

[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 01105310.0

C07K 14/47

C07K 16/18

C12N 15/10

C12N 15/11

C12N 15/12

C12N 15/63

C12P 21/00

[45] 授权公告日 2005 年 5 月 4 日

[11] 授权公告号 CN 1199998C

[22] 申请日 2001.2.8 [21] 申请号 01105310.0

[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有限公司

代理人 徐 迅

A61K 38/17

A61P 35/00

权利要求书 1 页 说明书 34 页

[54] 发明名称 具有抑制癌细胞生长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。
- 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: 14、17、20、23、26、29、32。
- 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)完全互补的多核苷酸。
- 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。
- 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所述的多肽以及药理学上可接受的载体。

具有抑制癌细胞生长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；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。

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。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

本发明采用大规模 cDNA 克隆转染癌细胞，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癌细胞(肝癌细胞)具有抑制克隆形成的作用，其抑制率在 50%或 50%以上。

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。

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。

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。

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。

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 PP9284 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 PP9320 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:6 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:5 的蛋白质，但与 SEQ ID NO:6 所示的编码区序列有差别的核酸序列。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%，更佳地至少 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 PP9284 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 上述提到的方法中，分离基因组 DNA 最不常用。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.,
35 Molecular Cloning, A Laboratory Manual, Cold Spring 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
5 产物。上述方法可单用，也可多种方法联合应用。

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
10 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。

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 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。
15

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。
20

本发明也涉及包含本发明的多核苷酸的载体，以及用本发明的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；
- 30 (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。

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
35

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并不限于: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。

5 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。

10 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。

15 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
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。

20 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
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。

25 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。

30 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。

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。

35 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
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。

- 5 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
- 10 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

15 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
- 20 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。

- 25 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.
- 30 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。

- 35 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基因的获得及对癌细胞克隆形成的抑制作用

PP9284、PP9320、PP10122、PP12744、PP13624、PP13671、PP13759、PP14328、PP14450、PP14733和PP14762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 cDNA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和空载体同时转染肝癌细胞系7721。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，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。发现以上克隆有抑制细胞克隆形成作用，结果如下表所示。

cDNA 克隆转染细胞(7721)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)		
PP9284	13	19	16	23	28	25
PP9320	15	19	12	23	28	25
PP10122	13	17	20	23	28	25
PP12744	9	15	14	23	28	25
PP13624	3	2	1	23	28	25
PP13671	0	0	0	23	28	25
PP13759	0	0	0	23	28	25
PP14328	7	3	5	23	28	25
PP14450	0	0	0	23	28	25
PP14733	12	10	9	23	28	25
PP14762	8	7	3	23	28	25

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 5 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、16、19、22、25、28、31)。

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 10 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。利用各个基因的转导引物(如下表所示), 按 97 °C 3' 1 个循环。94 °C 30", 60 °C 30", 72 °C 1', 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 15 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、23、26、29、32)。

基因特异引物

克隆名称	特异引物 1(5' → 3')	特异引物 2(5' → 3')
PP9284	TCTCCCAAATCTCCAGATG	CTCCAACGTGTTGGTCAGAG
PP9320	GCTTCTCCCCACGATTTT	GCTGGAGTCTCCCCAACAT
PP10122	GGAGGCCATTTATTGCTGA	TTAAGAGACGGGGTCTTGC
PP12744	CGCTGTTCCATTGTGAGCTA	CACTTCCTACCGGCCACAC
PP13624	GAGGCCTGAGTTGGGCTC	CGTGTCCATCTCTCCAGTCA
PP13671	CAGCTCAGAGCAGGGGTG	CAGGCTTCTCTTGACGCAGT
PP13759	AGCAGATCCATACTGGAAAATG	GCTGTTGTGCTCGCTCA
PP14328	AACTCATTGACTCCGCCAAC	TTCCCTGCCATTCTGGATAG
PP14450	GATGAGGTGCCTCCACAACT	GCTGCGTTTCACTCAGCAT
PP14733	CCTGATTCACTATTGCCTGG	CAAAATCAAAGTGGAAGAATAAGAAC
PP14762	AAGAAGCCCCTGGTACCCT	CTGCACCCCTACCACTTGTT

实施例 3：cDNA 克隆序列分析

1. PP9284

A：核苷酸序列 (SEQ ID NO:1) 长度：2394

5 1 GGGACCTCAG ACGTATAGTT TTCTCAGATT TCTGTGCTTT CTGGGGCTGG GCTACTAGTG
 61 GAAGAAAGCA GTCTATTCTG TCTTCTCCCA AATCTCCAG ATGCCAGTC TGTTGAAGGA
 121 GGAGCAGAAC CAGGGGGCCT TTCCCGCTGA GGCCCGACCT GTGTCTCCTT CAAATGACAC
 181 GCGGGACTCA GGGCCTTCCC ATGACCATGG GGCCAGGGG GCGTCACCTG GCCCAGGGCC
 241 CAGTGCTAGA AACAGATGAC CCCAGGAGGA GGAGGCAGGG CAGGAGGGAA GCTGGCAGGG
 301 CTGGGATGGT CAGCCAGGCT GAGGGGCGGA CTCGCACCAG GATGGAGCTA GGAAATGATC
 10 361 CAGGTGTGTT TGGCGGCTGC AGGTGGTCC GCATGGCTGT GCAGGGAGGG AAGGGCTGCG
 421 TGGCAGGAGA GCAGCCGGGG GAGGCCAGA CTCTGCTGAA GAGATGCCTG TTGTGCCGGC
 481 CTCCACATCC GCTGCCGGCT CCTTCCGGAG CTCCTGCCCC GCCATGCTCA GCCTGACTCT
 541 GACCAACACG TTGGAGAGAA GAATGATCCC TTTGTGCTAT TAAGCTTGCT TATTTGGTTT
 601 CTAAGTGCTT CATGCGAACC TAGAGGAAAA AATTATTTTC CACCTTTGTT TGTCTTAAGA
 15 661 AAATAACACA CTTTTTTTTT TCCTATTTGA ACAGGCAGAC GGCTAATCCA CATGGTCTTC
 721 GTCCTTGACG TCGTTTTACA AGAAAAAAT GGGGCTGGTT TTGCTTCCCC GTGCATGATT
 781 TACTCTTAGA GATGATTCAG AGGTCACTTC ATTTTTATTA AACAGTGAAC TTGTCTGGCT
 841 TTGGCACTCT CTGCCATTCT GTGCAGGCTG CAGTGGCTCC CCTGCCAGC CTGCTCTCCC
 901 TAACCCTTG TCCGCAAGGG GTGATGGCCG GCTGGTGTG GGCCTGGCG GTCAAGTGTG
 20 961 GAGGAGAGGG GTGGAGGCTG CCCCATTGAG ATCTTCTGCG TGAGTCCTTT CCAGGGGCCA
 1021 ATTTTGGATG AGCATGGAGC TGTCACCTCT CAGCTGCTGG ATGACTTGAG ATGAAAAAGG
 1081 AGAGACATGG AAAGGGAGAC AGCCAGGTGG CACCTGCAGC GGCTGCCCTC TGGGGCCACT
 1141 TGGTAGTGTC CCCAGCCTAC CTCTCCACAA GGGGATTTTG CTGATGGGTT CTTAGAGCCT
 1201 TAGCAGCCCT GGATGGTGGC CAGAAATAAA GGGACCAGCC CTTTCATGGT GGTGACGTGG
 25 1261 TAGTCACTTG TAAGGGGAAC AGAAACATTT TTGTTCTTAT GGGGTGAGAA TATAGACAGT
 1321 GCCCTTGGTG CGAGGGAAGC AATTGAAAAG GAACTGCCC TGAGCACTCC TGGTGCAGGT
 1381 CTCCACCTGC ACATTGGGTG GGGCTCCTGG GAGGGAGACT CAGCCTCCT CCTCATCCTC
 1441 CCTGACCCCTG CTCCTAGCAC CCTGGAGAGT GCACATGCCC CTTGGTCCTG GCAGGGCGCC
 1501 AAGTCTGGCA CCATGTTGGC CTCTCAGGC CTGCTAGTCA CTGGAAATTG AGGTCCATGG
 30 1561 GGGAAATCAA GGATGCTCAG TTTAAGGTAC ACTGTTTCCA TGTATGTTT CTACACATTG
 1621 CTACCTCAGT GCTCCTGGAA ACTTAGCTTT TGATGTCTCC AAGTAGTCCA CCTTCATTTA
 1681 ACTCTTTGAA ACTGTATCAT CTTTGCCAAG TAAGAGTGGT GGCCTATTTC AGCTGCTTTG
 1741 ACAAAATGAC TGGCTCCTGA CTTAACGTTT TATAAATGAA TGTGCTGAAG CAAAGTGCCC
 1801 ATGGTGGCGG CGAAGAAGAG AAAGATGTGT TTTGTTTTGG ACTCTCTGTG GTCCCTTCCA
 35 1861 ATGCTGTGGG TTTCCAACCA GGGGAAGGGT CCCTTTTGCA TTGCCAAGTG CCATAACCAT
 1921 GAGCACTACT CTACCATGGT TCTGCCTCCT GGCCAAGCAG GCTGGTTTGC AAGAATGAAA
 1981 TGAATGATTC TACAGCTAGG ACTTAACCTT GAAATGAAA GTCTTGCAAT CCCATTTGCA
 2041 GGATCCGTCT GTGCACATGC CTCTGTAGAG AGCAGCATT CCAGGGACCT TGGAAACAGT
 2101 TGGCACTGTA AGGTGCTTGC TCCCAAGAC ACATCCTAAA AGGTGTTGTA ATGGTGA AAA
 40 2161 CGTCTTCTT CTTTATTGCC CCTTCTTATT TATGTGAACA ACTGTTTGTG TTTTTTTGTA
 2221 TCTTTTTTAA ACTGTAAAGT TCAATTGTGA AAATGAATAT CATGCAAATA AATTATGCGA
 2281 TTTTTTTTTC AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
 2341 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAA

45 B：氨基酸序列 (SEQ ID NO:2) 长度：120

1 MTRGTQGLPM TMGPRGRHLA QGPVLEDDP RRRRQRREA GRAGMVSQAE GRTRTRMELG
 61 NDPGVFGGCR WVRMAVQGGK GCVAGEQPGE AQTLLKRCLL CRPPHPLPAP SGAPAPPCSA

50 C. 核苷酸及氨基酸组合序列 (SEQ ID NO:3) 克隆号：PP9284

起始编码子：174 ATG 终止编码子：534 TGA 蛋白质分子量：12778.12

(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)

1 GG GAC CTC AGA CGT ATA GTT TTC TCA GAT TTC TGT GCT TTC TGG GGC 47
 48 TGG GCT ACT AGT GGA AGA AAG CAG TCT ATT CTG TCT TCT CCC AAA TCT 95

	96	CCC AGA TGC CCA GTC TGT TGA AGG AGG AGC AGA ACC AGG GGG CCT TTC	143
	144	CCG CTG AGG CCC GAC CTG TGT CTC CTT CAA ATG ACA CGC GGG ACT CAG	191
	1	Met Thr Arg Gly Thr Gln	6
5	192	GGC CTT CCC ATG ACC ATG GGG CCC AGG GGG CGT CAC CTG GCC CAG GGC	239
	7	Gly Leu Pro Met Thr Met Gly Pro Arg Gly Arg His Leu Ala Gln Gly	22
	240	CCA GTG CTA GAA ACA GAT GAC CCC AGG AGG AGG AGG CAG GGC AGG AGG	287
10	23	Pro Val Leu Glu Thr Asp Asp Pro Arg Arg Arg Arg Gln Gly Arg Arg	38
	288	GAA GCT GGC AGG GCT GGG ATG GTC AGC CAG GCT GAG GGG CGG ACT CGC	335
	39	Glu Ala Gly Arg Ala Gly Met Val Ser Gln Ala Glu Gly Arg Thr Arg	54
	336	ACC AGG ATG GAG CTA GGA AAT GAT CCA GGT GTG TTT GGC GGC TGC AGG	383
15	55	Thr Arg Met Glu Leu Gly Asn Asp Pro Gly Val Phe Gly Gly Cys Arg	70
	384	TGG GTC CGC ATG GCT GTG CAG GGA GGG AAG GGC TGC GTG GCA GGA GAG	431
	71	Trp Val Arg Met Ala Val Gln Gly Gly Lys Gly Cys Val Ala Gly Glu	86
20	432	CAG CCG GGG GAG GCC CAG ACT CTG CTG AAG AGA TGC CTG TTG TGC CGG	479
	87	Gln Pro Gly Glu Ala Gln Thr Leu Leu Lys Arg Cys Leu Leu Cys Arg	102
	480	CCT CCA CAT CCG CTG CCC GCT CCT TCC GGA GCT CCT GCC CCG CCA TGC	527
25	103	Pro Pro His Pro Leu Pro Ala Pro Ser Gly Ala Pro Ala Pro Pro Cys	118
	528	TCA GCC TGA CTC TGA CCA ACA CGT TGG AGA GAA GAA TGA TCC CTT TGT	575
	119	Ser Ala ***	121
	576	GCT ATT AAG CTT GCT TAT TTG GTT TCT AAG TGC TTC ATG CGA ACC TAG	623
30	624	AGG AAA AAA TTA TTT TCC ACC TTT GTT TGT CTT AAG AAA ATA ACA CAC	671
	672	TTT TTT TTT TCC TAT TTG AAC AGG CAG ACC GCT AAT CCA CAT GGT CTT	719
	720	CGT CCT TGA CGT CGT TTT ACA AGA AAA CAA TGG GGC TGG TTT TGC TTC	767
	768	CCC GTG CAT GAT TTA CTC TTA GAG ATG ATT CAG AGG TCA CTT CAT TTT	815
	816	TAT TAA ACA GTG AAC TTG TCT GGC TTT GGC ACT CTC TGC CAT TCT GTG	863
35	864	CAG GCT GCA GTG GCT CCC CTG CCC AGC CTG CTC TCC CTA ACC CCT TGT	911
	912	CCG CAA GGG GTG ATG GCC GGC TGG TTG TGG GCA CTG GCG GTC AAG TGT	959
	960	GGA GGA GAG GGG TGG AGG CTG CCC CAT TGA GAT CTT CCT GCT GAG TCC	1007
	1008	TTT CCA GGG GCC AAT TTT GGA TGA GCA TGG AGC TGT CAC CTC TCA GCT	1055
	1056	GCT GGA TGA CTT GAG ATG AAA AAG GAG AGA CAT GGA AAG GGA GAC AGC	1103
40	1104	CAG GTG GCA CCT GCA GCG GCT GCC CTC TGG GGC CAC TTG GTA GTG TCC	1151
	1152	CCA GCC TAC CTC TCC ACA AGG GGA TTT TGC TGA TGG GTT CTT AGA GCC	1199
	1200	TTA GCA GCC CTG GAT GGT GGC CAG AAA TAA AGG GAC CAG CCC TTC ATG	1247
	1248	GGT GGT GAC GTG GTA GTC ACT TGT AAG GGG AAC AGA AAC ATT TTT GTT	1295
	1296	CTT ATG GGG TGA GAA TAT AGA CAG TGC CCT TGG TGC GAG GGA AGC AAT	1343
45	1344	TGA AAA GGA ACT TGC CCT GAG CAC TCC TGG TGC AGG TCT CCA CCT GCA	1391
	1392	CAT TGG GTG GGG CTC CTG GGA GGG AGA CTC AGC CTT CCT CCT CAT CCT	1439
	1440	CCC TGA CCC TGC TCC TAG CAC CCT GGA GAG TGC ACA TGC CCC TTG GTC	1487
	1488	CTG GCA GGG CGC CAA GTC TGG CAC CAT GTT GGC CTC TTC AGG CCT GCT	1535
	1536	AGT CAC TGG AAA TTG AGG TCC ATG GGG GAA ATC AAG GAT GCT CAG TTT	1583
50	1584	AAG GTA CAC TGT TTC CAT GTT ATG TTT CTA CAC ATT GCT ACC TCA GTG	1631
	1632	CTC CTG GAA ACT TAG CTT TTG ATG TCT CCA AGT AGT CCA CCT TCA TTT	1679
	1680	AAC TCT TTG AAA CTG TAT CAT CTT TGC CAA GTA AGA GTG GTG GCC TAT	1727
	1728	TTC AGC TGC TTT GAC AAA ATG ACT GGC TCC TGA CTT AAC GTT CTA TAA	1775
	1776	ATG AAT GTG CTG AAG CAA AGT GCC CAT GGT GGC GGC GAA GAA GAG AAA	1823
55	1824	GAT GTG TTT TGT TTT GGA CTC TCT GTG GTC CCT TCC AAT GCT GTG GGT	1871
	1872	TTC CAA CCA GGG GAA GGG TCC CTT TTG CAT TGC CAA GTG CCA TAA CCA	1919
	1920	TGA GCA CTA CTC TAC CAT GGT TCT GCC TCC TGG CCA AGC AGG CTG GTT	1967
	1968	TGC AAG AAT GAA ATG AAT GAT TCT ACA GCT AGG ACT TAA CCT TGA AAT	2015
	2016	GGA AAG TCT TGC AAT CCC ATT TGC AGG ATC CGT CTG TGC ACA TGC CTC	2063
60	2064	TGT AGA GAG CAG CAT TCC CAG GGA CCT TGG AAA CAG TTG GCA CTG TAA	2111
	2112	GGT GCT TGC TCC CCA AGA CAC ATC CTA AAA GGT GTT GTA ATG GTG AAA	2159

	2160	ACG TCT TCC TTC TTT ATT GCC CCT TCT TAT TTA TGT GAA CAA CTG TTT	2207
	2208	GTC TTT TTT TGT ATC TTT TTT AAA CTG TAA AGT TCA ATT GTG AAA ATG	2255
	2256	AAT ATC ATG CAA ATA AAT TAT GCG ATT TTT TTT TCA AAA AAA AAA AAA	2303
	2304	AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA	2351
5	2352	AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA A	2394

