

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

C12N 15/31

C07K 14/37 C12N 15/80

C12N 1/15 C12Q 1/02

C12P 19/42



[12] 发明专利申请公开说明书

[21] 申请号 02817625.1

[43] 公开日 2004 年 12 月 8 日

[11] 公开号 CN 1553956A

[22] 申请日 2002.8.9 [21] 申请号 02817625.1

[30] 优先权

[32] 2001.8.10 [33] DE [31] 10139454.3

[32] 2001.8.10 [33] DE [31] 10139455.1

[32] 2001.8.10 [33] DE [31] 10139457.8

[32] 2001.8.10 [33] DE [31] 10139458.6

[32] 2001.8.10 [33] DE [31] 10139459.4

[32] 2001.8.10 [33] DE [31] 10139460.8

[32] 2001.8.10 [33] DE [31] 10139461.6

[32] 2001.8.10 [33] DE [31] 10139462.4

[32] 2001.8.10 [33] DE [31] 10139463.2

[32] 2001.8.10 [33] DE [31] 10139464.0

[32] 2002.3.6 [33] DE [31] 10209819.0

[32] 2002.3.6 [33] DE [31] 10209816.6

[32] 2002.4.11 [33] DE [31] 10216033.3

[32] 2002.5.16 [33] DE [31] 10221911.7

[32] 2002.5.16 [33] DE [31] 10221928.1

[32] 2002.5.16 [33] DE [31] 10221909.5

[32] 2002.6.7 [33] DE [31] 10225390.0

[32] 2002.6.7 [33] DE [31] 10225392.7

[32] 2002.6.21 [33] DE [31] 10227796.6

[32] 2002.7.29 [33] DE [31] 10234455.8

[86] 国际申请 PCT/EP2002/008937 2002.8.9

[87] 国际公布 WO2003/014353 德 2003.2.20

[85] 进入国家阶段日期 2004.3.9

[71] 申请人 巴斯福股份公司

地址 德国路德维希港

[72] 发明人 M·考罗什 H·阿尔特赫费

B·克勒格尔

J·L·雷韦尔塔多瓦尔

[74] 专利代理机构 北京市中咨律师事务所

代理人 黄革生 隗永良

权利要求书 3 页 说明书 41 页 序列表 69 页

附图 10 页

[54] 发明名称 来自棉桃阿舒氏囊霉的与跨膜转运相关的新基因产物

[57] 摘要

本发明涉及来自棉桃阿舒氏囊霉的新颖多核苷酸；与之杂交的寡核苷酸；包含这些多核苷酸的表达盒和载体；用其转化的微生物；由这些多核苷酸编码的多肽；及该新颖多肽和多核苷酸作为靶点在改善阿舒氏囊霉属微生物的跨膜转运，且尤其是改善阿舒氏囊霉属微生物的维生素 B2 生产中的用途。

1. 可自棉桃阿舒氏囊霉分离的多核苷酸, 其编码与跨膜转运相关的蛋白和/或跨膜蛋白。

2. 权利要求 1 的多核苷酸, 其与跨膜转运相关和/或为跨膜蛋白并具有表 1 所示的结构和/或功能特性。

3. 包含 SEQ ID NO: 1, 5, 10, 14, 19, 23, 28, 33, 37 或 42 所示的核酸序列并优选自棉桃阿舒氏囊霉中分离的权利要求 1 或 2 的多核苷酸; 其互补多核苷酸; 以及通过遗传密码简并性衍生自这些多核苷酸的序列。

4. 权利要求 4 的多核苷酸, 其包含 SEQ ID NO: 3, 8, 12, 17, 21, 26, 31, 35, 40 或 44 显示的核酸序列或其片段。

5. 寡核苷酸, 其可以与任一前述权利要求的多核苷酸杂交, 且尤其是在严格的条件下杂交。

6. 多核苷酸, 其可以与权利要求 5 的寡核苷酸杂交, 且尤其是在严格的条件下杂交, 并编码阿舒氏囊霉属微生物的基因产物或该基因产物的功能等价物。

7. 多肽及其功能等价物, 特别是具有权利要求 2 所述的结构和/或功能特性之一的功能等价物, 其中, 所述多肽由包含权利要求 1 至 4 中任意一项的核酸序列或其片段的多核苷酸编码, 或由权利要求 6 的多核苷酸编码; 或具有包含 SEQ ID NO: 2, 4, 6, 7, 9, 11, 13, 15, 16, 18, 20, 22, 24, 27, 29, 30, 32, 34, 36, 38, 39, 41, 43 或 SEQ ID NO: 45 中所示的至少 10 个连续氨基酸残基的氨基酸序列。

8. 包含权利要求 1 至 6 中任意一项的核酸序列的表达盒, 其中, 所述核酸序列与至少一个调节核酸序列可操作地连接。

9. 包含至少一个权利要求 8 的表达盒的重组载体。

10. 由至少一个权利要求 9 的载体转化的原核或真核宿主。

11. 原核或真核宿主, 在该宿主中, 至少一个编码权利要求 7 的多肽的基因的功能性表达受到调节; 或权利要求 7 的多肽的生物活性被降低或

增加。

12. 权利要求 10 或 11 的宿主，其来自阿舒氏囊霉属。

13. 权利要求 8 的表达盒、权利要求 9 的载体或权利要求 10 至 12 之任一项的宿主在维生素 B2 和/或其前体和/或其衍生物的微生物生产中的用途。

14. 权利要求 8 的表达盒、权利要求 9 的载体或权利要求 10 至 12 之任一项的宿主在权利要求 7 的多肽的重组生产中的用途。

15. 检测可以用于调节维生素 B2 和/或其前体和/或其衍生物的微生物生产的效应物靶点的方法，其中，用效应物处理能够通过微生物法生产维生素 B2 和/或其前体和/或其衍生物的微生物，所述效应物可以与选自权利要求 7 的多肽或其核酸编码序列的靶点相互作用，尤其是结合；验证效应物对微生物法生产的维生素 B2 和/或其前体和/或其衍生物的产量的影响；并且适当时分离该靶点。

16. 调节维生素 B2 和/或其前体和/或其衍生物的微生物生产的方法，其中，用效应物处理能够以微生物法生产维生素 B2 和/或其前体和/或其衍生物的微生物，所述效应物可以与选自权利要求 7 的多肽或其核酸编码序列的靶点相互作用。

17. 针对选自权利要求 7 的多肽或其核酸编码序列的靶点的效应物，其中所述效应物选自：

- a) 抗体或其抗原结合片段；
- b) 可以与权利要求 7 的多肽相互作用的不同于 a) 的多肽配体；
- c) 可以调节权利要求 7 的多肽的生物活性的低分子量效应物；
- d) 反义核酸序列。

18. 微生物生产维生素 B2 和/或其前体和/或其衍生物的方法，其中在利于产生维生素 B2 和/或其前体和/或其衍生物的条件下培养权利要求 10 至 12 之任一项的宿主，自培养混合物中分离所需的产物。

19. 权利要求 18 的方法，其中用权利要求 17 的效应物在培养前和/或培养过程中处理宿主。

20. 权利要求 18 或 19 的方法，其中宿主选自阿舒氏囊霉属微生物。

21. 权利要求 18 至 20 之任一项的方法，其中微生物为权利要求 10 至 12 之任一项的宿主。

22. 权利要求 1 至 4 和 6 之任一项的多核苷酸或权利要求 7 的多肽作为靶点在调节阿舒氏囊霉属微生物的维生素 B2 和/或其前体和/或其衍生物的生产中的用途。

23. 权利要求 1 至 4 和 6 之任一项的多核苷酸或权利要求 7 的多肽作为靶点在培养阿舒氏囊霉属微生物生产维生素 B2 和/或其前体和/或其衍生物的过程中的用途，用于调节阿舒氏囊霉属微生物的跨膜转运、跨膜蛋白活性或与之相关的后续状态。

24. 权利要求 12 的宿主，其具有增强的对外界条件的细胞应答。

来自棉桃阿舒氏囊霉的与跨膜转运相关的新基因产物

技术领域

本发明涉及来自棉桃阿舒氏囊霉 (*Ashbya gossypii*) 的新颖多核苷酸; 与之杂交的寡核苷酸; 包含这些多核苷酸的表达盒和载体; 由该多核苷酸转化的微生物; 由这些多核苷酸编码的多肽; 及将该新颖多肽和多核苷酸用作调节阿舒氏囊霉属微生物中跨膜运输, 且尤其是提高维生素 B2 的产量的靶点。

背景技术

维生素 B2 (核黄素, 乳黄素) 是一种碱敏感和光敏感的维生素, 其在溶液中显示黄绿色荧光。维生素 B2 的缺乏可以导致外胚层的损害, 尤其是白内障、角膜炎、角膜血管化, 或导致自主泌尿生殖系统紊乱。维生素 B2 是 FAD 和 FMN 分子的前体, 这两种分子与 NAD^+ 和 NADP^+ 在生物学上对氢传递具有重要意义。维生素 B2 通过磷酸化 (FMN) 和随后的腺苷酸化 (FAD) 形成这两种分子。

维生素 B2 在植物, 酵母菌以及多种微生物中由 GTP 和核酮酸-5-磷酸合成。该反应途径首先打开 GTP 的嘧啶环并消除一个磷酸残基。保留的磷酸盐经脱氨, 还原, 和消除形成 5-氨基-6-核醇基 (ribityl) 氨基-2, 4-嘧啶酮。该化合物与 3, 4-二羟基-2-丁酮-4-磷酸反应产生双环分子 6, 7-二甲基-8-核醇基二氧四氢蝶啶。通过歧化反应该化合物转化成三环化合物核黄素, 该反应转移了一个 4-碳单位。

维生素 B2 存在于许多蔬菜和肉类中, 在谷类产品中的含量较低。成人每日的维生素 B2 需要量为约 1.4 到 2 毫克。核黄素又是辅酶 FMN 和 FAD 在人体中的主要分解产物并就此被排出。

维生素 B2 是人类和动物的重要营养物质。因此研究人员努力使维生素 B2 可工业规模产生，并建议由微生物的途径合成维生素 B2。可用于该目的的微生物是，例如，枯草杆菌 (*Bacillus subtilis*)，子囊菌阿舒假囊酵母菌 (*Eremothecium ashbyii*)，棉桃阿舒氏囊霉及酵母 *Candida flareri* 和酿酒酵母 (*Saccharomyces cerevisiae*)。为此使用的营养培养基包含作为碳源的糖蜜或植物油、无机盐、氨基酸、动物或植物胨和蛋白质及维生素添加剂。在无菌有氧浸没培养方法中，产率为每升培养液在几天内生成 10 克以上维生素 B2。所需条件为培养液良好的通风，小心搅动及设定低于约 30°C 的温度。去掉生物质、蒸发并干燥浓缩液得到富含维生素 B2 的产物。

例如，WO-A-92/01060, EP-A-0 405 370 和 EP-A-0 531 708 中描述了维生素 B2 的微生物生产。

在 Ullmann 氏工业化学百科全书 A27 卷, 521 页及以下等中，可以找到关于维生素 B2 的重要性，出现，生产，生物合成和使用的调查。

细胞膜在细胞中起多种作用。首先，膜将细胞内容物从周围环境中划分出来，使得细胞保持完整。膜还起屏障的作用，使得危险或不良化合物不能进入而所需的化合物不能流出。由于细胞膜的结构：双层脂分子，其中极性头部基团朝向外（于细胞外或细胞内部）而非极性尾部朝向双层中央并形成疏水核心，其对例如蛋白质，水分子和离子等亲水化合物的非易化扩散天然地不能通透（膜的结构和功能的概述见 Gennis, R.B. (1989) 生物膜，分子结构和功能 (*Biomembranes, Molecular Structure and Function*, Springer: Heidelberg)。由于有效阻止了化合物透过膜的扩散，该屏障使细胞较周围介质可以含有较高的所需化合物浓度和相对低的不需要的化合物的浓度。

然而，该膜也成为输入所需分子和输出废物分子的有效屏障。为克服这一困难，细胞膜含有多种能够促进不同类型化合物跨膜运输的转运蛋白：孔或通道和运载蛋白。前者形成一种通过膜的可调节孔道，是整合膜蛋白，偶尔为蛋白复合物。通常这种调节或“门控”对通过该孔或该通道

的待转运底物具特异性，因而这些跨膜结构对特定类别的底物特异；例如钾通道的构造使得只有具有与钾相似的电荷和大小的离子才能通过。通道和孔蛋白具有一定的疏水和亲水结构域，使得该蛋白的疏水部分能够结合于膜内，而亲水部分则构成通道的内侧，从而提供一种使所选亲水分子能够通过受保护的亲水环境。本领域已知许多这样的孔/通道，包括钾通道、钙通道、钠通道和氯离子通道等。

由孔和通道介导的这一系统限于非常小的分子，例如离子，因为使完整的蛋白可能经易化扩散通过的足够大的孔或通道将不能阻止更小的分子通过。由于转运的发生需要浓度梯度的驱动力，通过这种方法运输分子也偶尔被称为“易化扩散”。当较大分子（例如葡萄糖或其他糖）在膜一侧的浓度大于另一侧时，通透酶可以类似地促成比较大分子易化扩散进入细胞（也称为“单向转运”）。与孔或通道不同，这些整合蛋白质（常含有6至14个跨膜螺旋）不形成通过膜的开放通道，但其确实与膜表面的目标分子结合，并经过构象改变使目标分子在膜另一侧释放。

