

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷



[12] 发明专利说明书

[21] ZL 专利号 01105309.7

C07K 14/47

C07K 16/18 C07H 21/00

C12N 15/10 C12N 15/11

C12N 15/12 C12N 15/63

C12P 21/02 A61K 38/17

A61P 35/00

[45] 授权公告日 2004 年 10 月 6 日

[11] 授权公告号 CN 1169833C

[22] 申请日 2001.2.8 [21] 申请号 01105309.7

[71] 专利权人 上海市肿瘤研究所

地址 200032 上海市斜土路 2200 弄 25 号

[72] 发明人 顾健人

审查员 林峻凯

[74] 专利代理机构 上海专利商标事务所

代理人 徐 迅

权利要求书 1 页 说明书 32 页

[54] 发明名称 具有抑癌功能的新的人蛋白及其编码序列

[57] 摘要

本发明公开了一类新的具有抑癌功能的人蛋白，编码此多肽的多核苷酸和经重组技术产生该多肽的方法。本发明还公开了此多肽用于治疗多种疾病如癌症等的方法。本发明还公开了抗此多肽的拮抗剂及其治疗作用。本发明还公开了编码这类新的具有抑癌功能的人蛋白的多核苷酸的用途。

I S S N 1 0 0 8 - 4 2 7 4

-
- 1.一种分离的具有抑癌功能的人蛋白多肽，其特征在于，它是具有选自下组的氨基酸序列的多肽：SEQ ID NO: 2、5、8、11、14、17、20、23、26、29、32、35。
- 5 2.如权利要求1所述的多肽，其特征在于，该多肽的氨基酸序列选自下组：SEQ ID NO: 2、8、23、29。
- 3.一种分离的多核苷酸，其特征在于，选自下组：
(a)编码如权利要求1所述多肽的多核苷酸；
(b)与多核苷酸(a)互补的多核苷酸。
- 10 4.如权利要求3所述的多核苷酸，其特征在于，该多核苷酸编码的多肽具有选自下组的氨基酸序列：SEQ ID NO: 2、5、8、11、14、17、20、23、26、29、32、35。
- 5.如权利要求3所述的多核苷酸，其特征在于，该多核苷酸的序列选自下组：
SEQ ID NO: 3、6、9、12、15、18、21、24、27、30、33、36的编码区序列或全长序列。
- 6.一种载体，其特征在于，它含有权利要求3所述的多核苷酸。
- 15 7.一种遗传工程化的宿主细胞，其特征在于，它是选自下组的一种宿主细胞：
(a)用权利要求6所述的载体转化或转导的宿主细胞；
(b)用权利要求3所述的多核苷酸转化或转导的宿主细胞。
- 8.一种具有抑癌功能的人蛋白活性的多肽的制备方法，其特征在于，该方法包含：
(a)在适合表达具有抑癌功能的人蛋白的条件下，培养权利要求7所述的宿主细胞；
20 (b)从培养物中分离出具有抑癌功能的人蛋白活性的多肽。
- 9.一种能与权利要求1所述的具有抑癌功能的人蛋白多肽特异性结合的抗体。
- 10.一种药物组合物，其特征在于，它含有安全有效量的权利要求1所述的多肽以及药学上可接受的载体。

具有抑癌功能的新的人蛋白及其编码序列

技术领域

- 5 本发明属于生物技术领域，具体地说，本发明涉及新的编码具有抑癌功能的人蛋白的多核苷酸和此多核苷酸编码的多肽。本发明还涉及此多核苷酸和多肽的用途和制备。

背景技术

人基因组学研究目前是国际上的热点，除人染色体 DNA 大规模测序，表达序列测序(EST)的方法外，还缺少从功能开始的筛选具有功能基因的高通量的方法。

- 10 癌症是危害人类健康的主要疾病之一。为了有效地治疗和预防肿瘤，目前人们已越来越关注肿瘤的基因治疗。因此，本领域迫切需要开发研究具有抑癌功能的人蛋白及其激动剂/抑制剂。

发明内容

- 15 本发明的目的是提供一类新的具有抑癌功能的人蛋白多肽以及其片段、类似物和衍生物。

本发明的另一目的是提供编码这些多肽的多核苷酸。

本发明的另一目的是提供生产这些多肽的方法以及该多肽和编码序列的用途。

- 20 在本发明的第一方面，提供新颖的分离出的具有抑癌功能的蛋白多肽，它包含具有选自下组的氨基酸序列的多肽：SEQ ID NO: 2、5、8、11、14、17、20、23、26、29、32、35；或其保守性变异多肽、或其活性片段、或其活性衍生物。

较佳地，该多肽是具有选自下组的氨基酸序列的多肽：SEQ ID NO: 2、5、8、11、14、17、20、23、26、29、32、35。

- 25 在本发明的第二方面，提供了一种分离的多核苷酸，它包含一核苷酸序列，该核苷酸序列与选自下组的一种核苷酸序列有至少85%相同性：(a)编码上述的具有抑癌功能的蛋白多肽的多核苷酸；(b)与多核苷酸(a)互补的多核苷酸。较佳地，该多核苷酸编码的多肽具有选自下组的氨基酸序列：SEQ ID NO: 2、5、8、11、14、17、20、23、26、29、32、35。更佳地，该多核苷酸的序列选自下组：SEQ ID NO: 3、6、9、12、15、18、21、24、27、30、33、36的编码区序列或全长序列。

- 30 在本发明的第三方面，提供了含有上述多核苷酸的载体，以及被该载体转化或转导的宿主细胞或者被上述多核苷酸直接转化或转导的宿主细胞。

在本发明的第四方面，提供了制备具有抑癌功能的蛋白活性的多肽的制备方法，该方法包含：(a)在适合表达具有抑癌功能的蛋白的条件下，培养上述被转化或转导的宿主细胞；(b)从培养物中分离出具有抑癌功能的蛋白活性的多肽。

- 35 在本发明的第五方面，提供了与上述的具有抑癌功能的蛋白多肽特异性结合的抗体。还提供了可用于检测的核酸分子，它含有上述的多核苷酸中连续 10 个核苷酸至全长核苷酸，较佳地它含有连续的约 10-800 个核苷酸。

在本发明的第六方面，提供了一种药物组合物，它含有安全有效量的本发明的具有

抑癌功能的蛋白多肽以及药学上可接受的载体。这些药物组合物可治疗癌症以及细胞异常增殖等病症。

本发明的其它方面由于本文的公开内容，对本领域的技术人员而言是显而易见的。

5 3T3 细胞是一种小鼠成纤维细胞(J. Cell. Biol., 17:299, 1963)。在癌症研究领域，常将外源基因(尤其是人基因)引入 3T3 细胞，观察其对 3T3 细胞生长的影响情况。通常认为，对 3T3 细胞生长有影响的基因是癌症相关基因，其中对 3T3 细胞生长有抑制作用的基因大多是抑癌基因，而对 3T3 细胞生长有促进作用的基因大多是(原)癌基因。

10 本发明采用大规模 cDNA 克隆转染小鼠胚胎成纤维细胞，在获得具有抑癌作用的基础上，经测序证明为新的基因，进一步得到全长 cDNA 克隆。DNA 转染试验证明，本发明的具有抑癌功能的蛋白对 3T3 细胞具有抑制克隆形成的作用，其抑制率 $\geq 50\%$ 。

如本文所用，“分离的”是指物质从其原始环境中分离出来(如果是天然物质，原始环境即是天然环境)。如活体细胞内的天然状态下的多聚核苷酸和多肽是没有分离纯化的，但同样的多聚核苷酸或多肽如从天然状态中同存在的其他物质中分开，则为分离纯化的。

15 如本文所用，“分离的具有抑癌功能的蛋白或多肽”是指具有抑癌功能的蛋白多肽基本上不含天然与其相关的其它蛋白、脂类、糖类或其它物质。本领域的技术人员能用标准的蛋白质纯化技术纯化具有抑癌功能的蛋白。基本上纯的多肽在非还原聚丙烯酰胺凝胶上能产生单一的主带。

20 本发明的多肽可以是重组多肽、天然多肽、合成多肽，优选重组多肽。本发明的多肽可以是天然纯化的产物，或是化学合成的产物，或使用重组技术从原核或真核宿主(例如，细菌、酵母、高等植物、昆虫和哺乳动物细胞)中产生。根据重组生产方案所用的宿主，本发明的多肽可以是糖基化的，或可以是非糖基化的。本发明的多肽还可包括或不包括起始的甲硫氨酸残基。

25 本发明还包括具有抑癌功能的人蛋白的片段、衍生物和类似物。如本文所用，术语“片段”、“衍生物”和“类似物”是指基本上保持本发明的天然具有抑癌功能的人蛋白相同的生物学功能或活性的多肽。本发明的多肽片段、衍生物或类似物可以是(i)有一个或多个保守或非保守性氨基酸残基(优选保守性氨基酸残基)被取代的多肽，而这样的取代的氨基酸残基可以是也可以不是由遗传密码编码的，或(ii)在一个或多个氨基酸残基中
30 具有取代基团的多肽，或(iii)成熟多肽与另一个化合物(比如延长多肽半衰期的化合物，例如聚乙二醇)融合所形成的多肽，或(iv)附加的氨基酸序列融合到此多肽序列而形成的多肽(如前导序列或分泌序列或用来纯化此多肽的序列或蛋白原序列)。根据本文的教导，这些片段、衍生物和类似物属于本领域熟练技术人员公知的范围。

35 本发明的多核苷酸可以是 DNA 形式或 RNA 形式。DNA 形式包括 cDNA、基因组 DNA 或人工合成的 DNA。DNA 可以是单链的或是双链的。DNA 可以是编码链或非编码链。以 PP9457 蛋白(在本申请中，蛋白质的命名采用其克隆编号)为例，编码成熟多肽的编码区序列可以与 SEQ ID NO:3 所示的编码区序列相同或者是简并的变异体。如本文所用，“简并的变异体”在本发明中是指编码具有 SEQ ID NO:2 的蛋白质，但与 SEQ ID

NO:3 所示的编码区序列有差别的核酸序列。再以 PP9943 蛋白为例, 编码成熟多肽的编码区序列可以与 SEQ ID NO:6 所示的编码区序列相同或者是简并的变异体。对于其他具有抑癌功能的蛋白, 可依此类推。

5 编码成熟多肽的多核苷酸包括: 只编码成熟多肽的编码序列; 成熟多肽的编码序列和各种附加编码序列; 成熟多肽的编码序列(和任选的附加编码序列)以及非编码序列。

术语“编码多肽的多核苷酸”可以是包括编码此多肽的多核苷酸, 也可以是还包括附加编码和/或非编码序列的多核苷酸。

10 本发明还涉及上述多核苷酸的变异体, 其编码与本发明有相同的氨基酸序列的多肽或多肽的片段、类似物和衍生物。此多核苷酸的变异体可以是天然发生的等位变异体或非天然发生的变异体。这些核苷酸变异体包括取代变异体、缺失变异体和插入变异体。如本领域所知的, 等位变异体是一个多核苷酸的替换形式, 它可能是一个或多个核苷酸的取代、缺失或插入, 但不会从实质上改变其编码的多肽的功能。

15 本发明还涉及与上述的序列杂交且两个序列之间具有至少 50%, 较佳地至少 70%, 更佳地至少 80%相同性的多核苷酸。本发明特别涉及在严格条件下与本发明所述多核苷酸可杂交的多核苷酸。在本发明中, “严格条件”是指: (1)在较低离子强度和较高温度下的杂交和洗脱, 如 $0.2 \times \text{SSC}$, 0.1%SDS, 60 °C; 或(2)杂交时加有变性剂, 如 50%(v/v)甲酰胺, 0.1%小牛血清/0.1% Ficoll, 42 °C等; 或(3)仅在两条序列之间的相同性至少在 95%以上, 更好是 97%以上时才发生杂交。并且, 可杂交的多核苷酸编码的多肽与 SEQ ID NO: 2 所示的成熟多肽有相同的生物学功能(以 PP9457 蛋白为例)和活性。

20 本发明还涉及与上述的序列杂交的核酸片段。如本文所用, “核酸片段”的长度至少含 15 个核苷酸, 较好是至少 30 个核苷酸, 更好是至少 50 个核苷酸, 最好是至少 100 个核苷酸以上。核酸片段可用于核酸的扩增技术(如 PCR)以确定和/或分离编码具有抑癌功能的蛋白的多聚核苷酸。

本发明中的多肽和多核苷酸优选以分离的形式提供, 更佳地被纯化至均质。

25 本发明的 DNA 序列能用几种方法获得。例如, 用本领域熟知的杂交技术分离 DNA。这些技术包括但不局限于: 1)用探针与基因组或 cDNA 文库杂交以检出同源性核苷酸序列, 和 2)表达文库的抗体筛选以检出具有共同结构特征的克隆的 DNA 片段。

编码具有抑癌功能的蛋白的特异 DNA 片段序列产生也能用下列方法获得: 1)从基因组 DNA 分离双链 DNA 序列; 2)化学合成 DNA 序列以获得所需多肽的双链 DNA。

30 当需要的多肽产物的整个氨基酸序列已知时, DNA 序列的直接化学合成是经常选用的方法。如果所需的氨基酸的整个序列不清楚时, DNA 序列的直接化学合成是不可能的, 选用的方法是 cDNA 序列的分离。分离感兴趣的 cDNA 的标准方法是从高表达该基因的供体细胞分离 mRNA 并进行逆转录, 形成质粒或噬菌体 cDNA 文库。提取 mRNA 的方法已有多种成熟的技术, 试剂盒也可从商业途径获得(Qiagene)。而构建 cDNA 文库也是通常的方法(Sambrook, et al., Molecular Cloning, A Laboratory Manual, Cold Spring
35 Harbor Laboratory. New York, 1989)。还可得到商业供应的 cDNA 文库, 如 Clontech 公司的不同 cDNA 文库。当结合使用聚合酶反应技术时, 即使极少的表达产物也能克隆。

可用常规方法从这些 cDNA 文库中筛选本发明的基因。这些方法包括(但不限于): (1)DNA-DNA 或 DNA-RNA 杂交; (2)标志基因的功能出现或丧失; (3)测定具有抑癌功能的蛋白的转录本的水平; (4)通过免疫学技术或测定生物学活性, 来检测基因表达的蛋白产物。上述方法可单用, 也可多种方法联合应用。

5 在第(1)种方法中, 杂交所用的探针是与本发明的多核苷酸的任何一部分同源, 其长度至少 15 个核苷酸, 较好是至少 30 个核苷酸, 更好是至少 50 个核苷酸, 最好是至少 100 个核苷酸。此外, 探针的长度通常在 2kb 之内, 较佳地为 1kb 之内。此处所用的探针通常是在本发明的基因 DNA 序列信息的基础上化学合成的 DNA 序列。本发明的基因本身或者片段当然可以用作探针。DNA 探针的标记可用放射性同位素, 荧光素或酶(如碱性磷酸酶)等。

10 在第(4)种方法中, 检测具有抑癌功能的蛋白基因表达的蛋白产物可用免疫学技术如 Western 印迹法, 放射免疫沉淀法, 酶联免疫吸附法(ELISA)等。

应用 PCR 技术扩增 DNA/RNA 的方法(Saiki, et al. Science 1985;230:1350-1354)被优选用于获得本发明的基因。特别是很难从文库中得到全长的 cDNA 时, 可优选使用 RACE 15 法(RACE-cDNA 末端快速扩增法), 用于 PCR 的引物可根据本文所公开的本发明的序列信息适当地选择, 并可用常规方法合成。可用常规方法如通过凝胶电泳分离和纯化扩增的 DNA/RNA 片段。

如上所述得到的本发明的基因, 或者各种 DNA 片段等的核苷酸序列的测定可用常规方法如双脱氧链终止法(Sanger et al. PNAS, 1977, 74: 5463-5467)。这类核苷酸序列测定也可用商业测序试剂盒等。为了获得全长的 cDNA 序列, 测序需反复进行。有时需要测定多个克隆的 cDNA 序列, 才能拼接成全长的 cDNA 序列。

本发明也涉及包含本发明多核苷酸的载体, 以及用本发明载体或具有抑癌功能的蛋白编码序列经基因工程产生的宿主细胞, 以及经重组技术产生本发明所述多肽的方法。

25 通过常规的重组 DNA 技术(Science, 1984; 224: 1431), 可利用本发明的多聚核苷酸序列可用来表达或生产重组的具有抑癌功能的蛋白多肽。一般来说有以下步骤:

(1).用本发明的编码具有抑癌功能的人蛋白的多核苷酸(或变异体), 或用含有该多核苷酸的重组表达载体转化或转导合适的宿主细胞;

(2).在合适的培养基中培养的宿主细胞;

(3).从培养基或细胞中分离、纯化蛋白质。

30 本发明中, 具有抑癌功能的人蛋白多核苷酸序列可插入到重组表达载体中。术语“重组表达载体”指本领域熟知的细菌质粒、噬菌体、酵母质粒、植物细胞病毒、哺乳动物细胞病毒如腺病毒、逆转录病毒或其他载体。在本发明中适用的载体包括但不限于: 在细菌中表达的基于 T7 的表达载体(Rosenberg, et al. Gene, 1987, 56:125); 在哺乳动物细胞中表达的 pMSXND 表达载体(Lee and Nathans, J Bio Chem. 263:3521,1988)和在昆虫细胞 35 中表达的来源于杆状病毒的载体。总之, 只要能在宿主体内复制和稳定, 任何质粒和载体都可以用。表达载体的一个重要特征是通常含有复制起点、启动子、标记基因和翻译控制元件。

本领域的技术人员熟知的方法能用于构建含有具有抑癌功能的人蛋白编码 DNA 序列和合适的转录/翻译控制信号的表达载体。这些方法包括体外重组 DNA 技术、DNA 合成技术、体内重组技术等(Sambrook, et al.)。所述的 DNA 序列可有效连接到表达载体中的适当启动子上,以指导 mRNA 合成。这些启动子的代表性例子有:大肠杆菌的 lac 或 trp 启动子;λ 噬菌体 P_L 启动子;真核启动子包括 CMV 立即早期启动子、早期和晚期 SV40 启动子、反转录病毒的 LTRs 和其他一些已知的可控制基因在原核或真核细胞或其病毒中表达的启动子。表达载体还包括翻译起始用的核糖体结合位点和转录终止子。

此外,表达载体优选地包含一个或多个选择性标记基因,以提供用于选择转化的宿主细胞的表型性状,如真核细胞培养用的二氢叶酸还原酶、新霉素抗性以及绿色荧光蛋白(GFP),或用于大肠杆菌的四环素或氨基青霉素抗性。

包含上述的适当 DNA 序列以及适当启动子或者控制序列的载体,可以用于转化适当的宿主细胞,以使其能够表达蛋白质。

宿主细胞可以是原核细胞,如细菌细胞;或是低等真核细胞,如酵母细胞;或是高等真核细胞,如哺乳动物细胞。代表性例子有:大肠杆菌,链霉菌属;鼠伤寒沙门氏菌的细菌细胞;真菌细胞如酵母;植物细胞;果蝇 S2 或 Sf9 的昆虫细胞;CHO、COS 或 Bowes 黑素瘤细胞的动物细胞等。

本发明的多核苷酸在高等真核细胞中表达时,如果在载体中插入增强子序列时将会使转录得到增强。增强子是 DNA 的顺式作用因子,通常大约有 10 到 300 个碱基对,作用于启动子以增强基因的转录。可举的例子包括在复制起始点晚期一侧的 100 到 270 个碱基对的 SV40 增强子、在复制起始点晚期一侧的多瘤增强子以及腺病毒增强子等。

本领域一般技术人员都清楚如何选择适当的载体、启动子、增强子和宿主细胞。

用重组 DNA 转化宿主细胞可用本领域技术人员熟知的常规技术进行。当宿主为原核生物如大肠杆菌时,能吸收 DNA 的感受态细胞可在指数生长期后收获,用 CaCl₂ 法处理,所用的步骤在本领域众所周知。可供选择的是用 MgCl₂。如果需要,转化也可用电穿孔的方法进行。当宿主是真核生物,可选用如下的 DNA 转染方法:磷酸钙共沉淀法,常规机械方法如显微注射、电穿孔、脂质体包装等。

获得的转化子可以用常规方法培养,表达本发明的基因所编码的多肽。根据所用的宿主细胞,培养中所用的培养基可选自各种常规培养基。在适于宿主细胞生长的条件下进行培养。当宿主细胞生长到适当的细胞密度后,用合适的方法(如温度转换或化学诱导)诱导选择的启动子,将细胞再培养一段时间。

在上面的方法中的重组多肽可包被于细胞内、细胞外或在细胞膜上表达或分泌到细胞外。如果需要,可利用其物理的、化学的和其它特性通过各种分离方法分离和纯化重组的蛋白。这些方法是本领域技术人员所熟知的。这些方法的例子包括但不限于:常规的复性处理、用蛋白沉淀剂处理(盐析方法)、离心、渗透破菌、超处理、超离心、分子筛层析(凝胶过滤)、吸附层析、离子交换层析、高效液相层析(HPLC)和其它各种液相层析技术及这些方法的结合。

重组的具有抑癌功能的人蛋白或多肽有多方面的用途。这些用途包括(但不限于):

直接做为药物治疗具有抑癌功能的蛋白功能低下或丧失所致的疾病，和用于筛选促进或对抗具有抑癌功能的蛋白功能的抗体、多肽或其它配体。例如，抗体可用于激活或抑制具有抑癌功能的人蛋白的功能。用表达的重组具有抑癌功能的人蛋白筛选多肽库可用于寻找有治疗价值的能抑制或刺激具有抑癌功能的人蛋白功能的多肽分子。

5 本发明也提供了筛选药物以鉴定提高(激动剂)或阻遏(拮抗剂)具有抑癌功能的人蛋白的药剂的方法。激动剂提高具有抑癌功能的人蛋白刺激细胞增殖等生物功能，而拮抗剂阻止和治疗与细胞过度增殖有关的紊乱如各种癌症。例如，能在药物的存在下，将哺乳动物细胞或表达具有抑癌功能的人蛋白的膜制剂与标记的具有抑癌功能的人蛋白一起培养。然后测定药物提高或阻遏此相互作用的能力。

10 具有抑癌功能的人蛋白的拮抗剂包括筛选出的抗体、化合物、受体缺失物和类似物等。具有抑癌功能的人蛋白的拮抗剂可以与具有抑癌功能的人蛋白结合并消除其功能，或是抑制具有抑癌功能的人蛋白的产生，或是与多肽的活性位点结合使多肽不能发挥生物学功能。具有抑癌功能的人蛋白的拮抗剂可用于治疗用途。

在筛选作为拮抗剂的化合物时，可以将本发明蛋白加入生物分析测定中，通过测定
15 化合物影响具有抑癌功能的蛋白和其受体之间的相互作用来确定化合物是否是拮抗剂。用上述筛选化合物的同样方法，可以筛选出起拮抗剂作用的受体缺失物和类似物。

本发明的多肽可直接用于疾病治疗，例如，各种恶性肿瘤、和细胞异常增殖等。

本发明的多肽，及其片段、衍生物、类似物或它们的细胞可以用来作为抗原以生产
20 抗体。这些抗体可以是多克隆或单克隆抗体。多克隆抗体可以通过将此多肽直接注射动物的方法得到。制备单克隆抗体的技术包括杂交瘤技术，三瘤技术，人 B-细胞杂交瘤技术，EBV-杂交瘤技术等。

可以将本发明的多肽和拮抗剂与合适的药物载体组合后使用。这些载体可以是水、葡萄糖、乙醇、盐类、缓冲液、甘油以及它们的组合。组合物包含安全有效量的多肽或拮抗剂以及不影响药物效果的载体和赋形剂。这些组合物可以作为药物用于疾病治疗。