2. PP9320

A: 核苷酸序列 (SEQ ID NO:4) 长度: 2923

	1	GCCGAGCTGA GCCGATCAAA AGGCCATGGG TAGGAGGCTT GCAGGTGTGA CTTTAGGATC
10	61	ATGGCTTCTC CCCACGATTT TTTGGGGTTC GAGGGAAGCA GAGCATGGTG CTCCAATCCC
	121	AGAACCCAGC TGTCTCAAG CTCAGAGGAC CAGGTCCTGG CCTGGCTGCT GCTGCTGGGA
	181	CTCCTCCCTA TGGGCATCCC TGCATGTCTT GTGTGTCCCT GAGGAGGAC ATGGGAACT
	241	CAGGGGCCAC CTCTCCTCGA ACTGCGGGC CAGAGCAGAG AGCCCTTGCA CACCACCAGC
	301	CTCTCCTCCC TGTGCCCCAG GAGTTCATCC TTTGCTGCCT GGCCCGGGAC CCTGCCCGCC
15	361	GGCCCTCTGC CCACAGCCTC CTCTCCACC GCGTGTCTT CGAGGTGCAC TCGCTGAAGC
	421	TCTTGGCAGC CCACTGTCTC ATCCAGCACC AGTACCTCAT GCCTGAGAAT GTGGTGGAGG
	481	AGAAGACCAA GGCCATGGAC CTGCACGCGG TCTTGGCGGA GCTTCCCGG CCCCAGGCGC
	541	CCCCGCTGCA GTGGCGGTAC TCGGAAGTCT CTTTATGGA GCTGGACAAA TTCCTGGAGG
	601	ATGTCAGGAA TGGAACTAC CCACTGATGA ACTTTGCAGC CACTCGACCC CTGGGGCTGC
20	661	CCCGTGTGCT GGCCCCACCC CCGGAGGAGG TCCAAAAGGC CAAGACCCCG ACGCCAGAGC
	721	CCTTTGACTC TGAGACCAGA AAGGTCATCC AGATGCAGTG CAACCTGGAG AGAAGCGAGG
	781	ACAAGGCGCG CTGGCATCTC ACTCTGCTTC TGGTGTGGA AGACCGGCTG CACCCGCGAGC
	841	TGACCTACGA CCTGTCCCA ACGGACAGCG CCCAGGACCT CGCCTCGGAG CTCGTGCACT
	901	ATGGCTTCTT CCACGAGGAC GACCGGATGA AGCTGGCCGC CTTCTGGAG AGCACCTTCC
25	961	TCAAGTACCG TGGGACCCAG GCCTGACCCG GAGCCCCAGC CCCAGGGGAC CATGCCGGGG
	1021	TGCTGCCCGG GCAGGCCATG TTGGGGAGAC TCCAGCACCG TGGGGCTGCC CTCCTCCATG
	1081	CGCCTGGGAG CACAAAGGCC CCGGTAGTGA AGGAACCCCG CGTCTCCTGA GAGTGGGGCT
	1141	GACCCTGCCT TGGGCGCCGA GGGGTTTGGG GGGTGGGTGT GGGGGAGCCG TTAGCCTCCC
	1201	AGGTCCTTAG GATCAGGGTT GCCCCAGAAA CCCCTTCCCA TATCCTCCAT TCTCCGCCCT
30	1261	GAGTTCCTAC CCAGGCTGCC TGGCCGGGGC CACTGCCTCC TCAGCATGCA GGAGGCTGCC
	1321	CTGTAGGGAA CCCAGCTCT GGGGCTTGGG GGTGAGGGTC AGCCCTGGAC AGACCTCTGC
	1381	CCAGGGAAct GCTCCATGGG GTCTGGGAGA GCAGCCATCC CCTGCTGGCA CCATAGACCC
	1441	ACACAAGGAG CCTGCACAGC AAGCCAGCGG TGACACACCT GCAGGTGTCA GGCATGGCAC
	1501	TGGGCACAAC AGGGACCTGG CAGGAGAACA GACCACAGAG AGGTCTGGAG TTGAGGCTGT
35	1561	TGTCAGCAAA GCCCTGGTC CCACACAGCT CTGCCCTAGA GCCACCTCTT TGACCCTTTA
	1621	CCCACCTGA GACCAGAAct TGAGCCCTCT CTGCAGATCT CCTCTGGCCA CTGCAGCCCC
	1681	TCCAATGGGC TTTTCTCTC ATGCATTCCC TGGCCTGGAG GCGTCAGGGA CCCACATCC
	1741	TCCCTGCTCC TCAGACTCAC AGCCCTCCA TGTTACCTCC CGCACCTCCT CCCTGGGGCA
	1801	GCTGTCCCT GGGCCTCTGA GGATGTACG TCCTGGCTCC CTGCCTCTCT CCCACTCCAC
40	1861	TCCTGGCTCA GTCTTAGAGA TTCTATGCC CTCATGGATT CTACCCCTGC CTTCTGGCC
	1921	TCTTGATTCT TGGCTTGCCT CTCTCCAAT TCCAACTTA GTGAAATGGC CTTAAGCATT
	1981	TTAAACTGTA TGTATACATT AGCGATTCA TGCCTTCTA AACGCATTTC AAATGTCAAC
	2041	CAGGAAGGCA CACCACTGTA TTAGTTTAT ACTGCGCTG TAAATTTAC CACAACTTA
	2101	GTGACTTAAC ACAAATTTAT TGCAATTCTG TAGGCTGGAA GTCTGACTAT GGGTCTCACT
45	2161	GGACTAGAAT CAAGGCTGGC AGGCTGCCTT CCTTCTGGA GGTCTAGGG GAGACTCTGT
	2221	CTCCTGCTCC TTCAGGCTGC TGGCAGAATC CACATCCTT CGGTGGCAGG GCCAAGGTCC
	2281	CCACTTCTT GCTGACTGTA AACTAAGGCC ACTTCCAGCT TGTAGAGGCT GCCTACATTC
	2341	CTTGGCTCTT GGCCCTCC TCCATCTCA GAGCTAGCAG GTTCACTGTG TGTACGAAAC
	2401	CATTTCTCTG GTTCCCTGCA GACAGGAAAG GTTGTCCCTA AGGACTCATG AGATTAGGTT
50	2461	GGGCCAGCC AGATAATACA TGATAATCTC CCTCCTCAAG GTTTTAAATA TTAACACAT
	2521	CTGCAGGACA CATTTTGCCA TGATAACTAA CATTCACTGG TTCCAGGAT TAAGGAATGA
	2581	ACCTCTTTG TTGGGAAGG GTGGCATTCT GCTGACCACA GCACTCCAAC CAAAAGCCAA
	2641	AAACCAAAGC AAGACTTACT AACGCATATC AAATAAATTA AAGGTACAAA ATCGTGAATC
	2701	TCAGTTATCT TAAATATCC AATACTATT ACAAATTTAT TCAAATCTC ACGCCTCCA

2761 ACTCAAAATT AGCAATCTAA AGTAATTTCC ATATCCTAGA TGGAAACCCT CATGCTAAAC
 2821 TGTCTGATTA TGCATGGTTC TAAATGGTTT CAGTGGCAAA TACATAACAT TGTACTACTG
 2881 ATTAAACTGA ACTTAAAAGC ATCAAAAAAA AAAAAAAAAA AAA

5 B: 氨基酸序列 (SEQ ID NO:5) 长度: 293

1 MVLQSQNPV LKLRGPGPGL AAAAGTPPYG HPCMSCVSLR RDMGNSGATS PRTAGPEQRA
 61 LAHHQPLLPV PQEFILCCLA RDPARRPSAH SLLFHRVLFV VHSLKLLAAH CFIHQYQLMP
 121 ENVVEEKTKA MDLHAVLAEL PRPRRPPLQW RYSEVSFMEL DKFLEDVRNG IYPLMNFAAT
 181 RPLGLPRVLA PPPEEVQKAK TPTPEPFDSR TRKVIQMOCN LERSEDKARW HLTLLLVLED
 15 241 RLHRQLTYDL LPTDSAQDLA SELVHYGFLH EDDRMKLAFF LESTFLKYRG TQA

C. 核苷酸及氨基酸组合序列 (SEQ ID NO:6) 克隆号: PP9320

起始编码子: 105 ATG 终止编码子: 984 TGA 蛋白质分子量: 33139.57

1 GC CGA GCT GAG CCG ATC AAA AGG CCA TGG GTA GGA GGC TTG CAG GTG 47
 15 48 TGA CTT TAG GAT CAT GGC TTC TCC CCA CGA TTT TTT GGG GGT CGA GGG 95
 96 AAG CAG AGC ATG GTG CTC CAA TCC CAG AAC CCA GCT GTC CTC AAG CTC 143
 1 Met Val Leu Gln Ser Gln Asn Pro Ala Val Leu Lys Leu 13
 144 AGA GGA CCA GGT CCT GGC CTG GCT GCT GCT GCT GGG ACT CCT CCC TAT 191
 20 14 Arg Gly Pro Gly Pro Gly Leu Ala Ala Ala Ala Gly Thr Pro Pro Tyr 29
 192 GGG CAT CCC TGC ATG TCC TGT GTG TCC CTG AGG AGG GAC ATG GGG AAC 239
 30 Gly His Pro Cys Met Ser Cys Val Ser Leu Arg Arg Asp Met Gly Asn 45
 240 TCA GGG GCC ACC TCT CCT CGA ACT GCG GGG CCA GAG CAG AGA GCC CTT 287
 25 46 Ser Gly Ala Thr Ser Pro Arg Thr Ala Gly Pro Glu Gln Arg Ala Leu 61
 288 GCA CAC CAC CAG CCT CTC CTC CCT GTG CCC CAG GAG TTC ATC CTT TGC 335
 30 62 Ala His His Gln Pro Leu Leu Pro Val Pro Gln Glu Phe Ile Leu Cys 77
 336 TGC CTG GCC CGG GAC CCT GCC CGC CGG CCC TCT GCC CAC AGC CTC CTC 383
 78 Cys Leu Ala Arg Asp Pro Ala Arg Arg Pro Ser Ala His Ser Leu Leu 93
 384 TTC CAC CGC GTG CTC TTC GAG GTG CAC TCG CTG AAG CTC CTG GCA GCC 431
 35 94 Phe His Arg Val Leu Phe Glu Val His Ser Leu Lys Leu Leu Ala Ala 109
 432 CAC TGC TTC ATC CAG CAC CAG TAC CTC ATG CCT GAG AAT GTG GTG GAG 479
 110 His Cys Phe Ile Gln His Gln Tyr Leu Met Pro Glu Asn Val Val Glu 125
 480 GAG AAG ACC AAG GCC ATG GAC CTG CAC GCG GTC TTG GCG GAG CTT CCC 527
 40 126 Glu Lys Thr Lys Ala Met Asp Leu His Ala Val Leu Ala Glu Leu Pro 141
 528 CGG CCC CGC AGG CCC CCG CTG CAG TGG CGG TAC TCG GAA GTC TCC TTC 575
 45 142 Arg Pro Arg Arg Pro Pro Leu Gln Trp Arg Tyr Ser Glu Val Ser Phe 157
 576 ATG GAG CTG GAC AAA TTC CTG GAG GAT GTC AGG AAT GGA ATC TAC CCA 623
 158 Met Glu Leu Asp Lys Phe Leu Glu Asp Val Arg Asn Gly Ile Tyr Pro 173
 624 CTG ATG AAC TTT GCA GCC ACT CGA CCC CTG GGG CTG CCC CGT GTG CTG 671
 50 174 Leu Met Asn Phe Ala Ala Thr Arg Pro Leu Gly Leu Pro Arg Val Leu 189
 672 GCC CCA CCC CCG GAG GAG GTC CAA AAG GCC AAG ACC CCG ACG CCA GAG 719
 190 Ala Pro Pro Pro Glu Glu Val Gln Lys Ala Lys Thr Pro Thr Pro Glu 205
 720 CCC TTT GAC TCT GAG ACC AGA AAG GTC ATC CAG ATG CAG TGC AAC CTG 767
 55 206 Pro Phe Asp Ser Glu Thr Arg Lys Val Ile Gln Met Gln Cys Asn Leu 221
 768 GAG AGA AGC GAG GAC AAG GCG CGC TGG CAT CTC ACT CTG CTT CTG GTG 815
 60 222 Glu Arg Ser Glu Asp Lys Ala Arg Trp His Leu Thr Leu Leu Val 237
 816 CTG GAA GAC CGG CTG CAC CGG CAG CTG ACC TAC GAC CTG CTC CCA ACG 863
 238 Leu Glu Asp Arg Leu His Arg Gln Leu Thr Tyr Asp Leu Leu Pro Thr 253
 864 GAC AGC GCC CAG GAC CTC GCC TCG GAG CTC GTG CAC TAT GGC TTC CTC 911
 65 254 Asp Ser Ala Gln Asp Leu Ala Ser Glu Leu Val His Tyr Gly Phe Leu 269

	912	CAC GAG GAC GAC CGG ATG AAG CTG GCC GCC TTC CTG GAG AGC ACC TTC	959
	270	His Glu Asp Asp Arg Met Lys Leu Ala Ala Phe Leu Glu Ser Thr Phe	285
5	960	CTC AAG TAC CGT GGG ACC CAG GCC TGA CCC GGA GCC CCA GCC CCA GGG	1007
	286	Leu Lys Tyr Arg Gly Thr Gln Ala ***	294
	1008	GAC CAT GCC GGG GTG CTG CCC GGG CAG GCC ATG TTG GGG AGA CTC CAG	1055
	1056	CAC CGT GGG GCT GCC CTC CTC CAT GCG CCT GGG AGC ACA AAG GCC CCG	1103
10	1104	GTA GTG AAG GAA CCC CCC GTC TCC TGA GAG TGG GGC TGA CCC TGC CTT	1151
	1152	GGG CGC CGA GGG GTT TGG GGG GTG GGT GTG GGG GAG CCG TTA GCC TCC	1199
	1200	CAG GTC CTT AGG ATC AGG GTT GCC CCC AGA ACC CCT TCC CAT ATC CTC	1247
	1248	CAT TCT CCG CCC TGA GTT CCT ACC CAG GCT GCC TGG CCG GGG CCA CTG	1295
	1296	CCT CCT CAG CAT GCA GGA GGC TGC CCT GTA GGG AAC CCC AGC TCT GGG	1343
15	1344	GCT TGG GGG TGA GGG TCA GCC CTG GAC AGA CCT CTG CCC AGG GAA CTG	1391
	1392	CTC CAT GGG GTC TGG GAG AGC AGC CAT CCC CTG CTG GCA CCA TAG ACC	1439
	1440	CAC ACA AGG AGC CTG CAC AGC AAG CCA GCG GTG ACA CAC CTG CAG GTG	1487
	1488	TCA GGC ATG GCA CTG GGC ACA ACA GGG ACC TGG CAG GAG AAC AGA CCA	1535
	1536	CAG AGA GGT CTG GAG TTG AGG CTG TTG TCA GCA AAG CCC CTG GTC CCA	1583
20	1584	CAC AGC TCT GCC CTA GAG CCA CCT CTT TGA CCC TTT ACC CAC CCT GAG	1631
	1632	ACC AGA ACT TGC AGC CCC TCT GCA GAT CTC CTC TGG CCA CTG CAG CCC	1679
	1680	CTC CAA TGG GCT TTT TCT CTC ATG CAT TCC CTG GCC TGG AGG CGT CAG	1727
	1728	GGA CCC CAC ATC CTC CCT GCT CCT CAG ACT CAC AGC CCC TCC ATG TTA	1775
	1776	CCT CCC GCA CCT CCT CCC TGG GGC AGC TGC TCC CTG GGC CTC TGA GGA	1823
25	1824	TGT CAG CTC CTG GCT CCC TGC CTC TCT CCC ACT CCA CTC CTG GCT CAG	1871
	1872	TCT TAG AGA TTT CTA TGC CCT CAT GGA TTC TAC CCC TGC CTT CCT GGC	1919
	1920	CTC TTG ATT CTT GGC TTG CCT CTC CTC CAA TTC CAA ACT TAG TGA AAT	1967
	1968	GGC CTT AAG CAT TTT AAA CTG TAT GTA TAC ATT AGC GCA TTC ATG CCT	2015
	2016	TTC TAA ACG CAT TTC AAA TGT CAA CCA GGA AGG CAC ACC ACT GTA TTA	2063
30	2064	GTT TTA TAC TGC CGC TGT AAA ATT TAC CAC AAA CTT AGT GAC TTA ACA	2111
	2112	CAA ATT TAT TGC AAT TCT GTA GGC TGG AAG TCT GAC TAT GGG TCT CAC	2159
	2160	TGG ACT AGA ATC AAG GCT GGC AGG CTG CCT TCC TTC CTG GAG GTT CTA	2207
	2208	GGG GAG ACT CTG TCT CCT GCT CCT TCA GGC TGC TGG CAG AAT CCA CAT	2255
	2256	CCT TTC GGT GGC AGG GCC AAG GTC CCC ACT TTC TTG CTG ACT GTA AAC	2303
35	2304	TAA GGC CAC TTC CAG CTT GTA GAG GCT GCC TAC ATT CCT TGG CTC TTG	2351
	2352	GCC CCC TCC TCC ATC TTC AGA GCT AGC AGG TTC AGT CTG TGT CAC GAA	2399
	2400	CCA TTT CTC TGG TTC CCT GCA GAC AGG AAA GGT TGT CCC TAA GGA CTC	2447
	2448	ATG AGA TTA GGT TGG GCC CAG CCA GAT AAT ACA TGA TAA TCT CCC TCC	2495
	2496	TCA AGG TTT TTA ATA TTA AAC ACA TCT GCA GGA CAC ATT TTG CCA TGT	2543
40	2544	AAA CTA ACA TTC ACT GGT TCC AGG GAT TAA GGA ATG AAC CTC TTT TGT	2591
	2592	TGG GGA AGG GTG GCA TTC TGC TGA CCA CAG CAC TCC AAC CAA AAG CCA	2639
	2640	AAA ACC AAA GCA AGA CTT ACT AAC GCA TAT CAA ATA AAT TAA AGG TAC	2687
	2688	AAA ATC GTG AAT CTC AGT TAT CTT AAA TAT TCC AAT ACT ATT TAC AAA	2735
	2736	ATT ATT CAA ATT CTC ACG CCT TCC AAC TCA AAA TTA GCA ATC TAA AGT	2783
45	2784	AAT TTC CAT ATC CTA GAT GGA AAC CCT CAT GCT AAA CTG TCT GAT TAT	2831
	2832	GCA TGG TTC TAA ATG GTT TCA GTG GCA AAT ACA TAA CAT TGT ACT ACT	2879
	2880	GAT TAA ACT GAA CTT AAA AGC ATC AAA AAA AAA AAA AAA AA	2923

3. PP10122

50 A: 核苷酸序列 (SEQ ID NO:7) 长度: 1867

1 GCAAATCCAT TCTTAAACCT TGCAAGCTGA CTGCGGGAGC CCATGGATTG CAGGCCTCCC
61 CTGCAGCACT CCCTGGCTGC AGCCCCTTCC TGACCGTGCA GCTCCCTCGT GGAGCCCTGT
121 GTGGGACTCC CAGGGTGGAC TCTGGATGTG GGGTCCCAAG AATGCAGAGC CTTCTATGGG
181 TGACCCAGG TGGTGTCCCA GTCACTTTGC CATTGAGGGA TAGGTGCTGA GGGAGCCTGT
55 241 CCTCTCAGAA CCTCAGTCTC CCTGTCTGTA AATTTGGGGA ACCCAGTTGG CTGCACAACC
301 CCTGCAGAAC CCTGTTGTGT GGCTCTTAGC TGGTGGTCCC TAAGTGGGTT TCTCCTGCAG
361 TCCCTCTAA GCATCCACCC CTCCTCCCA AGCCTCCAAC CCCCACACC TGAGACTCCC

421 CACGTGGCCC TGCCTGGATG CCACCCACCT GTGTCCCTTC CTCTGTTCCCT ACTAGTTGTT
 481 GAAGACATCA ACAAAACGGCG GGAACCCATT CCCAGTCTGG AGGCCATTTA TTTGCTGAGC
 541 CCCACGGAGA AGGTGCCTAC ATGAGTGAGC GTGTGTGTAT GCGCGTGCAT GCGTGTACAT
 601 GTGCATGTGT GTGTATGTCT GCATGCATGT GATTGCATGT GTGCATGTGT ATACGTGTGC
 5 661 ATGTGTCCAT GTGTATGTGT GTGCATGTGT GTGTGCATCT GTGTATGCAT GTGTGTGCGT
 721 TTTGCATGTG TGTCTATGTA TGTGTCTGTG TGCATGTGCA TGTGTGTGCG TCTGTGTGTG
 781 CATTGTGTGT TATGTGTGTA TGCGTGTGTG TCTGTGGGTC TGTGTGTGCA TTTGTGTCTG
 841 TGCATGTGTG TATGCGTGTG TATGTATGTG TGTGCATCTG TGTGCATCTG TATGTGTGTG
 901 TGTGCGTCTG TCTGTGTGCA TGTGTGTATG CGTGTGTATG TATGTGTCTG CGTCTGTGTG
 10 961 TGGCTCTGTG TGTATCTGTG TGTGCATGTG TGTATGCGTG TATATGTATG TGCATCTGTG
 1021 TGCATGTGTC TATGTATGTG TGTGCATCTG TGTGTGCGTG TGTATGCGTG TGTGTATGCG
 1081 TCTGTGTGCA TGTGTGTATG CGTGTGTGTG CGCATCAGTG TCTGCATGTG TGTATATGTG
 1141 TGTATGTATG TGTGTGCGCG CGCATCTGTG TGTGTGCATG TGTGTATGTA TGTGTGTGCA
 1201 TCTCTGTGTG TGCATGTGTG TATGTGTGTG TGCATCTGTG TGTGTGTGCA TGTGTGCATG
 15 1261 CGTGTGTATG TGTGTGTGCG TGCGTGCATC TGTGTGTGTG CGCGTGTGCC CATGTGGGTG
 1321 CGACACTAGT GTGCATTTC ACATATACAT GTCCCGTCC GTCCATGTTT GCACATGGTG
 1381 GCAGATGGGG GGTGGCTGGG AGGCCTAGGC AGCCAATGAG CCTAGGTGTG CAGGCTCAGG
 1441 CCCAGAGAGT GATCCACCTT CCCCAGTCGG TTCAGGCCCT GATCAAAGAC TTCCAGGGGA
 1501 CCCCGACTTT CACCTACAAA GCGGCCATA TCTTCTTAC CGACACCTGC CCCGAGCCCC
 20 1561 TGTTCAAGTA GCTAGGCCG TCTCGTCTGG CAAAGGTGGT GAAGACGTTG AAGGAGATTC
 1621 ACCTTGCCCTT CCTCCCCTAC GAGGCCAGG TACGGCCCGG GCTCATCCTG GGCAGGGGGT
 1681 GGGGGTTTGT GACCAAATGT CCCCTGTTCC CTCAGAAACA GACACTGAGG CTGGGCGCAG
 1741 TGGCTCATAC CTGTAATCCC AGCGTGTGG GAGGCCAGG CAGGAGGATC ACTTGGGGCC
 1801 AGGAGTTTGA GACCAGCCTG GGTACAGAGC AAGACCCCGT CTCTTAAAAA AAAAAAAAAA
 25 1861 AAAAAAA