然而，细胞常常需要逆存在的浓度梯度输入或输出（“主动转运”）分子，在这种情况下易化扩散不能发生。细胞的这种膜运输通常有两种机制：同向转运或逆向转运，和能量偶联的转运，例如由ABC转运蛋白介导的转运。同向转运和逆向转运系统将两种不同分子的跨膜移动偶联在一起（通过对两种不同分子具有不同的结合位点的通透酶实现）；两种分子在同向转运中运输方向相同，而在逆向转运中，一种分子被输入而另一种被输出。因为两种分子之一顺浓度梯度移动，该转运在能量上是可能的，而且只有所需化合物逆优势浓度梯度同时移动才能实现这一能量上有利的过程。有些分子在能量驱动的过程中可逆浓度梯度跨膜转运，例如，通过ABC转运蛋白的转运。在这一系统中，位于膜上的转运蛋白具有ATP结合盒子，在与目标分子结合后ATP被转化成 $\text{ADP} + \text{P}_i$ ，产生的能量释放后促成目标分子在转运蛋白的协助下向膜对侧移动。所有转运系统的更具体的描述见 Bamberg, E.等(1993)“离子泵在脂双层膜上的电荷运输”，*Q. Rev. Biophys.* 26: 1-25; Findlay, J.B.C. (1991)“膜转运系统的结构和功能”，

Curr. Opin. Struct. Biol. 1: 804-810; Higgins, C.F. (1992) “从微生物到人类的 ABC 转运蛋白”, Ann. Rev. Cell. Biol. 8: 67-113; Gennis, R.B. (1989) “孔, 通道和转运蛋白”, 《Biomembranes, Molecular Structure and Function》, Springer: Heidelberg, 270-322 页; 和 Nikaido, H. 和 Saier, H. (1992) “细菌中的转运蛋白: 其设计中的共同主题”, Science 258: 936-942, 及这些文献中引用的参考文献。目前仍无利用跨膜转运基因产生具有改变的跨膜转运特性的微生物, 优选地为阿舒氏囊霉属, 尤其是棉桃阿舒氏囊霉株的描述。

发明概述

本发明的目的是提供一种可用于影响阿舒氏囊霉属微生物, 尤其是棉桃阿舒氏囊霉的跨膜转运的新颖靶点。具体目的为改善这种微生物的跨膜转运。进一步的目的是通过这种微生物提高维生素 B2 的产量。

我们发现尤其可以通过提供在棉桃阿舒氏囊霉的维生素 B2 生产过程中被向上或下调节的编码核酸序列实现该目的(基于试验部分详细描述的结果)。

我们发现具体地可以通过提供可分离自棉桃阿舒氏囊霉并编码跨膜转运相关蛋白和/或跨膜蛋白, 且尤其是具有表 1 中所示的结构(例如序列同源性)和/或功能特性(如酶活性)的多核苷酸实现该目的:

a) 优选被上调的核酸序列, 其编码具有线粒体能量传递蛋白的功能的蛋白。

本发明此方面的一个优选实施方案中, 分离了编码本发明核酸序列的特征性部分序列的 DNA 克隆, 其内部名称为“Oligo 19”。

在再一优选的实施方案中, 根据本发明分离了编码本发明的核酸的完全序列的 DNA 克隆, 其内部名称为“Oligo 19v”。

本发明一方面涉及包含 SEQ ID NO:1 所示的核酸序列的多核苷酸。本发明另一方面涉及包含 SEQ ID NO:3 所示的核酸序列或其片段的多核苷酸。这些多核苷酸优选地自阿舒氏囊霉属微生物, 尤其是棉桃阿舒氏囊霉

中分离。本发明还涉及与这些多核苷酸互补的多核苷酸，以及这些多核苷酸经遗传密码简并性衍生而成的序列。

“Oligo 19”和“Oligo 19v”的插入序列与来自酿酒酵母的 MIPS 标签“Ygr257c”有显著的同源性。这些插入序列具有 SEQ ID NO: 1 或 SEQ ID NO: 3 所示的核酸序列。来自 SEQ ID NO: 1 的相应互补序列或来自 SEQ ID NO: 3 所示的编码链的氨基酸序列或氨基酸部分序列与酿酒酵母线粒体能量传递蛋白具有显著的序列同源性。

b) 优选被上调的核酸序列，其编码具有与酿酒酵母 ABC 转运蛋白相应的功能的蛋白。ABC (ATP 结合盒子) 蛋白起运输系统的作用，并参与从细胞摄取或释放底物。ATP 水解在此驱动该运输过程。

本发明此方面的一个优选实施方案中，分离了编码本发明核酸序列的特征性部分序列的 DNA 克隆，其内部名称为“Oligo 24”。

在另一优选的实施方案中根据本发明分离了编码本发明的核酸的完全序列的 DNA 克隆，其内部名称为“Oligo 24v”。

本发明一方面涉及包含 SEQ ID NO: 5 所示的核酸序列的多核苷酸。本发明另一方面涉及包含 SEQ ID NO: 8 所示的核酸序列或其片段的多核苷酸。这些多核苷酸优选地自阿舒氏囊霉属微生物，尤其是棉桃阿舒氏囊霉中分离。本发明还涉及与这些多核苷酸互补的多核苷酸，以及这些多核苷酸经遗传密码简并性衍生而成的序列。

“Oligo 24”和“Oligo 24v”的插入序列与来自酿酒酵母的 MIPS 标签“Mdl2”有显著的同源性，这些插入序列具有 SEQ ID NO: 5 或 SEQ ID NO: 8 所示的核酸序列。来自 SEQ ID NO: 5 的相应互补序列或来自 SEQ ID NO: 8 所示的编码链的氨基酸序列或氨基酸部分序列与酿酒酵母 ABC 转运蛋白具有显著的序列同源性。

c) 优选被下调的核酸序列，其编码具有线粒体膜整合蛋白的功能的蛋白。

本发明此方面的一个优选实施方案中，分离了编码本发明核酸序列的特征性部分序列的 DNA 克隆，其内部名称为“Oligo 109”。

本发明另一优选的实施方案中根据本发明分离了编码本发明的核酸的完全序列的 DNA 克隆，其内部名称为“Oligo 109v”。

本发明一方面涉及包含 SEQ ID NO:10 所示的核酸序列的多核苷酸。本发明另一方面涉及包含 SEQ ID NO:12 所示的核酸序列或其片段的多核苷酸。该多核苷酸优选地自阿舒氏囊霉属微生物，尤其是棉桃阿舒氏囊霉中分离。本发明还涉及与这些多核苷酸互补的多核苷酸，以及这些多核苷酸经遗传密码简并性衍生而成的序列。

“Oligo 109”和“Oligo 109v”的插入序列与来自酿酒酵母的 MIPS 标签“Prp12”有显著的同源性，这些插入序列具有 SEQ ID NO: 10 或 SEQ ID NO:12 所示的核酸序列。来自该编码链的氨基酸序列或氨基酸部分序列与酿酒酵母线粒体膜整合蛋白具有显著的序列同源性。

d) 优选被下调的核酸序列，其编码具有线粒体内膜转运蛋白的功能的蛋白。

本发明此方面的一个优选实施方案中，分离了编码本发明核酸序列的特征性部分序列的 DNA 克隆，其内部名称为“Oligo 163”。

本发明另一优选的实施方案中根据本发明分离了编码本发明的核酸的完全序列的 DNA 克隆，其内部名称为“Oligo 163v”。

本发明一方面涉及包含 SEQ ID NO:14 所示的核酸序列的多核苷酸。本发明另一方面涉及包含 SEQ ID NO:17 所示的核酸序列或其片段的多核苷酸。这些多核苷酸优选地自阿舒氏囊霉属微生物，尤其是棉桃阿舒氏囊霉中分离。本发明还涉及与这些多核苷酸互补的多核苷酸，以及这些多核苷酸经遗传密码简并性衍生而成的序列。

“Oligo 163”和“Oligo 163v”的插入序列与来自酿酒酵母的 MIPS 标签“Flx1”有显著的同源性，这些插入序列具有 SEQ ID NO: 14 或 SEQ ID NO:17 所示的核酸序列。来自该编码链的氨基酸序列或氨基酸部分序列与酿酒酵母线粒体内膜转运蛋白具有显著的序列同源性。

e) 优选被下调的核酸序列，其编码具有非液泡 H^+ -ATP 酶 V0 结构域的 102kD 亚基的功能的蛋白。

本发明此方面的一个优选实施方案中，分离了编码本发明核酸序列的特征性部分序列的 DNA 克隆，其内部名称为“Oligo 31”。

本发明另一优选的实施方案中根据本发明分离了编码本发明的核酸的完全序列的 DNA 克隆，其内部名称为“Oligo 31v”。

本发明一方面涉及包含 SEQ ID NO:19 所示的核酸序列的多核苷酸。本发明另一方面涉及包含 SEQ ID NO:21 所示的核酸序列或其片段的多核苷酸。这些多核苷酸优选地自阿舒氏囊霉属微生物，尤其是棉桃阿舒氏囊霉中分离。本发明还涉及与这些多核苷酸互补的多核苷酸，以及这些多核苷酸经遗传密码简并性衍生而成的序列。

“Oligo 31 和 “Oligo 31v” 的插入序列与来自酿酒酵母的 MIPS 标签 “STV1” 有显著的同源性，这些插入序列具有 SEQ ID NO: 19 或 SEQ ID NO:21 所示的核酸序列。来自该编码链的氨基酸序列或氨基酸部分序列与酿酒酵母非液泡 H⁺-ATP 酶 V0 结构域的 102kD 亚基具有显著的序列同源性。

f) 优选被上调的核酸序列，其编码具有与粟酒裂殖酵母 (S. pombe.) 的 isp4 蛋白相类似的功能的蛋白。

本发明此方面的一个优选实施方案中，分离了编码本发明核酸序列的特征性部分序列的 cDNA 克隆，其内部名称为“Oligo 4”。

本发明另一优选的实施方案中根据本发明分离了编码本发明的核酸的完全序列的 DNA 克隆，其内部名称为“Oligo 4v”。

本发明一方面涉及包含 SEQ ID NO:23 所示的核酸序列的多核苷酸。本发明另一方面涉及包含 SEQ ID NO:26 所示的核酸序列或 SEQ ID NO: 25 所示的 SEQ ID NO: 26 的互补序列的多核苷酸。这些多核苷酸优选地自阿舒氏囊霉属微生物，尤其是棉桃阿舒氏囊霉中分离。本发明还涉及与这些多核苷酸互补的多核苷酸，以及这些多核苷酸经遗传密码简并性衍生而成的序列。

“Oligo 4 和 “Oligo 4v” 的插入序列与来自酿酒酵母的 MIPS 标签 “OPT2” 有显著的同源性，这些插入序列包含 SEQ ID NO: 23 或 SEQ ID

NO:25 所示的核酸序列。来自该编码链(包含 SEQ ID NO:26)的氨基酸序列或氨基酸部分序列与类似于粟酒裂殖酵母 *isp4* 蛋白的酿酒酵母蛋白具有显著的序列同源性。本发明的蛋白也因此被确定为具有寡肽转运蛋白的活性。

g) 优选被上调的核酸序列,其编码具有酿酒酵母 VAC1 蛋白的功能的蛋白, VAC1 蛋白是一种具有三个锌指的胞质膜外周蛋白。

本发明此方面的一个优选实施方案中,分离了编码本发明核酸序列的特征性部分序列的 DNA 克隆,其内部名称为“Oligo 6”。

本发明另一优选的实施方案中根据本发明分离了编码本发明的核酸的完全序列的 DNA 克隆,其内部名称为“Oligo 6v”。

本发明一方面涉及包含 SEQ ID NO:28 所示的核酸序列的多核苷酸。本发明另一方面涉及包含 SEQ ID NO:31 所示的核酸序列或其片段的多核苷酸。这些多核苷酸优选地自阿舒氏囊霉属微生物,尤其是棉桃阿舒氏囊霉中分离。本发明还涉及与这些多核苷酸互补的多核苷酸,以及这些多核苷酸经遗传密码简并性衍生而成的序列。

“Oligo 6”和“Oligo 6v”的插入序列与来自酿酒酵母的 MIPS 标签“VAC1”有显著的同源性,这些插入序列具有 SEQ ID NO: 28 或 SEQ ID NO:31 所示的核酸序列。来自 SEQ ID NO:28 的相应互补序列或来自 SEQ ID NO:3 所示序列的氨基酸序列或氨基酸部分序列与酿酒酵母的 VAC1 蛋白(一种具有三个锌指的胞质膜外周蛋白)具有显著的序列同源性。

h) 优选被上调的核酸序列,其编码具有 ATP 酶样功能的蛋白。

本发明此方面的一个优选实施方案中,分离了编码本发明核酸序列的特征性部分序列的 DNA 克隆,其内部名称为“Oligo 146”。

本发明另一优选的实施方案中根据本发明分离了编码本发明的核酸的完全序列的 DNA 克隆,其内部名称为“Oligo 146v”。

本发明一方面涉及包含 SEQ ID NO:33 所示的核酸序列的多核苷酸。本发明另一方面涉及包含 SEQ ID NO:35 所示的核酸序列或其片段的多核苷酸。这些多核苷酸优选地自阿舒氏囊霉属微生物,尤其是棉桃阿舒氏囊

霉中分离。本发明还涉及与这些多核苷酸互补的多核苷酸，以及这些多核苷酸经遗传密码简并性衍生而成的序列。

“Oligo 146”和“Oligo 146v”的插入序列与来自酿酒酵母的 MIPS 标签“Ymr162c”有显著的同源性，这些插入序列具有 SEQ ID NO: 33 或 SEQ ID NO:35 所示的核酸序列。来自 SEQ ID NO:33 的相应互补序列或来自 SEQ ID NO:35 所示编码链的氨基酸序列或氨基酸部分序列与酿酒酵母的具有 ATP 酶功能或具有 ATP 酶样功能的蛋白具有显著的序列同源性。

i) 优选被上调的核酸序列，其编码具有类似于酿酒酵母 PHO85 蛋白的功能的蛋白。PHO85 是一种激酶，参与多种细胞过程，包括调节 PHO 基因、糖原代谢、调节细胞周期和细胞形态。