25 本发明还提供含有一种或多种容器的药盒或试剂盒，容器中装有一种或多种本发明的药用组合物成分。与这些容器一起，可以有由制造、使用或销售药品或生物制品的政府管理机构所给出的指示性提示，该提示反映出生产、使用或销售的政府管理机构许可其在人体上施用。此外，本发明的多肽可以与其它的治疗化合物结合使用。

30 药物组合物可以以方便的方式给药，如通过局部、静脉内、腹膜内、肌内、皮下、鼻内或皮内的给药途径。具有抑癌功能的蛋白以有效地治疗和/或预防具体的适应症的量来给药。施用于患者的具有抑癌功能的蛋白的量和剂量范围将取决于许多因素，如给药方式、待治疗者的健康条件和诊断医生的判断。

具有抑癌功能的人蛋白的多聚核苷酸也可用于多种治疗目的。基因治疗技术可用于治疗由于具有抑癌功能的蛋白的无表达或异常/无活性的具有抑癌功能的蛋白的表达所
35 致的细胞增殖、发育或代谢异常。重组的基因治疗载体可用于治疗具有抑癌功能的蛋白表达或活性异常所致的疾病。来源于病毒的表达载体如逆转录病毒、腺病毒、腺病毒相关病毒、单纯疱疹病毒、细小病毒等可用于将具有抑癌功能的蛋白基因转移至细胞内。

构建携带具有抑癌功能的蛋白基因的重组病毒载体的方法可见于已有文献(Sambrook, et al.)。另外重组具有抑癌功能的人蛋白基因可包装到脂质体中转移至细胞内。

抑制具有抑癌功能的人蛋白 mRNA 的寡聚核苷酸(包括反义 RNA 和 DNA)以及核酶也在本发明的范围之内。核酶是一种能特异性分解特定 RNA 的酶样 RNA 分子,其作用机制是核酶分子与互补的靶 RNA 特异性杂交后进行核酸内切作用。反义的 RNA 和 DNA 及核酶可用已有的任何 RNA 或 DNA 合成技术获得,如固相磷酸酰胺化学合成法合成寡核苷酸的技术已广泛应用。反义 RNA 分子可通过编码该 RNA 的 DNA 序列在体外或体内转录获得。这种 DNA 序列已整合到载体的 RNA 聚合酶启动子的下游。为了增加核酸分子的稳定性,可用多种方法对其进行修饰,如增加两侧的序列长度,核糖核苷之间的连接应用磷酸硫酸酯键或肽键而非磷酸二酯键。

多聚核苷酸导入组织或细胞内的方法包括:将多聚核苷酸直接注入到体内组织中;或在体外通过载体(如病毒、噬菌体或质粒等)先将多聚核苷酸导入细胞中,再将细胞移植到体内等。

本发明的多肽还可用作肽谱分析,例如,多肽可用物理的、化学或酶进行特异性切割,并进行一维或二维或三维的凝胶电泳分析。

本发明还提供了针对具有抑癌功能的人蛋白抗原决定簇的抗体。这些抗体包括(但不限于):多克隆抗体、单克隆抗体、嵌合抗体、单链抗体、Fab 片段和 Fab 表达文库产生的片段。这些抗体可用常规方法制备。抗具有抑癌功能的人蛋白的抗体可用于免疫组织化学技术中,检测活检标本中的具有抑癌功能的人蛋白。

与具有抑癌功能的人蛋白结合的单克隆抗体也可用放射性同位素标记,注入体内可跟踪其位置和分布。本发明中的抗体可用于治疗或预防与具有抑癌功能的人蛋白相关的疾病。给予适当剂量的抗体可以刺激或阻断具有抑癌功能的人蛋白的产生或活性。

抗体也可用于设计针对体内某一特殊部位的免疫毒素。如具有抑癌功能的人蛋白高亲和性的单克隆抗体可与细菌或植物毒素(如白喉毒素,蓖麻蛋白,红豆碱等)共价结合。

多克隆抗体的生产可用具有抑癌功能的人蛋白或多肽免疫动物,如家兔,小鼠,大鼠等。多种佐剂可用于增强免疫反应,包括但不限于弗氏佐剂等。

具有抑癌功能的人蛋白单克隆抗体可用杂交瘤技术生产(Kohler and Milstein. Nature,1975, 256:495-497)。将人恒定区和非人源的可变区结合的嵌合抗体可用已有的技术生产(Morrison et al ,PNAS,1985,81:6851)。而已有的生产单链抗体的技术(U.S. Pat No.4946778)也可用于生产抗具有抑癌功能的人蛋白的单链抗体。

能与本发明蛋白结合的多肽分子可通过筛选由各种可能组合的氨基酸结合于固相物组成的随机多肽库而获得。筛选时,必须对具有抑癌功能的人蛋白分子进行标记。

本发明还涉及定量和定位检测具有抑癌功能的人蛋白水平的诊断试验方法。这些试验是本领域所熟知的,且包括 FISH 测定和放射免疫测定。试验中所检测的具有抑癌功能的人蛋白水平,可以用作解释具有抑癌功能的人蛋白在各种疾病中的重要性和用于诊断具有抑癌功能的蛋白起作用的疾病。

具有抑癌功能的蛋白的多聚核苷酸可用于具有抑癌功能的蛋白相关疾病的诊断和

治疗。在诊断方面，具有抑癌功能的蛋白的多聚核苷酸可用于检测具有抑癌功能的蛋白的表达与否或在疾病状态下具有抑癌功能的蛋白的异常表达。如具有抑癌功能的蛋白DNA序列可用于对活检标本的杂交以判断具有抑癌功能的蛋白的表达异常。杂交技术包括 Southern 印迹法, Northern 印迹法、原位杂交等。这些技术方法都是公开的成熟技术，相关的试剂盒都可从商业途径得到。本发明的多核苷酸的一部分或全部可作为探针固定在微阵列(Microarray)或 DNA 芯片(又称为“基因芯片”)上，用于分析组织中基因的差异表达分析和基因诊断。用具有抑癌功能的蛋白特异的引物进行 RNA-聚合酶链反应(RT-PCR)体外扩增也可检测具有抑癌功能的蛋白的转录产物。

检测具有抑癌功能的蛋白基因的突变也可用于诊断具有抑癌功能的蛋白相关的疾病。具有抑癌功能的蛋白突变的形式包括与正常野生型具有抑癌功能的蛋白 DNA 序列相比的点突变、易位、缺失、重组和其它任何异常等。可用已有的技术如 Southern 印迹法、DNA 序列分析、PCR 和原位杂交检测突变。另外，突变有可能影响蛋白的表达，因此用 Northern 印迹法、Western 印迹法可间接判断基因有无突变。

本发明的序列对染色体鉴定也是有价值的。这些序列会特异性地针对某条人染色体具体位置且并可以与其杂交。目前，需要鉴定染色体上的各基因的具体位点。然而现在只有很少的基于实际序列数据(重复多态性)的染色体标记物可用于标记染色体位置。为了将这些序列与疾病相关基因相关联。第一步就是将本发明DNA序列定位于染色体上。

简而言之，根据cDNA制备PCR引物(优选15-35bp)，可以将序列定位于染色体上。然后，将这些引物用于PCR筛选含各条人染色体的体细胞杂合细胞。只有那些含有相应于引物的人基因的杂合细胞会产生扩增的片段。

体细胞杂合细胞的PCR定位法，是将DNA定位到具体染色体的快捷方法。使用本发明的寡核苷酸引物，通过类似方法，可利用一组来自特定染色体的片段或大量基因组克隆而实现亚定位。可用于染色体定位的其它类似策略包括原位杂交、用标记的流式分选的染色体预筛选和杂交预选，从而构建染色体特异的cDNA库。

将cDNA克隆与中期染色体进行荧光原位杂交(FISH)，可以在一个步骤中精确地进行染色体定位。此技术的综述，参见Verma等，Human Chromosomes:a Manual of Basic Techniques,Pergamon Press, New York(1988)。

一旦序列被定位到准确的染色体位置，此序列在染色体上的物理位置就可以与基因图数据相关联。这些数据可见于例如，V.Mckusick,Mendelian Inheritance in Man(可通过与Johns Hopkins University Welch Medical Library联机获得)。然后可通过连锁分析，确定基因与业已定位到染色体区域上的疾病之间的关系。

接着，需要测定患病和未患病个体间的cDNA或基因组序列差异。如果在一些或所有的患病个体中观察到某突变，而该突变在任何正常个体中未观察到，则该突变可能是疾病的病因。比较患病和未患病个体，通常涉及首先寻找染色体中结构的变化，如从染色体水平可见的或用基于cDNA序列的PCR可检测的缺失或易位。

本发明的具有抑癌功能的蛋白核苷酸全长序列或其片段通常可以用PCR扩增法、重组法或人工合成的方法获得。对于PCR扩增法，可根据本发明所公开的有关核苷酸序列，

尤其是开放阅读框序列来设计引物,并用市售的cDNA库或按本领域技术人员已知的常规方法所制备的cDNA库作为模板,扩增而得有关序列。当序列较长时,常常需要进行两次或多次PCR扩增,然后再将各次扩增出的片段按正确次序拼接在一起。

一旦获得了有关的序列,就可以用重组法来大批量地获得有关序列。这通常是将其克隆入载体,再转入细胞,然后通过常规方法从增殖后的宿主细胞中分离得到有关序列。

此外,还可用人工合成的方法来合成有关序列,尤其是片段长度较短时。通常,通过先合成多个小片段,然后再进行连接可获得序列很长的片段。

目前,已经可以完全通过化学合成来编码本发明蛋白(或其片段,或其衍生物)的DNA序列。然后可将该DNA序列引入本领域中的各种DNA分子(如载体)和细胞中。此外,还可通过化学合成将突变引入本发明蛋白序列中。

此外,由于本发明的具有抑癌功能的蛋白具有源自人的天然氨基酸序列,因此,与来源于其他物种的同族蛋白相比,预计在施用于人时将具有更高的活性和/或更低的副作用(例如在人体内的免疫原性更低或没有)。

下面结合具体实施例,进一步阐述本发明。应理解,这些实施例仅用于说明本发明而不适用于限制本发明的范围。下列实施例中未注明具体条件的实验方法,通常按照常规条件如Sambrook等人,分子克隆:实验室手册(New York: Cold Spring Harbor Laboratory Press, 1989)中所述的条件,或按照制造厂商所建议的条件。

具体实施方式

实施例1: cDNA基因的获得及对3T3细胞克隆形成的抑制作用

PP9457、PP9943、PP9974、PP10443、PP10472、PP11662、PP11741、PP12301、PP12616、PP12723、PP14356和PP14737是通过用常规方法构建人胎盘cDNA文库获得的。取3、6、10月龄的胎盘组织,用Trizol试剂(GIBCO BRL公司)按厂方说明书提取总RNA,用mRNA提纯试剂盒(Pharmacia公司)提取mRNA。用pCMV-script TMXR cDNA文库构建试剂盒(Stratagene公司)构建上述mRNA的cDNA文库。其中反转录酶改用MMLV-RT-Superscript II(GIBCO BRL),反转录反应在42℃进行。转化XL 10-Gold感受细胞,获得了 1×10^6 cfu/ μ g滴度的cDNA文库。第一轮随机挑取cDNA克隆,其后以高丰度cDNA克隆和已证明有抑癌细胞生长功能的cDNA克隆为探针,杂交筛选cDNA文库,挑取弱阳性及阴性克隆。用Qiagen 96孔板质粒抽提试剂盒,按厂家说明书进行质粒DNA的提取。质粒DNA和空载体同时转染3T3细胞系。100ng DNA酒精沉淀干燥后,加6 μ l H₂O溶解,待转染。每份DNA样品中加0.74 μ l脂质体及9.3 μ l无血清培养液,混匀后,室温放置10分钟。每管中加150 μ l无血清培养液,均分加入3孔生长于96孔板的7721细胞中,37℃放置2小时,每孔再加50 μ l无血清培养液,37℃ 24小时。每孔换100 μ l全培养液,37℃ 24小时,换含G418的全培养液100 μ l,37℃ 24-48小时,边观察,边换G418浓度不等的培养液。约2-3次后,直到镜检细胞有克隆形成,计数。发现以上12个克隆有抑制3T3细胞克隆形成作用,结果如下表所示。

cDNA 克隆转染细胞(3T3)后的克隆形成情况

cDNA 克隆名称	cDNA 克隆数(三个重复)			空载体克隆数(三个重复)		
PP9457	2	5	1	27	29	30
PP9943	14	15	14	27	29	30
PP9974	10	9	7	27	29	30
PP10443	10	13	15	27	29	30
PP10472	12	13	15	27	29	30
PP11662	15	18	13	27	29	30
PP11741	14	19	15	27	29	30
PP12301	7	5	8	27	29	30
PP12616	14	17	11	27	29	30
PP12723	4	8	7	27	29	30
PP14356	16	19	15	27	29	30
PP14737	16	18	19	27	29	30

对 cDNA 克隆采用双脱氧终止法,在 ABI377 DNA 自动测序仪上测定其一端近 500bp 的核苷酸序列。分析后,确定为新基因克隆,进行另一端测序,仍未获得全长 cDNA 序列,设计引物,再次进行测序,直到获得全长序列(SEQ ID NO: 1、4、7、10、13、5 16、19、22、25、28、31、34)。

实施例 2: 从胎盘 cDNA 中 PCR 获得全长基因:

取 3、6、10 月龄的胎盘组织,用 Trizol 试剂(GIBCO BRL 公司)按厂方说明书提取总 RNA,用 mRNA 提纯试剂盒(Pharmacia 公司)提取 mRNA。用 MMLV-RT-Superscript II(GIBCO BRL),反转录酶在 42 °C 进行反转录反应,获得胎盘 cDNA。利用各个基因的转引物(如下表所示),按 97 °C 3' 1 个循环; 94 °C 30" → 60 °C 30" → 72 °C 1', 共 35 个循环; 72 °C 10', 1 个循环进行 PCR 扩增,获得含有完整开放阅读框序列的各蛋白基因的扩增产物。扩增产物经测序验证,与实施例 1 测得的序列相符,随后用常规技术将扩增产物转入宿主细胞,获得重组蛋白(SEQ ID NO: 2、5、8、11、14、17、20、15 23、26、29、32、35)。

基因特异引物

克隆名称	特异引物 1(5' → 3')	特异引物 2(5' → 3')
PP9457	CCCCACGTTCTCTCAGAT	GCTCAGGGAGGTGAAGTGAC
PP9943	GGTCGAGGCTTCAGTGAGTC	AAATTCCTCAAATTCCTCAAAGG
PP9974	GCAGCCCCGACTAGCAGTCTA	ACAAGATGGAGAGCCACGAC
PP10443	AGAGAGAGGACACCGAAGCC	TTGGTGGAAAAACAAGAGGC
PP10472	CTAGCTCCGCTTGGTGACAG	TTCGATGCTTTCTCAAGGCT
PP11662	GAGGTGTGGATCGAGGAGG	GCCTGTTTTGACAGGTCGTT
PP11741	CTAACCCAGCCCTCCGTG	GCAGTCGTGCTTGTGTCA
PP12301	TTAAAAACAATGAGCTGGGC	ATTTTATTTAGCCGGGCAC
PP12616	CCAAAGCACCTCGAATTAT	GAGGCGGATGGATTACTTGA
PP12723	GTGAAGCTACTGCCTTTGCC	TTGTACATCTGGCCCACTGA
PP14356	GAAGTTGCAGGGTACAAGGC	GGCCCTCTATTTTATTCCACTG
PP14737	GGTGGATGCAGAAACACAAG	AGATCTGCCTCCAAAGCTGA

实施例 3: cDNA 克隆序列分析

1. PP9457

A: 核苷酸序列 (SEQ ID NO:1) 长度: 2554

```

1   GCTCTGACAG ATGGCCTATT GAGGTCAACT TGAATGTGAG GGCTACGGTG TGGTTTCAAA
5   61   CATTTCATGAT GGATGTATTT TCCTACCCCT AACTTAAGGA GAAAAAAAAA AAGACTTCCT
121  TTTTTTTGCC AAAGTCCAGA AAGGGCCTT TAGCCTTTAG TAGGAGCTCA AATTGTTGGG
181  GCCCCTCTAC CTCTCTCAGG GCTAGAACTG CCTGACTCTT GGTGGACGAG CCCTTCAGGG
241  TTCTGCTTTC AGCCCCACCT GGACAGAGGC TTACAAGACT AGGGTCTGGA CCAGAATCTG
301  TGTATTTCTG TCTGGGACCA GGAAGCCGCA GCTGTCCCAT CATCCCCAGC AAATCCTAGA
10  361  AGTGGAGTCT GGATACTTCA AGGATAGAAG TGTGGCAGC CACAGCCATG GACCCAGCTG
421  AGCAGAGCAG ACGCTTTGCA GGCTGCCCTT GGCTTCTTCC TCCCTTTCCC GCTTCTGCTC
481  TCTTTATGGA CTGGTCAGAG GGTAGGTGGG AAAGAACAGA CAAGCCATGG GAAGTTGGCA
541  GTGGGGAGAT TTCCACTGTG GAAACCGCCT GGAATTCGG GCCAGCAGCT TCCTCCTTCA
601  GCCACCTGGC CATACCCCTT AAATAAGCCC CTCACCTTGC TGCCTCAGGA CCTCAAGAT
15  661  TCCATCTGTG GGCTGGCCGG CAAGATGGCA CCAGTGGGGA CCCACACCCT GGCTGGGCAG
721  AGGTGCTGCT AGCAACCTCT CTTCTCTAT AAGAGGAAAT GGAAAATGCA GGTGTGGAA
781  TTGCCCTTTG GGGTCCTTCC TTAATTGAAG GCCACCTTCT CACAGGTTTC ATTCTGCAGG
841  GATTTACTGG AATCTATTGG TGCTGCTGCA TGAGTCTGCT GACAACCTGA CTGCACAAGG
901  ACTGGGTAGC AGACTCCTCA GAGTCCTCTT GACACAAATG TCAGATTTGT GTCACTCTTC
20  961  TGCCTTCGTG AAAAGCCAAT AGCACTCTCA GATATCAGGG GATTTTAGTT CCAAGCAGGG
1021 ACCCTGGTTT CCATACTGCC CTCAGTGGGA GTTTGGATCC AAAGGCTCTG GCTAAGTCAT
1081 TATGTCATCT TTTCACAGGA ATGTAATTT GACTGTCACC TCTGAATTG TTCAGTGTCC
1141 CACCATGGTC TATGAGAAGT ACACTGGAAG CGTGGGGGGA ACACATGACA TGATTTGTGA
1201 ATATCATCAT CTTTGCCAGA CAAGTCTCCA GGGGATCCCT GTTCCCAAC TGAAAGGTGT
25  1261 GAACGGACAC ACACACAGCC TGGATGACGC CTTGGCTGTT CTAAGGGGCT GTAAGGTGGG
1321 CTCTGGGCCT TCCAGCTAGG CTCTCAAGCA CAGCAGAAGC CTCACTGGGC TGCTATGTCT
1381 CTGTATTTGT GGCTTGTGTG GTAGCCTCAG AAGCAGAGCT GTTTGGCAGA CTGGCTGGAG
1441 AAATCCCTC TAGGAGACTT GCCTGTGCTG TGCTTCCAGG TCACAGAGCC CCCCAGAAAC
1501 TCACAGGGGC CCTCTCCCA GAAAAGAATC TATTCTATCA CTTCAGAATC AGGACACTCA
30  1561 AGCTCTGGCA GAGGAAGGCC AAGTACTTT CATGGTCTTA CCCTCTGCTT TTCCCTTTT
1621 TGCAAAAAC CACTGGCCAA ATCCGAACCA TTGCCCTTGT TTCCCCACG TTCTCTCTCA
1681 GATCTTTGTC TCGAAGGGAA AACATAGTGG ATGAAAAGGT GTGGCAGGCT TTGGCACCTT
1741 GTAAAAATTT CTAGTCATCT GTGGATGTTA CCTTGCTTGT CCACAGCAGC CAGTCACCCCT
1801 GGCCAGTCCC ACTTCCTGGA TAATTCCTA CCCTCACCCC ACAGAGCCAT CTCTCTCCAG
35  1861 ACCAAAAGCT GGAAGGAGAG TTGCTTTGAG AGCTTGTTTT TACAACGCA TGTTTATTAT
1921 GATACTTTCT CTCCAAAGGA AACTTTTAAA TCAATGGGAA CAATTAGCAA CAGAAAGAGC
1981 ACAGTCCCCT CTTTGACTG GGTTCCTATT TTAAGCACA ATGAGAGCTC TGGAGCCAGA
2041 ATGCCAGGGT TCTAACTTCA GCATTCACTT ACTAGCTGTA TGATCTTGGC CAAGTCACTT
2101 CACCTCCCTG AGCCCAATT CCCAAGTTG TGAATGGCA ACAATACCTA TGTGTACTG
40  2161 GATTATTGGT TAAAACAGAA TGAGATTCCT TGTGTGAAAA TAGCTATTAT ACCTGACACA
2221 CTCATCGTAT GGGCTCTGCA AAGGGATATT CCCCAACCTG TCCTTCTGTA CAGGAAGCAT
2281 AGGGCACTGC AGATGGGGAA GCATGTCACC TTGGCAGTGA CTCGGTGGCT TCCCAAGCAG
2341 GAGTGTGAGG GGAACCATGA GAGAGAGTCT AGGAGCAAAC ACATACCAC CCTGAGCAGA
2401 TACAGGAGTG GGGAGGGGGC TGTAACCTAG TGAGTGGCTT CCAGGGGCCC CAGGCCCTGC
45  2461 TGGATGTGGG CCAAGCCCTA CAGCTTCCCT AGGCAGTAAG TAAAAACATT CTCCTAGCAT
2521 TAAAATGGTT TCCATAAAAA AAAAAAAAAA AAAA

```

B: 核苷酸序列 (SEQ ID NO:2) 长度: 114

```

1   MKRCGRWLHL VKISSHLWML PCLSTAASHP QQSHFLDNSL PSPHRAISLQ TKSWKESCFE
50  61   SLFLQLHVVY DTFSPKETFK SMTISNRKS TVPAFDWVPI LSTNESSGAR MPGF

```

C. 核苷酸及氨基酸组合序列 (SEQ ID NO:3) 克隆号: PP9457

起始编码子: 1711 ATG 终止编码子: 2053 TAA 蛋白质分子量: 12942.19

(注: (1)给出的是起始和终止编码子第一个核苷酸的位置, (2)分子量单位是道尔顿)

```

55  1   GCT CTG ACA GAT GGC CTA TTG AGG TCA ACT TGA ATG TGA GGG CTA CGG      48
49  49  TGT GGT TTC AAA CAT TCA TGA TGG ATG TAT TTT CCT ACC CCT AAC TTA      96
97  97  AGG AGA AAA AAA AAA AGA CTT CCT TTT TTT TGC CAA AGT CCA GAA AGG      144