B: 氨基酸序列 (SEQ ID NO:8) 长度: 400

1 MCIRVHVMC MCVHVCVHLC MHVCAFMCV YVCVCHVHV CASVCAFVCM CVCVCVCGSV
 61 CAFVSVHVM RYVVCVICV HLYVCVCL CACVYACVCM CLRLCVRLCV SVCACVYACI
 30 121 CMCICVHVM YVCASVCAV CVCVCHVHV CMRVCAHQCL HVCICVYVCV CARICVCACV
 181 YVCVVISVCA CVYVCVHLCV CACVHACVVCV CACVHLCVCA CAHVGA TLVC ICTYTCPRPS
 241 MFAHGRWGV AGRPRQPM SLGVQAQQRVI HLPQSVQALI KDFQGTPTFT YKAAHIFFTD
 301 TCPEPLFSEL GRSRLAKVVK TLKEIHLAFL PYEAQVRPGL ILGRWGFVT KCPLFPQKQT
 361 LRLGVAHTC NPSAVGGQGR RITWQEFET SLGTEQDPVS
 35

C: 核苷酸及氨基酸组合序列 (SEQ ID NO:9) 克隆号: PP10122

起始编码子: 645 ATG 终止编码子: 1845 TAA 蛋白质分子量: 43382.43

1 GC AAA TCC ATT CTT AAA CCT TGC AAG CTG ACT GCG GGA GCC CAT GGA 47
 48 TTG CAG GCC TCC CCT GCA GCA CTC CCT GGC TGC AGC CCC TTC CTG ACC 95
 40 96 GTG CAG CTC CCT CGT GGA GCC CTG TGT GGC ACT CCC AGG GTG GAC TCT 143
 144 GGA TGT GGG GTC CCA AGA ATG CAG AGC CTT CTA TGG GTG ACC CCA GGT 191
 192 GGT GTC CCA GTC ACT TTG CCA TTG AGG GAT AGG TGC TGA GGG AGC CTG 239
 240 TCC TCT CAG AAC CTC AGT CTC CCT GTC TGT AAA TTT GGG GAA CCC AGT 287
 288 TGG CTG CAC AAC CCC TGC AGA ACC CTG TTG TGT GGC TCT TAG CTG GTG 335
 45 336 GTC CCT AAG TGG GTT TCT CCT GCA GTC CCT CCT AAG CAT CCA CCC CTC 383
 384 CTC CCC AAG CCT CCA ACC CCC CAC ACC TGA GAC TCC CCA CGT GGC CCT 431
 432 GCC TGG ATG CCA CCC ACC TGT GTC CCT TCC TCT GTT CCT ACT AGT TGT 479
 480 TGA AGA CAT CAA CAA ACG GCG GGA ACC CAT TCC CAG TCT GGA GGC CAT 527
 528 TTA TTT GCT GAG CCC CAC GGA GAA GGT GCC TAC ATG AGT GAG CGT GTG 575
 50 576 TGT ATG CGC GTG CAT GCG TGT ACA TGT GCA TGT GTG TGT ATG TCT GCA 623
 624 TGC ATG TGA TTG CAT GTG TGC ATG TGT ATA CGT GTG CAT GTG TCC ATG 671
 1 Met Cys Ile Arg Val His Val Ser Met 9
 672 TGT ATG TGT GTG CAT GTG TGT GTG CAT CTG TGT ATG CAT GTG TGT GCG 719
 55 10 Cys Met Cys Val His Val Cys Val His Leu Cys Met His Val Cys Ala 25

	720	TTT TGC ATG TGT GTC TAT GTA TGT GTC TGT GTG CAT GTG CAT GTG TGT	767
	26	Phe Cys Met Cys Val Tyr Val Cys Val Cys Val His Val His Val Cys	41
5	768	GCC TCT GTG TGT GCA TTT GTG TGT ATG TGT GTA TGC GTG TGT GTC TGT	815
	42	Ala Ser Val Cys Ala Phe Val Cys Met Cys Val Cys Val Cys Val Cys	57
	816	GGG TCT GTG TGT GCA TTT GTG TCT GTG CAT GTG TGT ATG CGT GTG TAT	863
10	58	Gly Ser Val Cys Ala Phe Val Ser Val His Val Cys Met Arg Val Tyr	73
	864	GTA TGT GTG TGC ATC TGT GTG CAT CTG TAT GTG TGT GTG TGC GTC TGT	911
	74	Val Cys Val Cys Ile Cys Val His Leu Tyr Val Cys Val Cys Val Cys	89
	912	CTG TGT GCA TGT GTG TAT GCG TGT GTA TGT ATG TGT CTG CGT CTG TGT	959
15	90	Leu Cys Ala Cys Val Tyr Ala Cys Val Cys Met Cys Leu Arg Leu Cys	105
	960	GTG CGT CTG TGT GTA TCT GTG TGT GCA TGT GTG TAT GCG TGT ATA TGT	1007
	106	Val Arg Leu Cys Val Ser Val Cys Ala Cys Val Tyr Ala Cys Ile Cys	121
20	1008	ATG TGC ATC TGT GTG CAT GTG TCT ATG TAT GTG TGT GCA TCT GTG TGT	1055
	122	Met Cys Ile Cys Val His Val Ser Met Tyr Val Cys Ala Ser Val Cys	137
	1056	GCC TGT GTA TGC GTG TGT GTA TGC GTC TGT GTG CAT GTG TGT ATG CGT	1103
25	138	Ala Cys Val Cys Val Cys Val Cys Val Cys Val His Val Cys Met Arg	153
	1104	GTG TGT GCG CAT CAG TGT CTG CAT GTG TGT ATA TGT GTG TAT GTA TGT	1151
	154	Val Cys Ala His Gln Cys Leu His Val Cys Ile Cys Val Tyr Val Cys	169
	1152	GTG TGC GCG CGC ATC TGT GTG TGT GCA TGT GTG TAT GTA TGT GTG TGC	1199
30	170	Val Cys Ala Arg Ile Cys Val Cys Ala Cys Val Tyr Val Cys Val Cys	185
	1200	ATC TCT GTG TGT GCA TGT GTG TAT GTG TGT GTG CAT CTG TGT GTG TGT	1247
	186	Ile Ser Val Cys Ala Cys Val Tyr Val Cys Val His Leu Cys Val Cys	201
35	1248	GCA TGT GTG CAT GCG TGT GTA TGT GTG TGT GCG TGC GTG CAT CTG TGT	1295
	202	Ala Cys Val His Ala Cys Val Cys Val Cys Ala Cys Val His Leu Cys	217
	1296	GTG TGC GCG TGT GCC CAT GTG GGT GCG ACA CTA GTG TGC ATT TGC ACA	1343
40	218	Val Cys Ala Cys Ala His Val Gly Ala Thr Leu Val Cys Ile Cys Thr	233
	1344	TAT ACA TGT CCC CGT CCG TCC ATG TTT GCA CAT GGT GGC AGA TGG GGG	1391
	234	Tyr Thr Cys Pro Arg Pro Ser Met Phe Ala His Gly Gly Arg Trp Gly	249
	1392	GTG GCT GGG AGG CCT AGG CAG CCA ATG AGC CTA GGT GTG CAG GCT CAG	1439
45	250	Val Ala Gly Arg Pro Arg Gln Pro Met Ser Leu Gly Val Gln Ala Gln	265
	1440	GCC CAG AGA GTG ATC CAC CTT CCC CAG TCG GTT CAG GCC CTG ATC AAA	1487
	266	Ala Gln Arg Val Ile His Leu Pro Gln Ser Val Gln Ala Leu Ile Lys	281
50	1488	GAC TTC CAG GGG ACC CCG ACT TTC ACC TAC AAA GCG GCC CAT ATC TTC	1535
	282	Asp Phe Gln Gly Thr Pro Thr Phe Thr Tyr Lys Ala Ala His Ile Phe	297
	1536	TTC ACC GAC ACC TGC CCC GAG CCC CTG TTC AGT GAG CTA GGC CGC TCT	1583
55	298	Phe Thr Asp Thr Cys Pro Glu Pro Leu Phe Ser Glu Leu Gly Arg Ser	313
	1584	CGT CTG GCA AAG GTG GTG AAG ACG TTG AAG GAG ATT CAC CTT GCC TTC	1631
	314	Arg Leu Ala Lys Val Val Lys Thr Leu Lys Glu Ile His Leu Ala Phe	329
	1632	CTC CCC TAC GAG GCC CAG GTA CGG CCC GGG CTC ATC CTG GGC AGG GGG	1679
60	330	Leu Pro Tyr Glu Ala Gln Val Arg Pro Gly Leu Ile Leu Gly Arg Gly	345
	1680	TGG GGG TTT GTG ACC AAA TGT CCC CTG TTC CCT CAG AAA CAG ACA CTG	1727
	346	Trp Gly Phe Val Thr Lys Cys Pro Leu Phe Pro Gln Lys Gln Thr Leu	361
65	1728	AGG CTG GGC GCA GTG GCT CAT ACC TGT AAT CCC AGC GCT GTG GGA GGC	1775
	362	Arg Leu Gly Ala Val Ala His Thr Cys Asn Pro Ser Ala Val Gly Gly	377
	1776	CAG GGC AGG AGG ATC ACT TGG GGC CAG GAG TTT GAG ACC AGC CTG GGT	1823
	378	Gln Gly Arg Arg Ile Thr Trp Gly Gln Glu Phe Glu Thr Ser Leu Gly	393
70	1824	ACA GAG CAA GAC CCC GTC TCT TAA AAA AAA AAA AAA AAA AAA AA	1867

394 Thr Glu Gln Asp Pro Val Ser ***

401

4. PP12744

A: 核苷酸序列 (SEQ ID NO:10) 长度: 3094

5 1 GAGACACGCG GGGCTTCAGG CTGCTGCCCC ATTGGAAGAT TACTCCCCAG GCTTCCCTTG
61 CCCCAGCAG TGAGCTGACT GGAATGGTAC CCCGGGAGGC CCCTGAGTCT GCTCAGTGCC
121 TGTGCCCTTC CCTCACCATC CCAAATGCCA AGGATGTGCT TCGGAAGAGG CACAAGAGAA
181 GGAGCCGACA GCACCAGCGG TTCATGGCCC GGAAGCCTTG CTGCAGGAGC AGGGGCTGCT
241 GAGCATGCCT CCAGAGCCAG GGTCTCCCC ACTGCCACC CCTTTCGGGG CAGCGACACG
10 301 AACTGAAGCT GCCAGCAGTG GGAAGCAGTG TCTGAGGGCT GGATCTGGCA GTGCCCCATG
361 CAGCAGAAGG CCTGCTCCCG GAAAGCCTC AGGGCCCTTG CCCAGCAAGT GTGTGGCTAT
421 CGACTGTGAG ATGGTGGGCA CGGGACCCCG AGGCGGGTAA GCGAGCTGGC CCGCTGTTC
481 ATTTGAGCT ACCATGGCGA TGTCTCTAT GACAAGTACA TCAGGCCGTA GATGCCCATC
541 GCTGACTACC GTACCCGCTG GAGTGGCATC ACTCGGCAGC ACATGCCGAA GGCTGTCCCC
15 601 TTCCAGGTGG CCCAGAAAGA GATCCTAAG CTCCTGAAGG GCAAGGTGGT GGTGGGGCAC
661 GCGTGCACA ACGACTTCCA GCGCTCAAG TATGTCCACC CTGGAGCCA GACCCGGGAT
721 ACCACCTATG TCCAAACTT CTCAGCGAG CCCGGCCTCC ACACCCGGGC CCGGTCTCT
781 CTAAGGACC TGGCCCTGCA GCTGCTGCAC AAGAAGATCC AGGTGGGCCA GCACGGGCAC
841 TCATCAGTAG AAGATGCCAC GACAGCCATG GAGCTCTACC GGCTGGTGA GGTGCAGTGG
20 901 GAACAGCAGG AGGCCCGCAG CCTCTGGACC TGCCCCAGG ACAGAGAACC TGACAGCAGC
961 ACAGACATGG AACAGTACAT GGAGGACCAG TACTGGCCCG ATGACCTGGC CCACGGCAGC
1021 AGAGGAGGAG CCAGGGAGGC ACAGGACAGA AGGAATTGAG AAGGGGGCGG GGCTCCCTGG
1081 CTGGGCTTCC GGTGTGGCCG GTAGGAAGTG GGGGCCAGGA GAGCAGCGGG CACTCCTTCC
1141 TGGGCAGGGT GGGGCAGGAT GCAGTGAGCC AGCCCCAGGG CCAGAGGAGT TGGGGTCATC
25 1201 TGTTACCTTG ACACCCTCTG CACACAGCAT AGCCCTCTCT CTCTCCAGGG CTGTTGGTTC
1261 TTTCTCCTGA CTCCTGTGGT TTTGCTAATG GCACTTTACA GACTCCATGG AGATGTCAGG
1321 TGGACCATCT TCTAGGGCCC AGCAGGAGTA GGAATGTGC CAACAGACTG CCCAGGTTGC
1381 CGTGGCCTTC CCCACCCCCC AGATCTCCTG AGTCATCATG CTGTGCTAAT GAAAGGGATC
1441 ATATCATCCT CTCTGGGGAT GGTGGGTGGG GGGTGTCAAT ATCCTGGAGC TCCCTTACCC
30 1501 CAACTCAATG ACTTGGGGGT AAAGTCTCT TCCTTTTGT GCCTACCTCT TCCTCCACTC
1561 ATTTGGGTTT AGAATAAACA TGCCCTGAAG TTAAGAGGA GTTAAGTCTT AAAGGAGGCA
1621 TTTCTTCCCC ACCTCCCTGA CCTGGAATC TGGCTACAGC CATTGTAGGA ACTCCGTGCC
1681 TGGCCTGTCA GCTCCCTGCT AGGCTACAGT GGAATAGCAG AGCCACAGG CTTCTCGTGG
1741 GGAGTTGGCT CCTTAACATT TCTTGGCAAC AGAAAGCCCC AGGCACAGCT CAGGGAGGAG
35 1801 GGAAGGCAGG TAAGCTTGG ACGAGAACTG GCATATTTAT TTTGACCCAA ATCAAGGGAT
1861 TTCCCCAGTC CACCCAGTAC TGGGCTCTTA AGCAAAAGTC TGAGAAACAA GACAGTGGTT
1921 TGAATCTGG GGCCTTTGTG TAGGATTGTG CCTGACCTTT ATTTATTTAT TAACAGCAGG
1981 GCCACTCGTC TAGGGCAGTG GAGTCTGCGT GTCTCCTGGG GCTGGGGCAG GGCATTGGCA
2041 GTTACGCAGT GGCCCTGAAC CTGGTCTGGT GCCCCGAAAC CTGGTCTGAT GCCCCCTCAG
40 2101 CTCTTTGACA ATCACTGTGG CTGTTGGGTT TTCTCCTATT TCTAAAAATG TTCTTTCTT
2161 TCCTAAGTGA CAGTTTGA GATTGGATA ACCAAGAGCT CAGGTACAC AGACCTTGGG
2221 GCCCATCTT TGCTTGCAGC TTAGCTTTGA GATACTAAGC AAGCGAGGGA CTTCACTCTT
2281 GAGCCTGTTT TTTCTCATC TGTAAGTGG GGATAGTGGT ACGGCCTACC TCATAGGGTT
2341 GTAATGAGCA CTAATGTGA AGTCTTGGC ACAGTGCTG GCACATGGTG AGCCACAGCT
45 2401 ACTGTGAGTT TCTGTTTATC CCCCTTTTTT TTTTTAAGC CCATTGTTTA TTCTTTGAGA
2461 ATTTGTTGTA ACTTCTTCAG GATAACACCT GAGTCCACAG GCTGAGCAGC TGTGGCCCAG
2521 ACAGAACTGC TCCGGCTTGG CTGTTCCAGC AGGTGGGGCG CTGGCCTCGG TGAGGGCACA
2581 GCAGCAAGGT TCACGGATAT CCGTGTGTCT TGCTGTGGC CACCAGGCAC AGGTTTGGCT
2641 TCCGGTCAGT GTCCCGACAC TGTGCGGGAG GTGACAACAG AGCAAAGCAG CGCAGGGGTC
50 2701 AGGGAGGTGG AGACTGCT GAAATCACAC TACCCACCC TCAGCTGAAG CCCACGTTT
2761 CACAACTTG GGGTCATAGA TTGTCCAGTC ACTGGCTCCC TCCCTGTCAG CACAGCACAG
2821 AGGAAGGGC TAACTGAATC TTTTACCCT TCTGGCCTGG CTCCAGAACT TTGTTCTAGA
2881 TTCCTTAAAA GTCGGTAGCT GATGTCAAAC TCAATTGAGC AGTAGCTTTG ATCCCTTGGT
2941 CTGGGGTTCG AAGGAAGATG GCGCTGTAT CAGCGGGGAA ATGTACTATT TAAGATCAGC

3001 TTTGTTGTAA AACCATTTGT TCTAGAATAA AACTCAATTG GAAACGTGAA AAAAAAAAAA
 3061 AAAAAAAAAA AAAAAAAAAAC TCGAGACTAG TTCT

B: 氨基酸序列 (SEQ ID NO:11) 长度: 175

5 1 MPIADYRTRW SGITRQHMRK AVPFQVAQKE ILKLLKGVV VGHALHNDQFQ ALKYVHPRSQ
 61 TRDTTYVPNF LSEPGLHTRA RVSLKDLALQ LLHKKIQVCQ HGHSSVEDAT TAMELYRLVE
 121 VQWEQEQEARS LWTCPEDREP DSSTDMEQYM EDQYWPDDLA HGSRRGGAREA QDRRN

C. 核苷酸及氨基酸组合序列 (SEQ ID NO:12) 克隆号: PP12744

10 起始编码子: 532 ATG 终止编码子: 1057 TGA 蛋白质分子量: 20287.83

	1	GAG ACA CGC GGC GCT TCA GGC TGC CCC ATT GGA AGA TTA CTC CCC	48
	49	AGG CTT CCC TTG CCC CAA GCA GTG AGC TGA CTG GAA TGG TAC CCC GGG	96
	97	AGG CCC CTG AGT CTG CTC AGT GCC TGT GCC CTT CCC TCA CCA TCC CAA	144
	145	ATG CCA AGG ATG TGC TTC GGA AGA GGC ACA AGA GAA GGA GCC GAC AGC	192
15	193	ACC AGC GGT TCA TGG CCC GGA AGC CTT GCT GCA GGA GCA GGG GCT GCT	240
	241	GAG CAT GCC TCC AGA GCC AGG GTC CTC CCC ACT GCC CAC CCC TTT CGG	288
	289	GGC AGC GAC ACG AAC TGA AGC TGC CAG CAG TGG GAA GCA GTG TCT GAG	336
	337	GGC TGG ATC TGG CAG TGC CCC ATG CAG CAG AAG GCC TGC TCC CGG GAA	384
	385	AGC CTC AGG GCC CTT GCC CAG CAA GTG TGT GGC TAT CGA CTG TGA GAT	432
20	433	GGT GGG CAC GGG ACC CCG AGG CGG GTA AGC GAG CTG GCC CGC TGT TCC	480
	481	ATT GTG AGC TAC CAT GGC GAT GTC CTC TAT GAC AAG TAC ATC AGG CCT	528
	529	GAG ATG CCC ATC GCT GAC TAC CGT ACC CGC TGG AGT GGC ATC ACT CGG	576
	1	Met Pro Ile Ala Asp Tyr Arg Thr Arg Trp Ser Gly Ile Thr Arg	15
25	577	CAG CAC ATG CGC AAG GCT GTC CCC TTC CAG GTG GCC CAG AAA GAG ATC	624
	16	Gln His Met Arg Lys Ala Val Pro Phe Gln Val Ala Gln Lys Glu Ile	31
	625	CTT AAG CTC CTG AAG GGC AAG GTG GTG GTG GGG CAC GCG CTG CAC AAC	672
30	32	Leu Lys Leu Leu Lys Gly Lys Val Val Val Gly His Ala Leu His Asn	47
	673	GAC TTC CAG GCG CTC AAG TAT GTC CAC CCT CGG AGC CAG ACC CGG GAT	720
	48	Asp Phe Gln Ala Leu Lys Tyr Val His Pro Arg Ser Gln Thr Arg Asp	63
	721	ACC ACC TAT GTC CCA AAC TTC CTC AGC GAG CCC GGC CTC CAC ACC CGG	768
35	64	Thr Thr Tyr Val Pro Asn Phe Leu Ser Glu Pro Gly Leu His Thr Arg	79
	769	GCC CGG GTC TCT CTA AAG GAC CTG GCC CTG CAG CTG CTG CAC AAG AAG	816
	80	Ala Arg Val Ser Leu Lys Asp Leu Ala Leu Gln Leu Leu His Lys Lys	95
40	817	ATC CAG GTG GGC CAG CAC GGG CAC TCA TCA GTA GAA GAT GCC ACG ACA	864
	96	Ile Gln Val Gly Gln His Gly His Ser Ser Val Glu Asp Ala Thr Thr	111
	865	GCC ATG GAG CTC TAC CGG CTG GTG GAG GTG CAG TGG GAA CAG CAG GAG	912
45	112	Ala Met Glu Leu Tyr Arg Leu Val Glu Val Gln Trp Glu Gln Gln Glu	127
	913	GCC CGC AGC CTC TGG ACC TGC CCC GAG GAC AGA GAA CCT GAC AGC AGC	960
	128	Ala Arg Ser Leu Trp Thr Cys Pro Glu Asp Arg Glu Pro Asp Ser Ser	143
	961	ACA GAC ATG GAA CAG TAC ATG GAG GAC CAG TAC TGG CCC GAT GAC CTG	1008
50	144	Thr Asp Met Glu Gln Tyr Met Glu Asp Gln Tyr Trp Pro Asp Asp Leu	159
	1009	GCC CAC GGC AGC AGA GGA GGA GCC AGG GAG GCA CAG GAC AGA AGG AAT	1056
	160	Ala His Gly Ser Arg Gly Gly Ala Arg Glu Ala Gln Asp Arg Arg Asn	175
55	1057	TGA GAA GGG GGC GGG GCT CCC TGG CTG GGC TTC CGG TGT GGC CGG TAG	1104
	176	***	176
	1105	GAA GTG GGG GCC AGG AGA GCA GCG GGC ACT CCT TCC TGG GCA GGG TGG	1152
	1153	GGC AGG ATG CAG TGA GCC AGC CCC AGG GCC AGA GGA GTT GGG GTC ATC	1200
60	1201	TGT TAC CTT GAC ACC CTC TGC ACA CAG CAT AGC CCT CTC TCT CTC CAG	1248
	1249	GGC TGT TGG TTC TTT CTC CTG ACT CCT GTG GTT TTG CTA ATG GCA CTT	1296
	1297	TAC AGA CTC CAT GGA GAT GTC AGG TGG ACC ATC TTC TAG GGC CCA GCA	1344