本发明此方面的一个优选实施方案中，分离了编码本发明核酸序列的特征性部分序列的 DNA 克隆，其内部名称为“Oligo 56”。

本发明另一优选的实施方案中根据本发明分离了编码本发明的核酸的完全序列的 DNA 克隆，其内部名称为“Oligo 56v”。

本发明一方面涉及包含 SEQ ID NO:37 所示的核酸序列的多核苷酸。本发明另一方面涉及包含 SEQ ID NO:40 所示的核酸序列或其片段的多核苷酸。这些多核苷酸优选地自阿舒氏囊霉属微生物，尤其是棉桃阿舒氏囊霉中分离。本发明还涉及与这些多核苷酸互补的多核苷酸，以及这些多核苷酸经遗传密码简并性衍生而成的序列。

“Oligo 56”和“Oligo 56v”的插入序列与来自酿酒酵母的 MIPS 标签“Ypl110c”有显著的同源性，这些插入序列具有 SEQ ID NO: 37 或 SEQ ID NO:40 所示的核酸序列。来自 SEQ ID NO:37 的相应互补序列或 SEQ ID NO:40 所示编码链的氨基酸序列或氨基酸部分序列与酿酒酵母的 PHO85 蛋白具有显著的序列同源性。

k) 优选被上调的核酸序列，其编码具有类似于参与膜运输的酿酒酵母 p24 蛋白质的功能的蛋白。P24 蛋白家族成员是具有短的胞质 COOH 端的小的 I 型跨膜蛋白。其在早期分泌路径中起运输作用，并参与例如各种分泌蛋白由内质网向高尔基体的转运。

本发明此方面的一个优选实施方案中，分离了编码本发明核酸序列的特征性部分序列的 DNA 克隆，其内部名称为“Oligo 167”。

本发明另一优选的实施方案中根据本发明分离了编码本发明的核酸的完全序列的 DNA 克隆，其内部名称为“Oligo 167v”。

本发明一方面涉及包含 SEQ ID NO:42 所示的核酸序列的多核苷酸。本发明另一方面涉及包含 SEQ ID NO:44 所示的核酸序列或其片段的多核苷酸。这些多核苷酸优选地自阿舒氏囊霉属微生物，尤其是棉桃阿舒氏囊霉中分离。本发明还涉及与这些多核苷酸互补的多核苷酸，以及这些多核苷酸经遗传密码简并性衍生而成的序列。

“Oligo 167”和“Oligo 167v”的插入序列与来自酿酒酵母的 MIPS 标签“ERP5”有显著的同源性，这些插入序列具有 SEQ ID NO: 42 或 SEQ ID NO:44 所示的核酸序列。来自 SEQ ID NO:42 的相应互补序列和来自 SEQ ID NO:44 的编码链的氨基酸序列与参与膜运输的酿酒酵母 p24 蛋白具有显著的序列同源性。

本发明再一方面涉及尤其是在严格的条件下与以上多核苷酸之一杂交的寡核苷酸。

本发明还涉及可以与本发明的寡核苷酸之一杂交并编码来自阿舒氏囊霉属微生物的基因产物或该基因产物的功能等价物的多核苷酸。

本发明进一步涉及由上述多核苷酸编码的多肽或蛋白质；及其具有如下氨基酸序列的肽片段，该氨基酸序列包含 SEQ ID NO: 2, 4, 6, 7, 9, 11, 13, 15, 16, 18, 20, 22, 24, 27, 29, 30, 32, 34, 36, 38, 39, 41, 43 或 SEQ ID NO:45 中显示的至少 10 个连续的氨基酸残基；以及本发明的多肽或蛋白质的功能等价物。

在这方面，功能等价物与本发明中具体公开的产物相比在氨基酸序列上于至少一个，例如 1 到 30 或 1 到 20 或 1 到 10 个序列位置由于添加，插入，替代，缺失或倒位而不同（这可以经与其他蛋白质进行序列比对推断），但功能等价物不丧失最初观察到的蛋白功能。因此，与天然蛋白相比较功能等价物可能具有基本相同的、较高或较低的活性。

本发明其它方面涉及用于重组产生本发明蛋白的表达盒，其包含可操作地与至少一个调节核酸序列连接的上述核酸序列之一；并涉及包含至少一个这样的本发明表达盒的重组载体。

本发明还提供了由至少一个上述类型的载体转化的原核或真核宿主。一个优选的实施方案提供了原核或真核宿主，在宿主中至少一种编码上述本发明多肽的基因的功能性表达受到调节（例如被抑制或被过量表达）；或上述定义的多肽的生物活性被降低或增加。优选的宿主选自子囊菌，尤其是阿舒氏囊霉属，优选地为棉桃阿舒氏囊霉株。

上述意义的基因表达调节包括对表达的抑制（如通过阻断表达的某一阶段（尤其是转录或翻译）抑制表达）以及基因的特异过量表达（例如通过修饰调节序列或增加编码序列的拷贝数）。

本发明另一方面涉及本发明的表达盒、本发明的载体或本发明的宿主在维生素 B2 和/或其前体和/或其衍生物的微生物生产中的应用。

本发明另一方面涉及本发明表达盒、本发明的载体或本发明的宿主在上述本发明多肽的重组生产中的应用。

本发明还提供了一种检测或验证可以用于调节维生素 B2 和/或其前体和/或其衍生物的微生物生产的效应物靶点的方法。该方法涉及：使用能够与选自上述本发明多肽或其编码核酸序列的靶点相互作用（例如非共价结合）的效应物处理能够以微生物学方式生产维生素 B2 和/或其前体和/或其衍生物的微生物，验证该效应物对微生物生产的维生素 B2 和/或其前体和/或其衍生物的量的影响，以及在适当时分离靶点。在此优选地通过与在缺乏该效应物而其他条件相同时微生物的维生素 B2 产生直接比较来进行验证。

本发明另一方面涉及调节维生素 B2 和/或其前体和/或其衍生物的微生物生产（涉及产生的量和/或速率）的方法，其中用效应物处理能够以微生物学方式产生维生素 B2 和/或其前体和/或其衍生物的微生物，该效应物可以与选自上面定义的本发明多肽或编码该多肽的核酸序列的靶点相互作用。

应提及的上述效应物的优选例子为：

- a) 抗体或其抗原结合片段；
- b) 与 a) 不同并可以与本发明的多肽相互作用的多肽配体；
- c) 可以调节本发明多肽的生物活性的低分子量效应物；
- d) 可以与本发明的核酸序列相互作用的反义核酸序列。

本发明还涉及对至少一个以上定义的本发明靶点具有特异性的上述效应物。

本发明另一方面涉及微生物生产维生素 B2 和/或其前体和/或其衍生物的方法，其中在有利产生维生素 B2 和/或其前体和/或其衍生物的条件下培养以上定义的宿主，所需的产物自培养混合物中分离。在这方面，优选地用以上定义的效应物在培养前和/或培养过程中处理该宿主。在此优选的宿主选自阿舒氏囊霉属微生物，尤其是上述转化的宿主。

本发明最后的方面涉及利用本发明的多核苷酸或多肽作为调节阿舒氏囊霉属微生物生产维生素 B2 和/或其前体和/或其衍生物的靶点。

附图说明

图 1 显示本发明的氨基酸部分序列（相应于 SEQ ID NO:1 第 609 位至第 1 位的互补链）（上面的序列）和来自酿酒酵母的 MIPS 标签 Ygr257c 的部分序列（下面的序列）的比对。相同的序列位置在两个序列之间标出。相似的序列位置以“+”标出。

图 2 显示本发明的氨基酸部分序列（SEQ ID NO:6）（相应于 SEQ ID NO:5 第 1494 位至第 1387 位的互补链）（上面的序列）和来自酿酒酵母的 MIPS 标签 Mdl2 的部分序列（下面的序列）的比对。相同的序列位置在两个序列之间标出。相似的序列位置以“+”标出。

图 3 显示本发明的氨基酸部分序列（相应于 SEQ ID NO:10 第 15 位至第 455 位的编码链）（上面的序列）和来自酿酒酵母的 MIPS 标签 Prp12 的部分序列（下面的序列）的比对。相同的序列位置在两个序列之间标出。相似的序列位置以“+”标出。

图 4 显示本发明的氨基酸部分序列 (相应于 SEQ ID NO:14 第 246 位至第 1118 位的编码链) (上面的序列)和来自酿酒酵母的 MIPS 标签 Flx1 的部分序列 (下面的序列) 的比对。相同的序列位置在两个序列之间标出。相似的序列位置以“+”标出。

图 5 显示本发明的氨基酸部分序列 (相应于 SEQ ID NO:19 第 2 位至第 790 位的编码链) (上面的序列)和来自酿酒酵母的 MIPS 标签 STV1 的部分序列 (下面的序列) 的比对。相同的序列位置在两个序列之间标出。相似的序列位置以“+”标出。

图 6 显示本发明的氨基酸部分序列 (相应于 SEQ ID NO:23 第 869 位至第 522 位的互补链) (上面的序列)和来自酿酒酵母的 MIPS 标签 OPT2 的部分序列 (下面的序列) 的比对。相同的序列位置在两个序列之间标出。相似的序列位置以“+”标出。

图 7A 显示本发明的氨基酸部分序列 (相应于 SEQ ID NO:28 第 356 位至第 243 位的互补序列) (上面的序列)和来自酿酒酵母的 MIPS 标签 VAC1 的部分序列 (下面的序列) 的比对。相同的序列位置在两个序列之间标出。相似的序列位置以“+”标出。图 7B 显示本发明的氨基酸部分序列 (相应于 SEQ ID NO:28 第 166 位至第 2 位的互补序列) (上面的序列)和来自酿酒酵母的 MIPS 标签 VAC1 的部分序列 (下面的序列) 的比对。相同的序列位置在两个序列之间标出。相似的序列位置以“+”标出。

图 8 显示本发明的氨基酸部分序列 (相应于 SEQ ID NO:33 第 904 位至第 707 位的互补序列) (上面的序列)和来自酿酒酵母的 MIPS 标签 Ymr162c 的部分序列 (下面的序列) 的比对。相同的序列位置在两个序列之间标出。相似的序列位置以“+”标出。

图 9 显示本发明的氨基酸部分序列 (相应于 SEQ ID NO:37 第 898 位至第 5 位的互补序列) (上面的序列)和来自酿酒酵母的 MIPS 标签 Ypl110c 的部分序列 (下面的序列) 的比对。相同的序列位置在两个序列之间标出。相似的序列位置以“+”标出。

图 10 显示本发明的氨基酸部分序列 (相应于 SEQ ID NO:42 第 931

位至第 806 位的互补序列) (上面的序列)和来自酿酒酵母的 MIPS 标签 ERP5 的部分序列 (下面的序列) 的比对。相同的序列位置在两个序列之间标出。相似的序列位置以“+”标出。

发明详述

本发明的核酸分子编码的多肽或蛋白在此被称为跨膜转运蛋白 (例如具有与跨膜转运系统相关的活性) 或简称为“TMT 蛋白”。这些 TMT 蛋白具有例如控制膜相关转运系统的功能, 该转运系统消耗能量将所需蛋白质逆浓度梯度转运入细胞内。TMT 蛋白能够影响细胞对外界条件的反应, 并由此例如调节细胞的代谢。由于已掌握了可用于棉桃阿舒氏囊霉的克隆载体 (例如 Wright 和 Philipsen (1991) 基因 (Gene), 109, 99-105 中公开) 以及棉桃阿舒氏囊霉和相关的酵母物种的遗传操作技术, 本发明的核酸分子可以用于这些生物体, 尤其是棉桃阿舒氏囊霉的遗传操作, 从而使其更好并更有效的生产维生素 B2 和/或其前体和/或其衍生物。产量或生产效率的提高可以由操作本发明的基因的直接效应或该操作的间接效应导致。

本发明基于在此被称为 TMT 核酸和 TMT 蛋白的新颖分子的提供, 这些分子参与尤其是棉桃阿舒氏囊霉中的跨膜转运 (例如参与转运蛋白的合成或调节)。本发明的 TMT 分子在棉桃阿舒氏囊霉中的活性影响该微生物的维生素 B2 生产。优选对本发明的 TMT 分子的活性进行调节, 以便本发明的 TMT 蛋白参与的棉桃阿舒氏囊霉的代谢和/或能量途径在维生素 B2 的产率、产量和/或生产效率方面受到调节, 即直接或间接地调节棉桃阿舒氏囊霉中维生素 B2 的产率、产量和/或生产效率。

本发明提供的核酸序列可以自例如棉桃阿舒氏囊霉株基因组分离, 所述菌株可以从美国典型培养物保藏中心 (保藏号为 ATCC 10895) 自由获得。

维生素 B2 生产的改进

通过改变本发明的 TMT 蛋白的量和/或活性有多种可能的机制可以直接地影响棉桃阿舒氏囊霉株的维生素 B2 的产率、产量和/或生产效率。

因此，更有效的跨膜转运可提高细胞的应答，从而增加有价值的所需产物的生成。对本发明的一种或多种 TMT 蛋白的诱变也可以导致 TMT 蛋白的活性改变（增加或降低），间接地影响棉桃阿舒氏囊霉株中所需产物的生产。例如可通过 TMT 蛋白使细胞适应新的或改变的外部环境。通过改善这些修饰细胞的生长和繁殖，可以增加这些细胞在大规模培养中的生存力并提高分裂速率。最后，可由此增加这些细胞生产所需目标产物的产率。

多肽

本发明涉及包含上述的氨基酸序列或其特征性部分序列和/或由在此描述的核酸序列编码的多肽。

本发明还包括具体公开的新颖多肽的“功能等价物”。

具体公开的多肽的“功能等价物”或类似物为本发明的目的是指与具体公开的多肽不同但仍具有所需生物活性（例如底物特异性）的多肽。

“功能等价物”在本发明中尤其指在至少一个上述序列位置中具有与具体公开所不同的氨基酸但仍具有一种上述的生物活性的突变体。“功能等价物”因此包括可由一个或多个氨基酸添加，替换，缺失和/或倒位获得的突变体，只要生成的突变体能具有本发明的性质谱，所述的修饰可发生于任何序列位置。功能等价物尤其还可以是这样的形式，即突变体和未修饰的多肽就反应模式而言在性质上是一致的，即例如对同样的底物的转化速率不同。

