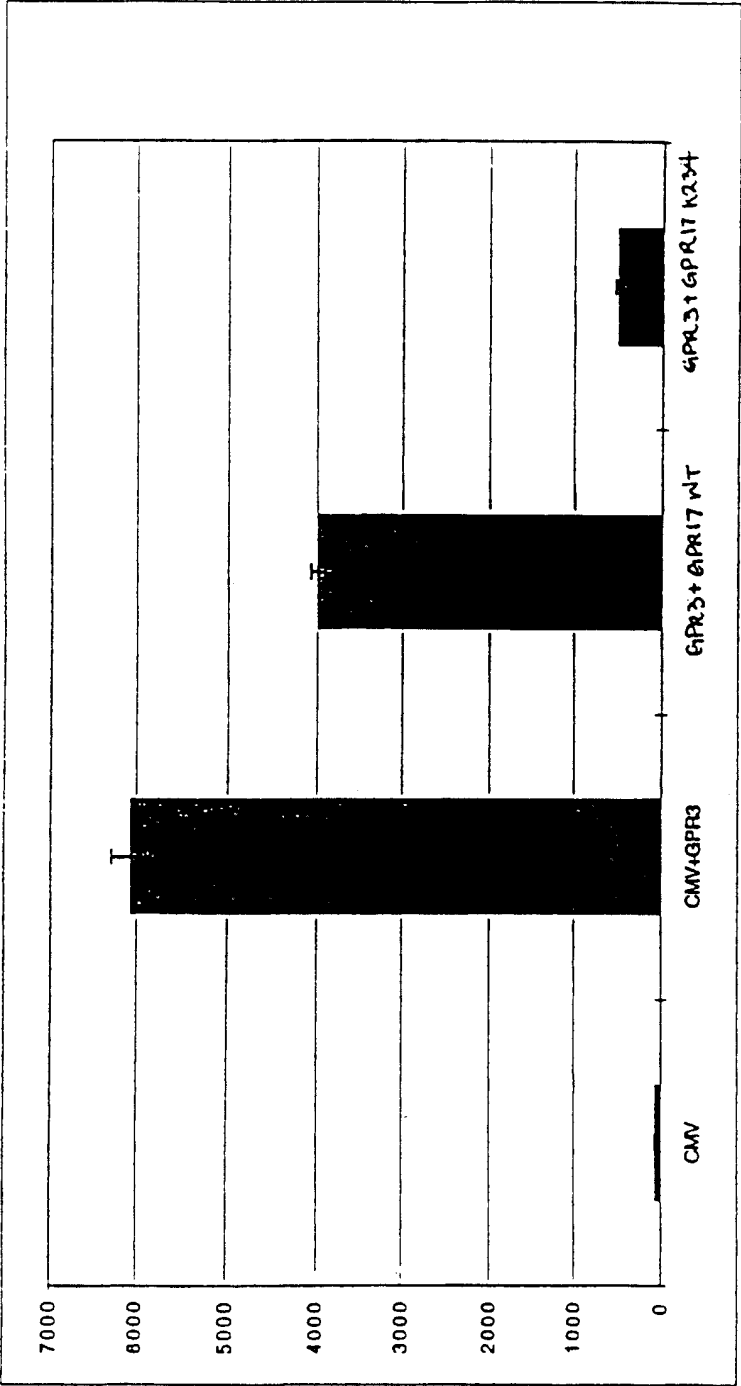


图5

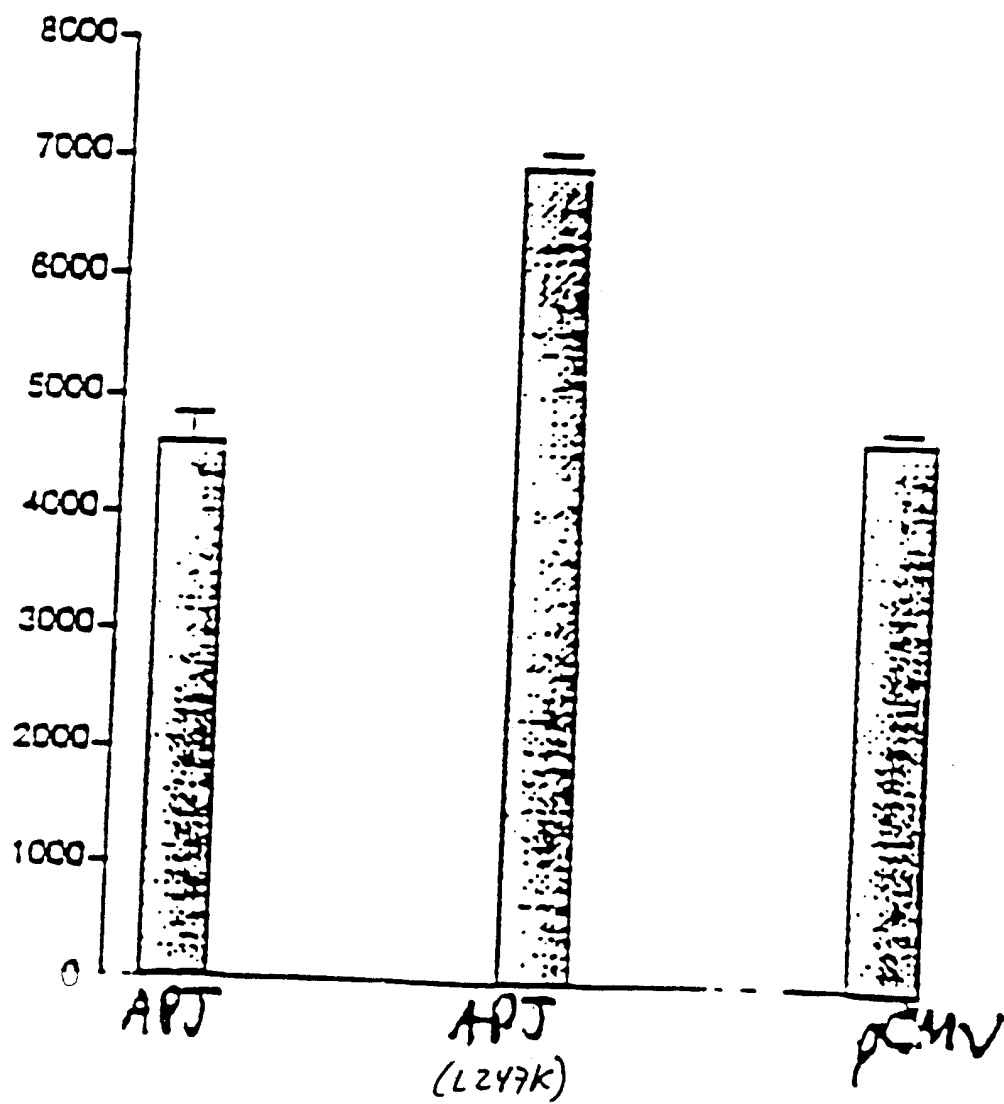


图6