	1345	GGA GTA GGG AAT GTG CCA ACA GAC TGC CCA GGT TGC CGT GGC CTT CCC	1392
	1393	CAC CCC CCA GAT CTC CTG AGT CAT CAT GCT GTG CTA ATG AAA GGG ATC	1440
	1441	ATA TCA TCC TCT CTG GGG ATG GTG GGT GGG GGG TGT CAA TAT CCT GGA	1488
	1489	GCT CCC TTA CCC CAA CTC AAT GAC TTG GGG GTA AAG TTC TCT TCC TTT	1536
5	1537	TGT TGC CTA CCT CTT CCT CCA CTC ATT TGG GTT CAG AAT AAA CAT GCC	1584
	1585	CTG AAG TTA AAG AGG AGT TAA GTC CTA AAG GAG GCA TTT CTT CCC CAC	1632
	1633	CTC CCT GAC CTG GAA CTC TGG CTA CAG CCA TTG TAG GAA CTC CGT GCC	1680
	1681	TGG CCT GTC AGC TCC CTG CTA GGC TAC AGT GGA ATA GCA GAG CCC ACA	1728
	1729	GGC TTC TCG TGG GGA GTT GGC TCC TTA ACA TTT CTT GGC AAC AGA AAG	1776
10	1777	CCC CAG GCA CAG CTC AGG GAG GAG GGA AGG CAG GTA AGC TTT GGA CGA	1824
	1825	GAA CTG GCA TAT TTA TTT TGA CCC AAA TCA AGG GAT TTC CCC AGT CCA	1872
	1873	CCC AGT ACT GGG CTC TTA AGC AAA AGT CTG AGA AAC AAG ACA GTG GTT	1920
	1921	TGA ATT CTG GGG CCT TTG TGT AGG ATT GTG CCT GAC CTT TAT TTA TTT	1968
	1969	ATT AAC AGC AGG GCC ACT CGT CTA GGG CAG TGG AGT CTG CGT GTC TCC	2016
15	2017	TGG GGC TGG GGC AGG GCA TTG GCA GTT ACG CAG TGG CCC TGA ACC TGG	2064
	2065	TCT GGT GCC CCC GAA CCT GGT CTG ATG CCC CCT CAG CTC TTT GAC AAT	2112
	2113	CAC TGT GGC TGT TGG GTT TTC TCC TAT TTC TAA AAA TGT TCT CTT CTT	2160
	2161	TCC TAA GTG ACA GTT TTG AAG TAT TGG ATA ACC AAG AGC TCA GGT CAC	2208
	2209	ACA GAC CTT GGA GCC CCA TCT TTG CTT GCA GCT TAG CTT TGA GAT ACT	2256
20	2257	AAG CAA GCG AGG GAC TTC ACT CTT GAG CCT GTT TTT TCC TCA TCT GTA	2304
	2305	AAG TGG GGA TAG TGG TAC GGC CTA CCT CAT AGG GTT GTA ATG AGC ACT	2352
	2353	AAA TGT GAA GTG CTT GGC ACA GTG CCT GGC ACA TGG TGA GCC ACA GCT	2400
	2401	ACT GTG AGT TTC TGT TTA TCC CCC TTT TTT TTT TTT AAG CCC ATT GTT	2448
	2449	TAT TCT TTG AGA ATT TGT TGT AAC TTC TTC AGG ATA ACA CCT GAG TCC	2496
25	2497	ACA GGC TGA GCA GCT GTG GCC CAG ACA GAA CTG CTC CGG CTT GGC TGT	2544
	2545	TCC AGC AGG TGG GGC GCT GGC CTC GGT GAG GGC ACA GCA GCA AGG TTC	2592
	2593	ACG GAT ATC CGT GTG TCT TGT CTG TGG CCA CCA GGC ACA GGT TTG GCT	2640
	2641	TCC GGT CAG TGT CCC GAC ACT GTG CGG GAG GTG ACA ACA GAG CAA AGC	2688
	2689	AGC GCA GGG GTC AGG GAG GTG GAG ACA CTG CTG AAA TCA CAC TAC CCC	2736
30	2737	ACC CTC AGC TGA AGC CCC ACG TTC CAC AAA CTT GGG GTC ATA GAT TGT	2784
	2785	CCA GTC ACT GGC TCC CTC CCT GTC AGC ACA GCA CAG AGG AAG GGG CTA	2832
	2833	ACT GAA TCT TTT ACC ACT TCT GGC CTG GCT CCA GAA CTT TGT TCT AGA	2880
	2881	TTC CTT AAA AGT CGG TAG CTG ATG TCA AAC TCA ATT GAG CAG TAG CTT	2928
	2929	TGA TCC CTT GGT CTG GGG GTC GAA GGA AGA TGG CGC TGT TAT CAG CGG	2976
35	2977	GGA AAT GTA CTA TTT AAG ATC AGC TTT GTT GTA AAA CCA TTT GTT CTA	3024
	3025	GAA TAA AAC TCA ATT GGA AAC GTG AAA AAA AAA AAA AAA AAA AAA	3072
	3073	AAA AAA ACT CGA GAC TAG TTC T	3094

5. PP13624

40	A: 核苷酸序列 (SEQ ID NO:13) 长度: 2140
	1 GCGGAGCCGG GCTCACCAGG TCGCTGCTGC GAGGGAGTTG CTGTGCTGGG GCCTGGGTGG
	61 CGGCTGGAGG CCTGAGTTGG GCTCGCGGCG GGGGTCGGCA GGGGCGGGG TGGCGGAATG
	121 ATGGAGGAGG AGGAACTGGA GTTCGTGGAG GAGCTGGAAG CCGTGCTGCA GCTCAGCCCC
	181 GAGGTGCAGC TGGCCATCGA GCAGGTGTTT CCAAGCCAGG ACCCTCTAGA TCGAGCAGAT
45	241 TTCAATGCTG TTGAGTATAT CAATACCCTG TTCCCAACCG AGCAATCTCT GGCGAACATA
	301 GACGAAGTCG TGAACAAAAT TAGGCTGAAA ATAAGGAGAC TGGATGACAA TATTCGAACT
	361 GTTGTAAAGG GTCAGACGAA CGTGGGGCAG GATGGACGGC AAGCGTTGA AGAGGCTCAG
	421 AAAGCTATCC AACAACTCTT TGGCAAAATC AAAGATATCA AAGACAAAGC TGA AAAATCA
	481 GAGCAAATGG TGA AAGAAAT CACCCGTGAT ATTAAGCAAT TAGATCACGC CAAACGCCAC
50	541 CTGACCACCT CAATACCAC ACTGAACCAC CTGCACATGC TGGCAGGAGG TGTGACTCC
	601 CTCGAAGCCA TGACCAGGCG AAGACAATAC GGAGAAGTTG CTAATCTCCT TCAGGGTGTG
	661 ATGAATGTCC TGGAGCACTT CCACAAGTAT ATGGGGATTC CGCAGATCCG GCAGCTTTC
	721 GAAAGAGTGA AGGCTGCACA GACTGAGTTA GGACAGCAA TCCTGGCAGA TTTTGAAGAA
	781 GCGTTTCCTT CCCAGGGCAC CAAGAGACCA GGAGGACCCA GCAATGTTCT ACGAGATGCA

841 TGTCTGCTTG CTAATATTCT AGATCCCAGG ATCAAACAGG AAATCATCAA AAAGTTTATT
 901 AAACAGCATC TGTCAGAGTA TCTGGTACTT TTCAAGAAA ACCAAGATGT TGCCTGGCTG
 961 GACAAAATCG ACAGACGCTA TGCCTGGATA AAACGCCAGC TTGTGGACTA TGAGGAGAAA
 1021 TACGGCCGCA TGTTCACAG TGAGTGGTGC ATGGCTGAGA GGATTGCGGT GGAATTTTGC
 5 1081 CATGTGACAA GGGCAGAACT TGCCAAGATT ATGCGTACCA GAGCGAAGGA AATTGAAGTG
 1141 AAATTGCTTC TTTTGTCTAT TCAAAGAACA ACTAACTTTG AGGGTTTCT TGCAAAACGC
 1201 TTCTCCGGCT GCACCCTGAC CGATGGGACC CTGAAAAAGC TTGAGTCTCC ACCCCCATCT
 1261 ACCAATCCCT TCCTGGAAGA TGAGCCAACA CCAGAGATGG AGGAACTGGC AACGGAGAAA
 1321 GGAGATTAG ATCAACCAAA GAAGCCTAAA GCCCCAGACA ATCCATTTCA TGGCATTGTT
 10 1381 TCCAAGTGT TTGAGCCTCA TCTCTACGTG TATATCGAAT CCCAAGACAA GAACCTCGGA
 1441 GAGCTGATAG ATCGGTTTGT GGCTGATTTT AAAGCCCAGG GGCCACCTAA GCCCAACACT
 1501 GATGAAGGGG GTGCCGTGCT CCCAGCTGC GCCGACCTCT TTGTCTACTA CAAGAAGTGC
 1561 ATGGTGCAAT GCTCTCAGCT CAGTACTGGG GAGCCCATGA TCGCCCTGAC CACCATTTTC
 1621 CAGAAGTACC TCCGAGAATA TGCCTGGAAA ATCCTCTCTG GCAACCTGCC CAAAACCACA
 15 1681 ACCAGCAGTG GAGGACTGAC TATCAGCAGC CTCCTCAAGG AAAAGGAGGG CTCAGAAGTA
 1741 GCCAAGTTCA CTCTGGAGGA GCTCTGCCTC ATCTGTAACA TCCTGAGCAC GGCAGAGTAC
 1801 TTGTCTGGCC ACCACCCAGC AGCTAGAAGA AAAACTCAAA GAAAAAGTGG ATGTAAGTCT
 1861 GATTGAACGA ATCAATCTGA CTGGAGAGAT GGACACGTTT AGCACCGTCA TCTCCAGCAG
 1921 TATTCAGCTG CTGGTTCAGG ATCTGGATGC TGCCTGTGAT CCTGCCCTGA CTGCCATGAG
 20 1981 CAAGATGCAG TGGCAGAACG TGGAGCACGT TGGTGACCAG AGCCCTACG TCACCTCTGT
 2041 CATTCTGCAC ATCAAGCAGA ACGTCCCAT CATCCGTGAC AACCTGGCCA ACATGGCAAA
 2101 ATCCCATCTC TACTAAAAAT ACAAAAAAAA AAAAAAAA

B: 氨基酸序列 (SEQ ID NO:14) 长度: 582

25 1 MMEEELEFV EEEAVLQLT PEVQLAIEQV FPSQDPLDRA DFNAVEYINT LFPTEQSLAN
 61 IDEVVKIRL KIRRLDDNIR TVVRGQTNVG QDGRQALEEA QKAIQQLFGK IKDIKDKAEK
 121 SEQMVKETR DIKQLDHAKR HLTTSITTLN HLHMLAGGVD SLEAMTRRRQ YGEVANLLQG
 181 VMNVLEHFHK YMGIPQIRQL SERVKAQTE LGQQILADFE EAFPSQGTKR PGGPSNVLRD
 241 ACLVANILDP RIKQEIIKF IKQHLSEYLV LQENQDVAW LDKIDRRYAW IKRQLVDYEE
 30 301 KYGRMPREW CMAERIAVEF CHVTRAEKAK IMRTRAKEIE VKLLFAIQR TTNFEGFLAK
 361 RFSGCTLTDG TLKKLESPPP STNPFLEDEP TPEMEELATE KGDLDPKPKP KAPDNPFHGI
 421 VSKCFEPHLY VYIESQDKNL GELIDRFVAD FKAQGPFPKN TDEGAVLPS CADLFFVYK
 481 CMVQCSQLST GEPMIALTTI FQKYLREYAW KILSGNLPKT TTSSGGLTIS SLLKEKEGSE
 541 VAKFTLEELC LICNILSTAE YLSGHHPAAR RKTQRKSGCK SD

C. 核苷酸及氨基酸组合序列 (SEQ ID NO:15) 克隆号: PP13624

起始编码子: 118 ATG 终止编码子: 1864 TGA 蛋白质分子量: 66318.71

1 GCG GAG CCG GGC TCA CCA GGT CGC TGC TGC GAG GGA GTT GCT GTG CTG 48
 49 GGG CCT GGG TGG CGG CTG GAG GCC TGA GTT GGG CTC GCG GCG GGG GTC 96
 40 97 GGC AGG GGG CCG GGT GGC GGA ATG ATG GAG GAG GAG GAA CTG GAG TTC 144
 1 Met Met Glu Glu Glu Glu Leu Glu Phe 9
 145 GTG GAG GAG CTG GAA GCC GTG CTG CAG CTC ACG CCC GAG GTG CAG CTG 192
 45 10 Val Glu Glu Leu Glu Ala Val Leu Gln Leu Thr Pro Glu Val Gln Leu 25
 193 GCC ATC GAG CAG GTG TTT CCA AGC CAG GAC CCT CTA GAT CGA GCA GAT 240
 26 Ala Ile Glu Gln Val Phe Pro Ser Gln Asp Pro Leu Asp Arg Ala Asp 41
 241 TTC AAT GCT GTT GAG TAT ATC AAT ACC CTG TTC CCA ACC GAG CAA TCT 288
 50 42 Phe Asn Ala Val Glu Tyr Ile Asn Thr Leu Phe Pro Thr Glu Gln Ser 57
 289 CTG GCG AAC ATA GAC GAA GTC GTG AAC AAA ATT AGG CTG AAA ATA AGG 336
 58 Leu Ala Asn Ile Asp Glu Val Val Asn Lys Ile Arg Leu Lys Ile Arg 73
 55 337 AGA CTG GAT GAC AAT ATT CGA ACT GTT GTA AGA GGT CAG ACG AAC GTG 384
 74 Arg Leu Asp Asp Asn Ile Arg Thr Val Val Arg Gly Gln Thr Asn Val 89
 385 GGG CAG GAT GGA CGG CAA GCG CTT GAA GAG GCT CAG AAA GCT ATC CAA 432

	90	Gly Gln Asp Gly Arg Gln Ala Leu Glu Glu Ala Gln Lys Ala Ile Gln	105
	433	CAA CTC TTT GGC AAA ATC AAA GAT ATC AAA GAC AAA GCT GAA AAA TCA	480
5	106	Gln Leu Phe Gly Lys Ile Lys Asp Ile Lys Asp Lys Ala Glu Lys Ser	121
	481	GAG CAA ATG GTG AAA GAA ATC ACC CGT GAT ATT AAG CAA TTA GAT CAC	528
	122	Glu Gln Met Val Lys Glu Ile Thr Arg Asp Ile Lys Gln Leu Asp His	137
	529	GCC AAA CGC CAC CTG ACC ACC TCA ATC ACC ACA CTG AAC CAC CTG CAC	576
10	138	Ala Lys Arg His Leu Thr Thr Ser Ile Thr Thr Leu Asn His Leu His	153
	577	ATG CTG GCA GGA GGT GTC GAC TCC CTC GAA GCC ATG ACC AGG CGA AGA	624
	154	Met Leu Ala Gly Gly Val Asp Ser Leu Glu Ala Met Thr Arg Arg Arg	169
15	625	CAA TAC GGA GAA GTT GCT AAT CTC CTT CAG GGT GTG ATG AAT GTC CTG	672
	170	Gln Tyr Gly Glu Val Ala Asn Leu Leu Gln Gly Val Met Asn Val Leu	185
	673	GAG CAC TTC CAC AAG TAT ATG GGG ATT CCG CAG ATC CGG CAG CTT TCC	720
20	186	Glu His Phe His Lys Tyr Met Gly Ile Pro Gln Ile Arg Gln Leu Ser	201
	721	GAA AGA GTG AAG GCT GCA CAG ACT GAG TTA GGA CAG CAA ATC CTG GCA	768
	202	Glu Arg Val Lys Ala Ala Gln Thr Glu Leu Gly Gln Gln Ile Leu Ala	217
	769	GAT TTT GAA GAA GCG TTT CCT TCC CAG GGC ACC AAG AGA CCA GGA GGA	816
25	218	Asp Phe Glu Glu Ala Phe Pro Ser Gln Gly Thr Lys Arg Pro Gly Gly	233
	817	CCC AGC AAT GTT CTA CGA GAT GCA TGT CTG GTT GCT AAT ATT CTA GAT	864
	234	Pro Ser Asn Val Leu Arg Asp Ala Cys Leu Val Ala Asn Ile Leu Asp	249
30	865	CCC AGG ATC AAA CAG GAA ATC ATC AAA AAG TTT ATT AAA CAG CAT CTG	912
	250	Pro Arg Ile Lys Gln Glu Ile Ile Lys Lys Phe Ile Lys Gln His Leu	265
	913	TCA GAG TAT CTG GTA CTT TTT CAA GAA AAC CAA GAT GTT GCC TGG CTG	960
35	266	Ser Glu Tyr Leu Val Leu Phe Gln Glu Asn Gln Asp Val Ala Trp Leu	281
	961	GAC AAA ATC GAC AGA CGC TAT GCC TGG ATA AAA CGC CAG CTT GTG GAC	1008
	282	Asp Lys Ile Asp Arg Arg Tyr Ala Trp Ile Lys Arg Gln Leu Val Asp	297
	1009	TAT GAG GAG AAA TAC GGC CGC ATG TTT CCA CGT GAG TGG TGC ATG GCT	1056
40	298	Tyr Glu Glu Lys Tyr Gly Arg Met Phe Pro Arg Glu Trp Cys Met Ala	313
	1057	GAG AGG ATT GCG GTG GAA TTT TGC CAT GTG ACA AGG GCA GAA CTT GCC	1104
	314	Glu Arg Ile Ala Val Glu Phe Cys His Val Thr Arg Ala Glu Leu Ala	329
45	1105	AAG ATT ATG CGT ACC AGA GCG AAG GAA ATT GAA GTG AAA TTG CTT CTT	1152
	330	Lys Ile Met Arg Thr Arg Ala Lys Glu Ile Glu Val Lys Leu Leu Leu	345
	1153	TTT GCT ATT CAA AGA ACA ACT AAC TTT GAG GGG TTT CTT GCA AAA CGC	1200
50	346	Phe Ala Ile Gln Arg Thr Thr Asn Phe Glu Gly Phe Leu Ala Lys Arg	361
	1201	TTC TCC GGC TGC ACC CTG ACC GAT GGG ACC CTG AAA AAG CTT GAG TCT	1248
	362	Phe Ser Gly Cys Thr Leu Thr Asp Gly Thr Leu Lys Lys Leu Glu Ser	377
	1249	CCA CCC CCA TCT ACC AAT CCC TTC CTG GAA GAT GAG CCA ACA CCA GAG	1296
55	378	Pro Pro Pro Ser Thr Asn Pro Phe Leu Glu Asp Glu Pro Thr Pro Glu	393
	1297	ATG GAG GAA CTG GCA ACG GAG AAA GGA GAT TTA GAT CAA CCA AAG AAG	1344
	394	Met Glu Glu Leu Ala Thr Glu Lys Gly Asp Leu Asp Gln Pro Lys Lys	409
60	1345	CCT AAA GCC CCA GAC AAT CCA TTT CAT GGC ATT GTT TCC AAG TGT TTT	1392
	410	Pro Lys Ala Pro Asp Asn Pro Phe His Gly Ile Val Ser Lys Cys Phe	425
	1393	GAG CCT CAT CTC TAC GTG TAT ATC GAA TCC CAA GAC AAG AAC CTC GGA	1440
65	426	Glu Pro His Leu Tyr Val Tyr Ile Glu Ser Gln Asp Lys Asn Leu Gly	441
	1441	GAG CTG ATA GAT CGG TTT GTG GCT GAT TTC AAA GCC CAG GGG CCA CCT	1488
	442	Glu Leu Ile Asp Arg Phe Val Ala Asp Phe Lys Ala Gln Gly Pro Pro	457
	1489	AAG CCC AAC ACT GAT GAA GGG GGT GCC GTG CTC CCC AGC TGC GCC GAC	1536
70	458	Lys Pro Asn Thr Asp Glu Gly Gly Ala Val Leu Pro Ser Cys Ala Asp	473