上述意义的“功能等价物”也可以是所述多肽的前体以及该多肽的功能衍生物和盐类。“盐类”一词指本发明蛋白分子的羧基的盐和氨基的酸加成盐。可由本领域已知方式制备羧基的盐，包括例如钠，钙，铵，铁和锌盐等无机盐以及与有机碱的盐，例如与胺类如三乙醇胺、精氨酸、赖氨酸、哌啶等的盐。本发明另一方面涉及酸加成盐，例如与无机酸（例如盐

酸或硫酸)的盐和与有机酸(例如乙酸和草酸)的盐。

本发明多肽的“功能衍生物”也可由已知技术在氨基酸侧链官能团上或在多肽的 N 末端或 C 末端制备。这样的衍生物包括例如羧基的脂族酯和可通过与氨或与伯胺或仲胺反应得到的羧基的酰胺;通过与酰基反应制备的游离氨基的 N-酰基衍生物;或通过与酰基反应制备的游离羟基的 O-酰基衍生物。

“功能等价物”自然地也包括可自其他生物体获得的多肽和自然出现的变体。例如,可以通过序列比对发现同源序列区并根据本发明的具体要求确立等价酶。

“功能等价物”还包括具有例如所需的生物学功能的片段,优选为本发明多肽的单结构域或序列基序。

“功能等价物”还可以是具有上述多肽序列之一或其功能等价物和至少一个另外的具有不同功能的异源序列的融合蛋白,该异源序列在所述多肽序列或其功能等价物的 N 或 C 末端进行功能性连接(即该融合蛋白的各部分在功能方面产生的相互损害是微不足道的)。这种异源序列的非限制性例子为例如信号肽,酶,免疫球蛋白,表面抗原,受体或受体配体。

根据本发明,“功能等价物”包括本文具体公开的蛋白的同系物。通过 Pearson 和 Lipman, Proc. Natl. Acad. Sci. (USA) 85(8), 1988, 2444-2448. 的算法计算,这些同系物与具体公开的序列之一具有至少 60%,优选地至少 75%,尤其至少 85%,例如 90%, 95%, 或 99%的同源性。

在可能存在蛋白糖基化的情况下,本发明的等价物包括以上定义的蛋白类型的去糖基化或糖基化形式,及可通过改变糖基化模式获得的修饰形式。

诱变可产生本发明的蛋白或多肽的同系物,例如通过蛋白点突变或截短实现。这里使用的“同系物”一词涉及可以对蛋白活性产生激动剂或拮抗剂作用的蛋白的变体形式。

可以通过筛选突变体(例如截短突变体)组合文库鉴定本发明蛋白的同系物。通过核酸水平的组合诱变,例如通过合成的寡核苷酸混合物的酶

促连接，可以产生一个多样化的蛋白质变体文库。大量方法可用于从简并寡核苷酸序列产生潜在同系物的文库。简并基因序列的化学合成可在自动DNA合成仪中进行，然后可将合成的基因与适合的表达载体连接。利用一组简并基因可在一个混合物中提供编码一组所需的潜在蛋白质序列的所有序列。本领域技术人员已知合成简并寡核苷酸的方法(例如 Narang, S.A. (1983) *Tetrahedron* 39:3; Itakura 等 (1984) *Annu. Rev. Biochem.* 53:323; Itakura 等 (1984) *Science* 198:1056; Ike 等 (1983) *Nucleic Acids Res.* 11:477)。

此外，也可以利用蛋白密码子的片段文库来生成多样的蛋白片段群用于筛选和随后选择本发明蛋白的同系物。在一个实施方案中，通过如下方法产生编码序列片段文库：在每个分子只发生大约一次切割的条件下，以核酸酶处理编码序列的双链PCR片段，变性双链DNA，之后使该DNA复性，从而形成可能包含来自不同切割产物的有义/反义对的双链DNA，经S1核酸酶处理自新形成的双链体中去除单链部分，并将产生的片段文库连接到表达载体中。可由这种方法得到编码本发明蛋白的不同大小的N末端、C末端和内部片段的表达文库。

现有技术中已知一些技术可以从点突变或截短产生的组合文库中筛选基因产物和从cDNA文库中筛选具有选定性质的基因产物。这些技术经改造后可用于快速筛选由本发明同系物的组合诱变产生的基因文库。最常用的高通量大型基因文库筛选分析技术包括将该基因文库克隆入可复制的表达载体中，使用产生的载体文库转化合适的细胞，并在检测所需的活性有助于分离如下载体的条件下表达该组合基因，所述载体编码的基因的产物是活性检测中被检出的产物。递归总体诱变 (Recursive ensemble mutagenesis, REM) 技术增加文库中功能性突变体的频率，因而可以与筛选实验结合用于鉴定同系物。(Arkin 和 Yourvan (1992) *PNAS* 89:7811-7815; Delgrave 等 (1993) *蛋白质工程 (Protein Engineering)* 6(3):327-331)。

可重组制备本发明的多肽(见以下部分)或可通过传统的生化技术自

微生物，特别是阿舒氏囊霉属微生物中分离自然形式的多肽。(见 Cooper, T.G., Biochemische Arbeitsmethoden, Verlag Walter de Gruyter, 柏林, 纽约或 Scopes, R., 蛋白质纯化 (Protein Purification), 施普林格, 纽约, 海德堡, 柏林)。

核酸序列

本发明还涉及编码上述多肽之一及其功能等价物的核酸序列 (单链和双链 DNA 和 RNA 序列, 例如 cDNA 和 mRNA), 该核酸序列可以例如利用人工核苷酸类似物获得。

本发明既涉及编码本发明的多肽或蛋白质或其生物活性部分的分离的核酸分子, 又涉及可用作例如鉴定或扩增本发明的编码核酸的杂交探针或引物的核酸片段。

本发明的核酸分子还可包含来自基因编码区的 3' 和/或 5' 末端的非翻译序列。

“分离的”核酸分子自存在于该核酸的天然来源中的其他核酸分子中分离出来, 如果该核酸分子由重组技术产生, 还可以基本不含其他细胞物质或培养基; 如果由化学方法合成, 则基本不含化学前体或其他化学物质。

通过分子生物学标准技术以及本发明提供的序列信息可以分离本发明的核酸分子。例如, 以一种具体公开的完全序列或其一部分为杂交探针, 通过标准杂交技术 (如 Sambrook, J., Fritsch, E.F. 和 Maniatis, T. 分子克隆: 实验室手册, 第二版, Molecular Cloning: A Laboratory Manual. 2nd edition, 美国冷泉港实验室 (Cold Spring Harbor Laboratory), 美国冷泉港实验室出版社, 冷泉港, 纽约, 1989 的描述), 可以从适宜的 cDNA 文库中分离 cDNA。使用基于一种公开序列而构建的寡核苷酸引物、通过聚合酶链反应也可以分离获得包含该序列或其一部分的核酸分子。由这种方法扩增的核酸分子可以被克隆入合适的载体中并由 DNA 序列分析法表征。由标准合成法, 例如使用自动 DNA 合成仪也可以产生相应于 TMT 核苷酸序列的本发明寡核苷酸。

本发明还包括与具体描述的核苷酸序列互补的核酸分子或其一部分。

本发明的核苷酸序列使得可产生能用于鉴定和/或克隆其他细胞类型和生物中的同源序列的探针和引物。这样的探针和引物通常包括如下核苷酸序列区域，该区域可以在严格条件下与本发明核酸序列的有义链或相应的反义链中的至少约 12，优选地约 25，例如约 40，50 或 75 个连续的核苷酸杂交。

本发明的其它核酸序列衍生自 SEQ ID NO: 1, 3, 5, 8, 10, 12, 14, 17, 19, 21, 23, 25, 26, 28, 31, 33, 35, 37, 40, 42 或 SEQ ID NO:44 并由于添加，替代、插入或缺失一个或多个核苷酸而与所衍生自的序列不同，但仍编码具有所需性质谱的多肽。

本发明还包括含有所谓的沉默突变或经过修饰（这是根据特定来源或宿主生物的密码子使用与具体公开的序列相比而言的）的核酸序列以及它们的自然存在的变体（例如拼接变体或等位基因变体）。本发明还涉及可通过保守核苷酸替代（即相关的氨基酸被具有同样电荷、大小、极性和/或溶解性的氨基酸替代）获得的序列。

本发明还涉及由序列多态性自具体公开的核酸衍生得到的分子。这些遗传多态性可由于自然变异存在于群体的个体中。这些自然变异通常导致基因的核苷酸序列出现 1%至 5%的变化。

本发明还包括与上述编码序列杂交或互补的核酸序列。对基因组或 cDNA 文库进行筛选可发现这些多核苷酸，适当时可以利用合适的引物通过 PCR 扩增这些多核苷酸，并例如使用合适的探针进行分离。另一种可能的方案是利用本发明的多核苷酸或载体转化合适的微生物，增殖该微生物并由此扩增这些多核苷酸然后分离之。再一可能方案是通过化学途径合成本发明的多核苷酸。

能与多核苷酸“杂交”的性质意味着多核苷酸或寡核苷酸在严格的条件下能与几乎互补的序列结合，而在此条件下，非互补序列之间不发生非特异的结合。为此这些序列应具有 70%至 100%，优选为 90%至 100%的互补性。互补序列能够特异性地互相结合的性质可用于例如 Northern 或

Southern 印迹杂交或在 PCR 或 RT-PCR 中用于引物的结合。为此目的通常使用长 30 个碱基对或以上的寡核苷酸。严格的条件指例如在 Northern 印迹杂交法中, 在 50–70°C, 优选地为 60–65°C 下进行洗涤, 如用含有 0.1x SSC 缓冲液和 0.1% SDS (20x SSC: 3M 氯化钠, 0.3M 柠檬酸钠, pH 7.0) 的洗涤溶液洗脱非特异杂交的 cDNA 探针或寡核苷酸。在此情况下, 如上述只有高度互补的核酸互相保持结合。本领域技术人员已知这种严格条件的建立, 并可参见例如 Ausubel 等的《Current Protocols in Molecular Biology》, John Wiley & Sons, 纽约 N.Y. (1989), 6.3.1-6.3.6. 中的描述。

本发明另一方面涉及反义核酸。其包含与有义编码核酸互补的核苷酸序列。该反义核酸可与整个编码链互补或仅与其中一部分互补。在另一个实施方案中, 该反义核酸分子与核苷酸序列编码链的非编码区反义。“非编码区”指序列中被称为 5'- 和 3'-非翻译区的部分。

反义寡核苷酸可以是例如大约 5, 10, 15, 20, 25, 30, 35, 40, 45 或 50 个核苷酸长。可由本领域已知的方法通过化学合成和酶促连接反应构建本发明的反义核酸。反义核酸可由天然存在的核苷酸或各种修饰的、设计以增强分子的生物稳定性或增强反义和有义核酸间形成的双链体的物理稳定性的核苷酸通过化学方式合成。可以使用的例子包括硫代磷酸酯衍生物和吡啶取代的核苷酸。可用于产生反义核酸的经修饰的核苷有例如: 5-氟尿嘧啶, 5-溴尿嘧啶, 5-氯尿嘧啶, 5-碘尿嘧啶, 次黄嘌呤, 黄嘌呤, 4-乙酰胞嘧啶, 5-(羧基羧基甲基)尿嘧啶, 5-羧基甲基氨基甲基-2-硫尿苷, 5-羧基甲基氨基甲基尿嘧啶, 二氢尿嘧啶, β -D-半乳糖基 queuosine, 肌苷, N6-异戊烯腺嘌呤, 1-甲基鸟嘌呤, 1-甲基肌苷, 2,2-二甲基鸟嘌呤, 2-甲基腺嘌呤, 2-甲基鸟嘌呤, 3-甲基胞嘧啶, 5-甲基胞嘧啶, N6-甲基腺嘌呤, 7-甲基鸟嘌呤, 5-甲基氨基甲基尿嘧啶, 5-甲氧基氨基甲基-2-硫尿嘧啶, β -D-甘露糖基 queuosine, 5-甲氧基羧基甲基尿嘧啶, 5-甲氧基尿嘧啶, 2-甲硫基-N6-异戊烯腺嘌呤, 尿嘧啶-5-氧基乙酸(v), Wybutoxosine, 假尿嘧啶, queuosine, 2-硫代胞嘧啶, 5-甲基-2-硫尿嘧啶, 2-硫尿嘧啶, 4-硫尿嘧啶, 5-甲基尿嘧啶, 尿嘧啶-5-氧乙酸甲基酯, 3-(3-氨基-3-羧基丙基)尿嘧啶, (acp3)w 和 2,6-二氨

基嘌呤等。也可利用其中以反义方向亚克隆了核酸的表达载体来生物产生反义核酸。

本发明的反义核酸分子通常被施用给细胞或在原位产生，这样它们可以与该细胞 mRNA 和/或编码 DNA 杂交或结合，通过例如抑制转录和/或翻译抑制蛋白的表达。

可以例如通过使反义核酸分子与能结合细胞表面受体或抗原的肽或抗体相连接，以修饰反义分子从而使之能与在所选的细胞表面上表达的受体或抗原特异性地结合。该反义核酸分子还可通过在此描述的载体被给予细胞。为使细胞内的反义分子达到足够的浓度，优选使用的载体构建体为其中反义核酸分子处于强的细菌、病毒或真核细胞启动子的控制之下的载体。

在另一个实施方案中，本发明的反义核酸分子是一种 α -异头核酸分子。 α -异头核酸分子 (α -anomeric nucleic acid molecule) 与互补的 RNA 形成特有的双链杂交体，与通常的 α 单元不同，这两条链的走向互相平行 (Gaultier 等, (1987) *Nucleic Acids Res.* 15:6625-6641)。该反义核酸分子还可以包含 2'-O-甲基核糖核苷酸(Inoue 等., (1987) *Nucleic Acids Res.* 15:6131-6148) 或嵌合 RNA-DNA 类似物(Inoue 等. (1987) *FEBS Lett.* 215:327-330)。