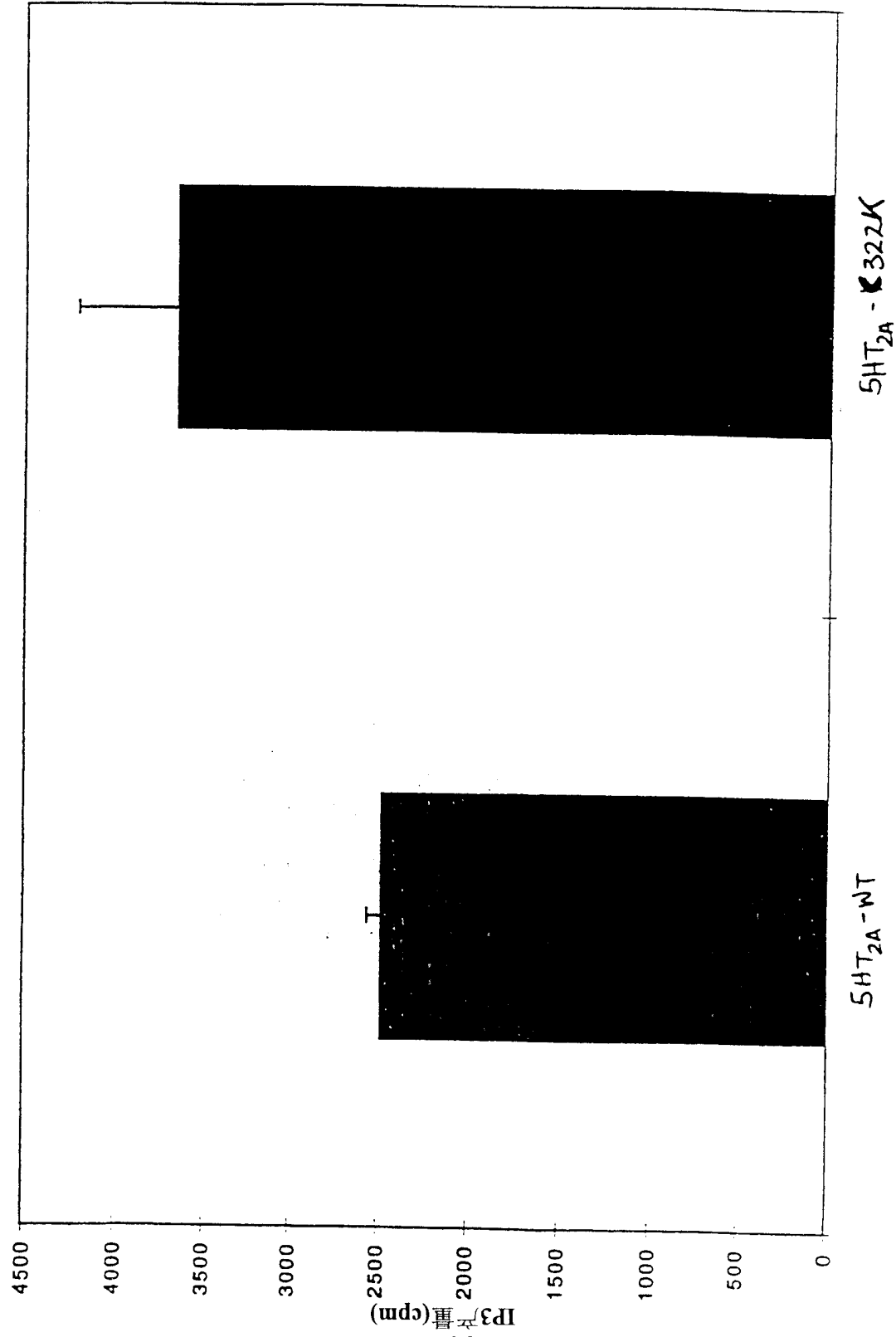


图7

图8A

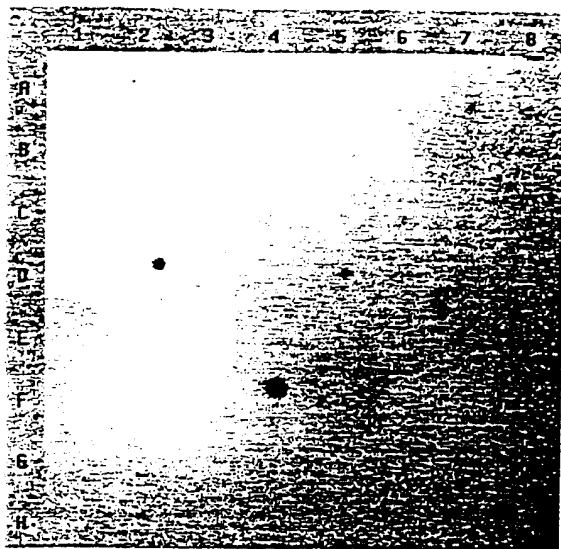