	1537	CTC TTT GTC TAC TAC AAG AAG TGC ATG GTG CAA TGC TCT CAG CTC AGT	1584
	474	Leu Phe Val Tyr Tyr Lys Lys Cys Met Val Gln Cys Ser Gln Leu Ser	489
5	1585	ACT GGG GAG CCC ATG ATC GCC CTG ACC ACC ATT TTC CAG AAG TAC CTC	1632
	1490	Thr Gly Glu Pro Met Ile Ala Leu Thr Thr Ile Phe Gln Lys Tyr Leu	505
	1633	CGA GAA TAT GCC TGG AAA ATC CTC TCT GGC AAC CTG CCC AAA ACC ACA	1680
	506	Arg Glu Tyr Ala Trp Lys Ile Leu Ser Gly Asn Leu Pro Lys Thr Thr	521
10	1681	ACC AGC AGT GGA GGA CTG ACT ATC AGC AGC CTC CTC AAG GAA AAG GAG	1728
	522	Thr Ser Ser Gly Gly Leu Thr Ile Ser Ser Leu Leu Lys Glu Lys Glu	537
	1729	GGC TCA GAA GTA GCC AAG TTC ACT CTG GAG GAG CTC TGC CTC ATC TGT	1776
15	538	Gly Ser Glu Val Ala Lys Phe Thr Leu Glu Glu Leu Cys Leu Ile Cys	553
	1777	AAC ATC CTG AGC ACG GCA GAG TAC TTG TCT GGC CAC CAC CCA GCA GCT	1824
	554	Asn Ile Leu Ser Thr Ala Glu Tyr Leu Ser Gly His His Pro Ala Ala	569
20	1825	AGA AGA AAA ACT CAA AGA AAA AGT GGA TGT AAG TCT GAT TGA ACG AAT	1872
	570	Arg Arg Lys Thr Gln Arg Lys Ser Gly Cys Lys Ser Asp ***	583
	1873	CAA TCT GAC TGG AGA GAT GGA CAC GTT CAG CAC CGT CAT CTC CAG CAG	1920
	1921	TAT TCA GCT GCT GGT TCA GGA TCT GGA TGC TGC CTG TGA TCC TGC CCT	1968
	1969	GAC TGC CAT GAG CAA GAT GCA GTG GCA GAA CGT GGA GCA CGT TGG TGA	2016
25	2017	CCA GAG CCC CTA CGT CAC CTC TGT CAT TCT GCA CAT CAA GCA GAA CGT	2064
	2065	CCC CAT CAT CCG TGA CAA CCT GGC CAA CAT GGC AAA ATC CCA TCT CTA	2112
	2113	CTA AAA ATA CAA AAA AAA AAA AAA A	2140

6. PP13671

30

A: 核苷酸序列 (SEQ ID NO:16) 长度: 2986

	1	GGGAGGACCC GCGACACAGC TCGGGATCCG CGCGCGCTGG GCCTTCGGCG CCGCTCGTGG
	61	GGTTCGGGAG GCCGGGACCC TGCCCGGGAG ATGTGGACGC CGGGCGGACC CCCGGGGTCC
	121	GCGGGCTGGG ACCGCCGTAG GTTGGGCGCG AGGTTGCCGG CGGCGTTCCG GGGGCTGCAG
	181	GAGCTGCAGG GGCTGCGAGC CACGACGAGC GAGCGGTAC GGGGCGCCCT GGCCCTGCAG
35	241	CCCCCGCCG CGCCCGCCG GCCCTGCGGC CCCACGGCC TCCACGGCCC CGAGCAGCAG
	301	CTGGAGGCGG CGCTGGCCCG GCTGCAGGAG CAGCTGTCCC GGCTGAGACA ACAGGACATC
	361	GGCCTGAAGA CCCACCTGGA CCAGCTGGAC CTGCAGATTA GCAAGCTGCA GCTGGATGTG
	421	GGCACAGCCT CAGGGGAGGC CCTGGACAGC GACAGCAGGC CCAGCTCAGG CTTTTACGAG
	481	ATGAGCGACG GTGGATCCTG CTCCTGTCC ACCTCCTGTG CCTCCGTCTG CAGTGACCAC
40	541	ATCTCTCCCT CGCTGGGCAG TTTGTTGCCT GTGGCCAGG CCCACAAGGC CAGGCCCAGC
	601	ATGGGGGACT GGAGGCCCGG GTCGGTTGAT GAGACTACTG TGCCAGCGTG GAGACCCAG
	661	GCTACCGAGG AGGGCCCCAG GCCCCAGGG AGCGTGAGG ATGCAGGCCA GCCGTGGGGC
	721	ACATTCTGGC CCAGGCCTGT GTCTACAGGT GATCTTGACA GAGCCCTGCC GCGGACACG
	781	GGGCTCCAGA AAGCCAGCGC GGACGCCGAG CTCCTCGGGC TCCTCTGCAG GGGGTGGATA
45	841	TCCCGCTGCA CGTGCCGGAC CCCAAGTATC GGCAGGACCT GGTGTCCAG GCGGCAGGG
	901	AGGTGTACCC GTACCCAGC CCCCTGCACG CCGTGGCTCT ACAGAGCCCC CTGTTGTGCC
	961	TGACTAAGGA AACCCACAG AGAGGTGGCC CTCGTTCCC TAGGGAGAGC CCCAGGGGCC
	1021	CCGCAGGTCT GAACACCATC CAGACTGGGC CGTCCTCGA GGCTGGCCCG GCCAGGGCCA
	1081	GAGCTTATAT CGACAGGTTG CTGCATCTGT GGGGCCGGGA GACCCAGCA AAGGTTAGCG
50	1141	AGGGAGAACA GGGACCTCTA AGGCATGCAG CGTCCCATC TCCACAGAGG CAGGTTGGCT
	1201	GGAGTACAGA CCGTGGAGGG CCACTGCTGG TCTTCGCCCC AGGGAGGGAG GACGAGGGAG
	1261	GGCCAGCTCA GAGCAGGGGT GCCGGCAGGG GCGGGCCCCA GCAGCAGGGA TACATGCCCC
	1321	TTGAGGTCC CCAGCAGTCT GGCAGCCTTC CAGAGGAGGG CTCCAAGCCC TCAAACAGCT
	1381	GTGTCCTCAG GGAGACCATG GTGCAGGCTT CTCCCAGCTC AAAGGCCAG CAGACACCCT
55	1441	CAGCTCAGGA CTATGGACGA GGCAACATCA TATCCCATC CAGGATGCTG GACAAGAGCC
	1501	CCTCACCGGC CTCTGGGCAC TTTGCCACC CATCCTTTGC TGCCAGCCTG AAAATGGGTC
	1561	CCCCAAGAG CAAGGCTGAA AAAATCAAGA GAAGTCCAT GGACAAGGTG CTGAGGTTTG
	1621	CAAGGCAGCC GCTGCTTCTA CTGACAGGC CTGAGGGAGC CCATGCAGCC CCCAGCCAT
	1681	CCCTGGAGTG GGACCCTGCC CACTGGCCCA CAGGGAGGG GGGCTCCAG CGGAGGCCAG

1741 CCCTGGCCTG GGAGGCACCC GGGCGCTCCT GTTCTGAGTC CACCCTCTAC CCCATGCCTG
 1801 TCCTCGTCCC CTGGCAGTG GCCCCGAGG AGAGCCACCG GACCTCAGCC CAAGCCCTGT
 1861 TCCCCTTTGA GGGCTCACTG CTCACCTCAG TGGCCAGGAG GAAGCATCGC CGCTGGCAGT
 1921 CCACCGTGA GATCTCGGCC CGGGCCCGCC TGGCCAGCTG TCCTGAGTCT AACCTGGGGC
 5 1981 CCCCAGGCC CGTGGCCAGG AGAGCAGGTG GCCCACTGGC CCGGGGCCGT CCCTCACTGG
 2041 TCCGCCAGGA CGCCTACACC AGGAGCGACT CAGAGCCCTC CAAGCACTCG GCCGAGTGTG
 2101 ACCCGCGGTT CCCGTCAGTC ATCCCGGAGA CCAGCGAGGG AGAGTCCAGT GACCACACCA
 2161 CCAACCGATT CGGAGACCGT GAGTCCAGCA GCAGCGACGA GGAGGGCGGC GCCCAGAGCA
 2221 GGGACTGTGA CCTGGCACTG GGCTATGTGG CGGCCGGGCA TGAGGAGCTG GCCTGGACCC
 10 2281 AGGAGGCCCC GGTCACTCG GGGCACTCC TGCCCCCGT GCCCAAGCTG TGCCGTATTA
 2341 AGGCCTCAA GGCCTGAAG AAGAAGATCC GCAGTTCCA GCCGACGGCC CTGAAGGTCA
 2401 TGACCATGGT GTGAGGTGCA GTGACTGCGT CAAGAGAAGC CTGCTGCACA CAGGGGCCCT
 2461 TTCCATGCAG TGTCCTGTGC ATGTCTGGGT TTCAGTGTG TCTTGCCAAG TCGCTTTATC
 2521 TTCCGAGGCC ACCGTTTAAC AGGAGTCCCC GCAGATGGT TCTGCCCTGT GGAGTATGGC
 15 2581 CGTGAGTCCC CGCCCCAGCA AAGACTTCCA TCGGAGCTC TGGTTTCAGA TGCCAGCAGC
 2641 CTCCCCGGCC GACCTAGTGT CTCTTTTCAA GGTCTGCCA GGATGCTGCT TCCCGCATG
 2701 CACCTTGTTT CCATGTTGTC ATCCTCCCAT GGCTTCCCT GCCCTCACTC CCCTCTGGCT
 2761 GGGGAAGTGT TTCCAATCC TTTGACCATT CCTCCCCAAA AGCATTTTAA AAAGCGGTT
 2821 GCAGGACTTT GAGCTGAAA TTGAATTTGA GTTAAACTC CACGTTAAGC TGTCAGCATA
 20 2881 CACAAACACC CAAATTATGC CATGATCTCA AATAAAAGGA ACATTTTTTT TCCCTCTTTT
 2941 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAA

B: 氨基酸序列 (SEQ ID NO:17) 长度: 366

1 MPLEGPQQSG SLPEEGSKPS NSCVLRETMV QASPSSKAQQ TPSAQDYGRG NIISPSRMLD
 25 61 KSPSPASGHF AHPFAASLK MGPPKSKAEK IKRSPMDKVL RFARQPLLLL DRPEGAHAAP
 121 QPSLEWPAH WPTGRGGLQR RPALAWEAPG RSCSESTLYP MPVLVPLAVA PQESHRTSAQ
 181 ALFPFEASLL TSVARRKHRR WQSTVEISAR ARLASCPESEN LGPPRPVARR AGGPLARGRP
 241 SLVRQDAYTR SDSEPSKHS A ECDPRFPSVI PETSEGESSD HTTNRFQDRE SSSSDEEGGA
 301 QSRDCDLALG YVAAGHEELA WTQEAPVSSG PLLSPVPKLC RIKASKALKK KIRRFQPTAL
 30 361 KVMTMV

C. 核苷酸及氨基酸组合序列 (SEQ ID NO:18) 克隆号: PP13671

起始编码子: 1314 ATG 终止编码子: 2412 TGA 蛋白质分子量: 39633.59

1 GG GAG GAC CCG CGA CAC AGC TCG GGA TCC GCG CGC GCT GGG CCT TCG 47
 35 48 GCG CCG CTC GTG GGG TTC GGG AGG CCG GGA CCC TGC CCG GGA GAT GTG 95
 96 GAC GCC GGG CGG ACC CCC GGG GTC CGC GGG CTG GGA CCG CCG TAG GTT 143
 144 GGG CGC GAG GTT GCG CGC GGC GTT CGC GGG GCT GCA GGA GCT GCA GGG 191
 192 GCT GCG AGC CAC GCA GCA GGA GCG GGT ACG GGG CGC CCT GGC CCT GCA 239
 240 GCC CCC GCC CGC GCC CGC CGC GCC CTG CGG CCC CCA CGG CCT CCA CGG 287
 40 288 CCC CGA GCA GCA GCT GGA GGC GGC GCT GGC CGC GCT GCA GGA GCA GCT 335
 336 GTC CCG GCT GAG ACA ACA GGA CAT CGG CCT GAA GAC CCA CCT GGA CCA 383
 384 GCT GGA CCT GCA GAT TAG CAA GCT GCA GCT GGA TGT GGG CAC AGC CTC 431
 432 AGG GGA GGC CCT GGA CAG CGA CAG CAG GCC CAG CTC AGG CTT TTA CGA 479
 480 GAT GAG CGA CGG TGG ATC CTG CTC CCT GTC CAC GTC CTG TGC CTC CGT 527
 45 528 CTG CAG TGA CCA CAT CTC TCC CTC GCT GGG CAG TTT GTT GCC TGT GGC 575
 576 CCA GGC CCA CAA GGC CAG GCC CAG CAT GGG GGA CTG GAG GCC CCG GTC 623
 624 GGT TGA TGA GAC TAC TGT GCC AGC GTG GAG ACC CCA GGC TAC CGA GGA 671
 672 GGG CGC CAG GCC CCC AGG GAG CGT GGA GGA TGC AGG CCA GCC GTG GGG 719
 720 CAC ATT CTG GCC CAG GCC TGT GTC TAC AGG TGA TCT TGA CAG AGC CCT 767
 50 768 GCC GGC GGA CAC GGG GCT CCA GAA AGC CAG CGC GGA CGC CGA GCT CCT 815
 816 CGG GCT CCT CTG CAG GGG GTG GAT ATC CCG CTG CAC GTG CCG GAC CCC 863
 864 AAG TAT CGG CAG GAC CTG GTG TCC CAG GGC GGC AGG GAG GTG TAC CCG 911
 912 TAC CCC AGC CCC CTG CAC GCC GTG GCT CTA CAG AGC CCC CTG TTT GTC 959
 960 CTG ACT AAG GAA ACC CCA CAG AGA GGT GGC CCC TCG TTC CCT AGG GAG 1007

	1008	AGC CCC AGG GGC CCC GCA GGT CTG AAC ACC ATC CAG ACT GGG CCG GTC	1055
	1056	CTC GAG GCT GGC CCG GCC AGG GCC AGA GCT TAT ATC GAC AGG TTG CTG	1103
	1104	CAT CTG TGG GGC CGG GAG ACC CCA GCA AAG GGT AGC GAG GGA GAA CAG	1151
	1152	GGA CCT CTA AGG CAT GCA GCG TCC CCA TCT CCA CAG AGG CAG GGT GGC	1199
5	1200	TGG AGT ACA GAC GGT GGA GGG CGA CTG CTG GTC TTC GCC CCA GGG AGG	1247
	1248	GAG GAC GAG GGA GGG CCA GCT CAG AGC AGG GGT GCC GGC AGG GGC GGG	1295
	1296	CCC CAG CAG CAG GGA TAC ATG CCC CTT GAG GGT CCC CAG CAG TCT GGC	1343
	1	Met Pro Leu Glu Gly Pro Gln Gln Ser Gly	10
10	1344	AGC CTT CCA GAG GAG GGC TCC AAG CCC TCA AAC AGC TGT GTC CTC AGG	1391
	11	Ser Leu Pro Glu Glu Gly Ser Lys Pro Ser Asn Ser Cys Val Leu Arg	26
	1392	GAG ACC ATG GTG CAG GCT TCT CCC AGC TCA AAG GCC CAG CAG ACA CCC	1439
15	27	Glu Thr Met Val Gln Ala Ser Pro Ser Ser Lys Ala Gln Gln Thr Pro	42
	1440	TCA GCT CAG GAC TAT GGA CGA GGC AAC ATC ATA TCC CCA TCC AGG ATG	1487
	43	Ser Ala Gln Asp Tyr Gly Arg Gly Asn Ile Ile Ser Pro Ser Arg Met	58
	1488	CTG GAC AAG AGC CCC TCA CCG GCC TCT GGG CAC TTT GCC CAC CCA TCC	1535
20	59	Leu Asp Lys Ser Pro Ser Pro Ala Ser Gly His Phe Ala His Pro Ser	74
	1536	TTT GCT GCC AGC CTG AAA ATG GGT CCC CCC AAG AGC AAG GCT GAA AAA	1583
	75	Phe Ala Ala Ser Leu Lys Met Gly Pro Pro Lys Ser Lys Ala Glu Lys	90
25	1584	ATC AAG AGA AGT CCC ATG GAC AAG GTG CTG AGG TTT GCA AGG CAG CCG	1631
	91	Ile Lys Arg Ser Pro Met Asp Lys Val Leu Arg Phe Ala Arg Gln Pro	106
	1632	CTG CTT CTA CTG GAC AGG CCT GAG GGA GCC CAT GCA GCC CCC CAG CCA	1679
30	107	Leu Leu Leu Leu Asp Arg Pro Glu Gly Ala His Ala Ala Pro Gln Pro	122
	1680	TCC CTG GAG TGG GAC CCT GCC CAC TGG CCC ACA GGG AGG GGC GGG CTC	1727
	123	Ser Leu Glu Trp Asp Pro Ala His Trp Pro Thr Gly Arg Gly Gly Leu	138
	1728	CAG CGG AGG CCA GCC CTG GCC TGG GAG GCA CCC GGG CGC TCC TGT TCT	1775
35	139	Gln Arg Arg Pro Ala Leu Ala Trp Glu Ala Pro Gly Arg Ser Cys Ser	154
	1776	GAG TCC ACC CTC TAC CCC ATG CCT GTC CTC GTC CCC TTG GCA GTG GCC	1823
	155	Glu Ser Thr Leu Tyr Pro Met Pro Val Leu Val Pro Leu Ala Val Ala	170
40	1824	CCG CAG GAG AGC CAC CGG ACC TCA GCC CAA GCC CTG TTC CCC TTT GAG	1871
	171	Pro Gln Glu Ser His Arg Thr Ser Ala Gln Ala Leu Phe Pro Phe Glu	186
	1872	GCG TCA CTG CTC ACC TCA GTG GCC AGG AGG AAG CAT CGC CGC TGG CAG	1919
45	187	Ala Ser Leu Leu Thr Ser Val Ala Arg Arg Lys His Arg Arg Trp Gln	202
	1920	TCC ACC GTG GAG ATC TCG GCC CGG GCC CGC CTG GCC AGC TGT CCT GAG	1967
	203	Ser Thr Val Glu Ile Ser Ala Arg Ala Arg Leu Ala Ser Cys Pro Glu	218
	1968	TCT AAC CTG GGG CCC CCC AGG CCC GTG GCC AGG AGA GCA GGT GGC CCA	2015
50	219	Ser Asn Leu Gly Pro Pro Arg Pro Val Ala Arg Arg Ala Gly Gly Pro	234
	2016	CTG GCC CGG GGC CGT CCC TCA CTG GTC CGC CAG GAC GCC TAC ACC AGG	2063
	235	Leu Ala Arg Gly Arg Pro Ser Leu Val Arg Gln Asp Ala Tyr Thr Arg	250
55	2064	AGC GAC TCA GAG CCC TCC AAG CAC TCG GCC GAG TGT GAC CCG CGG TTC	2111
	251	Ser Asp Ser Glu Pro Ser Lys His Ser Ala Glu Cys Asp Pro Arg Phe	266
	2112	CCG TCA GTC ATC CCG GAG ACC AGC GAG GGA GAG TCC AGT GAC CAC ACC	2159
60	267	Pro Ser Val Ile Pro Glu Thr Ser Glu Gly Glu Ser Ser Asp His Thr	282
	2160	ACC AAC CGA TTC GGA GAC CGT GAG TCC AGC AGC AGC GAC GAG GAG GGC	2207
	283	Thr Asn Arg Phe Gly Asp Arg Glu Ser Ser Ser Ser Asp Glu Glu Gly	298
	2208	GGC GCC CAG AGC AGG GAC TGT GAC CTG GCA CTG GGC TAT GTG GCG GCC	2255
65	299	Gly Ala Gln Ser Arg Asp Cys Asp Leu Ala Leu Gly Tyr Val Ala Ala	314
	2256	GGG CAT GAG GAG CTG GCC TGG ACC CAG GAG GCC CCG GTC AGC TCG GGG	2303
	315	Gly His Glu Glu Leu Ala Trp Thr Gln Glu Ala Pro Val Ser Ser Gly	330
70	2304	CCA CTC CTG TCC CCC GTG CCC AAG CTG TGC CGT ATT AAG GCC TCC AAG	2351