```

	145	GGC CTT TAG CCT TTA GTA GGA GCT CAA ATT GTT GGG GCC CCT CTA CCT	192
	193	CTC TCA GGG CTA GAA CTG CCT GAC TCT TGG TGG ACG AGC CCT TCA GGG	240
	241	TTC TGC TTT CAG CCC CAC CTG GAC AGA GGC TTA CAA GAC TAG GGT CTG	288
	289	GAC CAG AAT CTG TGT ATT TCT GTC TGG GAC CAG GAA GCC GCA GCT GTC	336
5	337	CCA TCA TCC CCA GCA AAT CCT AGA AGT GGA GTC TGG ATA CTT CAA GGA	384
	385	TAG AAG TGT TGG CAC GCA CAG CCA TGG ACC CAG CTG AGC AGA GCA GAC	432
	433	GCT TTG CAG GCT GCC CCT GGC TTC TTC CTC CCT TTC CCG CTT CTG CTC	480
	481	TCT TTA TGG ACT GGT CAG AGG GTA GGT GGG AAA GAA CAG ACA AGC CAT	528
	529	GGG AAG TTG GCA GTG GGG AGA TTT CCA CTG TGG AAA CCG CCT GGG AAT	576
10	577	TCC GGC CAG CAG CTT CCT CCT TCA GCC ACC TGG CCA TAC CCC TTA AAT	624
	625	AAG CCC CTC ACC TTG CTG CCT CAG GAC CTT CAA GAT TCC ATC TGT GGG	672
	673	CTG GCC GGC AAG ATG GCA CCA GTG GGG ACC CAC ACC CTG GCT GGG CAG	720
	721	AGG TGC TGC TAG CAA CCT CTC TTC CTC TAT AAG AGG AAA TGG AAA ATG	768
	769	CAG GGT GTG GAA TTG CCC TTT GGG GTC CTT CCT TAA TTG AAG GCC ACC	816
15	817	TTC TCA CAG GTT TCA TTC TGC AGG GAT TTA CTG GAA TCT ATT GGT GCT	864
	865	GCT GCA TGA GTC TGC TGA CAA CCT GAC TGC ACA AGG ACT GGG TAG CAG	912
	913	ACT CCT CAG AGT CCT CTT GAC ACA AAT GTC AGA TTT GTG TCA CTC TTC	960
	961	TGC CTT CGT GAA AAG CCA ATA GCA CTC TCA GAT ATC AGG GGA TTT TAG	1008
	1009	TTC CAA GCA GGG ACC CTG GTT TCC ATA CTG CCC TCA GCT GGA GTT TGG	1056
20	1057	ATC CAA AGG CTC TGG CTA AGT CAT TAT GTC ACT TTT TCA CAG GAA TGT	1104
	1105	AAA TTT GAC TGT CAC CTC TGA ATT TGT TCA GTG TCC CAC CAT GGT CTA	1152
	1153	TGA GAA GTA CAC TGG AAG CGT GGG GGG AAC ACA TGA CAT GAT TTG TGA	1200
	1201	ATA TCA TCA TCT TTG CCA GAC AAG TCT CCA GGG GAT CCC TGT TTC CCA	1248
	1249	ACT GAA AGG TGT GAA CGG ACA CAC ACA CAG CCT GGA TGA CGC CTT GGC	1296
25	1297	TGT TCT AAG GGG CTG TAA GGT GGG CTC TGG GCC TTC CAG CTA GGC TCT	1344
	1345	CAA GCA CAG CAG AAG CCT CAC TGG GCT GCT ATG TCT CTG TAT TTG TGG	1392
	1393	CTT GTG TGG TAG CCT CAG AAG CAG AGC TGT TTG GCA GAC TGG CTG GAG	1440
	1441	AAA TTC CCT CTA GGA GAC TTG CCT GTG CTG TGC TTC CAG GTC ACA GAG	1488
	1489	CCC CCC GGA AAC TCA CAG GGG CCC TCT TCC CAG AAA AGA ATC TAT TCT	1536
30	1537	ATC ACT TCA GAA TCA GGA CAC TCA AGC TCT GGC AGA GGA AGG CCA AGT	1584
	1585	TAC TTT CAT GGT CTT ACC CTC TGC TTT TCC CCT TTT TGG AAA AAA CCA	1632
	1633	CTG GCC AAA TCC GAA CCA TTG CCC TTG TTT CCC CCA CGT TCT CTC TCA	1680
	1681	GAT CTT TGT CTC GAA GGG AAA ACA TAG TGG ATG AAA AGG TGT GGC AGG	1728
35	1	Met Lys Arg Cys Gly Arg	6
	1729	CTT TGG CAC CTT GTT AAA ATT TCT AGT CAT CTG TGG ATG TTA CCT TGC	1776
	7	Leu Trp His Leu Val Lys Ile Ser Ser His Leu Trp Met Leu Pro Cys	22
40	1777	TTG TCC ACA GCA GCC AGT CAC CCT GGC CAG TCC CAC TTC CTG GAT AAT	1824
	23	Leu Ser Thr Ala Ala Ser His Pro Gly Gln Ser His Phe Leu Asp Asn	38
	1825	TCT CTA CCC TCA CCC CAC AGA GCC ATC TCT CTC CAG ACC AAA AGC TGG	1872
	39	Ser Leu Pro Ser Pro His Arg Ala Ile Ser Leu Gln Thr Lys Ser Trp	54
45	1873	AAG GAG AGT TGC TTT GAG AGC TTG TTT TTA CAA CTG CAT GTT TAT TAT	1920
	55	Lys Glu Ser Cys Phe Glu Ser Leu Phe Leu Gln Leu His Val Tyr Tyr	70
	1921	GAT ACT TTC TCT CCA AAG GAA ACT TTT AAA TCA ATG GGA ACA ATT AGC	1968
50	71	Asp Thr Phe Ser Pro Lys Glu Thr Phe Lys Ser Met Gly Thr Ile Ser	86
	1969	AAC AGA AAG AGC ACA GTC CCT GCT TTT GAC TGG GTT CCT ATT TTA AGC	2016
	87	Asn Arg Lys Ser Thr Val Pro Ala Phe Asp Trp Val Pro Ile Leu Ser	102
	2017	ACA AAT GAG AGC TCT GGA GCC AGA ATG CCA GGG TTC TAA CTT CAG CAT	2064
55	103	Thr Asn Glu Ser Ser Gly Ala Arg Met Pro Gly Phe ***	115
	2065	TCA CTT ACT AGC TGT ATG ATC TTG GCC AAG TCA CTT CAC CTC CCT GAG	2112
	2113	CCC CAA TTC CCA AGT TTG TGA AAT GGC AAC AAT ACC TAT GTG TCA CTG	2160
	2161	GAT TAT TGG TTA AAA CAG AAT GAG ATT CCT TGT GTG AAA ATA GCT ATT	2208
60	2209	ATA CCT GAC ACA CTC ATC GTA TGG GCT CTG CAA AGG GAT ATT CCC CAA	2256
	2257	CCT GTC CTT CCT GAC AGG AAG CAT AGG GCA CTG CAG ATG GGG AAG CAT	2304
	2305	GTC ACC TTG GCA GTG ACT CGG TGG CTT CCC AAG CAG GAG TGT CAG GGG	2352
	2353	AAC CAT GAG AGA GAG TCT AGG AGC AAA CAC ATC ACC ACC CTG AGC AGA	2400

2401	TAC AGG AGT GGG GAG GGG GCT GTA ACT CAG TGA GTG GCT TCC AGG GGC	2448
2449	CCC AGG CCC TGC TGG ATG TGG GCC AAG CCC TAC AGC TTC CCT AGG CAG	2496
2497	TAA GTA AAA ACA TTC TCC TAG CAT TAA AAT GGT TTC CAT AAA AAA AAA	2544
2545	AAA AAA AAA A	2554

5

2. PP9943

A: 核苷酸序列 (SEQ ID NO:4) 长度: 2039

1	GCTGAAAATA AAATCAGGCC TGTGTGCTGG GCCTGGCCTG TAGTAAGTAC TTCCAACCC	
61	CTTCTCTTGA CCCACAAATG TTGATGGAGC CCTTCTGTGC CAGGCCCTGA GCTGGGCTGA	
10	121 GGAGGCAGTG AGAGGCAGCT GACATTGTCC TTTCGTGGGA GAAGAGTTCA CTCAGGCTTC	
181	TTTCCACTGC GTTAACCCAG CCCGAGGGAG CCAAACCCCT GGAATCTATT CCCCTTGGCA	
241	AGAACTCCCA GGGTTCCAGG CAGGACTGAC CCCGCACTCC GAGGCATCAG GCCCAGAGCA	
301	CTTGGCTGAT CCAGGTGCCA TCCTTCCCGA AACACCACC ACAGGCAGAA AAGAGCCAAA	
361	CAGGCCAGGA AACCTGTCCC AGATTCTGGC TTCCAGCTGT GGAGGGAATA GCCAGGGCTG	
15	421 GTGCAGATGA GTCGGGAAAC CCAGCAGAGA TGGAGGGGG AGTCTTCGTC ATAACACGTG	
481	GCACTGGCAC ATTCAAGGTC GGTAGCATG GATCAGACAG GATCGTAGCA TGATGAGGGG	
541	TTAGCACAAG GACTAGAGTG GTTCTCAGCC CTGGCTGTAC ATGAGATTG CCTGGGCAAC	
601	TCGAAAATC CGGTGGCTGG GCCATAGACT CCTGAAGTCA GGCTCTCAGG AGCAGAAGGG	
661	CCATCAGTCA TCGTTTAAAGC TCCTCGGGTG CTCCAATGC ATAGTCAAGG TTGAGAGCTA	
20	721 CTGGCCTGGA GCCGTGCCTG GTAACAGTAG GTGCCACCTG CGTGTAGAT GTGGTTTATA	
781	ATGGAGACAC GGTGCACAGC TGCTCCAGAG AGACACATCT GGGCAGTTAC AGTCCAGCAC	
841	AGCAGTGGAG GCTTCCCGCG GCAGCAGGAG AGCCACAGG AAGCTTTCAG CTCAGTGAAG	
901	GGAGTCAAGG TCAGGCTTCT TGAAGGAGG TTGGACAGG CAGCTTTGGC CTCTGAGTCT	
961	TGGGGGCTC CCCATGTGGA ACTGTGCTCT TAAACCAGG CATCACCACC TCAACAGATG	
25	1021 CCAGGGTTTC CTGGTAAATG TAAATACAC AAGTCAGGCA GGGCACAGG GCCCACCCT	
1081	GTAATTCAG CACTTTGGGA GGACAAGGCA GGAGGATTGT TAGAGGTCAG GAGTTTGA	
1141	CCAACCTGGG CAATATAATG ATACCCCCCA TCTCTACAAA AAATAAAAAA ATTAGCCAGG	
1201	CATGGTGATG TGGGAGGATT GTTTGAGCTC AAGTGGTCGA GGCTTCAGTG AGTCATGGTT	
1261	GCGCCACTGC AGTCCAGCCT GGGCAACAGA GCGAGACCCC GTCTCTATAA AATAAAACAT	
30	1321 GCAAATCACT GTCCAGCCAA CACCCAGTC CAGATCCCTG CATTCCGATCC AAGGGAGGCA	
1381	GACTGCTGGG GGAAATTGAG AGTCCCTGAG GTGCCCTGG CAGTGAGCGG CCAGAAAGAG	
1441	AAGCAGGAAG GCGCCAGCAT CAGGGAAC TGCTTCCCTC AAAGTGCTGG GATTACAGGT	
1501	GTGAGCCACC GCGCCAGGCC TCAGCCCACT TCTTTTGGGT GGCAATGGTT TGGATATCGT	
1561	TTGTCTCAC TAAAATTCAT GTTGAGATTG GAGCCCACT GTGGCAGGTG TTGGGATGTG	
35	1621 GGGCCTCATA GGAGCTGTGT GGTCTTGGAG GTGGAGTCT CATGGATAGA TTAATGCCTG	
1681	CCTTAAGGGG TGAGTGAGTG CTCACCCTTG GGAATTTGGG AATTTGGTTT TCTCTCCCTT	
1741	GCTTCCCTCA GCACCATGTG ATCTCTGTGC ACACAGCTGC ATCCTTCTGC CCTCCCCAG	
1801	AAGCAGCAGC AGCCTCAGG ATGAGGGAGG TCCTCAGTGA GTGCACCTGC CCAGTCTTGA	
1861	ACCTTCCAGC CATCAGAATC TGAGCCAAAT CAACCTTTC CTTTTATAAA GACCCAGCCT	
40	1921 CAGGTCTTCT GTCAGAGCAA CACAAAATGG ACTCAGCAG GATCAAATG TGCTCCCCC	
1981	ACCCACAAA AAAACTTTAT ATTAAAATCC TAACCCCAAG CAAAAAAA AAAAAAAA	

B: 核苷酸序列 (SEQ ID NO:5) 长度: 139

1	MVAPLQSSLG NRARPRLYKI KHANHCPANT PVQIPAFDPR EADCWGLRV LEVPLAVSGQ	
45	61 KEKQEGASIT RNCFPQSAGI TGVSHRARPQ PTSFGWQWFG YRLSSLFML RFEPQCGRCW	
121	DVGPHRSCV LEVSSWID	

C. 核苷酸及氨基酸组合序列 (SEQ ID NO:6) 克隆号: PP9943

起始编码子: 1255 ATG 终止编码子: 1672 TAA 蛋白质分子量: 15661.18

50	1 GCT GAA AAT AAA ATC AGG CCT GTG TGC TGG GCC TGG CCT GTA GTA AGT	48
	49 ACT TCC CAA CCC CTT CTC TTG ACC CAC AAA TGT TGA TGG AGC CCT TCT	96
	97 GTG CCA GGC CCT GAG CTG GGC TGA GGA GGC AGT GAG AGG CAG CTG ACA	144
	145 TTG TCC TTT CGT GGG AGA AGA GTT CAC TCA GGC TTC TTT CCA CTG CGT	192
	193 TAA CCC AGC CCG AGG GAG CCA AAC CCC TGG AAT CTA TTC CCC TTG GCA	240
55	241 AGA ACT CCC AGG GTT CCA GGC AGG ACT GAC CCC GCA CTC CGA GGC ATC	288
	289 AGG CCC AGA GCA CTT GGC TGA TCC AGG TGC CAT CCT TCC CGA AAC ACC	336
	337 ACC CAC AGG CAG AAA AGA GCC AAA CAG GCC AGG AAA CCT GTC CCA GAT	384
	385 TCT GGC TTC CAG CTG TGG AGG GAA TAG CCA GGC CTG GTG CAG ATG AGT	432

	433	CGG GAA ACC CAG CAG AGA TGG AGG GGG GAG TCT TCG TCA TAA CAC GTG	480
	481	GCA CTG GCA CAT TCA AGG TCG GTT AGC ATG GAT CAG ACA GGA TCG TAG	528
	529	CAT GAT GAG GGG TTA GCA CAA GGA CTA GAG TGG TTC TCA GCC CTG GCT	576
5	577	GTA CAT GAG ATT TGC CTG GGC AAC TCG AAA ACT CCG GTG GCT GGG CCA	624
	625	TAG ACT CCT GAA GTC AGG CTC TCA GGA GCA GAA GGG CCA TCA GTC ATC	672
	673	GTT TAA GCT CCT CGG GTG CTC CCA ATG CAT AGT CAA GGT TGA GAG CTA	720
	721	CTG GCC TGG AGC CGT GCC TGG TAA CAG TAG GTG CCA CCT GCG TGT TAG	768
	769	ATG TGG TTT ATA ATG GAG ACA CGG TGC ACA GCT GCT CCA GAG AGA CAC	816
	817	ATC TGG GCA GTT ACA GTC CAG CAC AGC AGT GGA GGC TTC CCG CGG CAG	864
10	865	CAG GAG AGC CCA CAG GAA GCT TTC AGC TCA GTG AAG GGA GTC AAG GTC	912
	913	AGG CTT CTT GGA AGG AGG TTG GAC AGG GCA GCT TTG GCC TCT GAG TCT	960
	961	TGG GGG CCT CCC CAT GTG GAA CTG TGC TCT TAA ACC AGG GCA TCA CCA	1008
	1009	CCT CAA CAG ATG CCA GGG TTT CCT GGT AAA TGT AAA ATA CAC AAG TCA	1056
	1057	GGC AGG GCA CAG GGG CCC ACC CCT GTA ATT TCA GCA CTT TGG GAG GAC	1104
15	1105	AAG GCA GGA GGA TTG TTA GAG GTC AGG AGT TTG AGA CCA ACC TGG GCA	1152
	1153	ATA TAA TGA TAC CCC CCA TCT CTA CAA AAA ATA AAA AAA TTA GCC AGG	1200
	1201	CAT GGT GAT GTG GGA GGA TTG TTT GAG CTC AAG TGG TCG AGG CTT CAG	1248
	1249	TGA GTC ATG GTT GCG CCA CTG CAG TCC AGC CTG GGC AAC AGA GCG AGA	1296
20	1	Met Val Ala Pro Leu Gln Ser Ser Leu Gly Asn Arg Ala Arg	14
	1297	CCC CGT CTC TAT AAA ATA AAA CAT GCA AAT CAC TGT CCA GCC AAC ACC	1344
	15	Pro Arg Leu Tyr Lys Ile Lys His Ala Asn His Cys Pro Ala Asn Thr	30
	1345	CCA GTC CAG ATC CCT GCA TTC GAT CCA AGG GAG GCA GAC TGC TGG GGG	1392
25	31	Pro Val Gln Ile Pro Ala Phe Asp Pro Arg Glu Ala Asp Cys Trp Gly	46
	1393	AAA TTG AGA GTC CTC GAG GTG CCC CTG GCA GTG AGC GGC CAG AAA GAG	1440
	47	Lys Leu Arg Val Leu Glu Val Pro Leu Ala Val Ser Gly Gln Lys Glu	62
30	1441	AAG CAG GAA GGC GCC AGC ATC ACG AGG AAC TGC TTC CCT CAA AGT GCT	1488
	63	Lys Gln Glu Gly Ala Ser Ile Thr Arg Asn Cys Phe Pro Gln Ser Ala	78
	1489	GGG ATT ACA GGT GTG AGC CAC CGC GCC AGG CCT CAG CCC ACT TCT TTT	1536
35	79	Gly Ile Thr Gly Val Ser His Arg Ala Arg Pro Gln Pro Thr Ser Phe	94
	1537	GGG TGG CAA TGG TTT GGA TAT CGT TTG TCC TCA CTA AAA TTC ATG TTG	1584
	95	Gly Trp Gln Trp Phe Gly Tyr Arg Leu Ser Ser Leu Lys Phe Met Leu	110
	1585	AGA TTC GAG CCC CAG TGT GGC AGG TGT TGG GAT GTG GGG CCT CAT AGG	1632
40	111	Arg Phe Glu Pro Gln Cys Gly Arg Cys Trp Asp Val Gly Pro His Arg	126
	1633	AGC TGT GTG GTC TTG GAG GTG GAG TCC TCA TGG ATA GAT TAA TGC CTG	1680
	127	Ser Cys Val Val Leu Glu Val Glu Ser Ser Trp Ile Asp ***	140
45	1681	CCT TAA GGG GTG AGT GAG TGC TCA CCC TTG GGA ATT TGG GAA TTT GGT	1728
	1729	TTT CTC TCC CTT GCT TCC TTC AGC ACC ATG TGA TCT CTG TGC ACA CAG	1776
	1777	CTG CAT CCT TCT GCC CTC CCC CAG AAG CAG CAG CAG CCT CAG GGA TGA	1824
	1825	GGG AGG TCC TCA GTG AGT GCA CCT GCC CAG TCT TGA ACC TTC CAG CCA	1872
	1873	TCA GAA TCT GAG CCA AAT CAA CCT CTT CCT TTT ATA AAG ACC CAG CCT	1920
50	1921	CAG GTC TTC TGT CAG AGC AAC ACA AAA TGG ACT CAG CAC GGA TCA AAT	1968
	1969	TGT GTC TCC CCC ACC CCC ACA AAA AAC TTT ATA TTA AAA TCC TAA CCC	2016
	2017	CCA GCA AAA AAA AAA AAA AAA AA	2039

3. PP9974

55 A: 核苷酸序列 (SEQ ID NO:7) 长度: 2630

	1	GGCGAACACA GCACACACGA ACACAGCACA CACAGCACAC ACACAAACAC AGCACACACA
	61	TGCACACAGC ACATGCACAC ACAGCACACA CATGAACACA GCACACAGCA CACACATGCA
	121	CACAGCACAC ACGCATGCAC AGCACACATG AACACAGCAC ACACAAACAC ACAGCACACA
	181	CATGCACACA CAGCACACAG ACTCATGCGC AGCACATACA TGAACACAGC TCACAGCACA
60	241	CAAAACAGCA GCACACACGT TGCACACGCA AGCACCCACC TGCACACACA CATGGCCACA
	301	CACACGCACA CCCCCACAAA ATTTGGATGAA AACAAATAGC ATATCTAAGC AACTACGATA
	361	TCTGTATGGA TCAGGCCAAA GTCCCGCTAA GATTCTCCAA TGTTTTCATG GTCTGAGCCC
	421	CCCTCTGTT CCCATCTGCA CTGCCCTCG GCCCTGTCTG TGCCCTGCCT CTCAGAGGAG

481 GGGGCTCAGA TGGTGCGGCC TGAGTGTGCG GCCGGCGGCA TTTGGGATAC ACCCGTAGGG
 541 TGGGCGGGGT GTGTCCAGG CCTAATTCCA TCTTCCACC ATGACAGAGA TGCCCTGTG
 601 AGGCTGGCCT CCTTGGCGCC TGTCCCACG GCCCCGCGAG CGTGACGCAC GATGCTCCCC
 661 ATACCCACC CATTCCCGAT ACACCTTACT TACTGTGTGT TGGCCAGCC AGAGTGAGGA
 5 721 AGGAGTTTGG CCACATTGGA GATGGCGGTA GCTGAGCAGA CATGCCCCCA CGAGTAGCCT
 781 GACTCCCTGG TGTGCTCCTG GAAGGAAGAT CTTGGGGACC CCCCACCGG AGCACACCTA
 841 GGGATCATCT TTGCCCGTCT CCTGGGGACC CCCCAAGAAA TGTGGAGTCC TCGGGGGCCG
 901 TGCACTGATG CGGGGAGTGT GGGAACTCTG GCGGTTGGAG GGGTGGGTGG GGGGCAGTGG
 961 GGGCTGGCG GGGGAGTTC TGGGGTAGGA AGTGGTCCCG GGAGATTTG GATGGAAAAG
 10 1021 TCAGGAGGAT TGACAGCAGA CTTGCAGAAT TACATAGAGA AATTAGGAAC CCCCAAATTT
 1081 CATGTCAATT GATCTATTCC CCCTTTTGT TTCTTGGGC ATTTTTCCTT TTTTTTTTTT
 1141 TTTTGTTTTT TTTTACCC CTCTTAGCTT TATGCGCTCA GAAACCAAAT TAAACCCCC
 1201 CCCCATGTAA CAGGGGGGCA GTGACAAAAG CAAGAACGCA CGAAGCCAGC CTGAGACCA
 1261 CCACGTCCTG CCCCCGCCA TTTATCGCCC TGATTGGATT TTGTTTTTCA TCTGTCCCTG
 15 1321 TTGCTTGGGT TGAGTTGAGG GTGGAGCCTC CTGGGGGCA CTGGCCACTG AGCCCCCTG
 1381 GAGAAGTCAG AGGGGAGTGG AGAAGGCCAC TGTCGGCCT GGCTTCTGGG GACAGTGGCT
 1441 GGTCCCCAGA AGTCTGAGG GCGGAGGGG GGGTGGGCA GGGTCTCCTC AGGTGTCAGG
 1501 AGGGTGTCTG GAGGCCACAG GAGGGGGCTC CTGGCTGGCC TGAGGCTGGC CGGAGGGGAA
 1561 GGGGCTAGCA GGTGTGTAAG CAGAGGGTTC CATCAGGCTG GGGCAGGGTG GCCGCCTTCC
 20 1621 GCACACTTGA GGAACCCTCC CCTCTCCCTC GGTGACATCT TGCCCCCCCC TCAGCACCTC
 1681 GCCTTGTCTC CAGGAGGTCC GAAGCTCTGT GGGACCTCTT GGGGGCAAGG TGGGTTGAGG
 1741 CCGGGGAGTA GGGAGGTGAG GCGGGTCTGA GCCCACAGAG CAGGAGAGCT GCCAGGTCTG
 1801 CCCATCGACC AGGTTGCTTG GGGCCCGGAG CCCACGGGTC TGGTGATGCC ATAGCAGCCA
 1861 CCACCGCGGC GCCTAGGGCT GCGGCAGGGA CTCGGCCTCT GGGAGGTTTA CCTCGCCCC
 25 1921 ACTTGTGCCC CCAGTCTAGC CCCCTGACG GCAGCCCGAC TAGCAGTCTA GAGGCTGAG
 1981 GCTTCTGGGT CCTGGTGACG GGGCTGGCAT GACCCCGGGG GTCGTCCATG CCAGTCCGCC
 2041 TCAGTCCGAG AGGGTCCCTC GCAAGCGCC CTGTGAGTGG GCCATTCCGA ACATTGGACA
 2101 GAAGCCCAA GAGCCAAATT GTCACAATTG TGAACCCAC ATTGCCTGA GATCCAAAAC
 2161 GCTTCGAGGC ACCCCAAATT ACCTGCCCAT TCGTCAGGAC ACCCACCCAC CCAGTGTTAT
 30 2221 ATTCTGCCTC GCCGGAGTGG GTGTTCCCGG GGGCACTTGC CGACCAGCCC CTTGCGTCCC
 2281 CAGGTTTGCA GCTCTCCCT GGGCCACTAA CCATCCTGGC CCGGGCTGCC TGTCTGACCT
 2341 CCGTGCCTAG TCGTGGCTCT CCATCTTGT TCCCTCCCGT GTCCCAATG TCTTCAGTGG
 2401 GGGCCCCCT CTGGGTCCC CTCCTCTGCC ATCACCTGAA GACCCCCAGC CCAAACACTG
 2461 AATGTCACCT GTGCTGCG CCTCGGTCCA CCTTGGGGC CGTGTGTTGAC TCAACTCAGC
 35 2521 TCCTTAAACG CTAATATTTC CGGCAAAATC CCATGCTTGG GTTTTGTCTT TAACCTTGTA
 2581 ACGCTTGCAA TCCCAATAAA GCATTAAG TCAAAAAA AAAAAAAA