本发明还涉及核酶。其为具有核糖核酸酶活性的催化性 RNA 分子，能够切割具有与其互补的区域单链核酸例如 mRNA。因此核酶（例如锤头核酶(见 Haselhoff 和 Gerlach (1988) *Nature* 334:585-591)）可用于本发明转录物的催化性切割从而抑制相应核酸的翻译。对本发明的编码核酸具有特异性的核酶可以在例如本发明具体公开的 cDNA 的基础上形成。例如，可构建四膜虫-L-19 IVS RNA 的衍生物，其中活性位点的核苷酸序列与本发明的编码 mRNA 中待切割的核苷酸序列互补。(参考例如 US-A-4 987 071 和 US-A-5 116 742)。可替换地，可用 mRNA 自 RNA 分子池中选择具有特异的核糖核酸酶活性的催化性 RNA (见 例如 Bartel, D., 和 Szostak, J.W. (1993) *Science* 261:1411-1418)。

可替换地，本发明序列的基因表达可通过与本发明核苷酸序列的调节

区互补（例如与编码序列的启动子和/或增强子互补）的核苷酸序列的打靶来抑制，这种打靶可造成三股螺旋结构形成从而阻止目标细胞中相应基因的转录（Helene, C. (1991) *Anticancer Drug Res.* 6(6) 569-584; Helene, C. 等, (1992) *Ann. N. Y. Acad. Sci.* 660:27-36; 和 Maher., L.J. (1992) *Bioassays* 14(12):807-815)。

表达构建体和载体:

本发明还涉及包含在调节性核酸序列的遗传控制下的编码本发明多肽的核酸序列的表达构建体；以及包含至少一个这样的表达构建体的载体。本发明的这种表达构建体优选包含位于特定编码序列 5'端上游的启动子以及位于特定编码序列 3'端下游的终止序列和，适当时，其他通常的调节元件，这些元件每一个都与该编码序列可操作地连接。“可操作地连接”指启动子，编码序列，终止子和适当时其他调节元件的相继排列方式，这种方式使得每个调节元件都能按预期行使其功能用于编码序列的表达。可操作连接的序列的例子为前导序列和增强子、多腺苷酸化信号等。其他调节元件包括选择标记、扩增信号、复制起点等。适宜的调节序列在例如 Goeddel, 基因表达技术: 酶学方法 (Gene Expression Technology: Methods in Enzymology) 185, Academic Press, 圣地亚哥 San Diego, 加利福尼亚州 CA (1990) 中有所描述。

除了人工调节序列，天然的调节序列仍可以存在于实际的结构基因之前。这种天然调节可在适当时通过基因修饰而关闭，从而使基因的表达可被增加或降低。然而，基因构建体也可具有较为简单的结构，即在结构基因之前没有插入附加的调节信号，且具有调节功能的天然启动子未被删除。可替换地，可以使天然调节序列发生突变不再具有调节作用，从而增强或降低基因的表达。核酸序列在基因构建体中可有一个或多个拷贝。

可使用的启动子的例子包括：cos, tac, trp, tet, trp-tet, lpp, lac, lpp-lac, lacIq, T7, T5, T3, gal, trc, ara, SP6, λ -PR 或 λ -PL 启动子（它们优选用于革兰氏阴性细菌）；革兰氏阳性细菌启动子 amy 和 SPO2，酵母

启动子 ADC1, MF α , AC, P-60, CYC1, GAPDH 或植物启动子 CaMV/35S, SSU, OCS, lib4, usp, STLS1, B33, 或泛素或菜豆蛋白启动子。尤其优选使用可诱导的启动子, 例如光诱导的启动子, 更优选为温度诱导的启动子如 P_rP_i 启动子。原则上所有天然启动子以及它们的调节序列均可使用。此外, 使用合成启动子也可能是有利的。

所述的调节序列旨在用于实现核酸序列的特异表达。这意味着例如视宿主生物不同使基因只在诱导后才表达或过量表达或立即被表达和/或过量表达。

此外优选调节序列或因子可以正面地产生影响, 从而增加或减低表达。因此, 调节元件的增强作用可以有利地使用强转录信号如启动子和/或增强子在转录水平上发生。然而通过例如增加 mRNA 的稳定性, 也可加强其翻译。

融合适宜的启动子和适宜的核苷酸编码序列以及终止信号或多腺苷酸化信号可以产生表达盒。为此目的可以使用常规的重组和克隆技术, 参见: T. Maniatis, E.F. Fritsch 和 J. Sambrook, 分子克隆: 实验室手册 (Molecular Cloning: A Laboratory Manual), Cold Spring Harbor Laboratory, 冷泉港 Cold Spring Harbor, 纽约(1989); T.J. Silhavy, M.L. Berman 和 L.W. Enquist, 基因融合试验 (Experiments with Gene Fusions), Cold Spring Harbor Laboratory, Cold Spring Harbor, 纽约(1984); Ausubel, F.M. 等, Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley Interscience (1987)。

为在适宜的宿主生物中表达, 重组核酸构建体或基因构建体可以有利地插入宿主特异的载体中实现这些基因在宿主中的最佳表达。载体是本领域技术人员熟知的, 并可在例如“克隆载体”(Cloning Vectors) (Pouwels P. H. 等, eds, Elsevier, Amsterdam-New York-Oxford, 1985)中找到有关内容。载体不仅指质粒, 还包括本领域技术人员已知的所有其他载体, 例如噬菌体, 病毒如 SV40、CMV、杆状病毒和腺病毒, 转座子, IS 元件, 质粒, 粘粒和线性或环状 DNA。这些载体可以在宿主生物中自主复制或随染

染色体复制。

可以提及的适宜表达载体的例子有：

常规的融合表达载体如 pGEX (Pharmacia Biotech Inc; Smith, D.B.和 Johnson, K.S. (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) 和 pRIT 5 (Pharmacia, Piscataway, NJ), 使用这些载体可以分别使谷胱甘肽 S-转移酶(GST)、麦芽糖 E 结合蛋白和 A 蛋白 融合至重组目标蛋白。

非融合蛋白表达载体如 pTrc (Amann 等(1988) *Gene* 69:301-315) 和 pET 11d (Studier 等, 基因表达技术: 酶学方法 (*Gene Expression Technology: Methods in Enzymology*)185, Academic Press, 圣地亚哥, 加利福尼亚(1990) 60-89)。

用于酿酒酵母中表达的酵母表达载体如 pYepSec1 (Baldari 等 (1987) *Embo J.* 6:229-234), pMF α (Kurjan 和 Herskowitz (1982) *细胞 (Cell)* 30:933-943), pJRY88 (Schultz 等 (1987) *基因 (Gene)* 54:113-123)和 pYES2 (Invitrogen Corporation, 圣地亚哥, 加利福尼亚)。有关适用于其他真菌 (如丝状真菌) 的载体和载体构建方法的例子的详细描述见例如: van den Hondel, C.A.M.J.J. & Punt, P.J. (1991) “丝状真菌基因转移系统和载体的发展”, 《真菌的应用分子遗传学》 (*Applied Molecular Genetics of Fungi*), J.F. Peberdy 等 eds, pp. 1-28, Cambridge University Press: Cambridge。

现有用于培养的昆虫细胞(例如 Sf9 细胞)中蛋白的表达的杆状病毒载体包括: pAc 系列(Smith 等, (1983) *Mol. Cell Biol.* 3:2156-2165)和 pVL 系列(Lucklow 和 Summers (1989) *病毒学 (Virology)* 170:31-39)。

植物表达载体的例子的详细描述见例如: Becker, D., Kemper, E., Schell, J. 和 Masterson, R. (1992) “紧靠左边界带有选择标记的新植物二元载体”, *Plant Mol. Biol.* 20:1195-1197; 和 Bevan, M.W. (1984) “用于植物转化的农杆菌二元载体”, *Nucl. Acids Res.* 12:8711-8721。

哺乳动物表达载体例如 pCDM8 (Seed, B. (1987) Nature 329:840) 和 pMT2PC (Kaufman 等(1987) EMBO J. 6:187-195)。

适用于原核和真核细胞的其它表达系统在 Sambrook, J., Fritsch, E.F. 和 Maniatis, T., 分子克隆: 实验室手册 (Molecular cloning: A Laboratory Manual), 第二版, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 纽约, 1989. 的第 16 和 17 章中描述。

重组微生物

本发明的载体可用于产生重组微生物, 所述微生物由例如至少一个本发明的载体转化并可用于产生本发明的多肽。有利地, 可以将上述的本发明重组构建体引入适合的宿主系统中并在其中表达。本领域技术人员熟悉克隆和转染的方法, 例如: 共沉淀, 原生质体融合, 电穿孔法, 逆转录病毒转染法等, 并优选使用这些方法引起所述的核酸在特定的表达系统中表达。适宜的表达系统在例如《Current Protocols in Molecular Biology》(F. Ausubel 等, eds, Wiley Interscience, 纽约 1997) 或 Sambrook 等《分子克隆: 实验室手册》(Molecular Cloning: A Laboratory Manual) 第二版 (Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 纽约, 1989) 中有描述。

根据本发明还可以制备同源重组微生物。这需要产生包含至少一段本发明的基因或编码序列的载体, 在适当时, 向其中引入至少一个氨基酸缺失、添加或替代以修饰, 例如从功能上破坏本发明的序列 (剔除载体)。引入的序列也可为例如来自相关微生物的同系物或来自哺乳动物、酵母或昆虫。可替代地, 还可设计用于同源重组的载体以便同源重组过程中内源性基因发生突变或经过其他修饰但仍编码功能蛋白 (如可以修饰位于上游的调节序列以改变内源性蛋白的表达)。TMT 基因的修饰部分存在于同源重组载体内。适宜的同源重组载体的构建的描述见例如 Thomas, K.R. 和 Capecchi, M.R. (1987) Cell 51:503。

原则上适宜的宿主生物体可为所有能够表达本发明的核酸的生物体，其等位基因变体，其功能等价物或衍生物。宿主生物指例如细菌，真菌，酵母，植物或动物细胞。优选的生物为细菌，例如埃希氏杆菌属（*Escherichia*）如大肠杆菌，链霉菌属（*Streptomyces*）、芽孢杆菌属（*Bacillus*）或假单胞菌属（*Pseudomonas*），真核微生物如酿酒酵母，曲霉菌（*Aspergillus*），来自动物或植物的高等真核细胞，如 Sf9 或 CHO 细胞。优选的生物选自阿舒氏囊霉属，尤其是棉桃阿舒氏囊霉株。

成功转化的生物可通过同样存在于载体或表达盒中的标记基因选择。这样的标记基因为例如抗生素抗性基因和催化引起转化细胞染色的颜色形成反应的酶的基因。随后可以通过自动细胞分选法进行选择。由载体成功转化并包含适当的抗生素抗性基因（例如 G418 或潮霉素）的微生物可由适当的含抗生素的培养质或营养介质选择。利用细胞表面上存在的标记蛋白可以通过亲和层析法进行选择。

宿主生物与适合该生物的载体例如质粒、病毒或噬菌体，如具有 RNA 聚合酶/启动子系统的质粒、 λ 或 μ 噬菌体或其他温和噬菌体或转座子和/或其他有利调节序列的结合形成表达系统。“表达系统”一词指，例如，哺乳动物细胞如 CHO 细胞与适合哺乳动物细胞的载体如 pcDNA3neo 载体的结合。

如需要，基因产物还可在转基因生物体例如转基因动物（尤其为小鼠，绵羊）或转基因植物中表达。

多肽的重组生产

本发明还涉及重组产生本发明的多肽或其生物学活性功能片断的方法，该方法中培养可以产生多肽的微生物，在适当时诱导多肽的表达，并自培养物中分离多肽。如需要，该多肽可由此方法进行工业规模的生产。

可以利用已知的方法培养和发酵重组微生物。可以在例如 20 到 40°C 下，pH 为 6 到 9 的 TB 或 LB 培养基中培养细菌。适合的培养条件的描述见例如 T. Maniatis, E.F. Fritsch 和 J. Sambrook, *Molecular Cloning: A*

Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989)。

如果多肽不被分泌入培养基，则破坏细胞后，利用已知的蛋白分离法自裂解物中获得产物。备选地，可以通过高频率超声波，高压例如在弗氏细胞压碎器中，渗胀裂解（osmolysis），变性剂的作用，裂解酶或有机溶剂，匀浆或上述多种方法的组合来破碎细胞。

多肽的纯化可以利用已知的色谱法例如分子筛色谱(凝胶过滤法)，如 Q-Sepharose 色谱，离子交换色谱和疏水色谱以及其他常用方法如超滤，结晶，盐析，透析和非变性凝胶电泳。对适合的方法的描述见例如 Cooper, T.G., Biochemische Arbeitsmethoden, Verlag Walter de Gruyter, Berlin, New York 或 Scopes, R., Protein Purification, Springer Verlag, New York, Heidelberg, Berlin。

使用如下载体系统或寡核苷酸尤其利于分离重组蛋白，所述载体系统或寡核苷酸通过特定的核苷酸序列延长 cDNA，并由此编码可用于例如简化纯化过程的经修饰的多肽或融合蛋白。该类型的适当修饰包括例如所谓的标签——其起锚的作用（例如六组氨酸锚修饰法）或能被抗体识别为抗原的表位（描述见例如 Harlow, E.和 Lane, D., 1988, Antibodies: A Laboratory Manual. Cold Spring Harbor (N.Y.) Press）。这些锚可用于使蛋白结合在例如聚合物基质等固相载体上，该固相载体可装入例如层析柱中，或用于微量滴定板或其他载体上。