图8B

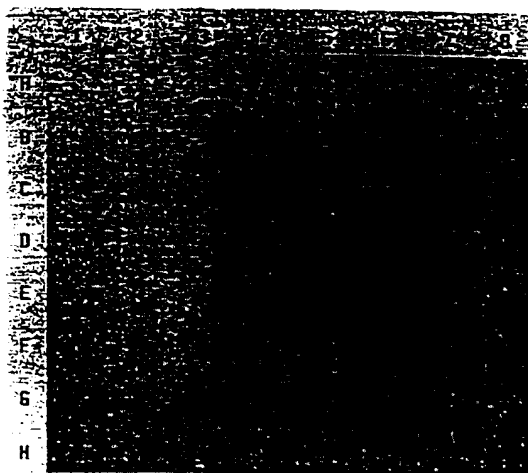
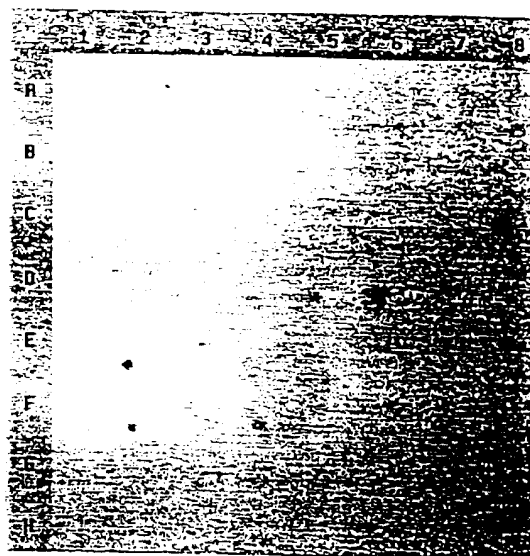


图8C