B: 核苷酸序列 (SEQ ID NO:8) 长度: 113

1 MTPGVVHASP PQSQRVPRQA PCEWAIRNIG QKPKEPNCHN CGTHIGLRSK TLRGTPNYLP
 40 61 IRQDTHPPSV IFCLAGVGP GGTCPAPCV PRFAALPWAT NHPGPGCLSD LRA

C. 核苷酸及氨基酸组合序列 (SEQ ID NO:9) 克隆号: PP9974

起始编码子: 2009 ATG 终止编码子: 2348 TAG 蛋白质分子量: 12086.38

1 G GCG AAC ACA GCA CAC ACG AAC ACA GCA CAC ACA GCA CAC ACA CAA 46
 45 47 ACA CAG CAC ACA CAT GCA CAC AGC ACA TGC ACA CAC AGC ACA CAC ATG 94
 95 AAC ACA GCA CAC AGC ACA CAC ATG CAC ACA GCA CAC ACG CAT GCA CAG 142
 143 CAC ACA TGA ACA CAG CAC ACA CAA ACA CAC AGC ACA CAC ATG CAC ACA 190
 191 CAG CAC ACA CAC TCA TGC GCA GCA CAT ACA TGA ACA CAG CTC ACA GCA 238
 239 CAC AAA CAC GCA GCA CAC ACG TTG CAC ACG CAA GCA CCC ACC TGC ACA 286
 50 287 CAC ACA TGC GCA CAC ACA CGC ACA CCC CCA CAA AAT TGG ATG AAA ACA 334
 335 ATA AGC ATA TCT AAG CAA CTA CGA TAT CTG TAT GGA TCA GGC CAA AGT 382
 383 CCC GCT AAG ATT CTC CAA TGT TTT CAT GGT CTG AGC CCC CCT CCT GTT 430
 431 CCC ATC TGC ACT GCC CCT CGG CCC TGT CTG TGC CCT GCC TCT CAG AGG 478
 479 AGG GGG CTC AGA TGG TGC GGC CTG AGT GTG CGG CCG GCG GCA TTT GGG 526
 55 527 ATA CAC CCG TAG GGT GGG CGG GGT GTG TCC CAG GCC TAA TTC CAT CTT 574
 575 TCC ACC ATG ACA GAG ATG CCC TTG TGA GGC TGG CCT CCT TGG CGC CTG 622
 623 TCC CCA CGG CCC CCG CAG CGT GAC GCA CGA TGC TCC CCA TAC CCC ACC 670
 671 CAT TCC CGA TAC ACC TTA CTT ACT GTG TGT TGG CCC AGC CAG AGT GAG 718

	719	GAA GGA GTT TGG CCA CAT TGG AGA TGG CGG TAG CTG AGC AGA CAT GCC	766
	767	CCC ACG AGT AGC CTG ACT CCC TGG TGT GCT CCT GGA AGG AAG ATC TTG	814
	815	GGG ACC CCC CCA CCG GAG CAC ACC TAG GGA TCA TCT TTG CCC GTC TCC	862
5	863	TGG GGA CCC CCC AAG AAA TGT GGA GTC CTC GGG GGC CGT GCA CTG ATG	910
	911	CGG GGA GTG TGG GAA GTC TGG CGG TTG GAG GGG TGG GTG GGG GGC AGT	958
	959	GGG GGC TGG GCG GGG GGA GTT CTG GGG TAG GAA GTG GTC CCG GGA GAT	1006
	1007	TTT TGA TGG AAA AGT CAG GAG GAT TGA CAG CAG ACT TGC AGA ATT ACA	1054
	1055	TAG AGA AAT TAG GAA CCC CCA AAT TTC ATG TCA ATT GAT CTA TTC CCC	1102
10	1103	CTC TTT GTT TCT TGG GGC ATT TTT CCT TTT TTT TTT TTT TTT GTT TTT	1150
	1151	TTT TTA CCC CTC CTT AGC TTT ATG CGC TCA GAA ACC AAA TTA AAC CCC	1198
	1199	CCC CCC ATG TAA CAG GGG GGC AGT GAC AAA AGC AAG AAC GCA CGA AGC	1246
	1247	CAG CCT GGA GAC CAC CAC GTC CTG CCC CCC GCC ATT TAT CGC CCT GAT	1294
	1295	TGG ATT TTG TTT TTC ATC TGT CCC TGT TGC TTG GGT TGA GTT GAG GGT	1342
	1343	GGA GCC TCC TGG GGG GCA CTG GCC ACT GAG CCC CCT TGG AGA AGT CAG	1390
15	1391	AGG GGA GTG GAG AAG GCC ACT GTC CGG CCT GGC TTC TGG GGA CAG TGG	1438
	1439	CTG GTC CCC AGA AGT CCT GAG GGC GGA GGG GGG GGT TGG GCA GGG TCT	1486
	1487	CCT CAG GTG TCA GGA GGG TGC TCG GAG GCC ACA GGA GGG GGC TCC TGG	1534
	1535	CTG GCC TGA GGC TGG CCG GAG GGG AAG GGG CTA GCA GGT GTG TAA ACA	1582
	1583	GAG GGT TCC ATC AGG CTG GGG CAG GGT GGC CGC CTT CCG CAC ACT TGA	1630
20	1631	GGA ACC CTC CCC TCT CCC TCG GTG ACA TCT TGC CCG CCC CTC AGC ACC	1678
	1679	CTG CCT TGT CTC CAG GAG GTC CGA AGC TCT GTG GGA CCT CTT GGG GGC	1726
	1727	AAG GTG GGG TGA GGC CGG GGA GTA GGG AGG TCA GGC GGG TCT GAG CCC	1774
	1775	ACA GAG CAG GAG AGC TGC CAG GTC TGC CCA TCG ACC AGG TTG CTT GGG	1822
	1823	CCC CGG AGC CCA CGG GTC TGG TGA TGC CAT AGC AGC CAC CAC CGC GGC	1870
25	1871	GCC TAG GGC TGC GGC AGG GAC TCG GCC TCT GGG AGG TTT ACC TCG CCC	1918
	1919	CCA CTT GTG CCC CCA GCT CAG CCC CCC TGC ACG CAG CCC GAC TAG CAG	1966
	1967	TCT AGA GGC CTG AGG CTT CTG GGT CCT GGT GAC GGG GCT GGC ATG ACC	2014
	1		Met Thr 2
30	2015	CCG GGG GTC GTC CAT GCC AGT CCG CCT CAG TCG CAG AGG GTC CCT CGG	2062
	3	Pro Gly Val Val His Ala Ser Pro Pro Gln Ser Gln Arg Val Pro Arg	18
	2063	CAA GCG CCC TGT GAG TGG GCC ATT CGG AAC ATT GGA CAG AAG CCC AAA	2110
35	19	Gln Ala Pro Cys Glu Trp Ala Ile Arg Asn Ile Gly Gln Lys Pro Lys	34
	2111	GAG CCA AAT TGT CAC AAT TGT GGA ACC CAC ATT GGC CTG AGA TCC AAA	2158
	35	Glu Pro Asn Cys His Asn Cys Gly Thr His Ile Gly Leu Arg Ser Lys	50
	2159	ACG CTT CGA GGC ACC CCA AAT TAC CTG CCC ATT CGT CAG GAC ACC CAC	2206
40	51	Thr Leu Arg Gly Thr Pro Asn Tyr Leu Pro Ile Arg Gln Asp Thr His	66
	2207	CCA CCC AGT GTT ATA TTC TGC CTC GCC GGA GTG GGT GTT CCC GGG GGC	2254
	67	Pro Pro Ser Val Ile Phe Cys Leu Ala Gly Val Gly Val Pro Gly Gly	82
45	2255	ACT TGC CGA CCA GCC CCT TGC GTC CCC AGG TTT GCA GCT CTC CCC TGG	2302
	83	Thr Cys Arg Pro Ala Pro Cys Val Pro Arg Phe Ala Ala Leu Pro Trp	98
	2303	GCC ACT AAC CAT CCT GGC CCG GGC TGC CTG TCT GAC CTC CGT GCC TAG	2350
50	99	Ala Thr Asn His Pro Gly Pro Gly Cys Leu Ser Asp Leu Arg Ala ***	114
	2351	TCG TGG CTC TCC ATC TTG TCT CCT CCC CGT GTC CCC AAT GTC TTC AGT	2398
	2399	GGG GGG CCC CCT CTT GGG TCC CCT CCT CTG CCA TCA CCT GAA GAC CCC	2446
	2447	CAC GCC AAA CAC TGA ATG TCA CCT GTG CCT GCC GCC TCG GTC CAC CTT	2494
	2495	GCG GCC CGT GTT TGA CTC AAC TCA GCT CCT TTA ACG CTA ATA TTT CCG	2542
55	2543	GCA AAA TCC CAT GCT TGG GTT TTG TCT TTA ACC TTG TAA CGC TTG CAA	2590
	2591	TCC CAA TAA AGC ATT AAA AGT CAA AAA AAA AAA AAA A	2630

4. PP10443

A: 核苷酸序列 (SEQ ID NO:10) 长度: 2525

60	1	GCCGGACCTT CAGGCCCTGG GGTGAATTCA GCTGCTCCTA CATCAGCTTC CGGAACCACC
	61	AAAAATTCAA ATTGGGATT TCCGAGTAA ACAAGAGCCT ATAGCCCTTT GCTCAATGCT
	121	GGATTTAATA CGTATATATT TTTAAGCGAG TTGGTTTTTT CCCCTTTGAT TTTTGATCTT
	181	CGCGACAGTT CCTCCCACGC ATATTATCGT TGTTGCCGTC GTTTCTCTC CCCGCGTGGC

241 TCCTTGACCT GCGAGGGAGA GAGAGGACAC CGAAGCCGGG AGCTCGCAGG GACCATGTAT
 301 CAGAGCTTGG CCATGGCCGC CAACCACGGG CCGCCCCCGG GTGCCTACGA GCGGGGCGGC
 361 CCCGGCGCCT TCATGCACGG CGCGGGCGCC GCGTCCTCGC CAGTCTACGT GCCCACACCG
 421 CGGGTGCCCT CCTCCGTGCT GGGCCTGTCC TACCTCCAGG GCGGAGGCGC GGGCTCTGCG
 5 481 TCCGGAGGCG CCTCGGGCGG CAGCTCCGGT GGGGCCGCGT CTGGTGCGGG GCCCGGGACC
 541 CAGCAGGGCA GCCCGGGATG GAGCCAGGGG GGAGCCGACG GAGCCGCTTA CACCCCGCCG
 601 CCGGTGTCCG CGCGCTTCTC CTTCCCGGGG ACCACCGGGT CCCTGGCGGC CGCCGCGGCC
 661 GCTGCGCGG CCCGGGAAGC TCGGCCTAC AGCAGTGGCG GCGGAGCGGC GGGTGCGGGC
 721 CTGGCGGGCC GCGAGCAGTA CGGGCGCGCC GGCTTCGCGG GCTCCTACTC CAGCCCCTAC
 10 781 CCGGCTTACA TGGCCGACGT GGGCGCGTCC TGGGCCGAG CCGCCGCGC CTCCCGCGGC
 841 CCCTTCGACA GCCCGGTCTT GCACAGCCTG CCCGGCCGGG CCAACCCGGC CGCCCGACAC
 901 CCCAATCTCG GTGAGTAGGA GCGCGAGGGC TGGGGCGCGT GAGGGCCGGG GCAGGGGCGG
 961 TCTTGAGCCC TGTCGAGGGC CTCTTGT TCCACCAACG CCTTCGTTGG GCTGGGGATG
 1021 GTGCTTCACT ACCTCGAGTT TCTAGGGAAG GCAGAAGCCA GTGCGGGGCT GCGGACATCA
 15 1081 CAGCCCCAGA AGACCGGCTT CTGTGAAGG GGCCGGGCTT GCCCGCCGGG GCCTCTTCTG
 1141 AGATGGTGTG AGGGTCGGAG TCGGCCTCC CCGCCATCCC AGACATCGAC CGTGGCCGCG
 1201 CTGCGCTGTG GGTGACGCGG GAGGACAGCG GGCTCCCTGG AGAGCCGGGG GCAGCGGCTT
 1261 GGGATTTCTT CGTGAAGGT GCTGGAGATT GCTGAGTTT TCGGCCCTT TCCTCCCGC
 1321 CCGCCCTCGG GCCTCCGAG GAACTGATT ACAATGGTTT GGACCGCAGA CCTTCTGGGC
 20 1381 CATTGGCGG CCCAGCTGGA GGATCCCTCG GGGTAGCTGA TGATTTTCCC GTCGGGGGTC
 1441 TCACACCGAG AACAAAGGAG GGATGGACAA AGGAGACGCC GGGGAGATGC GCGGAACAGG
 1501 AGCCGGCACT GTGCGGGTGC CACCCGGCCG AGCGCGTGGG CGCATCATGC GGGCAGCGGG
 1561 GGGGGGGGCG GCACACGCCC GGTCACTGTC CGGGAACATA GGGACCTCAA ACGGGCTTGT
 1621 TCATGACACC CGAGTTAAAT GGAGACTTTG CAGTCGCTTG CACGCGTGA GCCTCTCTT
 25 1681 CTGCGTGGG CCAGGGTTGG AAATAACCGT TGTGTAGGT TCCATGCACT GTTTCCTCG
 1741 GATGTCAGAC GGGGAGGGAC GGCAAACCTG TCTCAACCTC CACTGATTCA CAAATAAAGC
 1801 CAGCGGGATC TGAGAAGGGG CCTGAGTACA CGGGCCGGGG GAGAAAGGGA AGTGCCAACC
 1861 CCTAGTTCAA AATGCAAACG ACCTCTGGAA TTTCCGGGAA AGACGGAGGA GTGAGTTTGG
 1921 ATTGAGCCCA CCCTGTGGGG GAGGGGAAGC CCAGGCTTGA GAAGCAAAGC TCGCGTTTAT
 30 1981 TGACCACCTA CTAAGTGCTA AATCCTTCTG CATTGTCTT CACTTTGAAG AGTAAAGGCT
 2041 ACCCTGCCA GGTTCAGT CTTGGGCTGG CTCCGAGAAG GGCCTAGGCT TTGAAGCGCT
 2101 TTTAAATTGT CCTCTGGCC TGGGTGGCCA GGAAGGTTT CCGGGGGTGC AGCCAAATAC
 2161 ACATCGCCG CAAACTGATT AACCTGAAA GTGGCGCCA GGCCTGCTT GTGGCTCTT
 35 2221 GTGTTCCCTT CCCCCTACC CCTCTGGGG GAGCCAGACG GCCACCCCA GGGGAGGAGG
 2281 GGCCCGCCG AGCCTCGCCG CGACTTTTGC GAGATAGCGG GCGGACATGG CCACACAATG
 2341 GAGCCCGCAG GCGGGAGTGC GGGGCGGGC GCGGCGCCCT GGCCTTGGC GCTTACGGGG
 2401 TCCTCTCCAG GGCCTCTGG GGCCTCTGAC TTTAAATAGG GAGGACTGGG CCAGGAGATC
 2461 GAGACCATCC TGGCCAACCT GTTGAACCC CGCCTCTACT AAAATACAAA AAAAAAAAAA
 40 2521 AAAAA

B: 核苷酸序列 (SEQ ID NO:11) 长度: 207

1 MYQSLAMAAN HGPPPGAYEA GPGAFMHGA GAASSPYVYP TPRVPSSVLG LSYLQGGGAG
 61 SASGGASGGS SGGAAAGGAP GTQQSPGWS QAGADGAAAT PPPVSPRFSF PGTGSLAAA
 121 AAAAAAREAA AYSSGGGAAG AGLAGREQYG RAGFAGSYSS PYPAYMADVG ASWAAAAAAS
 45 181 AGPFDSPVLH SLPGRANPAA RHPNLGE

C. 核苷酸及氨基酸组合序列 (SEQ ID NO:12) 克隆号: PP10443

起始编码子: 295 ATG 终止编码子: 916 TAG 蛋白质分子量: 19350.81

50 1 GCC GGA CCT TCA GGC CCT GGG GTG AAT TCA GCT GCT CCT ACA TCA GCT 48
 49 TCC GGA ACC ACC AAA AAT TCA AAT TGG GAT TTT CCG GAG TAA ACA AGA 96
 97 GCC TAT AGC CCT TTG CTC AAT GCT GGA TTT AAT ACG TAT ATA TTT TTA 144
 145 AGC GAG TTG GTT TTT TCC CCT TTG ATT TTT GAT CTT CGC GAC AGT TCC 192
 193 TCC CAC GCA TAT TAT CGT TGT TGC CGT CGT TTT CTC TCC CCG CGT GGC 240
 241 TCC TTG ACC TGC GAG GGA GAG AGA GGA CAC CGA AGC CGG GAG CTC GCA 288
 55 289 GGG ACC ATG TAT CAG AGC TTG GCC ATG GCC GCC AAC CAC GGG CCG CCC 336
 1 Met Tyr Gln Ser Leu Ala Met Ala Ala Asn His Gly Pro Pro 14
 337 CCC GGT GCC TAC GAG GCG GGC GGC CCC GGC GCC TTC ATG CAC GGC GCG 384
 15 Pro Gly Ala Tyr Glu Ala Gly Gly Pro Gly Ala Phe Met His Gly Ala 30

	385	GGC GCC GCG TCC TCG CCA GTC TAC GTG CCC ACA CCG CGG GTG CCC TCC	432
	31	Gly Ala Ala Ser Ser Pro Val Tyr Val Pro Thr Pro Arg Val Pro Ser	46
5	433	TCC GTG CTG GGC CTG TCC TAC CTC CAG GGC GGA GGC GCG GGC TCT GCG	480
	47	Ser Val Leu Gly Leu Ser Tyr Leu Gln Gly Gly Gly Ala Gly Ser Ala	62
	481	TCC GGA GGC GCC TCG GGC GGC AGC TCC GGT GGG GCC GCG TCT GGT GCG	528
10	63	Ser Gly Gly Ala Ser Gly Gly Ser Ser Gly Gly Ala Ala Ser Gly Ala	78
	529	GGG CCC GGG ACC CAG CAG GGC AGC CCG GGA TGG AGC CAG GCG GGA GCC	576
	79	Gly Pro Gly Thr Gln Gln Gly Ser Pro Gly Trp Ser Gln Ala Gly Ala	94
	577	GAC GGA GCC GCT TAC ACC CCG CCG CCG GTG TCG CCG CGC TTC TCC TTC	624
15	95	Asp Gly Ala Ala Tyr Thr Pro Pro Pro Val Ser Pro Arg Phe Ser Phe	110
	625	CCG GGG ACC ACC GGG TCC CTG GCG GCC GCC GCC GCT GCC GCG GCC	672
	111	Pro Gly Thr Thr Gly Ser Leu Ala Ala Ala Ala Ala Ala Ala Ala	126
20	673	CGG GAA GCT GCG GCC TAC AGC AGT GGC GGC GGA GCG GCG GGT GCG GGC	720
	127	Arg Glu Ala Ala Ala Tyr Ser Ser Gly Gly Gly Ala Ala Gly Ala Gly	142
	721	CTG GCG GGC CGC GAG CAG TAC GGG CGC GCC GGC TTC GCG GGC TCC TAC	768
25	143	Leu Ala Gly Arg Glu Gln Tyr Gly Arg Ala Gly Phe Ala Gly Ser Tyr	158
	769	TCC AGC CCC TAC CCG GCT TAC ATG GCC GAC GTG GGC GCG TCC TGG GCC	816
	159	Ser Ser Pro Tyr Pro Ala Tyr Met Ala Asp Val Gly Ala Ser Trp Ala	174
	817	GCA GCC GCC GCC GCC TCC GCC GGC CCC TTC GAC AGC CCG GTC CTG CAC	864
30	175	Ala Ala Ala Ala Ala Ser Ala Gly Pro Phe Asp Ser Pro Val Leu His	190
	865	AGC CTG CCC GGC CGG GCC AAC CCG GCC GCC CGA CAC CCC AAT CTC GGT	912
	191	Ser Leu Pro Gly Arg Ala Asn Pro Ala Ala Arg His Pro Asn Leu Gly	206
35	913	GAG TAG GAG CGC GAG GGC TGG GGC CCG TGA GGG CCG GGG CAG GGG CCG	960
	207	Glu ***	208
	961	TCT TGA GCC CTG TCG AGG GCC TCT TGT TTT TCC ACC AAC GCC TTC GTT	1008
	1009	GGG CTG GGG ATG GTG CTT CAC TAC CTC GAG TTT CTA GGG AAG GCA GAA	1056
40	1057	GCC AGT GCG GGG CTG GCG ACA TCA CAG CCC CAG AAG ACC GGC TTC TGT	1104
	1105	GGA AGG GGC CGG GCC TGC CCG CCG GGG CCT CTT CTG AGA TGG TGT CAG	1152
	1153	GGT CGG AGT GCG GCC TCC CCG CCA TCC CAG ACA TCG ACC GTG GCC GCG	1200
	1201	CTG CGC TGT GGG TGA CGC GGG AGG ACA GCG GGC TCC CTG GAG AGC CGG	1248
	1249	GGG CAG CGG CCT GGG ATT TCC TCG TGG AAG GTG CTG GAG ATT GCT GAG	1296
45	1297	TTT CTG CGC CCC TTT CCT CCC CGC CCG CCC TCG GGC CTC CGC GAA	1344
	1345	CTG ATT ACA ATG GTT TGG ACC GCA GAC CTT CTG GGC CAT TTG GCG GCC	1392
	1393	CAG CTG GAG GAT CCC TCG GGG TAG CTG ATG ATT TTC CCG TCG GGG GTC	1440
	1441	TCA CAC CGA GAA CAA AGG AGG GAT GGA CAA AGG AGA CGC CGG GGA GAT	1488
	1489	GCG CGG AAC AGG AGC CGG CAC TGT GCG GGT GCC ACC CGG CCG AGC GCG	1536
50	1537	TGG GCG CAT CAT GCG GGC AGC GGG GGG GGG GGG CGC ACA CGC CCG GTC	1584
	1585	AGT GTC CGG GAA CAT AGG GAC CTC AAA CGG GCT TGT TCA TGA CAC CCG	1632
	1633	AGT TAA ATG GAG ACT TTG CAG TCG CTT GCA CGC GTG GAG CCT CCT CTT	1680
	1681	CTC GCG TGG GCC AGG GTT GGA AAT AAC CGT TGT GGT AGG TTC CAT GCA	1728
	1729	GTG TTT CCA TCG GAT GTC AGA CGG GGA GGG ACG GCA AAC CTG TCT CAA	1776
55	1777	CCT CCA CTG ATT CAC AAA TAA ACG CAG CGG GAT CTG AGA AGG GGC CTG	1824
	1825	AGT ACA CGG GCC GGG GGA GAA AGG GAA GTG GCA ACC CCT AGT TCA AAA	1872
	1873	TGC AAA CGA CCT CTG GAA TTT CGG GAA GAG ACG GAG GAG TGA GTT TGG	1920
	1921	ATT GAG CCC ACC CTG TGG GGG AGG GGA AGC CCA GGC TTG AGA AGC AAA	1968
	1969	GCT CGC GTT TAT TGA CCA CCT ACT AAG TGC TAA ATC CTT CTG CAT TTG	2016
60	2017	TTC TCA CTT TGA AGA GGA AAG GCT ACC CCT GCC AGG TTT CCA GTC TTG	2064
	2065	GGC TGG CTC CGA GAA GGG CCT AGG CTT TGA AGC GCT TTT AAA TTG TCC	2112
	2113	TCT GGC CCT GGG TGG CCA GGG AAG GTT CCC GGG GGT GCA GCC AAA TAC	2160
	2161	ACA TCG CCG GCA AAC TGA TTA ACC CTG AAA GTG GCG CCG AGG CCC GGT	2208
	2209	CTG TGG CTC TTG GTG TTC CCC TCC CCC TCA CCC CTC TTG GGG GAG CCA	2256
65	2257	GAC GGC CAC CCC CAG GGG AGG AGG GGC CCG GCC GAG CCT CGC CGC GAC	2304
	2305	TTT TGC GAG ATA GCG CGG CGA CAT GGC CAC ACA ATG GAG CCC GCA GGC	2352