同时这些锚也可用于蛋白的识别。还可以单独或与锚结合使用传统的标记物例如荧光染料，与底物反应后可形成可探测的反应产物的酶标记物或放射性标记物来衍生并由此识别蛋白质。

本发明还涉及维生素 B2 和/或其前体和/或其衍生物的微生物生产方法。

如果转化是用重组微生物进行的，则优选该微生物首先在有氧的复合培养基中培养，例如在约 20°C 或更高的温度下，pH 值约为 6 到 9 时培养直到细胞密度足够为止。为了能够更好的控制反应，优选使用可诱导的启

动子。对维生素 B2 产生实施诱导后，接着在有氧条件下继续培养 12 小时到 3 天。

以下非限制性实施例描述了本发明具体的实施方案。

一般性实验描述

a) 一般克隆方法

本发明利用的克隆步骤，例如限制性切割，琼脂糖凝胶电泳，DNA 片段的纯化，转移核酸至硝化纤维素膜和尼龙膜上，DNA 片段的连接，大肠杆菌细胞的转化，细菌的培养，噬菌体的复制，以及重组 DNA 的序列分析，均按照 Sambrook 等(1989)上述引文的描述进行。

b) 聚合酶链反应 (PCR)

按照标准方案在下面的标准混合物中进行 PCR:

8 μ l dNTP 混合液(200 μ M), 10 μ l 不含 $MgCl_2$ 的 Taq 聚合酶缓冲液 (10 x), 8 μ l $MgCl_2$ (25 mM), 每种引物各 1 μ l (0.1 μ M), 1 μ l DNA (待扩增), 2.5 U Taq 聚合酶 (MBI Fermentas, Vilnius, Lithuania), 去离子水加至 100 μ l.

c) 大肠杆菌的培养

在 LB-amp 培养基 (胰胨 10.0 g, 氯化钠 5.0 g, 酵母提取物 5.0 g, 氨苄青霉素 100 g/ml, H_2O 加至 1000 ml) 中、37°C 下培养重组大肠杆菌 DH5 α 菌株。为此每种情况下用接种环将一个菌落从琼脂平板转移至 5 ml LB-amp 中。在 220rpm 下振荡培养约 18 小时后，将 4ml 培养物接种至 2 升摇瓶中的 400ml 培养基中。当 OD578 值到达 0.8 和 1.0 之间后在 42°C 下热休克 3 至 4 个小时，诱导大肠杆菌中 P450 的表达。

d) 自培养物中纯化所需产物

可以利用多种本领域已知的方法将所需的产物自微生物或培养物上清中分离出来。如果所需的产物不被细胞分泌，则可通过慢速离心将细胞从培养物中分离出来，并可以由如机械作用力或超声作用等标准技术裂解细胞。

利用离心除去细胞碎片，得到包含可溶性蛋白的上清部分用于进一步纯化所需的化合物。如果该产物被细胞分泌，则由慢速离心除去培养物中的细胞，保留的上清部分用于进一步纯化。

在适合的树脂中层析从这两种纯化方法得到的上清部分，所需的分子以大于杂质的选择性保留于层析树脂上或通过之。如有必要可用相同或不同的层析树脂重复层析步骤。本领域技术人员能够选择适宜的层析树脂，并最为有效的利用其纯化特定的分子。纯化产物可以经过滤或超滤浓缩，并在该产物最为稳定的温度下储存。

本领域已知多种纯化方法。这些纯化技术的描述见例如，Bailey, J.E. & Ollis, D.F. *Biochemical Engineering Fundamentals*, McGraw-Hill: New York (1986)。

可以利用现有技术确定分离的化合物的成分和纯度。这些技术包括高效液相色谱法(HPLC)、光谱学方法、染色方法、薄层色谱法、NIRS、酶分析或微生物学分析法。这些分析方法的概述见：Patek 等 (1994) *Appl. Environ. Microbiol.* 60:133-140; Malakhova 等(1996) *Biotekhnologiya* 11 27-32; 和 Schmidt 等(1998) *Bioprocess Engineer.* 19:67-70; Ullmann's *Encyclopedia of Industrial Chemistry* (1996) Vol. A27, VCH: Weinheim, pp. 89-90, pp. 521-540, pp. 540-547, pp. 559-566, pp. 575-581 和 pp. 581-587; Michal, G (1999) *Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology*, John Wiley and Sons; Fallon, A. 等 (1987) HPLC 在生物化学中的应用，《*Laboratory Techniques in Biochemistry and Molecular Biology*》，Vol. 17。

e) MPSS 法、克隆鉴定和同源性检索的一般描述

MPSS 技术(Massive Parallel Signature Sequencing, 大规模平行信号测序, 参考 Brenner 等, *Nat. Biotechnol.* (2000) 18, 630-634)被应用于产生维生素 B2 的丝状真菌棉桃阿舒氏囊霉。利用该技术可以获得有关大量基因在真核生物体中的表达水平的高度准确的定量信息。该技术涉及在一个特定的时间 X 时分离生物体的 mRNA, 在逆转录酶的帮助下转录成 cDNA

并克隆入具有特定标签序列的特定载体中。选择数目足够多（大约高 1000 倍）的具有不同标签序列的载体，以便在统计学上每个 DNA 分子均被克隆入因其标签序列而独特的载体中。

该载体的插入序列随后与标签一起被切下。通过这种方法得到的 DNA 分子然后与具有所述标签的分子对应物的微珠共同孵育。孵育后可认为通过特定标签或对应物每个微珠仅加载了一种类型的 DNA 分子。将这些微珠转移至特殊的流式处理槽（flow cell）中并固定在那里，从而可能利用以荧光染料为基础的改造的测序法和数码彩色相机对所有的微珠进行群体测序。尽管这种方法可进行高数目的分析，却受到约 16 到 20 个碱基对的阅读宽度的限制。然而对大多数生物体而言，该序列长度却足以使序列和可能的基因之间建立明确的关系（20 bp 序列的频率为 $\sim 1 \times 10^{12}$ ；与之相比，人类基因组的大小“仅” $\sim 3 \times 10^9$ bp）。

计数相同序列的数目并相互比较频率以分析由此得到的数据。频繁出现的序列反映了高水平的表达，而只出现一次的序列反映了低水平的表达。如果在两个不同的时间点（X 和 Y）分离 mRNA，可构建单个基因按时间顺序的表达模式。

实施例 1：自棉桃阿舒氏囊霉中分离 mRNA

根据已知的技术培养棉桃阿舒氏囊霉（营养培养基：27.5 g/l 酵母提取物；0.5 g/l 硫酸镁；50 ml/l 大豆油；pH 7）。在发酵过程的不同时间（24 h，48 h 和 72 h）取出棉桃阿舒氏囊霉的菌丝样品，根据 Sambrook 等（1989）的方案分离相应的 RNA 或 mRNA。

实施例 2：MPSS 的应用

利用上述的 MPSS 分析法分析自棉桃阿舒氏囊霉分离的 mRNA。

利用统计分析法分析得到的数据并根据表达差异的显著性进行分类。这必需要进行关于表达水平的增加和降低的检测。对表达的改变进行分类：
a) 单调的改变(monotonic change)，b) 24 h 后改变，和 c) 48 h 后改变。

将由 MPSS 分析发现的反映了表达的改变的 20 bp 序列作为探针与来自棉桃阿舒氏囊霉的基因文库杂交,文库的插入序列的平均大小约为 1kb。在此处杂交温度为约 30 至 57°C。

实施例 3: 自棉桃阿舒氏囊霉构建基因组文库

为构建基因组 DNA 文库,首先通过 Wright 和 Philippsen(Gene (1991) 109:99-105) 及 Mohr (1995, PhD 论文, Bio Zentrum universität Basel, 瑞士)的方法分离染色体 DNA。

该 DNA 由 Sau3A 部分消化。为此,利用不同量(0.1 至 1 U) Sau3A 酶消化 6 µg 基因组 DNA。在蔗糖密度梯度中分级分离 DNA 片断。分离 1 kb 区并进行 QiaEx 提取。最大的片段与 BamHI 切割的载体 pRS416 (Sikorski 和 Hieter, Genetics (1988) 122; 19-27) 连接(90 ng BamHI 切割的脱磷酸化载体; 198 ng 插入 DNA; 5 ml 水; 2 µl 的 10× 连接缓冲液; 1 U 连接酶)。该连接混合物用于转化大肠杆菌实验室菌株 XL-1 blue, 产生的克隆用于鉴定该插入序列。

实施例 4: 有序基因文库的制备(芯片技术)

将棉桃阿舒氏囊霉基因文库的约 25,000 个菌落(相当于对该基因组的约 3 倍覆盖)有序地转移至尼龙膜然后由 Sambrook 等(1989)描述的菌落杂交法处理。从 Mpss 分析法发现的 20 bp 序列合成寡核苷酸并用放射性 ³²P 标记。在每种情况下组合 10 个标记的熔点相似的寡核苷酸并一起与该尼龙膜杂交。杂交并洗涤后,由放射自显影鉴定出阳性克隆,并直接由 PCR 测序法分析。

由此方法鉴定出包含内部名称为"Oligo 19"的插入序列并与来自酿酒酵母的 MIPS 标签"Ygr257c"具有显著同源性的克隆。该插入序列具有 SEQ ID NO:1 显示的核酸序列。

由此方法鉴定出包含内部名称为"Oligo 24"的插入序列并与来自酿酒酵母的 MIPS 标签“Mdl2”具有显著同源性的克隆。该插入序列具有 SEQ ID NO:5 显示的核酸序列。

由此方法还鉴定出包含内部名称为"Oligo 109"的插入序列并与来自酿酒酵母的 MIPS 标签“Prp12”具有显著同源性的克隆。该插入序列具有 SEQ ID NO:10 显示的核酸序列。

由此方法还鉴定出包含内部名称为"Oligo 163"的插入序列并与来自酿酒酵母的 MIPS 标签“Flx1”具有显著同源性的克隆。该插入序列具有 SEQ ID NO:14 显示的核酸序列。

由此方法还鉴定出包含内部名称为"Oligo 31"的插入序列并与来自酿酒酵母的 MIPS 标签“STV1”具有显著同源性的克隆。该插入序列具有 SEQ ID NO:19 显示的核酸序列。

由此方法还鉴定出包含内部名称为"Oligo 4"的插入序列并与来自酿酒酵母的 MIPS 标签“OPT2”具有显著同源性的克隆。该插入序列具有 SEQ ID NO:23 显示的核酸序列。

由此方法还鉴定出包含内部名称为"Oligo 6"的插入序列并与来自酿酒酵母的 MIPS 标签“VAC1”具有显著同源性的克隆。该插入序列具有 SEQ ID NO:28 显示的核酸序列。

由此方法还鉴定出包含内部名称为"Oligo 146"的插入序列并与来自酿酒酵母的 MIPS 标签“Ymr162c”具有显著同源性的克隆。该插入序列具有 SEQ ID NO:33 显示的核酸序列。

由此方法还鉴定出包含内部名称为"Oligo 56"的插入序列并与来自酿酒酵母的 MIPS 标签“Ypl110c”具有显著同源性的克隆。该插入序列具有 SEQ ID NO:37 显示的核酸序列。

由此方法还鉴定出包含内部名称为"Oligo 167”的插入序列并与来自酿酒酵母的 MIPS 标签“ERP5”具有显著同源性的克隆。该插入序列具有 SEQ ID NO:42 显示的核酸序列。

实施例 5: 由 BLASTX 查询法分析序列数据

对所得核酸序列的分析, 即通过功能性氨基酸序列推断出其功能, 由 BLASTX 查询法在序列数据库中进行。几乎所有发现的氨基酸序列同系物均与酿酒酵母(面包酵母)有关。由于该生物的全部序列已被测定, 有关这些基因的更详细的信息可查阅:

http://www.mips.gsf.de/proj/yeast/search/code_search.htm。

由此发现了与酿酒酵母氨基酸片段的如下同源性。相应的比对情况见附图 1 至 10。

a) 来自 SEQ ID NO:1 的相应互补链的氨基酸序列与酿酒酵母线粒体能量传递蛋白具有显著的序列同源性。图 1 描绘了来自该互补链(相应于 SEQ ID NO:1 第 609 位至第 1 位核苷酸)的氨基酸部分序列与该酿酒酵母蛋白的部分序列。SEQ ID NO:2 显示 N 端延长的氨基酸部分序列。

因此可以推断出发现的该棉桃阿舒氏囊霉核酸序列具有线粒体能量传递蛋白的功能。

b) 来自 SEQ ID NO:5 的相应互补链的氨基酸序列与酿酒酵母 ABC 转运蛋白具有显著的序列同源性。图 2 描绘了来自该互补链(相当于 SEQ ID NO:5 第 1494 位至 1387 位核苷酸)的氨基酸部分序列 (SEQ ID NO:6) 与该酿酒酵母蛋白的部分序列。SEQ ID NO:7 显示另一个本发明的氨基酸部分序列。

因此可以推断出发现的该棉桃阿舒氏囊霉核酸序列具有 ABC 转运蛋白的功能。

c) 来自 SEQ ID NO:10 的编码链的氨基酸序列与酿酒酵母线粒体膜整合蛋白具有显著的序列同源性。图 3 描绘了来自该编码链(相当于 SEQ ID NO:10 中的第 15 位至第 455 位核苷酸)的氨基酸部分序列与该酿酒酵母蛋白的部分序列。SEQ ID NO:11 显示一种 N 端延长的氨基酸部分序列。

因此可以推断出发现的该棉桃阿舒氏囊霉核酸序列具有线粒体膜整合蛋白的功能。

d) 来自 SEQ ID NO:14 的编码链的氨基酸序列与酿酒酵母线粒体内

膜转运蛋白具有显著的序列同源性。图 4 描绘了来自该编码链(相当于 SEQ ID NO:14 中的第 455 位至第 1215 位核苷酸)的氨基酸部分序列与该酿酒酵母酶的部分序列。SEQ ID NO:15 显示一种 N 端延长的氨基酸部分序列