	331	Pro Leu Leu Ser Pro Val Pro Lys Leu Cys Arg Ile Lys Ala Ser Lys	346
	2352	GCC CTG AAG AAG AAG ATC CGC AGG TTC CAG CCG ACG GCC CTG AAG GTC	2399
5	347	Ala Leu Lys Lys Lys Ile Arg Arg Phe Gln Pro Thr Ala Leu Lys Val	362
	2400	ATG ACC ATG GTG TGA GGT GCA GTG ACT GCG TCA AGA GAA GCC TGC TGC	2447
	363	Met Thr Met Val ***	367
	2448	ACA CAG GGG CCC TTT CCA TGC AGT GTC CTG TGC ATG TCT GGG TTT CAG	2495
10	2496	TGT CGT CTT GCC AAG TCG CTT TAT CTT CCG AGG CCA CCG TTT AAC AGG	2543
	2544	AGT CCC CGC AGA TGG GTT CTG CCC TGT GGA GTA TGG CCG TGA GTC CCC	2591
	2592	GCC CCA GCA AAG ACT TCC ATG CGG AGC TCT GGT TTC AGA TGC CAG CAG	2639
	2640	CCT CCC CGG CCG ACC TAG TGT CTC TTT TCA AGG TCT GCC CAG GAT GCT	2687
	2688	GCT TCC CGG CAT GCA CCT TGT TTC CAT GTT GTC ATC CTC CCA TGG CTT	2735
15	2736	CCC CTG CCC TCA CTC CCC TCT GGC TGG GGA AGT GTT TCC CAA TCC TTT	2783
	2784	GAC CAT TCC TCC CCA AAA GCA TTT TAA AAA GGC GGT TGC AGG ACT TTG	2831
	2832	AGC TGG AAA TTG AAT TTG AGT TTA AAC TCC ACG TTA AGC TGT CAG CAT	2879
	2880	ACA CAA ACA CCC AAA TTA TGC CAT GAT CTC AAA TAA AAG GAA CAT TTT	2927
	2928	TTT TCC CTC TTT TAA AAA AAA AAA AAA AAA AAA AAA AAA AAA	2975
20	2976	AAA AAA AAA AA	2986

7. PP13759

A: 核苷酸序列 (SEQ ID NO:19) 长度: 1581

	1	GAGGCGGAGA CACAAGGCTC CGTTCCTACTG GGCACGCCAG CGGTAGAGAC GGGGTTTCGC	
25	61	CATGTTGGCC AGGCTGGTCT CAAACTCCTG AGCTCGGGCG GTCTGCCCGC CTCGGCCTCC	
	121	CAAAGTGTG GGATTGCGGG CGTGAGCCAC TGCACCCGGT TCAAGCGATT CTCCTGCCTC	
	181	AGCCCCCTGA GTAGCTGGGA TTACAGGCAC GCGCCACCAC TCCCAGATCT TCACTTCAGG	
	241	TCAGCTCCAA GGAGGCTATC CTTGCCCATC TTCCTAAAAT AGAGCTCTCA AAACCTGGCT	
	301	GTTTTCTTC AGAGCACTGG CCTCCAATTG CTCGGTAGA TTGAAATGA CCTGCTTTCT	
30	361	TTCTGTTCCC GGGCGTTTGG ACCCCTGTCT TGGACCGCTG TCGGATAGTA AATCCCAAGT	
	421	AAGGTACTG CCGTCGGCAG ATTTGAGCTT TCTTCTTGA CACCTAATAC CCAGAGTCCT	
	481	CCAGGCTCCG GTAGATTGCA AATGACCTGC TTTCTTCTG TTCCCGGGCG GCATCGGACC	
	541	CGTCGGAGAG TAAATCCCAA GTAAGGTACC TGCCGTTGGC AGATTGAGC TTTCTTCTG	
	601	GACACCTAAT ACCCAGATC CTCCAGGCTC CGGTAGATTG CAAATGACCT GCTTCTTTC	
35	661	TGTTCCCGGG TGGCATCGAC CCGTCGGAGA GTAAATCCCA AGTAAGGTAC CTGCCGTTGG	
	721	CAGATTTGAG CTTTCTTCTT GGACACCTAA TACCCACAGT CCTCCAGTAA ATCCCAAGCA	
	781	AGGTACTGCG CGTCAGCAGA TTTGAGCTTT CTCTTGGAC ACCTAATACC CACAGTCTC	
	841	CAGGCTCCGG TAGATTGCAA ATGACCTGCT TTCTTCTGT TCCCGGGCTG CGTTCGGACC	
	901	CCTGTCGGAT AGTAACTCCC AAGTAAGGTA CCTGCCGTCG GCAGATTGGA GCTTCTTCT	
40	961	TGGAGACCTA ATACCCACAG TCCTCCAGAA AAACAACAA AGACATGGAT TTAAGTGCA	
	1021	TATTAGCAGA TCCATACTGG AAAATGCATG GAGGTTTCAT ATACACCACT TACAGTTTC	
	1081	AGCTCCTCAG TAGTGACAAA GCCATACCCA TCATTGTGGA TTCGATCAAC AATCTTCCCT	
	1141	AGCCTCTCCT TGCTCTCGTC CGGGGTGAGC TCGTCAAGT TCTTGGAGTC CTCCTTGCCC	
	1201	AGGAAGGCCT CGTGGTCGTA CTGGAAGCTC TGGTTGTCT CAGGGGGCCG CTCGCCCAGC	
45	1261	TCGAGTCCG GCCGCACCAC GCGCTCTTTG CGCACCGTGG GCTTGGCCCG CAGAACCCGC	
	1321	GGCGCCAGCA CCAGGCCAG CAGCAGCCCC AGGGCTAACC CCAGGCGGCC ACCGCGCGCC	
	1381	ATCGTCCCGA GGAGAGGGCG GCCGGGAGG AGACGCTGAG CGAGCGACAA CAGCGGCAGC	
	1441	TCGGGAATGG GGGCTCGGAG CGCGCGGCC AAGTTTTATG TTATGTATAT TTTACAAGTA	
	1501	AAAAAATTTT TTCACCTCAG CCTGAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA	
50	1561	AAAAAAAAAA AAAAAAAAAA A	

B: 氨基酸序列 (SEQ ID NO:20) 长度: 123

	1	MEVSYTPLTV FSSSVVTKPY PLSLIRSTIF PLSLLSSGV SSSKFLESSL PRKASWSYWK
	61	LWLSGGRRSP SSESGRTRRS LRTVGLARRT RGASTSASSS PRANPRRPPR AIVPRRGRPG
55	121	GRR

C. 核苷酸及氨基酸组合序列 (SEQ ID NO:21) 克隆号: PP13759
 起始编码子: 1048 ATG 终止编码子: 1417 TGA 蛋白质分子量: 13369.38

	1	GAG GCG GAG ACA CAA GGC TCC GTT CCA CTG GGC ACG CCA GCG GTA GAG	48
	49	ACG GGG TTT CGC CAT GTT GGC CAG GCT GGT CTC AAA CTC CTG AGC TCG	96
5	97	GGC GGT CTG CCC GCC TCG GCC TCC CAA AGT GCT GGG ATT GCG GGC GTG	144
	145	AGC CAC TGC ACC CGG TTC AAG CGA TTC TCC TGC CTC AGC CCC CTG AGT	192
	193	AGC TGG GAT TAC AGG CAC GCG CCA CCA CTC CCA GAT CTT CAC TTC AGG	240
	241	TCA GCT CCA AGG AGG CTA TCC TTG CCC ATC TTC CTA AAA TAG AGC TCT	288
	289	CAA AAC CTG GCT GTT TTC CTT CAG AGC ACT GGC CTC CAA TTG CTC CGG	336
10	337	TAG ATT GCA AAT GAC CTG CTT TCT TTC TGT TCC CGG GCG TTT GGA CCC	384
	385	CTG TCT TGG ACC GCT GTC GGA TAG TAA ATC CCA AGT AAG GTA CCT GCC	432
	433	GTC GGC AGA TTT GAG CTT TCT TCT TGG ACA CCT AAT ACC CAG AGT CCT	480
	481	CCA GGC TCC GGT AGA TTG CAA ATG ACC TGC TTT CTT TCT GTT CCC GGG	528
	529	CGG CAT CGG ACC CGT CGG AGA GTA AAT CCC AAG TAA GGT ACC TGC CGT	576
15	577	TGG CAG ATT TGA GCT TTC TTC TTG GAC ACC TAA TAC CCA CAG TCC TCC	624
	625	AGG CTC CGG TAG ATT GCA AAT GAC CTG CTT TCT TTC TGT TCC CGG GTG	672
	673	GCA TCG ACC CGT CGG AGA GTA AAT CCC AAG TAA GGT ACC TGC CGT TGG	720
	721	CAG ATT TGA GCT TTC TTC TTG GAC ACC TAA TAC CCA CAG TCC TCC AGT	768
	769	AAA TCC CAA GCA AGG TAC CTG CCG TCA GCA GAT TTG AGC TTT CTT CTT	816
20	817	GGA CAC CTA ATA CCC ACA GTC CTC CAG GCT CCG GTA GAT TGC AAA TGA	864
	865	CCT GCT TTC TTT CTG TTC CCG GGC TGC GTT CGG ACC CCT GTC GGA TAG	912
	913	TAA CTC CCA AGT AAG GTA CCT GCC GTC GGC AGA TTG GAG CTT TCT TCT	960
	961	TGG AGA CCT AAT ACC CAC AGT CCT CCA GAA AAA CAA ACA AAG ACA TGG	1008
	1009	ATT TAC TGT GCA TAT TAG CAG ATC CAT ACT GGA AAA TGC ATG GAG GTT	1056
25	1	Met Glu Val	3
	1057	TCA TAT ACA CCA CTT ACA GTT TTC AGC TCC TCA GTA GTG ACA AAG CCA	1104
	4	Ser Tyr Thr Pro Leu Thr Val Phe Ser Ser Ser Val Val Thr Lys Pro	19
30	1105	TAC CCA TCA TTG TCG ATT CGA TCA ACA ATC TTC CCT AGC CTC TCC TTG	1152
	20	Tyr Pro Ser Leu Ser Ile Arg Ser Thr Ile Phe Pro Ser Leu Ser Leu	35
	1153	CTC TCG TCC GGG GTG AGC TCG TCG AAG TTC TTG GAG TCC TCC TTG CCC	1200
35	36	Leu Ser Ser Gly Val Ser Ser Ser Lys Phe Leu Glu Ser Ser Leu Pro	51
	1201	AGG AAG GCC TCG TGG TCG TAC TGG AAG CTC TGG TTG TCC TCA GGG GGC	1248
	52	Arg Lys Ala Ser Trp Ser Tyr Trp Lys Leu Trp Leu Ser Ser Gly Gly	67
	1249	CGC TCG CCC AGC TCC GAG TCG GGC CGC ACC ACG CGC TCT TTG CGC ACC	1296
40	68	Arg Ser Pro Ser Ser Glu Ser Gly Arg Thr Thr Arg Ser Leu Arg Thr	83
	1297	GTG GGC TTG GCC CGC AGA ACC CGC GGC GCC AGC ACC AGC GCC AGC AGC	1344
	84	Val Gly Leu Ala Arg Arg Thr Arg Gly Ala Ser Thr Ser Ala Ser Ser	99
45	1345	AGC CCC AGG GCT AAC CCC AGG CGG CCA CCG CGC GCC ATC GTC CCG AGG	1392
	100	Ser Pro Arg Ala Asn Pro Arg Arg Pro Pro Arg Ala Ile Val Pro Arg	115
	1393	AGA GGG CGG CCG GGA GGG AGA CGC TGA GCG AGC GAC AAC AGC GGC AGC	1440
50	116	Arg Gly Arg Pro Gly Gly Arg Arg ***	124
	1441	TCG GGA ATG GGG GCT CGG AGC GCG GCG GCC AAG TTT TAT GTT ATG TAT	1488
	1489	ATT TTA CAA GTA AAA AAA TTT TTT CAC CTC AGC CTG AAA AAA AAA AAA	1536
	1537	AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA	1581

55 8. PP14328

A: 核苷酸序列 (SEQ ID NO:22) 长度: 1290

	1	GCGCGGGCGT TCCTGGTTCC CGCTCCCGCC CCCTCCGCC CGCCGACTCG CCGCTTCTC
	61	CGGAGCCCA GCGTCCCGAA CCAGGGACTG CATTCCCGG CAGGCGCCGC GGCCCGCGC
	121	CAATGAGGGG CTGAGGATTT GGCGCGGCG GCGCCCGAG AGTCGGGGTG GGGGGCTTT
60	181	GTGCGCGCG GCGGCGGGAG AGGCGCGGC GCGGCCAGC ACGGAGGCG AGGCCGAGG

241 GGCTGTGCAC AGGTCCGCCG GGAGAGGCGT GCGAATTCGG AGCCGAGCGC CGAGGACCGT
 301 GCTACCCAGG CCGGGCTGCC AGCCGCAGGC TCCTCTCTGG CAGCAGCGGC GGC CGGCGA
 361 CCCCCGTCCC TCGGCCTCCC CTTCATCC CACTCCCGA GCCTTCCTCT TCCCGCAGCA
 421 CGCCCGGCCG GGCCCGGCCG TGGCCCTCCT CAGTGCCGGC CGCCATGGCA GAGGCGTCCG
 5 481 GCGCGGGGAA AATCTAGCCC GGGGATTTCA TCGCGCCTAG CTCGGTTCGG CCTCCTCCTC
 541 GCGCGGCCCG AGCGGCTGCC CGCACCCAG CCCACTCCG GGCCTCCGTG TCTCTCCTGT
 601 GATCGCACTG ACACGGCCGG GGGGTTAGAA TGAACAAC TGAAGGCCG ATGAGAGAAA
 661 GGGAAAGTTA AGGATGCTGG AGCAGAACA TGGATTTCTC TTTCTTTTC ATGCAAGGGA
 721 TCATGGGAAA CACAATTCAG CAACCACCTC AACTCATGA CTCCGCCAAC ATCCGTCAGG
 10 781 AGGATGCCTT TGATAACAAC AGTGACATTG CTGAAGATGG TGGCCAGACA CCATATGAAG
 841 CTACTTTGCA GCAAGGCTTT CAGTACCCAG CTACAACAGA AGATCTTCCT CCACTCACAA
 901 ATGGGTATCC ATCATCAATC AGTGTGTATG AACTCAAAC CAAATACCAG TCATATAATC
 961 AGTATCCTAA TGGGTCAGCC AATGGCTTTG GTGCAGTTAG AACTTTAGC CCCACTGACT
 1021 ATTATCATTC AGAAATTTCA AACACAAGAC CACATGAAAT TCTGGAAAAA CCTTCCCCTC
 15 1081 CACAGCCACC ACCTCCTCCT TCGGTACCAC AACTGTGAT TCCAAAGAAG ACTGGCTCAC
 1141 CTGAAATTA ACTAAAAATA ACCAAAATA TCCAGAATGG CAGGGAATTG TTTGAGTCTT
 1201 CCCTTTGTGG AGACCTTTTA AATGAAGTAC AGGCAAGTGA GCACACGAAA TCAAAGCATG
 1261 AAAGCAGAAA AAAAAAAAAA AAAAAAAAAA

20 B: 氨基酸序列 (SEQ ID NO:23) 长度: 111

1 MPLITTVTL KVARHMKL LCSKAFSTQL QKIFLHSMK GIHQSVCMK LKPNSTSHIIS
 61 ILMGQPMALV QLETLAPLTI IIQKFQTDH MKFWKNLPLH SHHLLRLRYHK L

C. 核苷酸及氨基酸组合序列 (SEQ ID NO:24) 克隆号: PP14328

25 起始编码子: 784 ATG 终止编码子: 1117 TGA 蛋白质分子量: 12935.10

1 GCG CGG GCG TTC CTG GTT CGC GCT CCC GCC CCC TCC GCC CCG CCG ACT 48
 49 CGC CGC TTC CTC CGG AGC CCC AGC GTC CCG AAC CAG GGA CTG CAT TTC 96
 97 CCG GCA GGC GCC GCG GCC CGC GGC CAA TGA GGG GCT GAG GAT TTG GCG 144
 145 GCG GCG GCG CCC CGA GAG TCG GGG TGG GGG GGC TTT GTG CGC GGC GGC 192
 30 193 GGC GGG AGA GGC GGC GGC GGC GGC CAG CAC GGA GGC GGA GGC CGA GGG 240
 241 GGC TGT GCA CAG GTC GCC GCG GAG AGG CGT GCG AAT TCC GAG CCG AGC 288
 289 GCC GAG GAC CGT GCT ACC CAG GCC GGG CTG CCA GCC GCA GGC TCC TCT 336
 337 CTG GCA GCA GCG GCG GCG CGG CGA CCC CCG TCC CTC GGC CTC CCC TTC 384
 385 CCA TCC CAC CTC CCG AGE CTT CCT CTT CCC GCA GCA CGC CCG GCC CGG 432
 35 433 CCC GGC CGT GGC CCT CCT CAG TGC CGG CCG CCA TGG CAG AGG CGT CCG 480
 481 GCG CGG GGA AAA TCT AGC CCG GGG ATT TCA TGC GGC CTA GCT CGG TTC 528
 529 CGC CTC CTC CTC GCG CGG CCC CAG CGG CTG CCC GCA CCC CAG CCC CAC 576
 577 TCC GGG CCT CGG TGT CTC TCC TGT GAT CGC ACT GAC ACG GCC GGG GGG 624
 625 TTA GAA TGG AAC AAA CTG AAG GCC CGA TGA GAG AAA GGG AAA GTT AAG 672
 40 673 GAT GCT GGA GCA GAA CAA TGG ATT TCT CTT TCT CTT TCA TGC AAG GGA 720
 721 TCA TGG GAA ACA CAA TTC AGC AAC CAC CTC AAC TCA TTG ACT CCG CCA 768
 769 ACA TCC GTC AGG AGG ATG CCT TTG ATA ACA ACA GTG ACA TTG CTG AAG 816
 1 Met Pro Leu Ile Thr Thr Val Thr Leu Leu Lys 11
 45 817 ATG GTG GCC AGA CAC CAT ATG AAG CTA CTT TGC AGC AAG GCT TTC AGT 864
 12 Met Val Ala Arg His His Met Lys Leu Leu Cys Ser Lys Ala Phe Ser 27
 865 ACC CAG CTA CAA CAG AAG ATC TTC CTC CAC TCA CAA ATG GGT ATC CAT 912
 28 Thr Gln Leu Gln Gln Lys Ile Phe Leu His Ser Gln Met Gly Ile His 43
 50 913 CAT CAA TCA GTG TGT ATG AAA CTC AAA CCA AAT ACC AGT CAT ATA ATC 960
 44 His Gln Ser Val Cys Met Lys Leu Lys Pro Asn Thr Ser His Ile Ile 59
 961 AGT ATC CTA ATG GGT CAG CCA ATG GCT TTG GTG CAG TTA GAA ACT TTA 1008
 55 60 Ser Ile Leu Met Gly Gln Pro Met Ala Leu Val Gln Leu Glu Thr Leu 75
 1009 GCC CCA CTG ACT ATT ATC ATT CAG AAA TTC CAA ACA CAA GAC CAC ATG 1056