2353	GGG AGT GCG GGG CGG GGC GCG GCG CCC TGG CCT TGC GCG CTT ACG GGG	2400
2401	TCC TCT CCA GGG CCC TCT GGG GCC TCT GAC TTA AAA TAG GGA GGA CTG	2448
2449	GGC CAG GAG ATC GAG ACC ATC CTG GCC AAC CTG TTG AAA CCC CGC CTC	2496
2497	TAC TAA AAT ACA AAA AAA AAA AAA AA	2525

5

5. PP10472

A: 核苷酸序列 (SEQ ID NO:13) 长度: 2787

1	GTCGGGGCTT TCTATGAGAG GCTGTGGCCA GAGAGGACCC TGGGCTGCCT GGACAGTCAG
61	GACCTGGCTG TGGCCGATGG CCACCAACAC TCAGATCCGG CCGGAGGGAC ACAAAGGGAC
10	121 ATGGGAGCGA GGTCAGGGG CTCACAGGGG AGGAACGTGA GATGCCGGAG GAATCACGCA
181	CACCAGCAGG TGCCGGCATT TCTAAAGGCG CCTGGGCAGC CTGGCTCATC CACCCTCAGG
241	AGCGAGCATC GCCTGACACA GGGCACTTAG TGTCCACCAC CAAGTACCCT GTTGTCCCCA
301	CCAACCCCAA GACACAAGAA GTCACCCAG AGCCAGCCCC AGAGCGCTGG GAGGGGGGCA
361	CTCACCCCTG TGTGAGGTAG GGGTGAGGGG TGTGGGGGGA GCCACTTCTT GCGTCCCCCG
15	421 CAGGCGGCGG GAGGCTGTGG AAGGGAGGCC GCGGGTAGCC GGGTTCCGCG AGTGTCTCAG
481	CCGCTCCTCT CGGTCCCTGC GCTCCCGCTC GGCGTCATCG GCGGCCCGGC TGGCACCCTG
541	AGGAGGCAGG AGGTGAGGCC GGAATGGAAC CAGCTTTGGC AGAAAGCAAC CCTAGGTCCT
601	AGGCTGCCCC GGCCACGTGA GAAAATACCT ACCAAAGGGC ACTGGGACGA GACTTTTGCC
661	TGCAAAGCCC CCCACAGGCT GCACGTTTCT CCTCATGGGA TGACAGCATC ATCCCCACAA
20	721 GGGGTACAGT CCAGGCCTCT CAGCTCCCC AACAATCGAA AAAACCCAC TCGCCATGGC
781	ATCTCCCTGT GGGCTCAGAA CTTCCTTAAA CTTCCAAATG GGCAGCAGCT ACCCCCTGCT
841	GCTCTGTGAA CATGCTCACT AACACGTGCA GAAAGAGGAG AGGGACAGGA ACCATCGGAG
901	GCCAAAGACA GGCCCCAAT GGTGACAGCT GGGGAGGAAA GCGGGGTGGG GATGAGAGCG
961	TTTGGTCGGA CGAGGAGTAC ACTCAGGGTC CGGACCCTGC AGCCTCCAAC AGCCTGTGCA
25	1021 GGTGTGGTGG CGTGCCAGC AGCCGGCAGC CGGATCTGTG CTGCACTATC CAGATGCACG
1081	GGGGTGGGGA GGTCAGATTT TAAGCCGAGA CAGCAACCTA TGAAAAAGCA GATTCTAACT
1141	GCCAATGCTG TGCGGCAGGG GCCTGCAAAC TACAGCCCCG GGGCCAAACC CAGGTGCCAC
1201	TTCTTCTGTC AAGTACAGCT CATTGGACAC AGCCACGTTT ACCAGGCAAC ACACAGACCG
1261	ACACACCTGC GGCTGCAGCA CCGTGCAGCC AGCACCGCCC AGACTCCTCA TGCTGCCATT
30	1321 TACGGAAAAC GCTGGCCTAG CCTGCTCTGA GGGCCGGGTG ACTCTAATGC CGCAGGAGGG
1381	TCGCTCTGAT TCCTCCCGAG GCAGCACAGC CCCGGCAGAG GGACTGATCT GCCCTGCCGA
1441	GTGCACCCCT CCCGGCACCT GGGCCCGGAG GCTCGGGCCT GCCTCGCCTG CCATCTCTCC
1501	GGTGCCTCCA AGTACTCCC ACCTGATGG GCACTTGGCA AGTAAGTCAA GATGAACAGA
1561	GTGCACCCCT CCCAGCATCT GGAGACCCCG CCCCTCCCTG GCCCTGCAC GACTACTGGA
35	1621 CAAGGCTGCC GTTTTGTAC CCTCCAAAGA GGCAGTTCTG CTCTCCATC AGACTGCTGC
1681	CTGGAAAAGC TCCAAATCT ATTTAAAACT GACAACAAA CAGTAAGTTT TTAATGAAGG
1741	AACACAACAG TTCCACAGGA GCTTCATTC TGACAGCGGT GTCTGAAAAG TAATTCAGCT
1801	CAGACATGGA CCCAGGGCTG GGCTGCCTGT TTTTAGAGGC TCCACTGCAT ACTTCACACC
1861	GCATCAAGAA AGCCACAGGG AAAATCAGCT CATGACAGGG TCAGTCAGGC TAGACGGGTA
40	1921 GCACCTCACC AGGCCCAAC CCCACTCATC TCGTGGGCAC CACGACCATG TGGCACAGCG
1981	AATGGGAATG CGCACCCCA GCCCCACCC TCTGCCCA GGGCATGCTC TGGGCCCTAG
2041	CTCCGTTTGG TGACAAGGAG GGAAGGGGCC TCCTGCTGTC CACACCCAGG AATTCCATGG
2101	ACCCCTCTGC GTTCCCTAAT GCGCTCTTAC TTCTTGGCAC CCCTTTCTGG CAGGCAGACA
2161	CCCACTCAGT CCAGGTCCTT GCCCACTACA CACTTGGCCT TCACAACCAG AGACTCCC
45	2221 TATATGGTAC AGCCCTCAA CAAGACGGC GCGGCCTCT CACTGACAAG CCATCCCCCG
2281	CACACCACAC TCAGCGAAGC TGCTGGCAGT GCGACTTACA AATTTGAGCA TGTCCAGTC
2341	GAACACGTAG TCATAGGAGA AGCCCTGGCG ATGGAACAGA TTCGGAAAA GCTGCCGCAG
2401	GTACGAGTAG TCAGGCTTGT CGTCAAAACG CAAGGAACGG CAGAAATTCA GGTATGTGGC
2461	AAATTCGGCT ACAAACAAG AAATCAAAAG CTAATCATG AAACCCAACT GCGACTCAAG
50	2521 GTGTCTCTGC CAACGTCGCT GTCTACCTCC TGCTGCTGCA CACTCAAGGG GAGAAGGACA
2581	GATGCAAAAC ACCTGTCAGA TTTCTAAGAC CTGAAGCCTT GAGAAAGCAT CGAAAAGTAT
2641	TCAAGTCACG GCCGGGTGG GCGGCTCAG CCTGTACCC CAGCACTTA GGAGGCTGAG
2701	GAGGGCGGAT CACGAGGTCA GGAGATCAAG ACCATCCTGG CTAACACAGT GAAACCCCAT
2761	CTCTACTGAA AAAAAAAAA AAAAAAA

55

B: 核苷酸序列 (SEQ ID NO:14) 长度: 162

1	MASYFLAPLS GRQTPTQSRS LPTTHLPFTT RDPYMVQPL NKTAAGLSLT SHPPHTTLE
61	RAGSATYKFE HVPVEHVVIG EALAMEQIPE KLPQVRVRL VVKTQGTAEI QVCGKFGYKT

121 RNSKLTHETQ LRLKVS LPTS LSTSCCTLK GRRTDAKHLS DF

C. 核苷酸及氨基酸组合序列 (SEQ ID NO:15) 克隆号: PP10472

起始编码子: 2119 ATG 终止编码子: 2605 TAA 蛋白质分子量: 17887.73

5	1	GTC GGG GCT TTC TAT GAG AGG CTG TGG CCA GAG AGG ACC CTG GGC TGC	48
	49	CTG GAC AGT CAG GAC CTG GCT GTG GCC GAT GGC CAC CAA CAC TCA GAT	96
	97	CCG GCC GGA GGG ACA CAA AGG GAC ATG GGA GCG AGG TCA AGG GGC TCA	144
	145	CAG GGG AGG AAC GTG AGA TGC GGG AGG AAT CAC GCA CAC CAG CAG GTG	192
	193	CCG GCA TTT CTA AAG GCG CCT GGC CAG CCT GGC TCA TCC ACC CTC AGG	240
10	241	AGC GAG CAT CGC CTG ACA CAG GGC ACT TAG TGT CCA CCA CCA AGT ACC	288
	289	CTG TTG TCC CCA CCA ACC CCA AGA CAC AAG AAG TCA CCC CAG AGC CAG	336
	337	CCC CAG AGC GCT GGG AGG GGG GCA CTC ACC CGT GTG TGA GGT AGG GGT	384
	385	GAG GGG TGT GGG GGG AGC CAC TTC CTG CGT CCC CCG CAG GCG GCC GGA	432
	433	GGC TGT GGA AGG GAG GCC GCG GGT AGC CGG GTT CCG CGA GTG TCT CAG	480
15	481	CCG CTC CTC TCG GTC CCT GCG CTC CCG CTC GGC GTC ATC GGC GGC CCG	528
	529	GCT GGC ACC CTG AGG AGG CAG GAG GTG AGG CCG GAA TGG AAC CAG CTT	576
	577	TGG CAG AAA GCA ACC CTA GGT CCT AGG CTG CCC CGG CCA CGT GAG AAA	624
	625	ATA CCT ACC AAA GGG CAC TGG GAC GAG ACT TTT GCC TGC AAA GCC CCC	672
	673	CAC AGG CTG CAC GTT TCT CCT CAT GGG ATG ACA GCA TCA TCC CCA CAA	720
20	721	GGG GTC AGA GCC AGG CCT CTC AGC TCC CCC AAC AAT CGA AAA AAC CCC	768
	769	ACT CGC CAT GGC ATC TCC CTG TGG GCT CAG AAC TTC CTT AAA CTT CCA	816
	817	AAT GGG CAG CAG CTA CCC CCT GCT GCT CTG TGA ACA TGC TCA CTA ACA	864
	865	CGT GCA GAA AGA GGA GAG GGA CAG GAA CCA TCG GAG GCC AAA GAC AGG	912
	913	CCC CAA ATG GTG ACA GCT GGG GAG GAA AGC GGG GTG GGG ATG AGA GCG	960
25	961	TTT GGT CGG ACG AGG AGT ACA CTC AGG GTC CGG ACC CTG CAG CCT CCA	1008
	1009	ACA GCC TGT GCA GGT GTG GTG GCG TGG CCA GCA GCC GGC AGC CGG ATC	1056
	1057	TGT GCT GCA CTA TCC AGA TGC ACG GGG GTG GGG AGG TCA GAT TTT AAG	1104
	1105	CCG AGA CAG CAA CCT ATG AAA AAG CAG ATT CTA ACT GCC AAT GCT GTG	1152
	1153	CGG CAG GGG CCT GCA AAC TAC AGC CCC GGG GCC AAA CCC AGG TGC CAC	1200
30	1201	TTC TTC CTG CAA GTA CAG CTC ATT GGA CAC AGC CAC GTT CAC CAG GCA	1248
	1249	ACA CAC AGA CCG ACA CAC CTG CCG CTG CAG CAC CGT GCG GCC AGC ACC	1296
	1297	GCC CAG ACT CCT CAT GCT GCC ATT TAC GGA AAA CGC TGG CCT AGC CTG	1344
	1345	CTC TGA GGG CCG GGT GAC TCT AAT GCC GCA GGA GGG TCG CTC TGA TTC	1392
	1393	CTC CCG AGG CAG CAC AGC CCC GGC AGA GGG ACT GAT CTG CCC TGC CGA	1440
35	1441	GTG CAC CCC TCC CCG CAC CTG GGC CCC GAG GCT CGG GCC TGC CTC GCC	1488
	1489	TGC CAT CTC TCC GGT GCC TCC AAG TAC TCC CCA CGC TGA TGG GCA CTT	1536
	1537	GGC AAG TAA GTC AAG ATG AAC AGA GTG CAC CCC TCC CAG CAT CTG GAG	1584
	1585	ACC CCG CCC CTC CCT GGC CCC TGC ACG ACT ACT GGA CAA GGC TGC CGT	1632
	1633	TTT TGT ACC CTC CCA AGA GGC AGT TCT GCT CCT CCA TCA GAC TGC TGC	1680
40	1681	CTG GAA AAG CTC CAA ATT CTA TTT AAA ACT GAC AAC AAA ACA GTA AGT	1728
	1729	TTT TAA TGA AGG AAC ACA ACA GTT CCA CAG GAG CTT CAT TTC TGA CAG	1776
	1777	CGG TGT CTG AAA AGT AAT TCA GCT CAG ACA TGG ACC CAG GGC TGG GCT	1824
	1825	GCC TGT TTT TAG AGG CTC CAC TGC ATA CTT CAC ACC GCA TCA AGA AAG	1872
	1873	CCA CAG GGA AAA TCA GCT CAT GAC AGG GTC AGT CAG GCT AGA CCG GTA	1920
45	1921	GCA CCT CAC CAG GCC CCA ACC CCA CTC ATC TCG TGG GCA CCA CGA CCA	1968
	1969	TGT GGC ACA GCG AAT GGG AAT GCG CAC CCC CAG CCC CCA CCC TCT GCC	2016
	2017	CCC AGG GCA TGC TCT GGG CCC TAG CTC CGC TTG GTG ACA GGG AGG GAA	2064
	2065	GGG GCC TCC TGC TGT CCA CAC CCA GGA ATT CCA TGG ACC CCT CTG CGT	2112
	2113	TCC CTA ATG GCG TCT TAC TTC TTG GCA CCC CTT TCT GGC AGG CAG ACA	2160
50	1	Met Ala Ser Tyr Phe Leu Ala Pro Leu Ser Gly Arg Gln Thr	14
	2161	CCC ACT CAG TCC AGG TCC TTG CCC ACT ACA CAC TTG CCC TTC ACA ACC	2208
	15	Pro Thr Gln Ser Arg Ser Leu Pro Thr Thr His Leu Pro Phe Thr Thr	30
55	2209	AGA GAC ACT CCC TAT ATG GTA CAG CCC CTC AAC AAG ACG GCC GCC GGC	2256
	31	Arg Asp Thr Pro Tyr Met Val Gln Pro Leu Asn Lys Thr Ala Ala Gly	46
	2257	CTC TCA CTG ACA AGC CAT CCC CCG CAC ACC ACA CTC AGC GAA CGT GCT	2304
60	47	Leu Ser Leu Thr Ser His Pro Pro His Thr Thr Leu Ser Glu Arg Ala	62
	2305	GGC AGT GCG ACT TAC AAA TTT GAG CAT GTT CCA GTC GAA CAC GTA GTC	2352

	63	Gly Ser Ala Thr Tyr Lys Phe Glu His Val Pro Val Glu His Val Val	78
	2353	ATA GGA GAA GCC CTG GCG ATG GAA CAG ATT CCG GAA AAG CTG CCG CAG	2400
5	79	Ile Gly Glu Ala Leu Ala Met Glu Gln Ile Pro Glu Lys Leu Pro Gln	94
	2401	GTA CGA GTA GTC AGG CTT GTC GTC AAA ACG CAA GGA ACG GCA GAA ATT	2448
	95	Val Arg Val Val Arg Leu Val Val Lys Thr Gln Gly Thr Ala Glu Ile	110
	2449	CAG GTA TGT GGC AAA TTC GGC TAC AAA ACA AGA AAC TCA AAG CTA ACT	2496
10	111	Gln Val Cys Gly Lys Phe Gly Tyr Lys Thr Arg Asn Ser Lys Leu Thr	126
	2497	CAT GAA ACC CAA CTG CGA CTC AAG GTG TCT CTG CCA ACG TCG CTG TCT	2544
	127	His Glu Thr Gln Leu Arg Leu Lys Val Ser Leu Pro Thr Ser Leu Ser	142
	2545	ACC TCC TGC TGC TGC ACA CTC AAG GGG AGA AGG ACA GAT GCA AAA CAC	2592
15	143	Thr Ser Cys Cys Cys Thr Leu Lys Gly Arg Arg Thr Asp Ala Lys His	158
	2593	CTG TCA GAT TTC TAA GAC CTG AAG CCT TGA GAA AGC ATC GAA AAG TAT	2640
20	159	Leu Ser Asp Phe ***	163
	2641	TCA AGT CAC GGC CGG GTG CGG CGG CTC ACG CCT GTC ACC CCA GCA CTC	2688
	2689	TAG GAG GCT GAG GAG GGC GGA TCA CGA GGT CAG GAG ATC AAG ACC ATC	2736
	2737	CTG GCT AAC ACA GTG AAA CCC CAT CTC TAC TGA AAA AAA AAA AAA	2784
25	2785	AAA	2787

6. PP11662

A: 核苷酸序列 (SEQ ID NO:16) 长度: 2067

	1	GTGGAAAATG GGCATCTCTC CCTCCCATGT TAAGCTTTAA CCTCTGTAAT CTGCCTGTAT
	61	CTATAGGTGG GCATCTCACT CCATCAAAGG AGCCACCT CTCTTTGTCC CTCTATCCAT
30	121	GCAACAGTCT TCTCTGTGCA TTTCCCAAG CTGGGCCCTC TTCTACTCTC CATTAGGCC
	181	TGTTGATAAC TCCATTACCC ACCCATCACT GCTGTTCCCTC CAGGGCCAGC ACTCGGGCGA
	241	GGCAGGGGAG CTGCCTTCGG TACATAATTT GAAGGGGCAC TCCCTCTTGG GCACATGCCG
	301	GCCCTGAGTG CCTCCCTTGC CTCACTCTGA TCCCAGCCCC ATAATGTCCT CAGTGGAAAGG
35	361	TGACTGGGGG CCGGTGCTGT GGGGAGAGTA GAAAGAGGGG TTGGCATGAC TAAAAATACC
	421	AGTATGTGTA TTAAGTATTT TGAGAATGAA ATGCCAAGGA GTGCCTACTG TATGCCAGCT
	481	CTGCTCTAGG AATGGAGTAG ACAGTGGACA CAAGAAGGAC TTACGCCCTG AGCACAGGTG
	541	CCAACAGTGA CAAGACTGGC AAGACGTGAG GGCATGAATG GTTCATTGAG GCAGCTGCTG
	601	CAGGTGTGGT CACCTGGTGC CATCTGCTGC TCCCTTTTCC ACTTTTCTAT GTCCTCCTTC
	661	CACCCCAAGT CCCGGATCAC TCGTGTGTTT CTGGCTAGCT CTTGGCATCT CCATCTGAGC
40	721	CTAAAGTTGC CCACTGGCAC CAATAGATTC TGTTTGACCT GCTGTGCCCA TGCTCATCTT
	781	TGTCGGGGCT TGGCTTGCCA GCGAGTGAGG CTGGCATGGG GTGTGAAAGT GGGGTCTCAG
	841	TGAGGGGTGA CAGCGTGTG GCAGTCTCA CAGCCCTCGC TCGCTCTCGG CGCCTCCTCT
	901	GCCTGGGCTC CCACTTTGGC GGCACCTGAG GAGCCCTTCA GCCCACCCTG GCACCTGTGG
	961	GAGCCCTTTT CTGGGCTGGC CAAGGCCAGA GCAGGCTCCC TCAGCTTGCA GGGAGGTGTG
45	1021	GATCGAGGAG GCGCGAAGCG GGAACCGGGG CTGCGTGCAG CGCTTGGGGG GGGAGGTGAG
	1081	TTCCGGGTGG GCATGGGGCT TGGCGGGCCT GCACTCGGAG CAGCCGGCCG GCCCTTCCGG
	1141	CCCCGGGCAG TGAGGGACTT GGCACCCGGG CCAGCGGGCT GCAGAGGGTG TACTGGGTCC
	1201	CCCAGCAGTG CCGGCCACCC TGGCTGTGTC TTGATTTCTC GCTGGGCCTT AGCTGCCTTC
	1261	CCGCGGGGCA GGGCTCGGGA CCTGCAGCCC GCCGTGCCCTG AGCCTCCCAC CCGCTCCATG
50	1321	GGCTCCTGTG GGGCCCGAGC CTCCCCGACG AGTACCACCC CCTGCTCCAG GGCGCCCAGT
	1381	CCCATCGACC ACCCAAGGGC TGAGGAGTGC GAGCGCACGG CCGGGACTG GCAGGCAGCT
	1441	CCACCTGCAG CCCCGGTGGC GGATCCACTA GGTGAAGCCA GCTGGGCTCC TGAGTCTGAT
	1501	GGGACGTGG AGAGTCCTTA TGTCCTGCTC AGGGATTGTA AACACACCAA TCAGCACCTT
	1561	GTGTTTAGCT CAAGGTTTGT GAGTGCACCA GTCGACACTC TGATCTAGC TGCTCTGGTG
55	1621	GGGCCTCGGA GAACCTTTAT ATCTAGCTCA GGGATTGTAA ATACACCCAT CGGCACTGTG
	1681	TATCTAGCTC AAGGTTTATA AACACACCAA TCAACACCCCT GTGTCTAGCT CAGGGTTTGT
	1741	GAGTGCACCA ATCAACACTC TGTATCTAGC TGCTCTGGTG GGGCCTTGGG GAACCTGTGT
	1801	GTCGAAACTC TGTATCTAAC TAATCTGATG GGGACGTGGA GAACCTTTGT ATCTAGCTCA
	1861	GGGATTGTAA ACGCACCAAT CAACGACCTG TCAAAAACAGG CCACTCGGCT CTACCAATCA
60	1921	GCAGGATGTG GGTGGGGCCA GATAAGAGAA TAAAAGCAGG CTGCCCGAGC CAGCATTGGC
	1981	AACTCGCTCG GGTCCCTTC CACGCTGTGG GAGCTTTGTT CTTTGCAATA AATCTTGCTA
	2041	CTGCTCAAAA AAAAAAAAAA AAAAAA