因此可以推断出发现的该棉桃阿舒氏囊霉核酸序列具有线粒体内膜转运蛋白的功能。

e) 来自 SEQ ID NO:19 的编码链的氨基酸序列与酿酒酵母非液泡 H^+ -ATP 酶 V0 结构域的 102 Kd 亚基具有显著的序列同源性。图 5 描绘了来自该编码链(相当于 SEQ ID NO:19 中的第 2 位至第 790 位核苷酸)的氨基酸部分序列与此酿酒酵母酶的部分序列。SEQ ID NO:20 显示一种 N 端延长的氨基酸部分序列。

因此可以推断发现的该棉桃阿舒氏囊霉核酸序列具有非液泡 H^+ -ATP 酶 V0 结构域的 102 Kd 亚基的功能。

f) 来自 SEQ ID NO:23 的相应互补链的氨基酸序列与酿酒酵母中和粟酒裂殖酵母的 isp4 蛋白相似的蛋白具有显著的序列同源性。图 6 描绘了来自该互补链(相当于 SEQ ID NO:23 中的第 869 位至第 522 位核苷酸)的氨基酸部分序列与此酿酒酵母酶的部分序列。SEQ ID NO:24 显示一种 N 端延长的氨基酸部分序列。

因此可以推断出发现的该棉桃阿舒氏囊霉核酸序列具有与粟酒裂殖酵母 isp4 蛋白相似的蛋白的功能,因而具有寡肽转运蛋白的活性。

g) 来自 SEQ ID NO:28 的相应互补链的氨基酸序列与酿酒酵母 VAC1 蛋白(一种具有三个锌指结构的胞质膜外周蛋白)具有显著的序列同源性。图 7A 描绘了来自该互补链(相当于 SEQ ID NO:28 中的第 356 位至第 243 位核苷酸)的氨基酸部分序列与该酿酒酵母蛋白的部分序列。图 7B 描绘了来自该互补链(相当于 SEQ ID NO:28 中的第 166 位至第 2 位核苷酸)的氨基酸部分序列与该酿酒酵母蛋白的部分序列。SEQ ID NO:29 和 SEQ ID NO:30 各显示一种 N 端延长的氨基酸部分序列。

因此可以推断出发现的该棉桃阿舒氏囊霉核酸序列具有 VAC1 蛋白(一种具有三个锌指结构的胞质膜外周蛋白)的功能。

h) 来自 SEQ ID NO:33 的相应互补链的氨基酸序列与酿酒酵母的具有 ATP 酶样功能的蛋白具有显著的序列同源性。图 8 描绘了来自该互补链（相当于 SEQ ID NO:33 中的第 904 位至第 707 位核苷酸）的氨基酸部分序列与该酿酒酵母酶的部分序列。SEQ ID NO:34 显示一种 N 端延长的氨基酸部分序列。

因此可以推断出发现的该棉桃阿舒氏囊霉核酸序列具有 ATP 酶样蛋白的功能。

i) 来自 SEQ ID NO:37 的相应互补链的氨基酸序列与酿酒酵母 PHO85 蛋白具有显著的序列同源性。图 9 描绘了来自该互补链（相当于 SEQ ID NO:37 中的第 898 位至第 5 位核苷酸）的氨基酸部分序列与该酿酒酵母酶的部分序列。SEQ ID NO:38 和 SEQ ID NO:39 显示的氨基酸序列相应于分别来自 SEQ ID NO:37 第 950 位至第 900 位和第 898 位至第 5 位的互补链的氨基酸部分序列。

因此可以推断出发现的该棉桃阿舒氏囊霉核酸序列具有 PHO85 蛋白的功能。

k) 来自 SEQ ID NO:42 的相应互补链的氨基酸序列与酿酒酵母中参与膜运输的 p24 蛋白具有显著的序列同源性。图 10 描绘了来自该互补链（相当于 SEQ ID NO:42 中的第 931 位至第 806 位核苷酸）的氨基酸部分序列与该酿酒酵母蛋白的部分序列。SEQ ID NO:24 显示一种 N 端延长的氨基酸部分序列。

因此可以推断出发现的该棉桃阿舒氏囊霉核酸序列具有参与膜运输的 p24 蛋白的功能。

实施例 6: 全长 DNA 的分离

a) 构建棉桃阿舒氏囊霉基因文库

从 100ml 在液体 MA2 培养基（10 克葡萄糖, 10 克蛋白胨, 1 克酵母提取物, 0.3 g 肌醇, 加至 1000 ml）中生长两天的棉桃阿舒氏囊霉培养物制备高分子量细胞总 DNA。滤出菌丝体, 并用蒸馏水洗两遍, 悬浮于 10ml

包含 1 M 山梨糖醇、20 mM EDTA 和 20 mg 酶解酶 20T 的溶液中，并于 27°C 下孵育、轻振 30 至 60 分钟。将原生质体悬液调整至包含 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 100 mM EDTA 和 0.5% 浓度的十二烷基硫酸钠(SDS)，并于 65°C 下孵育 20 分钟。经过两次酚/氯仿(体积比 1:1)萃取，用异丙醇沉淀 DNA，悬浮于 TE 缓冲液中，经 RNA 酶的处理后再次用异丙醇沉淀并悬浮于 TE 缓冲液中。

根据大小选出并由 Sau3A 部分消化的基因组 DNA 与粘粒载体 Super-Cos1 (Stratagene)的去磷酸化臂连接产生棉桃阿舒氏囊霉粘粒基因文库。Super-Cos1 载体在两个 cos 位点间经 XbaI 消化打开并用牛小肠碱性磷酸酶(Boehringer)去磷酸化，接着用 BamHI 打开克隆位点。该连接在 20 μ l 包含 2.5 μ g 部分消化的染色体 DNA, 1 μ g Super-Cos1 载体臂, 40 mM Tris-HCl, pH 7.5, 10 mM MgCl₂, 1 mM 二硫苏糖醇, 0.5 mM ATP 和 2 Weiss 单位 T4-DNA 连接酶(Boehringer) 的混合液中、15°C 下过夜进行。利用提取物和 Stratagene 的方案(Gigapack II Packaging Extract)在体外包装连接产物。将包装产物用于感染大肠杆菌 NM554 (recA13, araD139, Δ (ara,leu)7696, Δ (lac)17A, galU, galK, hsrR, rps(str^r), mcrA, mcrB) 并涂布在含有氨苄青霉素 (50 μ g/ml)的 LB 板上。获得含有长度平均为 30-45 kb 的棉桃阿舒氏囊霉插入片段的转化体。

b) 粘粒基因文库的储存和筛选

将共 4×10^4 个新鲜单菌落各自单独地接种于 96 孔微量滴定板 (Falcon, No. 3072) 孔中的 100 μ l LB 培养基中，所述 LB 培养基补充有冻存液(36 mM K₂HPO₄/13.2 mM KH₂PO₄、1.7 mM 柠檬酸钠、0.4 mM MgSO₄、6.8 mM (NH₄)₂SO₄、4.4% (w/v) 甘油)和氨苄青霉素 (50 μ g/ml)，37°C 下振摇过夜生长之后在 -70°C 下冷冻。快速融解滴定板并在新鲜培养基中用经酒精浴消毒随后在热板上蒸发去除酒精的 96 孔复制板进行复制。在冷冻前和解冻后(采取其它所有措施前)，在微量滴定板振荡器 (Infors) 中短暂振荡这些板子以保证该细胞悬液的均匀。利用能够将少量液体自 96 孔微量滴定板转移至尼龙膜的自动机械系统(Bio-Robotics) 将单克隆置于尼龙膜上 (GeneScreen Plus, New England Nuclear)。当培养物自 96 孔微

量滴定板转移后 (1920 个克隆), 该膜被置于 22 x 22 厘米培养皿(Nunc)中含有氨苄青霉素(50 µg/ml) 的 LB 琼脂表面上并在 37°C 下孵育过夜。细胞汇合前, 用 Herrmann, B. G., Barlow, D. P 和 Lehrach, H. (1987), Cell 48, pp. 813-825 描述的方法处理这些膜, 包括第一个变性步骤后的附加处理, 即将滤膜置于沸水浴上饱含变性溶液的垫子上方暴露于蒸气 5 分钟。

利用随机六聚体引物标记法(Feinberg, A. P. 和 Vogelstein, B. (1983), Anal. Biochem. 132, pp. 6-13)、通过 $[\alpha\text{-}^{32}\text{P}]\text{dCTP}$ 的摄取, 得到高比活的标记双链探针。这些膜经过预杂交后在 42°C 下、于含有 50% (vol/vol) 甲酰胺、600 mM 磷酸钠、pH 7.2、1 mM EDTA、10%硫酸葡聚糖、1% SDS 和 10x Denhardt 氏液、鲑鱼精子 DNA (50 µg/ml)以及 ^{32}P 标记的探针 ($0.5\text{-}1 \times 10^6 \text{cpm/ml}$)的溶液中杂交 6 至 12 小时。典型地, 在 55 至 65°C 下、13 至 30 mM NaCl、1.5 至 3 mM 柠檬酸钠、pH 6.3、0.1% SDS 中洗涤约 1 小时, 滤膜在-70°C 用柯达增感屏放射自显影 12 至 24 小时。至今, 每张膜均已被成功重复使用 20 次以上。在 95°C 下、2 mM Tris-HCl, pH 8.0, 0.2 mM EDTA, 0.1% SDS 中孵育 2 x 20 分钟剥离滤膜上的探针。

c) 从存储的基因文库中回收阳性菌落

用无菌一次性柳叶刀将微量滴定板孔中的冰冻细菌培养物刮出, 并在含有氨苄青霉素(50 µg/ml)的 LB 琼脂培养皿上划线。然后, 将单菌落接种于液体培养基中, 通过碱裂解法产生 DNA(Birnboim, H. C. 和 Doly, J. (1979), Nucleic Acids Res. 7, pp. 1513-1523)。

d) 全长 DNA

如上述可鉴定出包含具有正确的完全序列的插入片段的克隆。这些克隆的内部名称为:

“Oligo 19v”, 包含完全序列的插入片断具有 SEQ ID NO:3 中所示的核酸序列。

“Oligo 24v”, 包含完全序列的插入片断具有 SEQ ID NO:8 中所示的核酸序列。

“Oligo 109v”, 包含完全序列的插入片断具有 SEQ ID NO:12 中所示的核酸序列。

“Oligo 163v”，包含完全序列的插入片段具有 SEQ ID NO:17 中所示的核酸序列。

“Oligo 31v”，包含完全序列的插入片段具有 SEQ ID NO:21 中所示的核酸序列。

“Oligo 4v”，包含完全序列的插入片段具有 SEQ ID NO:26 中所示的核酸序列。

“Oligo 6v”，包含完全序列的插入片段具有 SEQ ID NO:31 中所示的核酸序列。

“Oligo 146v”，包含完全序列的插入片段具有 SEQ ID NO:35 中所示的核酸序列。

“Oligo 56v”，包含完全序列的插入片段具有 SEQ ID NO:40 中所示的核酸序列。

“Oligo 167v”，包含完全序列的插入片段具有 SEQ ID NO:44 中所示的核酸序列。

本发明所有部分序列和完全序列的综览如下表 1:

表 1: 序列综览

<u>SEQ ID NO:</u>	<u>Oligo</u>	<u>序列说明</u>	<u>序列同系物</u>
1	019	DNA 部分序列	酿酒酵母线粒体能量传递蛋白
2	019	来自 SEQ ID NO:1 的互补链的氨基酸部分序列	
3	019	DNA 全长序列	
4	019	相应于 SEQ ID NO:3 第 112 位至第 294 位的编码区的氨基酸序列	
5	024	DNA 部分序列	酿酒酵母 ABC 转运蛋白
6	024	来自 SEQ ID NO:5 的互补链的氨基酸部分序列	
7	024	来自 SEQ ID NO:5 的互补链的氨基酸部分序列	
8	024	DNA 全长序列	
9	024	相应于 SEQ ID NO:8 第 820 位至第 3081 位的编码区的氨基酸序列	
10	109	DNA 部分序列	酿酒酵母线粒体膜整合蛋白
11	109	来自 SEQ ID NO:10 的编码链的氨基酸部分序列	
12	109	DNA 全长序列	
13	109	相应于 SEQ ID NO:12 第 502 位至第 2919 位的编码区的氨基酸序列	
14	163	DNA 部分序列	酿酒酵母线粒体内膜转蛋白
15	163	来自 SEQ ID NO:14 的编码链的氨基酸部分序列	
16	163	来自 SEQ ID NO:14 的编码链的氨基酸部分序列	
17	163	DNA 全长序列	

<u>SEQ ID NO:</u>	<u>Oligo</u>	<u>序列说明</u>	<u>序列同系物</u>
18	163	相应于 SEQ ID NO:17 第 329 位至第 1207 位的编码区的氨基酸序列	
19	31	DNA 部分序列	酿酒酵母非液泡 H ⁺ -ATP 酶 V0 结构域 102 kD 亚基
20	31	来自 SEQ ID NO:19 的编码链的氨基酸部分序列	
21	31	DNA 全长序列	
22	31	相应于 SEQ ID NO:21 第 623 位至第 3253 位的编码区的氨基酸序列	
23	4	DNA 部分序列	粟酒裂殖酵母 isp4 蛋白具有类似功能的酿酒酵母蛋白
24	4	来自 SEQ ID NO:23 的互补链的氨基酸部分序列	
25	4	DNA 全长序列(互补链)	
26	4	具有 ORF 区的 DNA 全长序列	
27	4	相应于 SEQ ID NO:26 第 738 位至第 1037 位的编码区的氨基酸序列	
28	6	DNA 部分序列	来自酿酒酵母的 VAC1 蛋白,一种具有三个锌指的胞质膜外周蛋白
29	6	来自 SEQ ID NO:28 的互补链的氨基酸部分序列	
30	6	来自 SEQ ID NO:28 的互补链的氨基酸部分序列	
31	6	DNA 全长序列	
32	6	相应于 SEQ ID NO:31 第 428 位至第 1993 位的编码区的氨基酸序列	