	76	Ala Pro Leu Thr Ile Ile Ile Gln Lys Phe Gln Thr Gln Asp His Met	91
	1057	AAA TTC TGG AAA AAC CTT CCC CTC CAC AGC CAC CAC CTC CTC CTT CGG	1104
5	92	Lys Phe Trp Lys Asn Leu Pro Leu His Ser His His Leu Leu Leu Arg	107
	1105	TAC CAC AAA CTG TGA TTC CAA AGA AGA CTG GCT CAC CTG AAA TTA AAC	1152
	108	Tyr His Lys Leu ***	112
	1153	TAA AAA TAA CCA AAA CTA TCC AGA ATG GCA GGG AAT TGT TTG AGT CTT	1200
10	1201	CCC TTT GTG GAG ACC TTT TAA ATG AAG TAC AGG CAA GTG AGC ACA CGA	1248
	1249	AAT CAA AGC ATG AAA GCA GAA AAA AAA AAA AAA AAA AAA	1290

9. PP14450

A: 核苷酸序列 (SEQ ID NO:25) 长度: 1816

15	1	GAGACGGAGT CTCACTCTGT CGCCCAGGCT GGAGTACAAT GGCATGATCT CGGCTCACTG
	61	CAACCTACAG CTCCCAGGTT CAAGCGATT CTTGCGCTCA GCCTCCCAGC TCCTGCTCGC
	121	CGGACGGCTC CCAGGGAGAG CAGACGCGCC AGACGCGCCA CCCTCGGGGC GCCGACGGTC
	181	ACGGAGCATG GGGTCGGCCT TTGAGCGGGT AGTCCGGAGA GTGGTCCAGG AGCTGGACCA
	241	TGGTGGGGAG TTCATCCCTG TGACCAGCCT GCAGAGCTCC ACTGGCTTCC AGCCCTACTG
20	301	CCTGGTGGTT AGGAAGCCCT CAAGCTCATG GTTCTGAAA CCCCGTTATA AGTGTGTCAA
	361	CCTGTCTATC AAGGACATCC TGGAGCCGGA TGCCGCGGAA CCAGACGTGC AGCGTGGCAG
	421	GAGCTTCCAC TTCTACGATG CCATGGATGG GCAGATACAG GGCAGCGTGG AGCTGGCAGC
	481	CCCAGGACAG GCAAAGATCG CAGGCGGGGC CGCGGTGTCT GACAGCTCCA GCACCTCAAT
	541	GAATGTGTAC TCGCTGAGTG TGGACCTAA CACCTGGCAG ACTCTGCTCC ATGAGAGGCA
25	601	CCTGCGGCAG CCAGAACACA AAGTCTGCA GCAGCTGCGC AGCCGCGGGG ACAACGTGTA
	661	CGTGGTGA CTGAGTCTGC AGACACAGAA GGAGGTGGAA GTCACGCGCA CCCACAAGCG
	721	GGAGGGCTCG GGCCGTTTTT CCCTGCCCGG AGCCACGTGC TTGCAGGGTG AGGGCCAGGG
	781	CCATCTGAGC CAGAAGAAGA CGGTCACCAT CCCCTCAGGC AGCACCCCTCG CATTCGGGT
	841	GGCCAGCTG GTTATTGACT CTGACTTGA CGTCTTCTC TTCCCGGATA AGAAGCAGAG
30	901	GACCTTCCAG CCACCCGCGA CAGGCCACAA GCGTTCCAG AGCGAAGGCG CCTGGCCACA
	961	GCTGCCCTCT GGCCTCTCCA TGATGAGGTG CCTCCACAAC TTCCTGACAG ATGGGGTCCC
	1021	TGCGGAGGGG GCGTTCCTG AAGACTTCCA GGGCTACCG GCAGAGGTGG AGACCATCTC
	1081	CAAGGAACTG GAGCTTTTGG ACAGAGAGCT GTGCCAGCTG CTGCTGGAGG GCCTGGAGGG
	1141	GGTGTGCGG GACCAGCTGG CCCTGCGAGC CTGGAGGAG GCGCTGGAGC AGAGCCAGAG
35	1201	CCTGGGCGG GTGGAGCCCC TGGACGGTCC AGCAGGTGCT GTCTGGAGT GCCTGGTGT
	1261	GTCCTCCGGA ATGCTGGTGC CGAACTCGC TATCCCTGTT GTCTACCTGC TGGGGGCACT
	1321	GACCATGCTG AGTGAACGC AGCACAAGCT TGCTGGCGGA GGCCTGGAG TCGCAGACCC
	1381	TGTTGGGGCC GCTCGAGCTG GTGGGCAGCC TCTTGGAGCA GAGTGCCCGG TGGCAGGAGC
	1441	GCAGCACCAT GTCCCTGCC CCCGGGCTCC TGGGGAACAG CTGGGGCGAA GGAGCACCAG
40	1501	CCTGGGTCTT GCTGGACGAG TGTGGCCTAG AGCTGGGGGA GGACTCTCC CACGTGTGCT
	1561	GGGAGCCGCA GGCCAGGGC CGCATGTGTG CACTTACGC CTCCTGGCA CTGCTATCAG
	1621	GACTGAGCCA GGAGCCCCAC TAGCCTGTGC CCGGCATGG CCTGGCAGCT CTCAGCAGG
	1681	GCAGAGTGTG TGCCACCAG CTGCTAGCCC TAGGAAGGCC AGGAGCCCAG TAGCCATGTG
	1741	GCCAGTCTAC CATGGGGCCC AGGAGTTGGG GAAACACAAT AAAGGTGGCA TACGAAGGAA
45	1801	AAAAAAAAA AAAAAA

B: 氨基酸序列 (SEQ ID NO:26) 长度: 103

	1	MGSLRRGRSL KTSRAYGQRW RPSRPNWSFW TESCASCWR AWRGCCGTSW PCEPWRRRWS
50	61	RARALGRWSP WTVQQVLSWS AWCCPPECWC RNSLSLLSTC WGH

C. 核苷酸及氨基酸组合序列 (SEQ ID NO:27) 克隆号: PP14450

起始编码子: 1011 ATG 终止编码子: 1320 TGA 蛋白质分子量: 12272.40

	1	GA GAC GGA GTC TCA CTC TGT CGC CCA GGC TGG AGT ACA ATG GCA TGA	47
	48	TCT CGG CTC ACT GCA ACC TAC ACC TCC CAG GTT CAA GCG ATT CTC TTG	95
55	96	CCT CAG CCT CCC AGC TCC TGC TCG CCG GAC GGC TCC CAG GGA GAG CAG	143

	144	ACG CGC CAG ACG CGC CAC CCT CGG GGC GCC GAC GGT CAC GGA GCA TGG	191
	192	GGT CGG CCT TTG AGC GGG TAG TCC GGA GAG TGG TCC AGG AGC TGG ACC	239
	240	ATG GTG GGG AGT TCA TCC CTG TGA CCA GCC TGC AGA GCT CCA CTG GCT	287
	288	TCC AGC CCT ACT GCC TGG TGG TTA GGA AGC CCT CAA GCT CAT GGT TCT	335
5	336	GGA AAC CCC GTT ATA AGT GTG TCA ACC TGT CTA TCA AGG ACA TCC TGG	383
	384	AGC CGG ATG CCG CGG AAC CAG ACC TGC AGC GTG GCA GGA GCT TCC ACT	431
	432	TCT ACG ATG CCA TGG ATG GGC AGA TAC AGG GCA GCG TGG AGC TGG CAG	479
	480	CCC CAG GAC AGG CAA AGA TCG CAG GCG GGG CCG CGG TGT CTG ACA GCT	527
	528	CCA GCA CCT CAA TGA ATG TGT ACT CGC TGA GTG TGG ACC CTA ACA CCT	575
10	576	GGC AGA CTC TGC TCC ATG AGA GGC ACC TGC GGC AGC CAG AAC ACA AAG	623
	624	TCC TGC AGC AGC TGC GCA GCC GCG GGG ACA ACG TGT ACG TGG TGA CTG	671
	672	AGG TGC TGC AGA CAC AGA AGG AGG TGG AAG TCA CGC GCA CCC ACA AGC	719
	720	GGG AGG GCT CGG GCC GGT TTT CCC TGC CCG GAG CCA CGT GCT TGC AGG	767
	768	GTG AGG GCC AGG GCC ATC TGA GCC AGA AGA AGA CGG TCA CCA TCC CCT	815
15	816	CAG GCA GCA CCC TCG CAT TCC GGG TGG CCC AGC TGG TTA TTG ACT CTG	863
	864	ACT TGG ACG TCC TTC TCT TCC CGG ATA AGA AGC AGA GGA CCT TCC AGC	911
	912	CAC CCG CGA CAG GCC ACA AGC GTT CCA CGA GCG AAG GCG CCT GGC CAC	959
	960	AGC TGC CCT CTG GCC TCT CCA TGA TGA GGT GCC TCC ACA ACT TCC TGA	1007
20	1008	CAG ATG GGG TCC CTG CGG AGG GGG CGT TCA CTG AAG ACT TCC AGG GCC	1055
	1	Met Gly Ser Leu Arg Arg Gly Arg Ser Leu Lys Thr Ser Arg Ala	15
	1056	TAC GGG CAG AGG TGG AGA CCA TCT CCA AGG AAC TGG AGC TTT TGG ACA	1103
	16	Tyr Gly Gln Arg Trp Arg Pro Ser Pro Arg Asn Trp Ser Phe Trp Thr	31
25	1104	GAG AGC TGT GCC AGC TGC TGC TGG AGG GCC TGG AGG GGG TGC TGC GGG	1151
	32	Glu Ser Cys Ala Ser Cys Cys Trp Arg Ala Trp Arg Gly Cys Cys Gly	47
	1152	ACC AGC TGG CCC TGC GAG CCT TGG AGG AGG CGC TGG AGC AGA GCC AGA	1199
30	48	Thr Ser Trp Pro Cys Glu Pro Trp Arg Arg Arg Trp Ser Arg Ala Arg	63
	1200	GCC TTG GGC CGG TGG AGC CCC TGG ACG GTC CAG CAG GTG CTG TCC TGG	1247
	64	Ala Leu Gly Arg Trp Ser Pro Trp Thr Val Gln Gln Val Leu Ser Trp	79
	1248	AGT GCC TGG TGT TGT CCT CCG GAA TGC TGG TGC CGG AAC TCG CTA TCC	1295
35	80	Ser Ala Trp Cys Cys Pro Pro Glu Cys Trp Cys Arg Asn Ser Leu Ser	95
	1296	CTG TTG TCT ACC TGC TGG GGG CAC TGA CCA TGC TGA GTG AAA CGC AGC	1343
	96	Leu Leu Ser Thr Cys Trp Gly His ***	104
40	1344	ACA AGC TTG CTG GCG GAG GCG CTG GAG TCG CAG ACC CTG TTG GGG CCG	1391
	1392	CTC GAG CTG GTG GGC AGC CTC TTG GAG CAG AGT GCC CCG TGG CAG GAG	1439
	1440	CGC AGC ACC ATG TCC CTG CCC CCC GGG CTC CTG GGG AAC AGC TGG GGC	1487
	1488	GAA GGA GCA CCG GCC TGG GTC TTG CTG GAC GAG TGT GGC CTA GAG CTG	1535
	1536	GGG GAG GAC ACT CCC CAC GTG TGC TGG GAG CCG CAG GCC CAG GGC CGC	1583
45	1584	ATG TGT GCA CTC TAC GCC TCC CTG GCA CTG CTA TCA GGA CTG AGC CAG	1631
	1632	GAG CCC CAC TAG CCT GTG CCC GGG CAT GGC CTG GCA GCT CTC CAG CAG	1679
	1680	GGC AGA GTG TTT GCC CAC CAG CTG CTA GCC CTA GGA AGG CCA GGA GCC	1727
	1728	CAG TAG CCA TGT GGC CAG TCT ACC ATG GGG CCC AGG AGT TGG GGA AAC	1775
50	1776	ACA ATA AAG GTG GCA TAC GAA GGA AAA AAA AAA AAA AAA AA	1816

10. PP14733

A: 核苷酸序列 (SEQ ID NO:28) 长度: 2638

	1	GAGCCAGAAC CATGTTTTTT GGGAAATGGC TCAATAAACC TTAATTTATT GTTGCCCTTT
	61	GTGAAGACTA TTAAGAGCAA TACAAAAAAA AAAAATAGTT GATGTTGCTT TGTGTAGGCT
55	121	ACTTTAGAAT CTCATTTTGT CTATTTGCTT ATCCTAAAAA TTAAGTATAT TAAAATATTT
	181	ATCATATACT TGCCCTTATT TGTGAACATT GAGAAGTGGG TGCTTAACGA GCATTTATTA
	241	AGAGGGAGAT GCCATAAAG TTTGGACTAG TAGTATATAT CCACTTCAAT ATGTGGTCTG
	301	TTGATGACAC AGGTAATGAG TTTAGAAGGT TAGTAATGTA GCTCGAGTAT CCTTAGTATC
	361	CCTTAATTAT CTAGTATACA GTGATTATTT ATTAATATAG TCCCCTTTTT ATAAAACCTT

421 TTA AACCATC TTTAATTAGC ATCACTTAAT ATGGAACAAG CTCCTTTCTC CATTAGAATA
 481 ATATTTGTTT GAAAACCACC ACTATTATCT CTTCTAGCTT CACCATAAAT ATTTTATAAT
 541 TTGGCAAACA AAATTATTAT GACCATAGTT TTTATAATCA AACAACATTT TTTCTTTCTT
 601 AATATACATA GAATAGAATG AAGAACAATT TTTTCCCAT TACTGTCTT AGATGAATGT
 5 661 CTATGACTTT ATAACATTGA ATATCTAGTT ATTCCCAAAC TGTTAAAAAG TTATTCATAC
 721 TAATGATAAA ATTTTAGCAT ACCATTGTCT TAACTCTGA CTGAAGGCCT TTTTAAAAA
 781 ATTATTTTCA GGGGCCAGGC ACAGTGGCTC ACGCCTGTAA TCCCAGCGTT TTGGGAGGCC
 841 GAAGCAGGTG GATCGCCTCA GGTCAAGGAGT GCGAGACCAG CCTGGCCAAC ATGGTGAAAC
 901 CCCGTCTCCA CTAAAAATAC AAAAAATTA GCTAGGTATG GTGGCGGGCA CCTGTAATTC
 10 961 CAGCTACTTG GGAGGCTGAG GCAGGAGAAT CATTGACCC AGGAGTCGGA GGTGTCAGTG
 1021 AGCCAAGATC ACGCCACTGC ACTCCAGCCT GGGCAACAAG CACAAAATC CGTCTCAAAA
 1081 AAAAAATTAT TTTTAGGAGC ATTAGAGAAA ACCCTTGAAA TGTTTCCACT GTAACATTTA
 1141 CTGGTTTCAT GAACCTCAAA TTGATGAAA GGTCAATGCA TAGTAAAAGT AAACCTGATT
 1201 CACTATTGCC TGGTAATGTC AGATATAGTT ATAAATGAAG TCAGTTTCAT TGAAGTTAA
 15 1261 ATGAAGGTTA GGAAAAATAA TGATGCTTAT TTAGTAGACA AGAACAAAAT TAATATGGAT
 1321 TGTTTCATTA GCTGTTTCTT TAAAAAGATG TTAACCACCC TCATGTTTTT ACATTGAGGT
 1381 ATCCTTAGTC TGTTGGAGCA TGGAGAAGAG TACACATTTT CTCTACCCTG TGCATATGCT
 1441 CGGTCAATTT TGACTGTTCC TTGGGTAGAA CTGGGTGGCA AAGTCAGTGT CAACTGTGCA
 1501 AAAACTGGAT ATTCAGCCAG CATCACTTTT CATACCAAGC CATTITATGG TGGCAAACCTG
 20 1561 CATCGGTAAG AACATTTTAA TGATTATCTA GTTCTTATTC TTCCACTTTG ATTTTGTATT
 1621 ACAACTCCTT ATCTTTTTTC TGCTCTGTTT TATGTGAAGT TGGATTGTA TTAATGTTTA
 1681 TTTCTTATTT AAAATATTTT ATCTTTAAAG CATTCTGAAA GTCATGCTCT GTTTTTAAAA
 1741 GTAGTAACAT TAAGATATCT ACTCTTAAAA ATTTTCTGTA TGAAGTAACT AGTGAGCTTT
 1801 ATAGTAGAAT TTCTAACAAG TGAACATGCA GTCTCTGGCC CAATTCTGCC TTTAAGGGAT
 25 1861 ACAGATTGG ATCTGGACCT CAGTGAAGGC AGCTATAGCT ATAAGTTAGG GTGGAAATAT
 1921 ATCTAATCAT GAAGAATGGT TCAAGAAGCA CAATGATTAA AGCCAGTGTA TGTGCTTCAT
 1981 TAGATGCTTA GGTGGAAGCA ACTTGCAGTG ATAAATTTCT AGATCCAGTA CTTCTTATTG
 2041 TCAATGATAT TGATGCATTT TTTTTTTTGG TCATTTGTTT TTTTGAGACA GAGTCTCACT
 2101 CTGTCGCCTA GGCTGGAGTG CAGTGGCACA ATCTTGGCTC CTTGCAGCCG CCTCTGGGT
 30 2161 TCAAGCGATT TTTGTGCTC AGCCTCCCAA GTAGCTGTGA TTACGGCATG TGCCACCATG
 2221 CCTGGCTGCT TTTTGTATTT TTAGTTGAGA TGGGCTTCA CCATGTTGGC CAGGCTGGTC
 2281 TCAAACCTCT GGCCTTAAGT GATCTGCCTG CCTCAACCTC CCAAAGTGCT GGGAGCCACC
 2341 ACACCTGGCT AATATTGATG TATTTCTTAG TCACATGTGA TTTTTAGTC ATGGGTGGAT
 2401 CACCTGAGGT CAGGATTCG AGACCAGCCT AGCCAACAAG ACGAGACTCC GTCTCTACTA
 35 2461 AAAATACAAA ATTAGCTGAG TATGGTGGCG CATGCCTGTA ATCCCAGCTA TTCAGGAGGC
 2521 AAGGCTGGAG GATCACTTGA ACCCAGTAGG CAGAGGTTGC AGTGAGCTGA GGTGACCA
 2581 CTGCTCTCCA GCCGGGTGA CAAGAGTGAA ACTCCGTCTC AAAAAAAAAA AAAAAAAAAA

B: 氨基酸序列 (SEQ ID NO:29) 长度: 102

40 1 MKVRKNNDAY LLDKNKINMD CFISCFPKKM LTTLMFHSYG ILSLLEHGEE YTFSLPCAYA
 61 RSILTPWVE LGGKVSINCA KTGYSASITF HTPPFYGGKL HR

C. 核苷酸及氨基酸组合序列 (SEQ ID NO:30) 克隆号: PP14733

起始编码子: 1261 ATG 终止编码子: 1567 TAA 蛋白质分子量: 11561.97

45 1 GAG CCA GAA CCA TGT TTT TTG GGA AAT GGC TCA ATA AAC CTT AAT TTA 48
 49 TTG TTG CCC TTT GTG AAG ACT ATT AAG AGC AAT ACA AAA AAA AAA AAT 96
 97 AGT TGA TGT TGC TTT GTG TAG GCT ACT TTA GAA TCT CAT TTT GTC TAT 144
 145 TTG CTT ATC CTA AAA ATT AAC TAT ATT AAA ATA TTT ATC ATA TAC TTG 192
 193 CCC TTA TTT GTG AAC ATT GAG AAC TGA GTG CTT AAC GAG CAT TTA TTA 240
 50 241 AGA GGG AGA TGC CAT AAA AGT TTG GAC TAG TAG TAT ATA TCC ACT TCA 288
 289 ATA TGT GGT CTG TTG ATG ACA CAG GTA ATG AGT TTA GAA GGT TAG TAA 336
 337 TGT AGC TCG AGT ATC CTT AGT ATC CCT TAA TTA TCT AGT ATA CAG TGA 384
 385 TTA TTT ATT AAT ATA GTC CCC TTT TTA TAA AAC CTT TTA AAC CAT CTT 432
 433 TAA TTA GCA TCA CTT AAT ATG GAA CAA GCT CCT TTC TCC ATT AGA ATA 480