B: 核苷酸序列 (SEQ ID NO:17) 长度: 202

1 MGLGGPALGA AGRPFRPRAY RDLAPGPAGC RGCTGSPSSA GPPALCLISR WALAAFPGR
 5 61 ARDLQPAVPE PPTRSMGSCG ARASPTSTTP CSRAPSPIDH PRAEERCERTA RDWQAAPPAA
 121 PYRDPLGEAS WAPESDGDVE SPYVLLRDCK HTNQHPVFSS RFVSAPVDL YLAALVGP
 181 TFISSGIVN TPIGTVYLAQ GL

C. 核苷酸及氨基酸组合序列 (SEQ ID NO:18) 克隆号: PP11662

起始编码子: 1093 ATG 终止编码子: 1699 TAA 蛋白质分子量: 21042.59

10 1 GTG GAA AAT GGG CAT CTC TCC CTC CCA TGT TAA GCT TTA ACC TCT GTA 48
 49 ATC TGC CTG TAT CTA TAG GTG GGC ATC TCA CTC CAT CAA AGG AGC CCA 96
 97 GCC TCT CTT TGT CCC TCT ATC CAT GCA ACA GTC TTC TCT GTG CAT TTC 144
 145 CCC AAG CTG GGC CCT CTT CTA CTC TCC ATT TAG GCC TGT TGA TAA CTC 192
 193 CAT TAC CCA CCC ATC ACT GCT GTT CCT CCA GGG CCA GCA CTC GGG CGA 240
 15 241 GGC AGG GGA GCT GCC TTC GGT ACA TAA TTT GAA GGG GCA CTC CCT CTT 288
 289 GGC CAC ATG CCG GCC CTG AGT GCC TCC CTT GCC TCA CTC TGA TCC CAG 336
 337 CCC CAT AAT GTC CTC AGT GGA AGG TGA CTG GGG GCC GGT GCT GTG GGG 384
 385 AGA GTA GAA AGA GGG GTT GGC ATG ACT AAA AAT ACC AGT ATG TGT ATT 432
 433 AAG TAT TTT GAG AAT GAA ATG CCA AGG AGT GCC TAC TGT ATG CCA GCT 480
 20 481 CTG CTC TAG GAA TGG AGT AGA CAG TGG ACA CAA GAA GGA CTT ACG CCC 528
 529 TGA GCA CAG GTG CCA ACA GTG ACA AGA CTG GCA AGA CGT GAG GGC ATG 576
 577 AAT GGT TCA TTC AGG CAG CTG CTG CAG GTG TGG TCA CCT GGT GCC ATC 624
 625 TGC TGC TCC CTT TTC CAC TTT TCT ATG TCC TCC TTC CAC CCC AAG TCC 672
 673 CGG ATC ACT CGC TGT TTT CTG GCT AGC TCT TGG CAT CTC CAT CTG AGC 720
 25 721 CTA AAG TTG CCC ACT GGC ACC AAT AGA TTC TGT TTG ACC TGC TGT GCC 768
 769 CAT GCT CAT CTT TGT CGG GGC TTG GCT TGG CAG CGA GTG AGG CTG GCA 816
 817 TGG GGT GCT GAA GTG GGG TCT CAG TGA GGG GTG ACA GCG TGC TGG CAG 864
 865 TCC TCA CAG CCC TCG CTC GCT CTC GGC GCC TCC TCT GCC TGG GCT CCC 912
 913 ACT TTG GCG GCA CTT GAG GAG CCC TTC AGC CCA CCG CTG CAC TGT GGG 960
 30 961 GAG CCC CTT TCT GGG CTG GCC AAG GCC AGA GCA GGC TCC CTC AGC TTG 1008
 1009 CAG GGA GGT GTG GAT CGA GGA GGC GCG AAG CGG GAA CCG GGG CTG CGT 1056
 1057 GCA GCG CTT GCG GGC CAG CTG GAG TTC CGG GTG GGC ATG GGG CTT GGC 1104
 1 Met Gly Leu Gly 4
 35 1105 GGG CCT GCA CTC GGA GCA GCC GGC CGG CCC TTC CGG CCC CGG GCA GTG 1152
 5 Gly Pro Ala Leu Gly Ala Ala Gly Arg Pro Phe Arg Pro Arg Ala Val 20
 1153 AGG GAC TTG GCA CCC GGG CCA GCG GGC TGC AGA GGG TGT ACT GGG TCC 1200
 21 Arg Asp Leu Ala Pro Gly Pro Ala Gly Cys Arg Gly Cys Thr Gly Ser 36
 40 1201 CCC AGC AGT GCC GGC CCA CCT GCG CTG TGC TTG ATT TCT CGC TGG GCC 1248
 37 Pro Ser Ser Ala Gly Pro Pro Ala Leu Cys Leu Ile Ser Arg Trp Ala 52
 1249 TTA GCT GCC TTC CCG CGG GGC AGG GCT CGG GAC CTG CAG CCC GCC GTG 1296
 45 53 Leu Ala Ala Phe Pro Arg Gly Arg Ala Arg Asp Leu Gln Pro Ala Val 68
 1297 CCT GAG CCT CCC ACC CGC TCC ATG GGC TCC TGT GGG GCC CGA GCC TCC 1344
 69 Pro Glu Pro Pro Thr Arg Ser Met Gly Ser Cys Gly Ala Arg Ala Ser 84
 50 1345 CCG ACG AGT ACC ACC CCC TGC TCC AGG GCG CCC AGT CCC ATC GAC CAC 1392
 85 Pro Thr Ser Thr Thr Pro Cys Ser Arg Ala Pro Ser Pro Ile Asp His 100
 1393 CCA AGG GCT GAG GAG TGC GAG CGC ACG GCG CGG GAC TGG CAG GCA GCT 1440
 55 101 Pro Arg Ala Glu Glu Cys Glu Arg Thr Ala Arg Asp Trp Gln Ala Ala 116
 1441 CCA CCT GCA GCC CCG GTG CGG GAT CCA CTA GGT GAA GCC AGC TGG GCT 1488
 117 Pro Pro Ala Ala Pro Val Arg Asp Pro Leu Gly Glu Ala Ser Trp Ala 132
 1489 CCT GAG TCT GAT GGG GAC GTG GAG AGT CCT TAT GTC CTG CTC AGG GAT 1536
 60 133 Pro Glu Ser Asp Gly Asp Val Glu Ser Pro Tyr Val Leu Leu Arg Asp 148
 1537 TGT AAA CAC ACC AAT CAG CAC CCT GTG TTT AGC TCA AGG TTT GTG AGT 1584
 149 Cys Lys His Thr Asn Gln His Pro Val Phe Ser Ser Arg Phe Val Ser 164

	1585	GCA CCA GTC GAC ACT CTG TAT CTA GCT GCT CTG GTG GGG CCT CGG AGA	1632
	165	Ala Pro Val Asp Thr Leu Tyr Leu Ala Ala Leu Val Gly Pro Arg Arg	180
	1633	ACC TTT ATA TCT AGC TCA GGG ATT GTA AAT ACA CCC ATC GGC ACT GTG	1680
5	181	Thr Phe Ile Ser Ser Ser Gly Ile Val Asn Thr Pro Ile Gly Thr Val	196
	1681	TAT CTA GCT CAA GGT TTA TAA ACA CAC CAA TCA ACA CCC TGT GTC TAG	1728
	197	Tyr Leu Ala Gln Gly Leu ***	203
10	1729	CTC AGG GTT TGT GAG TGC ACC AAT CAA CAC TCT GTA TCT AGC TGC TCT	1776
	1777	GGT GGG GCC TTG GAG AAC CTG TGT GTC GAA ACT CTG TAT CTA ACT AAT	1824
	1825	CTG ATG GGG ACG TGG AGA ACC TTT GTA TCT AGC TCA GGG ATT GTA AAC	1872
	1873	GCA CCA ATC AAC GAC CTG TCA AAA CAG GCC ACT CGG CTC TAC CAA TCA	1920
	1921	GCA GGA TGT GGG TGG GGC CAG ATA AGA GAA TAA AAG CAG GCT GCC CGA	1968
15	1969	GCC AGC ATT GGC AAC TCG CTC GGG TCC CCT TCC ACG CTG TGG GAG CTT	2016
	2017	TGT TCT TTG CAA TAA ATC TTG CTA CTG CTC AAA AAA AAA AAA AAA AAA	2064
	2065	AAA	2067

7. PP11741

20	A:	核苷酸序列 (SEQ ID NO:19) 长度: 3222
	1	GTTGGGAAGG ACTTGGGGAC TCCACCTCTT GGCCTGGTTC CCATGGCCCC TAGGGGTCTA
	61	ACCTACCTGA TCCTGGGTTG GAGAAGCAGG ACCTTCTGGA AGCCCATGTT CCAGCGCTTG
	121	GGCCCCAGTC CTGATGGGGA TCAGGAATAG GCCCCCTTGC ACAGACCCCC ACCTGGCCAG
	181	CTCAGTGCCA GCAGACATCT TTCTACCTCA TGGTTCATCT ATGGAATTCT GGGGGTCCAC
25	241	TGGATCCCTC ATCTTCCCTG GCTATGCCTG CTTCTATGCC AGGCACACCA TCCTCCAAGT
	301	GACCACAGAC AGACCTGAGT GTATGTGGCT GCGGCACACA CACCCTGTG ACCAAGTGCA
	361	CTTCTGTGCA TTGGGGCACA GATTGTGTGT TCATCGTGTG AGTGCATGTG TGGGTCTGTG
	421	ACACGTGGCC GTTGCTCCTC TGTGAGCCAT CTGCTGATGA GCTACATCTC TGCCTTTACG
	481	CGTGACTTTG TATCCATGTC CAGGTGTCCA TGGCCTGTG TGTGCAGATG TGTCCCTGTCC
30	541	CTGGGTGTGT CTGTGCCTGT GTGTGCCTTG ATATCTGTGC CTGCCCTCTT ACACATAGGT
	601	GGGAATGCAG AGTGTGTATT TTTGTATGC ACCTGTACAC AGGCTGGGCG GGCAGTGGAG
	661	GATGCGTATG TGGGTGGCTG TGTCTGTATC TGCATTTGCA TGGTGTATTG GAGATTGGAG
	721	CTGTGTGTCT GTGCGTGTGT GGGTAGTGTG TACCGTGTGC ACATGTATGT GTGTGCCTGT
	781	GGACCAGCAC CTGTGTTGCC ACATTTGGGT GACGGTACAT CCATGCACTG GGTCTGCAGG
35	841	TGTATTTGCG AGTGCCTGTG TCTGTCTAAC AACTCTGTA GATGTCGCCG CCTGAATGAG
	901	AGCCAGAGCA GAGCTCTCCC CAGCCCTTCC CAAGTACTGT TCCCCTCTAC CGACGACTCC
	961	CCAGTTCTCT CCTTCCCTGA TGCAATGCAC GCCTAGTGGG CTACGTGTGC CAACCTCCA
	1021	GGCCTTCTCC TGCCACAGGC TCTGTCTCTG TCCCGTCGCT GTGCCCTCTG CCCCTGCTAA
	1081	CCCAGCCCTC CGTGCCCTGG ATGCGCCCGG ACATGGCCAC CCTGCATGTG TGTCCGCTG
40	1141	CCCTCCGCCC CCACCCTGCC CCTCAGATCC ACCTCAGAGG CCCTGGGTCT GGTTTTCACT
	1201	GCTGTCCCTC CCTGGTCCCT GCCCTCCACT CCCACCCCTT GGTCCCTTTT CTCCTTCTCT
	1261	CTCTCGTTTC AGAAAAGAGG CCAGCTCTGC AGCCCCCTCG GGACGCCCC ACCAGCTCAA
	1321	ATTCGGAGTG CAGAAAAGAA ACCGGACCCC CCAGTGAGGC CTGCCAGGCC TCCCGGACCC
	1381	CTTGTTACTC AGGAACCTCA CCTGGACGG AATGGGATGG GGGCTTCGGT GCCCAACAAC
45	1441	CCCCACCTC CACTCTGCCA TTCCGGCCCA CCTCCCTCTG GCCGGACAGA ACTGGTGCTC
	1501	TCTTCTCCC ACTGTGCCCC TCCGCGGACC GGGGACCCTT CCCCGTGCC TACCCCTCC
	1561	CATTTTGATG GTGTCTGTGA CATTTCTGT TCGGAAGTAA AAGAGGGACC CCTGCGTCTC
	1621	GTCCTTTCT CTTGCAGTCT GCTTGTCCAT GTGTCTATT TTCCACTTGT CCATCTGTCC
	1681	AGTAAGGAGG TAATGGTGAC CCATGCTGGT CCGCTCGGAC CCCCATCCTC CTCTCCCCAC
50	1741	AGCTGGGCCC TGCTCAGCTG GTCTGTCTGG ATGCAGTGCC CCTTTGGACC CCCTTGTCGC
	1801	CTGGCCTCCC ACTCCCGGCC ACCTGGCCTG GCAGCTCTGC AGCTGTTGCT CTGTCCCGGT
	1861	GGGAGAAAGT CCAGCCAGGC CTCCACTTCC CCTGGGGCAT CCCAACCTCA GCCCTGCACA
	1921	GACCCAGCGT CTGCCCTTAT TTGCTCCCT GCAGACAAGG ACGAGTGCTC CAAGGATAAC
	1981	GGCGGCTGCC AGCAGGACTG CGTCAACAG TTCCGGCAGT ATGAGTGCCA ATGCCGAGT
55	2041	GGCTTCTGTC CTCCATGACA ACAAGCACGA CTGCAAAGAG GTACGGGCTG CATGCCAGGG
	2101	GCATCTGGGC TTGAAGGACC TGCTGCTGGG AGTGGAAGCC CAGAGGATCA TCTCCATCAG
	2161	CCTCCGTTTA GATTCTGGGG CCTGTATGAT CTCTAAGGCC CCCTCATTCC CCCACCCACC
	2221	TGGTACATAG ATCCTCCTCC CGCAGTGTG GAGCCACAG CTACTIONTGT TTGCCACAT
	2281	TAGGAGAGGA GGAGCTCACG GCTGCTCTGA GGCTTCATTT TCAGACTTTT CTAGTTTGTG
60	2341	CTTCTTAAAA CCATGTCCAA GTCTCTTGT GACTTCTTAC CCGCACAGAG GCGAGAAAGG

2401 ATAGTGGGTA AGAGCACAGC CTCCAGAGCC AGGTGGCCTG GCCCAGCCCT GCCGCTTCAC
 2461 TAGCTGTGTG CCTTTGAGTG ACTCCTTGAC CCTCTCTGTG CCTTGGTTTG CTCCTGGGTA
 2521 AAATGTAGAT AATAATAATA GTACCCACCT CAGAGAGCAT TGGGAGCTTG AAGGGCCCCA
 2581 AGCAGGGCCT GGCTCACCGC AAGCACAATG GCATCTGTTA CCACTATCAG TCACACTTAT
 5 2641 AACCCCTGGG GATGGTCAGA ACAGGGCCAC TTGCTTCCTG GCAGTGAAGC CCTGCACACA
 2701 TTTAAAAATA GCTCTAGGCC GGGCAGCATG GCTCATGCCT GTAATCCCAG CACTTTGGGA
 2761 GGCCGAGGCA AGGAGGATCG CTTGAGCTCA GGCATTCAAG ACCAGCCTGA GCAACATAGT
 2821 GAGACTCTGT CTCTATTTTA CTTTTAAACA GGTITTTAAA CTAGAAATAA AGTAAAAATAG
 2881 GCCAGTGGCG TGGTTCACAC CTGTAATCCC AGCACTTTGG GAGGCTGAGG TGGGCATGAT
 10 2941 TGGCTCAGCC CAGGAGTTTA AGACCAACCT GGGCAACATG GCGAAACCCC ATCTCTACAA
 3001 TAAATACAAA AAATTAGCTG GATGTGGTGG TGCATGCCTA TAGTCCCAGC TAGGATCGAG
 3061 GCTGAGGTGG GAGAATCACC TGAGTCTGGG GAGGCTGAGG CTGCATGAGC CGAGATCACA
 3121 CCACTGTACT CCAGCCTGGG TGACAGAGTG GGAATAAACA AATAAAAAATA TAAATAAATA
 3181 AATAAAGCAA AATAAAACCA AAAAAAAAAA AAAAAAAAAA AA

15

B: 核苷酸序列 (SEQ ID NO:20) 长度: 310

1 MCVRLPSRPT LPLRSTSEAL GLVFTAVPPW SLPSTPTPWS LFSFSLSFQK RGQLCSPLGT
 61 PPPAQIPSAE KKPDPVVRPA RPPGPLVTQE PHLGRNGMGA SVPTNPPPPL CHSGPPPSGR
 121 TELVLSSPHC ARPRTGDPSP CPTPSHFDGV CDISCCVKE GPLRPAPFSC SLLVHVSIFP
 181 LVHLSSEKVM VTHAGPLGPP SSSPHSWALL SWSWMCQCPF GPPCRLASHS RPPGLAALQL
 20 241 LLCPGGRKSS QASTSPGASQ PQPCTDPASA LICSPADKDE CSKDNNGCQQ DCVNTFGSYE
 301 CQCRSGFCPP

20

C. 核苷酸及氨基酸组合序列 (SEQ ID NO:21) 克隆号: PP11741

25 起始编码子: 1126 ATG 终止编码子: 2056 TGA 蛋白质分子量: 32605.64
 1 GTT GGG AAG GAC TTG GGG ACT CCA CCT CTT GGC CTG GTT CCC ATG GCC 48
 49 CCT AGG GGT CTA ACC TAC CTG ATC CTG GGT TGG AGA AGC AGG ACC TTC 96
 97 TGG AAG CCC ATG TTC CAG CGC TTG GGC CCC AGT CCT GAT GGG GAT CAG 144
 145 GAA TAG GCC CCC TTG CAG AGA CCC CCA CCT GGC CAG CTC AGT GCC AGC 192
 30 193 AGA CAT CTT TCT ACC TCA TGG TTC ATC TAT GGA ATT CTG GGG GTC CAC 240
 241 TGG ATC CCT CAT CTT CCC TGG CTA TGC CTG CTT CTA TGC CAG GCA CAC 288
 289 CAT CCT CCA AGT GAC CAC AGA CAG ACC TGA GTG TAT GTG GCT GCG GCA 336
 337 CAC ACA CCC CTG TGA CCA AGT GCA CTT CTG TGC ATT GGG GCA CAG ATT 384
 385 GTG TGT TCA TCG TGT GAG TGC ATG TGT GGG TCT GTG ACA CGT GGC CGT 432
 35 433 TGC TCC TCT GTG AGC CAT CTG CTG ATG AGC TAC ATC TCT GCC TTT ACG 480
 481 CGT GAC TTT GTA TCC ATG TCC AGG TGT CCA TGC GCC TGT GTG TGC AGA 528
 529 TGT GTC CTG TCC CTG GGT GTG TCT GTG CCT GTG TGT GCG TTG ATA TCT 576
 577 GTG CCT GCC TCT TCA CAC ATA GGT GGG AAT GCA GAG TGT GTA TTC TTT 624
 625 GTA TGC ACC TGT ACA CAG GCT GGG CGG GCA AGT GAG GAT GCG TAT GTG 672
 40 673 GGT GGC TGT GTC TGT ATC TGC ATT TGC ATG GTG TAT TGG AGA TTG GAG 720
 721 CTG TGT GTC TGT GCG TGT GTG GGT AGT GTG TAC CGT GTG CAC ATG TAT 768
 769 GTG TGT GCC TGT GGA CCA GCA CCT GTG TTG CCA CAT TTG GGT GAC GGT 816
 817 ACA TCC ATG CAC TGG GTC TGC AGG TGT ATT TGC GAG TGC GTG TGT CTG 864
 865 TCT AAC ACA CTC CCT TGT AGA TGT CGC CGC CTG AAT GAG AGC CAG AGC AGA 912
 45 913 GCT CTC CCC AGC CCT TCC CAA GTA CTG TTC CCC TCT ACC GAC GAC TCC 960
 961 CCA GTT CTC TCC TTC CCT GAT GCA ATG CAC GCC TAG TGG GCT ACG TGT 1008
 1009 GCC AAC CCT CCA GGC CTT CTC CTG CCA CAG GCT CTG TCT CTG TCC CGT 1056
 1057 CGC TGT GCC TCC TGC CCC TGC TAA CCC AGC CCT CCG TGC CCT GGA TGC 1104
 1105 GCC CGG ACA TGG CCA CCC TGC ATG TGT GTC CGC CTG CCC TCC CGC CCC 1152
 50 1 Met Cys Val Arg Leu Pro Ser Arg Pro 9
 1153 ACC CTG CCC CTC AGA TCC ACC TCA GAG GCC CTG GGT CTG GTT TTC ACT 1200
 10 Thr Leu Pro Leu Arg Ser Thr Ser Glu Ala Leu Gly Leu Val Phe Thr 25
 55 1201 GCT GTC CCT CCC TGG TCC CTG CCC TCC ACT CCC ACC CCT TGG TCC CTT 1248
 26 Ala Val Pro Pro Trp Ser Leu Pro Ser Thr Pro Thr Pro Trp Ser Leu 41
 1249 TTC TCC TTC TCT CTC TCG TTT CAG AAA AGA GGC CAG CTC TGC AGC CCC 1296
 60 42 Phe Ser Phe Ser Leu Ser Phe Gln Lys Arg Gly Gln Leu Cys Ser Pro 57