	481	ATA TTT GTT TGA AAA CCA CCA CTA TTA TCT CTT CTA GCT TCA CCA TAA	528
	529	ATA TTT TAT AAT TTG GCA AAC AAA ATT ATT ATG ACC ATA GTT TTT ATA	576
	577	ATC AAA CAA CAT TTT TTC TTT CTT AAT ATA CAT AGA ATA GAA TGA AGA	624
	625	ACA ATT TTT TCC CAT TAT ACT GTC TTA GAT GAA TGT CTA TGA CTT TAT	672
5	673	AAC ATT GAA TAT CTA GTT ATT CCC AAA CTG TTA AAA AGT TAT TCA TAC	720
	721	TAA TGA TAA AAT TTT AGC ATA CCA TTG TCT TAA ACT CTG ACT GAA GGC	768
	769	CTT TTT TAA AAA ATT ATT TTC AGG GGC CAG GCA CAG TGG CTC ACG CCT	816
	817	GTA ATC CCA GCG TTT TGG GAG GCC GAA GCA GGT GGA TCG CCT CAG GTC	864
	865	AGG AGT GCG AGA CCA GCC TGG CCA ACA TGG TGA AAC CCC GTC TCC ACT	912
10	913	AAA AAT ACA AAA AAA TTA GCT AGG TAT GGT GGC GGG CAC CTG TAA TTC	960
	961	CAG CTA CTT GGG AGG CTG AGG CAG GAG AAT CAT TCG ACC CAG GAG TCG	1008
	1009	GAG GTT GCA GTG AGC CAA GAT CAC GCC ACT GCA CTC CAG CCT GGG CAA	1056
	1057	CAA GCA CAA AAC TCC GTC TCA AAA AAA AAA TTA TTT TTA GGA GCA TTA	1104
	1105	GAG AAA ACC CTT GAA ATG TTT CCA CTG TAA CAT TTA CTG GTT TCA TGA	1152
15	1153	ACC TCA AAT TGA TGG AAA GGT CAT GTC ATA GTG AAA GTA AAC CTG ATT	1200
	1201	CAC TAT TGC CTG GTA ATG TCA GAT ATA GTT ATA AAT GAA GTC AGT TTC	1248
	1249	ATT GAA AGT TAA ATG AAG GTT AGG AAA AAT AAT GAT GCT TAT TTA CTA	1296
	1	Met Lys Val Arg Lys Asn Asn Asp Ala Tyr Leu Leu	12
20	1297	GAC AAG AAC AAA ATT AAT ATG GAT TGT TTC ATT AGC TGT TTC TTT AAA	1344
	13	Asp Lys Asn Lys Ile Asn Met Asp Cys Phe Ile Ser Cys Phe Phe Lys	28
	1345	AAG ATG TTA ACC ACC CTC ATG TTT TCA CAT TCA GGT ATC CTT AGT CTG	1392
25	29	Lys Met Leu Thr Thr Leu Met Phe Ser His Ser Gly Ile Leu Ser Leu	44
	1393	TTG GAG CAT GGA GAA GAG TAC ACA TTT TCT CTA CCC TGT GCA TAT GCT	1440
	45	Leu Glu His Gly Glu Glu Tyr Thr Phe Ser Leu Pro Cys Ala Tyr Ala	60
	1441	CGG TCA ATT TTG ACT GTT CCT TGG GTA GAA CTG GGT GGC AAA GTC AGT	1488
30	61	Arg Ser Ile Leu Thr Val Pro Trp Val Glu Leu Gly Gly Lys Val Ser	76
	1489	GTC AAC TGT GCA AAA ACT GGA TAT TCA GCC AGC ATC ACT TTT CAT ACC	1536
	77	Val Asn Cys Ala Lys Thr Gly Tyr Ser Ala Ser Ile Thr Phe His Thr	92
35	1537	AAG CCA TTT TAT GGT GGC AAA CTG CAT CGG TAA GAA CAT TTT AAT GAT	1584
	93	Lys Pro Phe Tyr Gly Gly Lys Leu His Arg ***	103
	1585	TAT CTA GTT CTT ATT CTT CCA CTT TGA TTT TGT ATT ACA ACT CCT TAT	1632
	1633	TCT TTT TCT GCT CTG TTC TAT GTG AAG TTG GAT TTG TAT TAA TGT TTA	1680
40	1681	TTT CTT ATT TAA AAT ATT TCA TCT TTA AAG CAT TCT GAA AGT CAT GCT	1728
	1729	CTG TTT TTA AAA GTA GTA ACA TTA AGA TAT CTA CTC TTA AAA ATT TTC	1776
	1777	CTG ATG AAG TAA CTA GTG AGC TTT ATA GTA GAA TTT CTA ACA AGT GAA	1824
	1825	CAT GCA GTC TCT GGC CCA ATT CTG CCT TTA AGG GAT ACA GAT TTG GAT	1872
	1873	CTG GAC CTC AGT GAA GGC AGC TAT AGC TAT AAG TTA GGG TGG AAA TAT	1920
45	1921	ATC TAA TCA TGA AGA ATG GTT CAA GAA GCA CAA TGA TTA AAG CCA GTG	1968
	1969	TAT GTG CTT CAT TAG ATG CTT AGG TGG AAG CAA CTT GCA GTG ATA AAT	2016
	2017	TTC TAG ATC CAG TAC TTC TTA TTG TCA ATG ATA TTG ATG CAT TTT TTT	2064
	2065	TTT TTG TCA TTT GTT TTT TTG AGA CAG AGT CTC ACT CTG TCG CCT AGG	2112
	2113	CTG GAG TGC AGT GGC ACA ATC TTG GCT CCT TGC AGC CGC CTC CTG GGT	2160
50	2161	TCA AGC GAT TTT TGT GCC TCA GCC TCC CAA GTA GCT GTG ATT ACG GCA	2208
	2209	TGT GCC ACC ATG CCT GGC TGC TTT TTG TAT TTT TAG TTG AGA TGG GCT	2256
	2257	TTC ACC ATG TTG GCC AGG CTG GTC TCA AAC TCC TGG CCT TAA GTG ATC	2304
	2305	TGC CTG CCT CAA CCT CCC AAA GTG CTG GGA GCC ACC ACA CCT GGC TAA	2352
	2353	TAT TGA TGT ATT TCT TAG TCA CAT GTG ATT TTT TAG TCA TGG GTG GAT	2400
55	2401	CAC CTG AGG TCA GGA TTT CGA GAC CAG CCT AGC CAA CAT GAC GAG ACT	2448
	2449	CCG TCT CTA CTA AAA ATA CAA AAT TAG CTG AGT ATG GTG GCG CAT GCC	2496
	2497	TGT AAT CCC AGC TAT TCA GGA GGC AAG GCT GGA GGA TCA CTT GAA CCC	2544
	2545	AGT AGG CAG AGG TTG CAG TGA GCT GAG GTT GCA CCA CTG CTC TCC AGC	2592
	2593	CGG GGT GAC AAG AGT GAA ACT CCG TCT CAA AAA AAA AAA AAA A	2638
60			

11. PP14762

A: 核苷酸序列 (SEQ ID NO:31) 长度: 2205

```

1   GATCTCACAG GGATTGAGTG GGCCCAAAC GGCTGTGTGC TGGCAGTGTG GGACACCTGC
61  TTGGAGGTAT GAAGATGACC AGGTGACATG TTTGCTTTCG AGAGCTTTTC CCGGAATTTT
5   121  ATGCTTTCTG AATGCTGCAA TATAGTTTGC TTTATCTTAT TTCAGATAAC AAGTTGAGCT
181  TTGTTTTAAA AGACAGGGTC TTGCTCTGTC ACCCAGGCTG GAGTGCAGTG GTGTGATCAC
241  GGCCCACTGC AGCCTCGACC TCCTGGACTC AAGTGATCCT CCTACCTCAG CCTCCCAAGT
301  AGCTGGGACC ACAAGCATGC ACCACCACGC CCGGCTAATT TTTGTATTCT TTGTAGAGAT
361  GGGGTTTAC  CGTGTGCCC  AGACTGCTCT TGAACTTCTG TGTTC AAGTG  ATCTGCCAGC
10  421  TTCGGCCTCC CAAAGTGTG GGATTACAGG CGTGAGCCAC AGTGCCCGGC CGAGTTGAAC
481  TTTATCAGC  GAGCCATTAG GTTGACCTCC CAATTTCCCC TTTTCCTCA TGCTCTCACT
541  GCCTCTCCGT GACTATCTCA GTACTAGCTG GCAGTCCAG CTTTTCCAG GCCACGGGA
601  CAGTCTGCC  AGCTTTGTGC TATCTGCACC AATGTCACTC TGCAGCACCC GCGTCTGTGG
661  TGTGTTCTG  GCGACGGGG  TGTATTCAAT ATTAAGTAA ATGTCTCATC TGCAGTACAA
15  721  GATTCTGTG  TACTCATTGG ATGGCCGGTT GTTGTCCAG TACAGCGCTT ACGAGTGGTC
781  CTTGGGCATC AAGTCTGTGG CCTGGAGCCC CAGCAGTCAG TTCCTGGCAG TTGGGAGCCG
841  TCCCCAGCC  GCCCAACCTC CTAAGCGCTG AAGGAATTAT TTTTGGGAGT CATGACCATA
901  GCTGCCTGCT TCCCCAACCC CTCGACGACG AAAAGGACTT CGGCCCCCGG CCGGGCGGCT
961  CCGGGGAAA  AAAGAGAGAG CGACCTCCGC CCCGCGCTCA GGGCCACTCT AGAGGGAGAA
20  1021 GCTGCCCGA  GGCTGGCAGA GCGCCAGCC GCGGTGAGA CCGGAAGCCT TCTGGAGCCC
1081 AAGGCTGTG  ACGTCCCTG  TGCTGATTCT CTGCCTAGGA AAGGACCATG CAGCTGGAGA
1141 TCAAAGTAGC CCTGAAGTTC ATCATCTTCT ACTTGTAACA CAAGCTGCTC TGGCGCCTCT
1201 GAAAAAGAAA TAAGAAGCCC ACTGGTACCC TGACAAACCA CTGAAGGGCT CTGGCTTCCA
1261 CACCGAGGAG ATGGTGGACC CCGTTGGGA  GCTGGCCGCT AAGCGAAGTG GCCTGACAGT
25  1321 GGAAGATGTG CGGGCCAATG TGCCTGAGGA GCTGAGCATA TGGATTGACC CCTTCGGGGT
1381 GTCCTACCAG ATTGGTGAGA AGGGAGCAGT GAAAGTGTG TACCTGGATG ACAGTGACGG
1441 CTGTGGGGCC CCGGAGCTGG ACATGAAGAT CAAGAGCAGT TTCACTCTG ACGACCAGAT
1501 GCTCTTCTC  CTCGGAAGCC AGGACAGCTC CTTGTCCAAC TCCCCGTCG TATGCTTTGC
1561 CCATCACCCA GCCCCACCTT CATCTCTGC CTGCCCAGC CCATCACCTT CACCATGGAC
30  1621 TCCTTTGCTG CCACCAAATT AGGCTCCACT AAGATGAAGA AGAGGGGTGG GGCAACAAGT
1681 GGTAGGGGTG CAGCCAGCGG TGAAGCAGGT GGCCAGCAGC CAGTACAGCA GCCTCCACGG
1741 CCCGCTTCAA CACCAACAGC TTGTTGAAAC ACAGGAGTCT TTCTCTGTCT GTGCATTCAT
1801 TGAAGTTCAT CCCAGCCAAC CGGCTCTCA ATCCCAGCTC TCGCCCCAAG GCCACGTACA
1861 GTCTCAAGTA CAACAGTGGT GGCTTGCCCA GTTTCTTCTT TGATGGGGCC AATGGCCAGG
35  1921 GTAGTGACAC CCGAGCCCTC TTTGGAGGCA GTGGGACTGT CACCTGCAAC GCCAGCAGCT
1981 TCTACATGGT CTAGGTATTT GGAGTGGTG CCACCGCCTC TTCCTGGAGA AGGCACCCTT
2041 CATGGAAGAC CTCAGCTACA ACCTGAACAC CATGCAATAT CCTAGTCAAC CGTTCACGTC
2101 CGTGGTGTG  GCCAACTGAC CGTCTACCTG CCCTGGGGCC GGGAGCACCC AAGACCACAG
2161 AAAAGAGAAA GAAAAGGAAA GGCCAAAAAA AAAAAAAAAA AAAAA

```

B: 氨基酸序列 (SEQ ID NO:32) 长度: 123

```

1   MVDPVGELAA KRSGLTVEDV RANVPEELSI WIDPFGVSYQ IGEKGAVKVL YLDDSDGCGA
61  PELDMKIKSS FTPDDQMLFL LGSQDSSLN SPSLCFAHHP APPSSASPS PSPSPWTPLL
121  PPN

```

C. 核苷酸及氨基酸组合序列 (SEQ ID NO:33) 克隆号: PP14762

起始编码子: 1271 ATG 终止编码子: 1640 TAG 蛋白质分子量: 13037.87

```

1   G ATC TCA CAG GGA TTG AGT GGG CCC CAA ACG GCT GTG TGC TGG CAG      46
47  TGT GGG ACA CCT GCT TGG AGG TAT GAA GAT GAC CAG GTG ACA TGT TTG      94
50  95  CTT TCG AGA GCT TTT CCC GGA ATT TCA TGC TTT CTG AAT GCT GCA ATA     142
143 TAG TTT GCT TTA TCT TAT TTC AGA TAA CAA GTT GAG CTT TGT TTT AAA     190
191 AGA CAG GGT CTT GCT CTG TCA CCC AGG CTG GAG TGC AGT GGT GTG ATC     238
239 ACG GCC CAC TGC AGC CTC GAC CTC CTG GAC TCA AGT GAT CCT CCT ACC     286

```

	287	TCA GCC TCC CAA GTA GCT GGG ACC ACA AGC ATG CAC CAC CAC GCC CGG	334
	335	CTA ATT TTT GTA TTC TTT GTA GAG ATG GGG TTT CAC CGT GTT GCC CAG	382
	383	ACT GCT CTT GAA CTT CTG TGT TCA AGT GAT CTG CCA GCT TCG GCC TCC	430
	431	CAA AGT GCT GGG ATT ACA GGC GTG AGC CAC AGT GCC CGG CCG AGT TGA	478
5	479	ACT TTT ATC AGC GAG CCA TTA GGT TGA CCT CCC AAT TTC CCC TTT TTC	526
	527	CTC ATG CTC TCA CTG CCT CTC CGT GAC TAT CTC AGT ACT AGC TGG CAG	574
	575	TTC CAG CTT TTC CAC GGC CAC GCG GAC AGT CCT GCC AGC TTT GTG CTA	622
	623	TCT GCA CCA ATG TCA CTC TGC AGC ACC CGC GTC TGT GGT GTG TTC CTG	670
	671	GCG ACG GGG GTG TAT TCA TTA TTA ACT GAA ATG TCT CAT CTG CAG TAC	718
10	719	AAG ATT CTG CTG TAC TCA TTG GAT GGC CGG TTG TTG TCC ACG TAC AGC	766
	767	GCT TAC GAG TGG TCC TTG GGC ATC AAG TCT GTG GCC TGG AGC CCC AGC	814
	815	AGT CAG TTC CTG GCA GTT GGG AGC CGT CCC CCA GCC GCC CAA CCT CCT	862
	863	AAG CGC TGA AGG AAT TAT TTT TGG GAG TCA TGA CCA TAG CTG CCT GCT	910
	911	TCC CCC AAC CCT CGA CGA CGA AAA GGA CTT CGG CCC CCG GCC GGG CGG	958
15	959	CTC CGG GGG AAA AAA GAG AGA GCG ACC TCC GCC CCG CGC TCA GGG CCA	1006
	1007	CTC TAG AGG GAG AAG CTG CCC CGA GGC TGG CAG AGC GCC CAG CCG CGG	1054
	1055	TGG AGA CCG GAA GCC TTC TGG AGC CCA AGG CTG TGC ACG TCC CCT GTG	1102
	1103	CTG ATT CTC TGC CTA GGA AAG GAC CAT GCA GCT GGA GAT CAA AGT AGC	1150
	1151	CCT GAA CTT CAT CAT CTT CTA CTT GTA CAA CAA GCT GCT CTG GCG CCT	1198
20	1199	CTG AAA AAG AAA TAA GAA GCC CAC TGG TAC CCT GAC AAA CCA CTG AAG	1246
	1247	GGC TCT GGC TTC CAC ACC GAG GAG ATG GTG GAC CCC GTT GGG GAG CTG	1294
	1	Met Val Asp Pro Val Gly Glu Leu	8
	1295	GCC GCT AAG CGA AGT GGC CTG ACA GTG GAA GAT GTG CGG GCC AAT GTG	1342
25	9	Ala Ala Lys Arg Ser Gly Leu Thr Val Glu Asp Val Arg Ala Asn Val	24
	1343	CCT GAG GAG CTG AGC ATA TGG ATT GAC CCC TTC GGG GTG TCC TAC CAG	1390
	25	Pro Glu Glu Leu Ser Ile Trp Ile Asp Pro Phe Gly Val Ser Tyr Gln	40
30	1391	ATT GGT GAG AAG GGA GCA GTG AAA GTG CTG TAC CTG GAT GAC AGT GAC	1438
	41	Ile Gly Glu Lys Gly Ala Val Lys Val Leu Tyr Leu Asp Asp Ser Asp	56
	1439	GGC TGT GGG GCC CCG GAG CTG GAC ATG AAG ATC AAG AGC AGT TTC ACT	1486
35	57	Gly Cys Gly Ala Pro Glu Leu Asp Met Lys Ile Lys Ser Ser Phe Thr	72
	1487	CCT GAC GAC CAG ATG CTC TTT CTC CTC GGA AGC CAG GAC AGC TCC TTG	1534
	73	Pro Asp Asp Gln Met Leu Phe Leu Leu Gly Ser Gln Asp Ser Ser Leu	88
	1535	TCC AAC TCC CCG TCG CTA TGC TTT GCC CAT CAC CCA GCC CCA CCT TCA	1582
40	89	Ser Asn Ser Pro Ser Leu Cys Phe Ala His His Pro Ala Pro Pro Ser	104
	1583	TCC TCT GCC TCG CCC AGC CCA TCA CCT TCA CCA TGG ACT CCT TTG CTG	1630
	105	Ser Ser Ala Ser Pro Ser Pro Ser Pro Ser Pro Trp Thr Pro Leu Leu	120
45	1631	CCA CCA AAT TAG GCT CCA CTA AGA TGA AGA AGA GGG GTG GGG CAA CAA	1678
	121	Pro Pro Asn ***	124
	1679	GTG GTA GGG GTG CAG CCA GCG GTG AAG CAG GTG GCC AGC AGC CAG TAC	1726
	1727	AGC AGC CTC CAC GGC CCG CTT CAA CAC CAA CAG CTT GTT GAA ACA CAG	1774
50	1775	GAG TCT TTC TCT GTC TGT GCA TTC ATT GAA CTT CAT CCC AGC CAA CCG	1822
	1823	GCC TCT CAA TCC CAG CTC TCG CCC CAA GGC CAC GTA CAG TCT CAA GTA	1870
	1871	CAA CAG TGG TGG CTT GCC CAG TTT CTT CCT TGA TGG GGC CAA TGG CCA	1918
	1919	GGG TAG TGA CAC CCG AGC CCT CTT TGG AGG CAG TGG GAC TGT CAC CTG	1966
	1967	CAA CGC CAG CAG CTT CTA CAT GGT CTA GGT ATT TGG AGG TGG TGC CAC	2014
55	2015	CGC CTC TTC CTG GAG AAG GCA CCC TTC ATG GAA GAC CTC AGC TAC AAC	2062
	2063	CTG AAC ACC ATG CAA TAT CCT AGT CAA CCG TTC CAG TCC GTG GTG CTG	2110
	2111	GCC AAC TGA CCG TCT ACC TGC CCT GGG GCC GGG AGC ACC CAA GAC CAC	2158
	2159	AGA AAA GAG AAA GAA AAG GAA AGG CCA AAA AAA AAA AAA AAA AA	2205