	1297	CTC GGG ACG CCC CCA CCA GCT CAA ATT CCG AGT GCA GAA AAG AAA CCG	1344
	58	Leu Gly Thr Pro Pro Pro Ala Gln Ile Pro Ser Ala Glu Lys Lys Pro	73
5	1345	GAC CCC CCA GTG AGG CCT GCC AGG CCT CCC GGA CCC CTT GTT ACT CAG	1392
	74	Asp Pro Pro Val Arg Pro Ala Arg Pro Pro Gly Pro Leu Val Thr Gln	89
	1393	GAA CCT CAC CTT GGA CGG AAT GGG ATG GGG GCT TCG GTG CCC ACC AAC	1440
	90	Glu Pro His Leu Gly Arg Asn Gly Met Gly Ala Ser Val Pro Thr Asn	105
10	1441	CCC CCA CCT CCA CTC TGC CAT TCC GGC CCA CCT CCC TCT GGC CGG ACA	1488
	106	Pro Pro Pro Pro Leu Cys His Ser Gly Pro Pro Pro Ser Gly Arg Thr	121
	1489	GAA CTG GTG CTC TCT TCT CCC CAC TGT GCC CGT CCG CGG ACC GGG GAC	1536
15	122	Glu Leu Val Leu Ser Ser Pro His Cys Ala Arg Pro Arg Thr Gly Asp	137
	1537	CCT TCC CCG TGC CCT ACC CCC TCC CAT TTT GAT GGT GTC TGT GAC ATT	1584
	138	Pro Ser Pro Cys Pro Thr Pro Ser His Phe Asp Gly Val Cys Asp Ile	153
20	1585	TCC TGT TGC GAA GTA AAA GAG GGA CCC CTG CGT CCT GCT CCT TTC TCT	1632
	154	Ser Cys Cys Glu Val Lys Glu Gly Pro Leu Arg Pro Ala Pro Phe Ser	169
	1633	TGC AGT CTG CTT GTC CAT GTG TCT ATT TTT CCA CTT GTC CAT CTG TCC	1680
	170	Cys Ser Leu Leu Val His Val Ser Ile Phe Pro Leu Val His Leu Ser	185
25	1681	AGT AAG GAG GTA ATG GTG ACC CAT GCT GGT CCG CTC GGA CCC CCA TCC	1728
	186	Ser Lys Glu Val Met Val Thr His Ala Gly Pro Leu Gly Pro Pro Ser	201
	1729	TCC TCT CCC CAC AGC TGG GCC CTG CTC AGC TGG TCC TGC TGG ATG CAG	1776
30	202	Ser Ser Pro His Ser Trp Ala Leu Leu Ser Trp Ser Cys Trp Met Gln	217
	1777	TGC CCC TTT GGA CCC CCT TGT CGC CTG GCC TCC CAC TCC CGG CCA CCT	1824
	218	Cys Pro Phe Gly Pro Pro Cys Arg Leu Ala Ser His Ser Arg Pro Pro	233
35	1825	GGC CTG GCA GCT CTG CAG CTG TTG CTC TGT CCC GGT GGG AGA AAG TCC	1872
	234	Gly Leu Ala Ala Leu Gln Leu Leu Leu Cys Pro Gly Gly Arg Lys Ser	249
	1873	AGC CAG GCC TCC ACT TCC CCT GGG GCA TCC CAA CCT CAG CCC TGC ACA	1920
	250	Ser Gln Ala Ser Thr Ser Pro Gly Ala Ser Gln Pro Gln Pro Cys Thr	265
40	1921	GAC CCA GCG TCT GCC CTT ATT TGC TCC CCT GCA GAC AAG GAC GAG TGC	1968
	266	Asp Pro Ala Ser Ala Leu Ile Cys Ser Pro Ala Asp Lys Asp Glu Cys	281
	1969	TCC AAG GAT AAC GGC GGC TGC CAG CAG GAC TGC GTC AAC ACG TTC GGC	2016
45	282	Ser Lys Asp Asn Gly Gly Cys Gln Gln Asp Cys Val Asn Thr Phe Gly	297
	2017	AGT TAT GAG TGC CAA TGC CGC AGT GGC TTC TGT CCT CCA TGA CAA CAA	2064
	298	Ser Tyr Glu Cys Gln Cys Arg Ser Gly Phe Cys Pro Pro ***	311
50	2065	GCA CGA CTG CAA AGA GGT ACG GGC TGC ATG CCA GGG GCA TCT GGG CTT	2112
	2113	GAA GGA CCT GCT GCT GGG AGT GGA AGC CCA GAG GAT CAT CTC CAT CAG	2160
	2161	CCT CCG TTT AGA TTC TGG GGC CTG TAT GAT CTC TAA GGC CCC CTC ATT	2208
	2209	CCC CCA CCC ACC TGG TAC ATA GAT CCT CCT CCC GCA GTG TCG GAG CCA	2256
	2257	CAC GCT CAC TTC TGT TGC CCA CAT TAG GAG AGG AGG AGC TCA CGG CTG	2304
55	2305	CTC TGA GGC TTC ATT TTC AGA CTT TTC TAG TTT GTG CTT CTT AAA ACC	2352
	2353	ATG TCC AAG TCT CTT TGT GAC TTC CTA CCG GCA CAG AGG CGA GAA AGG	2400
	2401	ATA GTG GGT AAG AGC ACA GCC TCC AGA GCC AGG TGG CCT GGC CCA GCC	2448
	2449	CTG CCG CTT CAC TAG CTG TGT GCC TTT GAG TGA CTC CTT GAC CCT CTC	2496
	2497	TGT GCC TTG GTT TGC TCC TGG GTA AAA TGT AGA TAA TAA TAA TAG TAC	2544
	2545	CCA CCT CAG AGA GCA TTG GGA GCT TGA AGG GCC CCA AGC AGG GCC TGG	2592
60	2593	CTC ACC GCA AGC ACA ATG GCA TCT GTT ACC ACT ATC AGT CAC ACT TAT	2640
	2641	AAC CCC TGG GGA TGG TCA GAA CAG GGC CAC TTG CTT CCT GGC AGT GAA	2688
	2689	GCC CTG CAC ACA TTT AAA AAT AGC TCT AGG CCG GGC ACG ATG GCT CAT	2736
	2737	GCC TGT AAT CCC AGC ACT TTG GGA GGC CGA GGC AAG GAG GAT CGC TTG	2784
	2785	AGC TCA GGC ATT CAA GAC CAG CCT GAG CAA CAT AGT GAG ACT CTG TCT	2832
65	2833	CTA TTT TAC TTT TAA ACA GGT TTT TAA ACT AGA AAT AAA GTA AAA TAG	2880
	2881	GCC AGT GCG GTG GTT CAC ACC TGT AAT CCC AGC ACT TTG GGA GGC TGA	2928
	2929	GGT GGG CAT GAT TGC GTC AGC CCA GGA GTT TAA GAC CAA CCT GGG CAA	2976

2977 CAT GGC GAA ACC CCA TCT CTA CAA TAA ATA CAA AAA ATT AGC TGG ATG 3024
 3025 TGG TGG TGC ATG CCT ATA GTC CCA GCT ACC CTG GAG GCT GAG GTG GGA 3072
 3073 GAA TCA CCT GAG TCT GGG GAG GCT GAG GCT GCA TGA GCC GAG ATC ACA 3120
 3121 CCA CTG TAC TCC AGC CTG GGT GAC AGA GTG GGA AAA AAC AAA TAA AAA 3168
 5 3169 TAT AAA TAA ATA AAT AAA GCA AAA TAA AAC CAA AAA AAA AAA AAA AAA 3216
 3217 AAA AAA 3222

8. PP12301

A: 核苷酸序列 (SEQ ID NO:22) 长度: 1404

10 1 GCCGGATTGC AGCACCTGGG ATTGGAGTTT GGAATCCTGC CTTCGATGTC ACCCCCCACG
 61 ACCTCATCAC TGGTGGCATC ATCACAGAAC TGGGGGTCTT TGCCCCTGAG GAGCTCCGGA
 121 CAGCCCTAAC CACCACCATC TCTTCCAGGG ATGGAACCCT AGATGGACCC CAGATGTAAC
 181 CAACTCAGCT CTCCCTAGCC TGCCTCTCTA GGTTTTTCAA TACATTTCTT GAATGGCTAC
 241 CCAAAAGCTG ACCGTCCAGC CCCTGACCAC ACTTGTTCCT AGTGCAGGGA GCTCAGACAG
 15 301 GGCCTTCCAT CTAGAGCCCA GCACCTAGAG CCAGGCTGCC CAGATTCAA TCCTGACTCC
 361 GCCACTTTTC CCAAAGTGCT GGGACTACAG GCATGAGCCA CTGTGCCTGA CCTTACAGCA
 421 GTATTTTITA AAAATCAAAA TTAATGCAAA AATCCATGAT GAGGCCAGGC TTGGTGGCTC
 481 ATGCCTGTGA TTCCAGCACT TTGGGAGGAT CCCTTGAGCC CAGGTGTTG AGACCAGTTT
 541 GGGTAACATA GGAAGAGCTT GTCTCTACAA AAAAAATTT AAAACAATG AGCTGGGCAT
 20 601 GGAATCACA CCTCTAGCCC CAGCTACTCG GGCTGAGGTG GGAGGATGGC TTGAGCCCAG
 661 GAGTTCAAGG ATGCAGTGTG CTTTGATTGC ACTACTGCGC TTCACCCTGG GCAACAGAGT
 721 GAGACCCCAT CTCTTAAAA AAAAAAAAAA ATCCATGATG AACAAAAACA GTATTTGTTT
 781 GAGACGGGT CTCGCTCTGT CGCCAGGCT GGAGTGCAGT GGTGCCATCT CAGCTCACTG
 841 CAACCTCCGC CTTCTGGGTT CAAGCGACTC TCCTGCCTCA GCCTCCCGAG TAGCTGGGAT
 25 901 TGCAGGAGCC TGGCACTATG CCTGGCTAAT TTTTGTATTT TTATTGGAGA CACCATGTTG
 961 GTCGTGGCTG GTCTCCAACCT CCTGACCTCG GGTGATCCGC GCCCCGCGC CTCGAAAGT
 1021 GCTGGGATTA CAGGCGTGAG CCACCGTGCC CGGCCTAAAT AAAATTTTGA AGAGGCTAGA
 1081 ACCCCGCACT TGTGCCTTGA GCTTACTGAC CTCAACACCC TGGTTCCACT AAAACTTTAT
 1141 TTACAAAATT ATGCTGCCGG TCTGCAGGAT GTAGTTAAC AATTTGATTA TTTTGGATAT
 30 1201 TAAAAATTT TTTAAGTCTT GAAAAATTT TTTACTATTA CCAAAGTTT TGAACCTTCT
 1261 TAAATCTGT ACCAGAGGTG AGTGCCACAC CCTAACCTAG TCTGAGTTT TATGTGTA
 1321 GATCATAAAA CGTTGTAAGT TTTCTAAAAC ACATAAGCTC TCAATAAACG TTAGCTTATC
 1381 ATTGAAAAA AAAAAAAAAA AAAA

35 B: 核苷酸序列 (SEQ ID NO:23) 长度: 128

1 MELTPLAPAT RAEVGGWLEP RSSRMQCALI ALLRFTLGNR VRPHLFKFKK KSMMNKTSIC
 61 LRRGLALSPR LECSGAISAH CNLRLLGSSD SPASASRVAG IAGAWHYAWL IFVFLLETPC
 121 WSWLVNS

40 C. 核苷酸及氨基酸组合序列 (SEQ ID NO:24) 克隆号: PP12301

起始编码子: 599 ATG 终止编码子: 983 TGA 蛋白质分子量: 14138.86

1 G CCG GAT TGC AGC ACC TGG GAT TGG AGT TTG GAA TCC TGC CTT CGA 46
 47 TGT CAC CCC CCA CGA CCT CAT CAC TGG TGG CAT CAT CAC AGA ACT GGG 94
 95 GGT CTT TGC CCC TGA GGA GCT CCG GAC AGC CCT AAC CAC CAC CAT CTC 142
 45 143 TTC CAG GGA TGG AAC CCT AGA TGG ACC CCA GAT GTA ACC AAC TCA GCT 190
 191 CTC CCT AGC CTG CCT CTC TAG GTT TTT CAA TAC ATT TCT TGA ATG GCT 238
 239 ACC CAA AAG CTG ACC GTC CAG CCC CTG ACC ACA CTT GTT CCT AGT GCA 286
 287 GGG AGC TCA GAC AGG GCC TTC CAT CTA GAG CCC AGC ACC TAG AGC CAG 334
 335 GCT GCC CAG ATT CAA ATC CTG ACT CCG CCA CTT TTC CCA AAG TGC TGG 382
 50 383 GAC TAC AGG CAT GAG CCA CTG TGC CTG ACC TTA CAG CAG TAT TTT TTA 430
 431 AAA ATC AAA ATT AAT GCA AAA ATC CAT GAT GAG GCC AGG CTT GGT GGC 478
 479 TCA TGC CTG TGA TTC CAG CAC TTT GGG AGG ATC CCT TGA GCC CAG GTG 526
 527 TTT GAG ACC AGT TTG GGT AAC ATA GGA AGA GCT TGT CTC TAC AAA CAA 574
 575 AAA TTT AAA AAC AAT GAG CTG GGC ATG GAA CTC ACA CCT CTA GCC CCA 622
 55 1 Met Glu Leu Thr Pro Leu Ala Pro 8
 623 GCT ACT CGG GCT GAG GTG GGA GGA TGG CTT GAG CCC AGG AGT TCA AGG 670
 9 Ala Thr Arg Ala Glu Val Gly Gly Trp Leu Glu Pro Arg Ser Ser Arg 24

	671	ATG CAG TGT GCT TTG ATT GCA CTA CTG CGC TTC ACC CTG GGC AAC AGA	718
	25	Met Gln Cys Ala Leu Ile Ala Leu Leu Arg Phe Thr Leu Gly Asn Arg	40
5	719	GTG AGA CCC CAT CTC TTT AAA AAA AAA AAA TCC ATG ATG AAC AAA	766
	41	Val Arg Pro His Leu Phe Lys Lys Lys Lys Ser Met Met Asn Lys	56
	767	ACA AGT ATT TGT TTG AGA CGG GGT CTC GCT CTG TCG CCC AGG CTG GAG	814
	57	Thr Ser Ile Cys Leu Arg Arg Gly Leu Ala Leu Ser Pro Arg Leu Glu	72
10	815	TGC AGT GGT GCC ATC TCA GCT CAC TGC AAC CTC CGC CTT CTG GGT TCA	862
	73	Cys Ser Gly Ala Ile Ser Ala His Cys Asn Leu Arg Leu Leu Gly Ser	88
	863	AGC GAC TCT CCT GCC TCA GCC TCC CGA GTA GCT GGG ATT GCA GGA GCC	910
15	89	Ser Asp Ser Pro Ala Ser Ala Ser Arg Val Ala Gly Ile Ala Gly Ala	104
	911	TGG CAC TAT GCC TGG CTA ATT TTT GTA TTT TTA TTG GAG ACA CCA TGT	958
	105	Trp His Tyr Ala Trp Leu Ile Phe Val Phe Leu Leu Glu Thr Pro Cys	120
	959	TGG TCG TGG CTG GTC TCC AAC TCC TGA CCT CGG GTG ATC CGC GCC CCG	1006
20	121	Trp Ser Trp Leu Val Ser Asn Ser ***	129
	1007	CGG CCT CCG AAA GTG CTG GGA TTA CAG GCG TGA GCC ACC GTG CCC GGC	1054
	1055	CTA AAT AAA ATT TTG AAG AGG CTA GAA CCC CGC ACT TGT GCC TTG AGC	1102
	1103	TTA CTG ACC TCA ACA CCC TGG TTC CAC TAA AAC TTT ATT TAC AAA ATT	1150
25	1151	ATG CTG CCG GTC TGC AGG ATG TAG TTT AAC AAT TTG ATT ATT TTG GAT	1198
	1199	ATT AAA ATA TTT TTT AAG TCT TGA AAA TAT TGT TTA CTA TTA CCA AAG	1246
	1247	TTT TTG AAC CTT CTT AAA TTC TGT ACC AGA GGT GAG TGC CAC ACC CTA	1294
	1295	ACC TAG TCT GAG TTT TTA TGT GTA AAG ATC ATA AAA CGT TGT AAG TTT	1342
	1343	TCT AAA ACA CAT AAG CTC TCA ATA AAC GTT AGC TTA TCA TTG AAA AAA	1390
30	1391	AAA AAA AAA AAA AA	1404

9. PP12616

A: 核苷酸序列 (SEQ ID NO:25) 长度: 993

	1	GCCGCAGTCC CATCATTCAG TTCCTAGGG TCACCGGCGC GGCAGTGGCC TCGCAGGGCG	
35	61	CTGGGTCCCT CTCCCAGCT CTCCTCCCC TGGCCCCGTC GCCCGCCCT CGCCGGGCTG	
	121	GGCTGCGGG TCAGGGGCG AGCGGAGAGG GGTGAGTATT CCCACAGCC CTTGCCGTT	
	181	GCCTCCTCCC GGCTCTGCTT CCCACACGGT CCTTGCCCA CTCCTAGGAC AGGGAGGAAG	
	241	GGCAGCGCG GGTAGGCGG AACAGCCA GTCTGAACA AAAGCGGG GAAGCGGGTC	
	301	CCCGCCGTA ACTGCAGGCC TGTGCTGGCC GCCGCGAGCG GGAAGGGCGA GGACACTCCC	
40	361	TCCTCGGGA CCCGGTCCC GTCGCGACA CGGTGTCACC AGCACACCTG GCCCAGTACC	
	421	CAAAGCACCC TCGAATTATC ATTTAACATG GAAGAAGATG AGTTCATTGG AGAAAAACA	
	481	TTCCAACGTT ATTGTGCAGA ATTCATTAAC CATTACAAC AGATAGGTGA TAGTTGGGAA	
	541	TGGAGACCAT CAAAGGACTG TTCTGATGGC TACATGTGCA AAATACACTT TCAAATTAAG	
45	601	AATGGGTCTG TGATGTGACA TCTAGGAGCA TCTACCCATG GACAGACATG TCTTCCCATG	
	661	GAGGTGAAGT CTGCTCTGT CACCCAGGCT GGAGTGCAGT TGCGTGATCT CAGTCACTG	
	721	CAACCTCCGC CTTCTGGTT CAAGCAGTTC TCCTGCCTCA GCCTCCGAG TAACTGGGAC	
	781	TACAGGGTT CACCACTACA CCTGGCTAAT TTTTGTATT TTAGTAGAG ATAAGGTTTC	
	841	ACCATGTTGG TCAGGCTGGT CTTGAACTCC TGACATCAAG TAATCCATCC GCCTCAGCCT	
	901	CCCAAAGTGC TGAGATTACA GGCATGAGCC ACTGTGCCTG GCCCTGAGCT TAAAATAAAA	
50	961	GTAAATTAA AAAAAAAAAA AAAAAAAAAA AAA	

B: 核苷酸序列 (SEQ ID NO:26) 长度: 125

	1	MEDEFI G E K T F Q R Y C A E F I K H S Q Q I G D S W E W R P S K D C S D G Y M C K I H F Q I K N G S V M S H L G	
55	61	A S T H G Q T C L P M E V K S C S V T Q A G V Q L R D L S S L Q P P S G F K Q F S C L S L P S N W D Y R G S P L H L A	
	121	N F L Y F	

C. 核苷酸及氨基酸组合序列 (SEQ ID NO:27) 克隆号: PP12616

起始编码子: 448 ATG 终止编码子: 823 TAG 蛋白质分子量: 14216.38

	1	GCC GCA GTC CCA TCA TTC AGT TCC GTA GGG TCA CCG GCG CCG CAG TGG	48
60	49	CCT CGC AGG GCG CTG GGT CCC TCT CCC CAG CTC TCC TCC ECC TGG CCC	96
	97	CGT CGC CCC GCC CTC GCC GGG CTG GGC TGC GGG GTC AGG GGC CGA GCG	144

	145	GAG AGG GGT GAG TAT TCC CCA CAG CCC TTG CCG GTT GCC TCC TCC CGG	192
	193	CTC TGC TTC CCA CAC GGT CCT TGC CCC ACT CCT AGG ACA GGG AGG AAG	240
	241	GGC ACG CGC GGG TAG GCG GGA AAC AGC CCA GTC CTG AAC AAA AGG CCG	288
	289	GGG AAG CGG GTC CCC GCC GGT AAC TGC AGG CCT GTG CTG GCC GCC GCG	336
5	337	AGC GGG AAG GGC GAG GAC ACT CCC TCC TCG GGG ACC CGG TCC CCG TCG	384
	385	CGC ACA CGG TGT CAC CAG CAC ACC TGG CCC AGT ACC CAA AGC ACC CTC	432
	433	GAA TTA TCA TTT AAC ATG GAA GAA GAT GAG TTC ATT GGA GAA AAA ACA	480
	1	Met Glu Glu Asp Glu Phe Ile Gly Glu Lys Thr	11
10	481	TTC CAA CGT TAT TGT GCA GAA TTC ATT AAA CAT TCA CAA CAG ATA GGT	528
	12	Phe Gln Arg Tyr Cys Ala Glu Phe Ile Lys His Ser Gln Gln Ile Gly	27
	529	GAT AGT TGG GAA TGG AGA CCA TCA AAG GAC TGT TCT GAT GGC TAC ATG	576
15	28	Asp Ser Trp Glu Trp Arg Pro Ser Lys Asp Cys Ser Asp Gly Tyr Met	43
	577	TGC AAA ATA CAC TTT CAA ATT AAG AAT GGG TCT GTG ATG TCA CAT CTA	624
	44	Cys Lys Ile His Phe Gln Ile Lys Asn Gly Ser Val Met Ser His Leu	59
	625	GGA GCA TCT ACC CAT GGA CAG ACA TGT CTT CCC ATG GAG GTG AAG TCT	672
20	60	Gly Ala Ser Thr His Gly Gln Thr Cys Leu Pro Met Glu Val Lys Ser	75
	673	TGC TCT GTC ACC CAG GCT GGA GTG CAG TTG CGT GAT CTC AGC TCA CTG	720
	76	Cys Ser Val Thr Gln Ala Gly Val Gln Leu Arg Asp Leu Ser Ser Leu	91
25	721	CAA CCT CCG CCT TCT GGG TTC AAG CAG TTC TCC TGC CTC AGC CTT CCG	768
	92	Gln Pro Pro Pro Ser Gly Phe Lys Gln Phe Ser Cys Leu Ser Leu Pro	107
	769	AGT AAC TGG GAC TAC AGG GGT TCA CCA CTA CAC CTG GCT AAT TTT TTG	816
30	108	Ser Asn Trp Asp Tyr Arg Gly Ser Pro Leu His Leu Ala Asn Phe Leu	123
	817	TAT TTT TAG TAG AGA TAA GGT TTC ACC ATG TTG GTC AGG CTG GTC TTG	864
	124	Tyr Phe ***	126
	865	AAC TCC TGA CAT CAA GTA ATC CAT CCG CCT CAG CCT CCC AAA GTG CTG	912
35	913	AGA TTA CAG GCA TGA GCC ACT GTG CCT GGC CCT GAG CTT AAA ATA AAA	960
	961	GTT AAA TTA AAA AAA AAA AAA AAA AAA AAA	993

10. PP12723

A: 核苷酸序列 (SEQ ID NO:28) 长度: 1548

40	1	GCAGAGTTTG GGGGTTGTTT GGGATTAGTG AAGCTACTGC CTTTGCCGCC AGCGCAGCCT
	61	CAGAGTTTGA TTATTTGCAA TGTCAGGCTT TGAAACTTA AACACGGATT TCTACCAGAC
	121	AAGTTACAGC ATCGATGATC AGTCACAGCA GTCCTATGAT TATGGAGGAA GTGGAGGACC
	181	CTATAGCAAA CAGTATGCTG GCTATGACTA TTCGCAGCAA GGCAGATTTG TCCCTCCAGA
	241	CATGATGCAG CCACAACAGC CATAACCCGG GCAGATTTAC CAGCCAACTC AGGCATATAC
45	301	TCCAGCTTCA CCTCAGCCTT TCTATGAAAA CAACTTTGAG GATGAGCCAC CTTTATTAGA
	361	AGAGTTAGGT ATCAATTTTG ACCACATCTG GCAAAAAACA CTAACAGTAT TACATCCGTT
	421	AAAAGTAGCA GATGCCAGCA TCATGAATGA AACTGATTTG GCAGGTCCAA TGGTTTTTTG
	481	CCTTGCTTTT GGAGCCACAT TGCTACTGGC TGGCAAAATC CAGTTTGGCT ATGTATACGG
50	541	GATCAGTGCA ATTGGATGTC TAGGAATGTT TTGTTTATTA AACTTAATGA GTATGACAGG
	601	TGTTTTCAATTT GGTGTGTGG CAAGTGTCTT TGGATATTGT CTTCTGCCCA TGATCCTACT
	661	TTCCAGCTTT GCAGTGATAT TTTCTTTGCA AGGAATGGTA GGAATCATT CACTGCTGG
	721	GATTATTGGA TGGTGTAGTT TTTCTGCTTC CAAAATATTT ATTTCTGCAT TAGCCATGGA
	781	AGGACAGCAA CTTTTAGTAG CATATCCTTG CGCTTTGTTA TATGGAGTCT TTGCCCTGAT
	841	TTCCGTCTTT TGAAAATTTA TCTGGGATGT GGACATCAGT GGGCCAGATG TACAAAAGG
55	901	ACCTTGAAC TTTACATTGG ACCAGCAAAC TGCTGCAGCG CAACTCTCAT GCAGATTTTA
	961	CATTTGACTG TTGGAGCAAT GAAAGTAAAC GTGTATCTCT TGTTCAATTT TATAGAACTT
	1021	TTGCATACTA TATTGGATTT ACCTGCGGTG TACTAGCTT TAAATGTTTG TGTTTATACA
	1081	GATAAGAAAT GCTATTTCTT TCTGGTTCCT GCAGCCATTG AAAAACTTT TTCCTTGCAA
	1141	ATTATAATGT TTTTGATAGA TTTTATCAA CTGTGGGAAA CCAACACAA AGCTGATAAC
60	1201	CTTTCTTAAA AACGACCCAG TCACAGTAAA GAAGACACAA GACGGCCGGG CGTGGTAGCT
	1261	CACGCTGTA ATCCCAGCAC TTTGGGAGGC CGAGGCGGGC GGATCACAA GGCAGGAGAT
	1321	CGAGACCATC CTGGTTAACA CGGTGAAACC CCGACTCTAC TAAAATACA AAAAAATTAG
	1381	CTGGGCGTGG TGGCGGGCCG CTGTAGTCCC AGTACTCAG GAGGCTGAGG CAGGAGAAAA

1441 GTGTGAACCC AGGAGGCGGA GCTTGCAGTG AGCCGAGATC ACACCACTGC ACTCCATACC
 1501 AGCCTGGGTG ACAGGGTGAG ACTCTGTTTC AAAAAAAAAA AAAAAAAAAA

B: 核苷酸序列 (SEQ ID NO:29) 长度: 257

5 1 MSGFENLNTD FYQTSYSIDD QSQQSYDYGG SGGPYSKQYA GYDYSQQGRF VPPDMMQPQQ
 61 PYTGQIYQPT QAYTPASPQP FYGNFEDEP PLLEELGINF DHIWQKTLTV LHPLKVADGS
 121 IMNETDLAGP MVFCLAFGAT LLLAGKIQFG YVYGISAIGC LGMFCLLNLMTGVSFGCV
 181 ASVLGYCLLP MILLSSFAVI FSLQGMVGII LTAGIIGWCS FSASKIFISA LAMEGQQLLV
 241 AYPCALLYGV FALISVF

10

C. 核苷酸及氨基酸组合序列 (SEQ ID NO:30) 克隆号: PP12723

起始编码子: 80 ATG 终止编码子: 851 TGA 蛋白质分子量: 27987.81

1 G CAG AGT TTG GGG GTT GTT TGG GAT TAG TGA AGC TAC TGC CTT TGC 46
 47 CGC CAG CGC AGC CTC AGA GTT TGA TTA TTT GCA ATG TCA GGC TTT GAA 94
 15 1 Met Ser Gly Phe Glu 5
 95 AAC TTA AAC ACG GAT TTC TAC CAG ACA AGT TAC AGC ATC GAT GAT CAG 142
 6 Asn Leu Asn Thr Asp Phe Tyr Gln Thr Ser Tyr Ser Ile Asp Asp Gln 21
 20 143 TCA CAG CAG TCC TAT GAT TAT GGA GGA AGT GGA GGA CCC TAT AGC AAA 190
 22 Ser Gln Gln Ser Tyr Asp Tyr Gly Gly Ser Gly Gly Pro Tyr Ser Lys 37
 191 CAG TAT GCT GGC TAT GAC TAT TCG CAG CAA GGC AGA TTT GTC CCT CCA 238
 25 38 Gln Tyr Ala Gly Tyr Asp Tyr Ser Gln Gln Gly Arg Phe Val Pro Pro 53
 239 GAC ATG ATG CAG CCA CAA CAG CCA TAC ACC GGG CAG ATT TAC CAG CCA 286
 54 Asp Met Met Gln Pro Gln Gln Pro Tyr Thr Gly Gln Ile Tyr Gln Pro 69
 287 ACT CAG GCA TAT ACT CCA GCT TCA CCT CAG CCT TTC TAT GGA AAC AAC 334
 30 70 Thr Gln Ala Tyr Thr Pro Ala Ser Pro Gln Pro Phe Tyr Gly Asn Asn 85
 335 TTT GAG GAT GAG CCA CCT TTA TTA GAA GAG TTA GGT ATC AAT TTT GAC 382
 86 Phe Glu Asp Glu Pro Pro Leu Leu Glu Glu Leu Gly Ile Asn Phe Asp 101
 35 383 CAC ATC TGG CAA AAA ACA CTA ACA GTA TTA CAT CCG TTA AAA GTA GCA 430
 102 His Ile Trp Gln Lys Thr Leu Thr Val Leu His Pro Leu Lys Val Ala 117
 431 GAT GGC AGC ATC ATG AAT GAA ACT GAT TTG GCA GGT CCA ATG GTT TTT 478
 40 118 Asp Gly Ser Ile Met Asn Glu Thr Asp Leu Ala Gly Pro Met Val Phe 133
 479 TGC CTT GCT TTT GGA GCC ACA TTG CTA CTG GCT GGC AAA ATC CAG TTT 526
 134 Cys Leu Ala Phe Gly Ala Thr Leu Leu Leu Ala Gly Lys Ile Gln Phe 149
 527 GGC TAT GTA TAC GGG ATC AGT GCA ATT GGA TGT CTA GGA ATG TTT TGT 574
 45 150 Gly Tyr Val Tyr Gly Ile Ser Ala Ile Gly Cys Leu Gly Met Phe Cys 165
 575 TTA TTA AAC TTA ATG AGT ATG ACA GGT GTT TCA TTT GGT TGT GTG GCA 622
 166 Leu Leu Asn Leu Met Ser Met Thr Gly Val Ser Phe Gly Cys Val Ala 181
 50 623 AGT GTC CTT GGA TAT TGT CTT CTG CCC ATG ATC CTA CTT TCC AGC TTT 670
 182 Ser Val Leu Gly Tyr Cys Leu Leu Pro Met Ile Leu Leu Ser Ser Phe 197
 671 GCA GTG ATA TTT TCT TTG CAA GGA ATG GTA GGA ATC ATT CTC ACT GCT 718
 55 198 Ala Val Ile Phe Ser Leu Gln Gly Met Val Gly Ile Ile Leu Thr Ala 213
 719 GGG ATT ATT GGA TGG TGT AGT TTT TCT GCT TCC AAA ATA TTT ATT TCT 766
 214 Gly Ile Ile Gly Trp Cys Ser Phe Ser Ala Ser Lys Ile Phe Ile Ser 229
 767 GCA TTA GCC ATG GAA GGA CAG CAA CTT TTA GTA GCA TAT CCT TGC GCT 814
 60 230 Ala Leu Ala Met Glu Gly Gln Gln Leu Leu Val Ala Tyr Pro Cys Ala 245
 815 TTG TTA TAT GGA GTC TTT GCC CTG ATT TCC GTC TTT TGA AAA TTT ATC 862
 246 Leu Leu Tyr Gly Val Phe Ala Leu Ile Ser Val Phe *** 258
 65 863 TGG GAT GTG GAC ATC AGT GGG CCA GAT GTA CAA AAA GGA CCT TGA ACT 910
 911 CTT ACA TTG GAC CAG CAA ACT GCT GCA GCG CAA CTC TCA TGC AGA TTT 958
 959 TAC ATT TGA CTG TTG GAG CAA TGA AAG TAA ACG TGT ATC TCT TGT TCA 1006

1007 TTT TTA TAG AAC TTT TGC ATA CTA TAT TGG ATT TAC CTG CGG TGT GAC 1054
 1055 TAG CTT TAA ATG TTT GTG TTT ATA CAG ATA AGA AAT GCT ATT TCT TTC 1102
 1103 TGG TTC CTG CAG CCA TTG AAA AAC CTT TTT CCT TGC AAA TTA TAA TGT 1150
 1151 TTT TGA TAG ATT TTT ATC AAC TGT GGG AAA CCA AAC ACA AAG CTG ATA 1198
 5 1199 ACC TTT CTT AAA AAC GAC CCA GTC ACA GTA AAG AAG ACA CAA GAC GGC 1246
 1247 CGG GCG TGG TAG CTC ACG CCT GTA ATC CCA GCA CTT TGG GAG GCC GAG 1294
 1295 GCG GGC GGA TCA CAA GGG CAG GAG ATC GAG ACC ATC CTG GTT AAC ACG 1342
 1343 GTG AAA CCC CGA CTC TAC TAA AAC TAC AAA AAA ATT AGC TGG GCG TGG 1390
 1391 TGG CGG GCG CCT GTA GTC CCA GCT ACT CAG GAG GCT GAG GCA GGA GAA 1438
 10 1439 AAG TGT GAA CCC AGG AGG CGG AGC TTG CAG TGA GCC GAG ATC ACA CCA 1486
 1487 CTG CAC TCC ATA CCA GCC TGG GTG ACA GGG TGA GAC TCT GTT TCA AAA 1534
 1535 AAA AAA AAA AAA AA 1548

11. PP14356

15 A: 核苷酸序列 (SEQ ID NO:31) 长度: 1250

1 GAAAATCCTC AACAAAAATT TGCAAGTAGA ATCCAATAGT GTATAAAAAA AATTATACAC
 61 CACAACCAAA TACTGTAGCA TACCAAGTGT GTAGTATCCC AAGTATGTAA AACTGGTTCA
 121 GCATTCTAAA GTCAAATTAAT GTCATCCATC ACATCGACAA ACTAAGGAAG AAAAAATGAC
 181 ATGATCATAT CAATAGATAC AGACAAAGTA TTTGTCAAAA TCCAACATCT ATATATGACA
 20 241 AAAGCTCTCA GTAAACTAGA AATATAGGGG AACTTCCTCA ACTTGATAAG GAATGTCTAT
 301 TAAAAAACT ACAGTTAACA TCATACTTAA TGACGATAAA CTGTACAGAA GCTTTCTCAC
 361 TAAGATCAGA AACCAAGCCAA ATATGTCATC TCTAATAAAC TTCTTTTCAA TGCCATGCTA
 421 GAAGCCCTCG CTAATACAAT AAATAAGACA AGTAAAGAAA ATAACAGGTA TACTGATTGG
 481 AAAGGAAGAA ATAAAACTG TCTTTGTCA CAGATGTCAT GATTGTCTAT GTAGAAAATC
 25 541 AAAAGAATTA ACAACAACAA CAACCAAAAA AAAAACTCCT GAAACTAATA AGTGAGTATA
 601 GTGAAGTTGC AGGGTACAAG GCTAATACAC AAAAGTCAAG TGCTTTCCTA TGTACCAACA
 661 ATGAACAAGT AGGATTGAA TTAAACACA TAATGCCATT TACATTAGCA CCACAAGAAT
 721 ACACAACACT TAGATATAAA AATCTAACAA CATATGTACA AGATTTATGT GAGGAAAAGT
 781 GCAAAACTCT AATGAAAGAA ATTAGAGATG ATCTAAATAA ATGGAGAGAT ATTCAACATT
 30 841 CATGGATAGG AAGACTCAAT ATTGTCAAGA TGTCAGTTCT TCCCAACTTG ATGTATATAT
 901 TCAATACAAT ACTAATCAAA ATCCTACAGA GTTATTTTGT GGATATTGAC AAAGTCTC
 961 TAAAGTTTAT ATGTAGAGAC AAAAGACCCA GAATAGCCAA CGCAGTACTG AAGGAGAAGA
 1021 ACAAGCCAG AGGACTGGCA CTACTAACG TCAAGACTTA CCGTAAAGCG ACAGTAATCA
 1081 CGACAGTGTG GTATTTGTGA AAGAATAGAC CTCTACCAAT AGTTCAGTGG AATAAAATAG
 35 1141 AGGCCTAGA AATAGACCCA CATAAATATA GTCAACTGAC CTTTGACAAA CGAGAAAACA
 1201 AAAGTACTG GAACAACTGG ACAACCACAT GCAAAAAAAA AAAAAAAA

B: 核苷酸序列 (SEQ ID NO:32) 长度: 135

1 MPFTLAPQEY TTLRYKNLTT YVQDLCEEK KTLMKRIRDD LNKWRDIQHS WIGRLNIVKM
 40 61 SVLPNLMYIF NTILIKILQS YFVDIDKLVL KFICRDKRPR IANAVLKEKN KARGLALLNV
 121 KTYRKATVIT TVWYL

C. 核苷酸及氨基酸组合序列 (SEQ ID NO:33) 克隆号: PP14356

起始编码子: 693 ATG 终止编码子: 1098 TGA 蛋白质分子量: 16072.35

45 1 GA AAA TCC TCA ACA AAA ATT TGC AAG TAG AAT CCA ATA GTG TAT AAA 47
 48 AAA AAT TAT ACA CCA CAA CCA AAT ACT GTA GCA TAC CAA GTG TGT AGT 95
 96 ATC CCA AGT ATG TAA AAC TGG TTC AGC ATT CTA AAG TCA ATT AAT GTC 143
 144 ATC CAT CAC ATC GAC AAA CTA AGG AAG AAA AAA TGA CAT GAT CAT ATC 191
 192 AAT AGA TAC AGA CAA AGT ATT TGT CAA AAT CCA ACA TCT ATA TAT GAC 239
 50 240 AAA AGC TCT CAG TAA ACT AGA AAT ATA GGG GAA CTT CCT CAA CTT GAT 287
 288 AAG GAA TGT CTA TTA AAA AAA CTA CAG TTA ACA TCA TAC TTA ATG ACG 335
 336 ATA AAC TGT ACA GAA GCT TTC TCA CTA AGA TCA GAA ACA AGC CAA ATA 383
 384 TGT CAT CTC TAA TAA ACT TCT TTT CAA TGC CAT GCT AGA AGC CCT CGC 431
 432 TAA TAC AAT AAA TAA GAC AAG TAA AGA AAA TAA CAG GTA TAC TGA TTG 479
 55 480 GAA AGG AAG AAA TAA AAA CTG TCT TTG TTC ACA GAT GTC ATG ATT GTC 527
 528 TAT GTA GAA AAT CAA AAG AAT TAA CAA CAA CAA CCA AAA AAA AAA 575
 576 CTC CTG AAA CTA ATA AGT GAG TAT AGT GAA GTT GCA GGG TAC AAG GCT 623
 624 AAT ACA CAA AAG TCA AGT GCT TTC CTA TGT ACC AAC AAT GAA CAA GTA 671

	672	GGA TTT GAA TTT AAA CAC ATA ATG CCA TTT ACA TTA GCA CCA CAA GAA	719
	1	Met Pro Phe Thr Leu Ala Pro Gln Glu	9
5	720	TAC ACA ACA CTT AGA TAT AAA AAT CTA ACA ACA TAT GTA CAA GAT TTA	767
	10	Tyr Thr Thr Leu Arg Tyr Lys Asn Leu Thr Thr Tyr Val Gln Asp Leu	25
	768	TGT GAG GAA AAG TGC AAA ACT CTA ATG AAA GAA ATT AGA GAT GAT CTA	815
	26	Cys Glu Glu Lys Cys Lys Thr Leu Met Lys Glu Ile Arg Asp Asp Leu	41
10	816	AAT AAA TGG AGA GAT ATT CAA CAT TCA TGG ATA GGA AGA CTC AAT ATT	863
	42	Asn Lys Trp Arg Asp Ile Gln His Ser Trp Ile Gly Arg Leu Asn Ile	57
	864	GTC AAG ATG TCA GTT CTT CCC AAC TTG ATG TAT ATA TTC AAT ACA ATA	911
15	58	Val Lys Met Ser Val Leu Pro Asn Leu Met Tyr Ile Phe Asn Thr Ile	73
	912	CTA ATC AAA ATC CTA CAG AGT TAT TTT GTG GAT ATT GAC AAA CTA GTT	959
	74	Leu Ile Lys Ile Leu Gln Ser Tyr Phe Val Asp Ile Asp Lys Leu Val	89
	960	CTA AAG TTT ATA TGT AGA GAC AAA AGA CCC AGA ATA GCC AAC GCA GTA	1007
20	90	Leu Lys Phe Ile Cys Arg Asp Lys Arg Pro Arg Ile Ala Asn Ala Val	105
	1008	CTG AAG GAG AAG AAC AAA GCC AGA GGA CTG GCA CTA CTT AAC GTC AAG	1055
	106	Leu Lys Glu Lys Asn Lys Ala Arg Gly Leu Ala Leu Leu Asn Val Lys	121
25	1056	ACT TAC CGT AAA GCG ACA GTA ATC ACG ACA GTG TGG TAT TTG TGA AAG	1103
	122	Thr Tyr Arg Lys Ala Thr Val Ile Thr Thr Val Trp Tyr Leu ***	136
	1104	AAT AGA CCT CTA CCA ATA GTT CAG TGG AAT AAA ATA GAG GGC CTA GAA	1151
	1152	ATA GAC CCA CAT AAA TAT AGT CAA CTG ACC TTT GAC AAA GCA GAA AAC	1199
30	1200	AAA AGG TAC TGG AAC AAC TGG ACA ACC ACA TGC AAA AAA AAA AAA AAA	1247
	1248	AAA	1250

12. PP14737

A: 核苷酸序列 (SEQ ID NO:34) 长度: 1037

35	1	GGGCGGCCCC AACGGCCGTC CAGGACGTGG AGCGAAATGA GCTCGGGCGC TGACGGCGGC
	61	GGTGGCGCTG CCGTGGCGGC GCGGTCGGAC AAGGGCAGTC CCGGGGAGGA CCGTTTCGTC
	121	CCGTCGGCGC TGGGACCCG CGAGCATTGG GATGCTGTCT ATGAGAGAGA ACTGCAAAC
	181	TTCCGAGAAT ATGGAGATAC AGGTGAAATC TGGTTGGAG AAGAGAGTAT GAATCGACTA
	241	ATAAGTGGA TGCAGAAACA CAAGATTCCA CTGGATGCTT CAGTGCTTGA TATTGAACT
40	301	GGAAATGGTG TTTTCCTGGT TGAACCTGTT GGCAAAATTT GGTTCCTCTA ATATTACTGG
	361	AATTGATTAC TCTCCTCTG CAATTCAGCT TTCTGGAAGT ATTATAGAAA AAGAAGGTTT
	421	ATCTAACATT AAGTTAAAGG TAGAAGACTT TTTGAATCTC TCCACACAGC TGTCTGGATT
	481	TCATATTTGT ATTGACAAAG GGACTTTTGA TGCCATAAGC CTTAATCCTG ACAATGCAAT
	541	TGAGAAGAGG AAGCAATATG TGAATCTCT CTCCAGGGTG TTGAAAGTAA AAGGCTTTTT
45	601	TTCTAATAAC GTCATGTAAT TGGACCAAG GAAGAGTTGC TAAATGAATT CAGTGAAGGA
	661	TTTGAACCTC TCGAAGAGCT ACCAACACCC AAGTTCAGCT TTGGAGGCAG ATCTGGAAC
	721	AGTGTAGCAG CATTGGTTTT CCAAAAAATG TGAGACTTTT TCTTGGACGA ATTCAGGTAG
	781	CTACACAGAA TCTACACAGC AAAGTTAAACC TGACACAGAA AATCCTTGTG CAAATAAATG
	841	CTTAGTAAGT ACACAGGATG CACATGTTGA ATAGAGTATA CTGGATTGGT GAAAGAAAAT
50	901	AATAATAATG AGCATCTAAG TGGTTGGGTT TTAGAGATCA ATCAAGAATA ATTTAATTT
	961	TCTTTTGTAT TTGAAATGTA AATAGTTTTT TTTTCGATTA AAAAAATTC CTATAAAAAA
	1021	AAAAAAAAAA AAAAAA

B: 核苷酸序列 (SEQ ID NO:35) 长度: 114

55	1	MLQCLILELE MVFSLNLLA KGFNSNITGI DYSPSAIQLS GSIIEKEGLS NIKLKVEDFL
	61	NLSTQLSGFH ICIDKGFDA ISLNPDAIE KRKQYVKSLS RVLKVKGFFS NNVM

C. 核苷酸及氨基酸组合序列 (SEQ ID NO:36) 克隆号: PP14737

起始编码子: 275 ATG 终止编码子: 617 TAA 蛋白质分子量: 12777.24

60	1	G GGC GGC CCC AAC GGC CGT CCA GGA CGT GGA GCG AAA TGA GCT CGG	46
	47	GCG CTG ACG GCG GCG GTG GCG CTG CGG TGG CGG GGT CGG ACA AGG	94
	95	GCA GTC CCG GGG AGG ACG GTT TCG TCC CGT CGG CGC TGG GGA CCC GCG	142

	143	AGC ATT GGG ATG CTG TCT ATG AGA GAG AAC TGC AAA CTT TCC GAG AAT	190
	191	ATG GAG ATA CAG GTG AAA TCT GGT TTG GAG AAG AGA GTA TGA ATC GAC	238
	239	TAA TAA GGT GGA TGC AGA AAC ACA AGA TTC CAC TGG ATG CTT CAG TGC	286
5	1	Met Leu Gln Cys	4
	287	TTG ATA TTG GAA CTG GAA ATG GTG TTT TCC TGG TTG AAC TTG TTG GCA	334
	5	Leu Ile Leu Glu Leu Glu Met Val Phe Ser Trp Leu Asn Leu Leu Ala	20
10	335	AAA TTT GGT TTC TCT AAT ATT ACT GGA ATT GAT TAC TCT CCT TCT GCA	382
	21	Lys Phe Gly Phe Ser Asn Ile Thr Gly Ile Asp Tyr Ser Pro Ser Ala	36
	383	ATT CAG CTT TCT GGA AGT ATT ATA GAA AAA GAA GGT TTA TCT AAC ATT	430
	37	Ile Gln Leu Ser Gly Ser Ile Ile Glu Lys Glu Gly Leu Ser Asn Ile	52
15	431	AAG TTA AAG GTA GAA GAC TTT TTG AAT CTC TCC ACA CAG CTG TCT GGA	478
	53	Lys Leu Lys Val Glu Asp Phe Leu Asn Leu Ser Thr Gln Leu Ser Gly	68
20	479	TTT CAT ATT TGT ATT GAC AAA GGG ACT TTT GAT GCC ATA AGC CTT AAT	526
	69	Phe His Ile Cys Ile Asp Lys Gly Thr Phe Asp Ala Ile Ser Leu Asn	84
	527	CCT GAC AAT GCA ATT GAG AAG AGG AAG CAA TAT GTG AAA TCT CTC TCC	574
	85	Pro Asp Asn Ala Ile Glu Lys Arg Lys Gln Tyr Val Lys Ser Leu Ser	100
25	575	AGG GTG TTG AAA GTA AAA GGC TTT TTT TCT AAT AAC GTC ATG TAA TTG	622
	101	Arg Val Leu Lys Val Lys Gly Phe Phe Ser Asn Asn Val Met ***	115
	623	GAC CCA AGG AAG AGT TGC TAA ATG AAT TCA GTG AAG GAT TTG AAC TTC	670
	671	TCG AAG AGC TAC CAA CAC CCA AGT TCA GCT TTG GAG GCA GAT CTG GAA	718
30	719	ACA GTG TAG CAG CAT TGG TTT TCC AAA AAA TGT GAG ACT TTT TCT TGG	766
	767	ACG AAT TCA GGT AGC TAC ACA GAA TCT ACA CAG CAA AGT TAA CCT GAC	814
	815	ACA GAA AAT CCT TGT GCA AAT AAA TGC TTA GTA AGT ACA CAG GAT GCA	862
	863	CAT GTT GAA TAG AGT ATA CTG GAT TGG TGA AAG AAA ATA ATA ATA ATG	910
	911	AGC ATC TAA GTG GTT GGG TTT TAG AGA TCA ATC AAG AAT AAT TTT AAT	958
35	959	TTT CTT TTG TAT TTG AAA TGT AAA TAG TTT TCT TTT CGA TTA AAA AAA	1006
	1007	TTT CCT ATA AAA AAA AAA AAA AAA AAA A	1037