<u>SEQ ID</u> <u>NO:</u>	<u>Oligo</u>	<u>序列说明</u>	<u>序列同系物</u>
33	146	DNA 部分序列	来自酿酒酵母的具有 ATP 酶功能或 ATP 酶 样功能的蛋白
34	146	来自 SEQ ID NO:33 的互补链的氨基酸部分序列	
35	146	DNA 全长序列	
36	146	相应于 SEQ ID NO:35 第 537 位至第 1034 位的编码区的氨基酸序列	
37	56	DNA 部分序列	酿酒酵母 PHO85 蛋白
38	56	来自 SEQ ID NO:37 的互补链的氨基酸部分序列	
39	56	来自 SEQ ID NO:37 的互补链的氨基酸部分序列	
40	56	DNA 全长序列	
41	56	相应于 SEQ ID NO:40 第 426 位至第 4388 位的编码区的氨基酸序列	
42	167	DNA 部分序列	酿酒酵母 p24 蛋白
43	167	来自 SEQ ID NO:42 的互补链的氨基酸部分序列	
44	167	来自 SEQ ID NO:42 的互补链的氨基酸部分序列	
45	167	DNA 全长序列	
46	167	相应于 SEQ ID NO:45 第 563 位至第 1216 位的编码区的氨基酸序列	

```

<110> 巴斯福股份公司 (BASF Aktiengesellschaft)
<120> 来自棉桃阿舒氏囊霉的与跨膜运输相关的新基因产物
<130> M/42291_PCT
<160> 45
<170> PatentIn version 3.1

<210> 1
<211> 1253
<212> DNA
<213> 棉桃阿舒氏囊霉 (Ashbya gossypii)

<220>
<221> misc_feature
<223> Oligo 19

<220>
<221> misc_feature
<222> (579)..(579)
<223> 未知核苷酸

<400> 1
gatctccaag cctttaaata gcgcacggta acccataacc gaaacctcat gccgcatctc      60
tcgcagcagg tcgcctatca gatatatgtg acgctctgtg tctcttgccg ggggcacact      120
ctggagccgt gtgcgcaata gctccagcgg cgcaatagta gttgcagcca atatccgcgc      180
aaatgctcca cacactagtg ggtttgccac aggtagccgg gatgccaagg gcgagttatc      240
acgcagtgtt tcatagccgg aaaagtacac cacgttagcc ggcaccgcca ttacgagcgt      300
aatccccagc cctcgccaca gagtcggcag accctctagt tgggctatct tcctgagccc      360
ctccagagtt ccctgcagcc tcgcagcagg ctctcggcag ccgacgtttg caaagcactc      420
atcctgccaa aacactttcc ctgccggttt cgacagctgt cctgtacatg tacaacttgg      480
aagcatttct tgctgttgaa gccgcacgcg gaccacgtcc agtggcgtga gaaacagtga      540
cgtgaccagc gagccagcag aggcgcttac catgcgttnc ttcagcgtaa ccttctttcc      600
gggttcctta ctcatggccc caaatTTgca cctttcagcg ctaaccccg tccaatagtac      660
tgccctcttc tgttagccta tcttgccgct cgagagtaac ttatgttttg atcgaagctt      720
tcttgggggc gattattcgc tgtgtatggg acttagtttt actatgcgct aaaatgttgt      780
cggggcccca aagccgttgt acggcgtcag cgttgatgtg caccgccgtg ctatttaata      840
atatactata taaatactaa aactattaag tattaatcat gatgatgtgg aatcagtatt      900
tcagtaagat gtagcatcag agttgtccaa tgggacctcc gggctttcct cgtatagtct      960
taaaaccatg gaggttatcc gtttcagctg gaattgtaat agttcgtcga cgggttgccat     1020

```

```

attaatctgg gtcttctccc cgcgctggaa caaagacatg tcaaattcat cctcctccca 1080
gttcatgtac tcgtcaatgc tatagcagtc gtgagtcagg gtactcttcg gtgaacaata 1140
tgtattggaa cattttctat ccattgggta accttctcgc ggtggggcat gttcagcgct 1200
gcgcgcggat taaaataggc aggaaagctc gactgtggaa catcgtaggg atc 1253

```

```

<210> 2
<211> 203
<212> PRT
<213> 棉桃阿舒氏囊霉

```

```

<220>
<221> misc_feature
<223> Oligo 19

```

```

<220>
<221> misc_feature
<222> (11)..(11)
<223> 未知氨基酸

```

```

<400> 2

```

```

Lys Glu Pro Gly Lys Lys Val Thr Leu Lys Xaa Arg Met Val Ser Ala
1          5          10          15

Ser Ala Gly Ser Leu Val Thr Ser Leu Phe Leu Thr Pro Leu Asp Val
20          25          30

Val Arg Val Arg Leu Gln Gln Gln Glu Met Leu Pro Ser Cys Thr Cys
35          40          45

Thr Gly Gln Leu Ser Lys Pro Ala Gly Lys Val Phe Trp Gln Asp Glu
50          55          60

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

Thr Leu Glu Gly Leu Arg Lys Ile Ala Gln Leu Glu Gly Leu Pro Thr
85          90          95

Leu Trp Arg Gly Leu Gly Ile Thr Leu Val Met Ala Val Pro Ala Asn
100         105         110

Val Val Tyr Phe Ser Gly Tyr Glu Ala Leu Arg Asp Asn Ser Pro Leu
115         120         125

Ala Ser Arg Leu Pro Val Ala Asn Pro Leu Val Cys Gly Ala Phe Ala
130         135         140

Arg Ile Leu Ala Ala Thr Thr Ile Ala Pro Leu Glu Leu Leu Arg Thr
145         150         155         160

Arg Leu Gln Ser Val Pro Arg Ala Arg Asp Thr Glu Arg Thr Ile Tyr
165         170         175

Leu Ile Gly Asp Leu Leu Arg Glu Met Arg His Glu Val Ser Val Met
180         185         190

```

Gly Tyr Arg Ala Leu Phe Lys Gly Leu Glu Ile
195 200

<210> 3

<211> 2020

<212> DNA

<213> 棉桃阿舒氏囊霉

<220>

<221> CDS

<222> (112)..(294)

<223>

<220>

<221> misc_feature

<223> Oligo 19

<400> 3

aatcccccta tattcaacaa cagtcgcgag tatttcccag atctgaacta tttatctcca 60

tggcacagtt attcaaacgg tgtcggcacc ccaataccaa ttaacagttc c atg att 117
Met Ile
1

ggg aac cca aac ctt tcc aac gtc ccc atg acc aag acg tac gat ccc 165
Gly Asn Pro Asn Leu Ser Asn Val Pro Met Thr Lys Thr Tyr Asp Pro
5 10 15

tac gat gtt cca cag tcg agc ttt cct gcc tat ttt aat ccg cgc gca 213
Tyr Asp Val Pro Gln Ser Ser Phe Pro Ala Tyr Phe Asn Pro Arg Ala
20 25 30

gcg ctg aac atg ccc cac cgc gag aag gtt aac caa tgg ata gaa aat 261
Ala Leu Asn Met Pro His Arg Glu Lys Val Asn Gln Trp Ile Glu Asn
35 40 45 50

gtt cca ata cat att gtt cac cga aga gta ccc tgactcacga ctgctatagc 314
Val Pro Ile His Ile Val His Arg Arg Val Pro
55 60

attgacgagt acatgaactg ggaggaggat gaatttgaca tgtctttgtt ccagcgcggg 374

gagaagaccc agattaatat ggcaaccgtc gacgaactat tacaattcca gctgaaacgg 434

ataacctcca tggttttaag actatacgag gaaagcccgg aggtcccatt ggacaactct 494

gatgctacat cttactgaaa tactgattcc acatcatcat gattaatact taatagtttt 554

agtattttata tagtatatta ttaaatagcc acgggggtgca catcaacgct gacgccgtac 614

aacgggttttc gggcccccgc aacatttttag cgcatagtaa aactaagtcc catacacagc 674

gaataatcgc ccccaagaaa gcttcgatca aaacataagt tactctcgag cggcaagata 734

ggctaacagg aagaggcagt actattggac gggggttagcg ctgaaagggtg caaatttggg 794

gccatgagta aggaaccggg aaagaagggt acgctgaagg aacgcatggt aagcgcctct 854

gctggctcgc tggtcacgtc actgtttctc acgccactgg acgtggtccg cgtgcccgtt 914

```

caacagcaag aaatgcttcc aagttgtaca tgtacaggac agctgtcgaa accggcaggg      974
aaagtgtttt ggcaggatga gtgcttttgca aacgtcggct gccgagagcc tgctgcgagg      1034
ctgcagggaa ctctggaggg gctcaggaag atagcccaac tagagggctct gccgactctg      1094
tggcgagggc tggggattac gctcgtaatg gcggtgccgg ctaacgtggt gtacttttcc      1154
ggctatgaag cactgctga taactcgccc ttggcatccc ggctacctgt ggcaaaccga      1214
ctagtgtgtg gagcatttgc gcggatattg gctgcaacta ctattgcgcc gctggagcta      1274
ttgcgcacac ggctccagag tgtgccccgc gcaagagaca cagagcgtac aatatactctg      1334
ataggcgacc tgctgcgaga gatgcggcat gaggtttcgg ttatgggtta ccgtgcgcta      1394
tttaaaggct tggagatcac tttatggagg gacgtgccct tcagcgcaat ctattgggga      1454
acatacgagt tctgtaaaac ccagttctgg gcccgccatg ccgcaaccga taatgcatca      1514
aactgggacc atttcatcgg cagttttgcc tgcggtagca tgggcggtgc tgttgagca      1574
cttttgacac atccttttga tgtgggcaag acccgcatgc agattgcgat tgccagtcca      1634
cagcagctaa ctgtgggggg aaaagctacg aaaactgatg actcaagagg catgttctca      1694
tttttgaatg ccattaggaa atcagaaggt attagagcgc tatataccgg cctattacct      1754
agggtgatga agattgcacc aagttgcgcc ataatgatct cgacttatga actgtcgaag      1814
aagttcttca ctagttgaac tttgtttctt cctgtatata ccaagtaatt cacactgttg      1874
atacacgata cagcattttt cttaaaaccc tgtgatatac gcttgctgta gtacacgccc      1934
aaacggtgag gtctataatt ttacatgtcc gcgagatggg attggatcgg cagaatttta      1994
gagctgccgc tgtagcccat ccgtcg                                     2020

```

```

<210> 4
<211> 61
<212> PRT
<213> 棉桃阿舒氏囊霉

```

```

<220>
<221> misc_feature
<223> Oligo 19

```

```

<400> 4

```

```

Met Ile Gly Asn Pro Asn Leu Ser Asn Val Pro Met Thr Lys Thr Tyr
1           5           10           15

```

```

Asp Pro Tyr Asp Val Pro Gln Ser Ser Phe Pro Ala Tyr Phe Asn Pro
20           25           30

```

```

Arg Ala Ala Leu Asn Met Pro His Arg Glu Lys Val Asn Gln Trp Ile
35           40           45

```

```

Glu Asn Val Pro Ile His Ile Val His Arg Arg Val Pro
50           55           60

```

<210> 5
 <211> 1495
 <212> DNA
 <213> 棉桃阿舒氏囊霉

<220>
 <221> misc_feature
 <223> Oligo 24

<220>
 <221> misc_feature
 <222> (1386)..(1386)
 <223> 未知核苷酸

<400> 5
 gatcgccacc tgggtccctgt tcagcgcgcc ggacgtctgc tgcgcctcca gcgggtccag 60
 gaacttcagc atcgcgggct cgatctgcga aaagcacttg tacagcgggt agaagaggcc 120
 acgcccgtcc tggatccagt cgaagtactg tggcatgtgc gacaccacca cgtccctaga 180
 aaacgtcccc ttcagcgaca ccttcagatc ctgcagcagg ttgaccaatc gcaacaggtc 240
 ctgcttgatc tcctccgtgc tcaggttctt gtttccgatt cgcagcatca catcctccag 300
 tgagtcagct ccgggcccta gtgtggtcgc taaggggggc gagataaacg cgtaggttac 360
 ctgtatcatg ctgcggacac cacctcactt tgaaatatta caactgtgct atcagctgcc 420
 ctttataagc agcgtgtct ccatacaatt gcgctgattt tatgtagttg atgcgtgtaa 480
 tctgcttttt ccctattgag acgtcagcat ctctagcaa ggctccgtgt gcttacttaa 540
 gaaaagatgg aatcgatgga atcgatgaaa ctaaaacatt ttatcaaata tcagttctat 600
 atatatgtag ttattgggag ataatacggg tgaatctggg gcatgtctta aggcatggg 660
 gcacggggccg tgtcctccag cacatgcttt atgacctctt catttaggcg aagatttccg 720
 ccctcgatat cgcggggaat gtctgcggcc tgtggagcgg aagcctcttc ctgggcagga 780
 gctggagacg cctcgggctt cgcaggctcc gacgcgtgtt cgttcaatag cttgaacaac 840
 tcgctgtcct cgtcacggta taggtcggca aacttgccgg cctcgatgac ggagccgtca 900
 ttgccgagaa ccacgatgtt ctccggagcgg cgtatgggtgc tgaggcgggtg ggcgatgctc 960
 acgatggtga tctccttgct gcgcattagc ctgccgaggg tgtagttgat ggcgcctcgc 1020
 ctttcaacgt ccaacgccga ggtggcctcg tcgaggatta ggattttcgcg tttcttgatg 1080
 agcgcgcggg cgatggcaat gcgctggcgc tggccccgcg tgagtagggc gcctctaggc 1140
 ccaatcacgg tgtcatagcc gtccgggaac ttggtgatga agttatggca gaagcacttc 1200
 ttggcaactg cgcgtatctc atccatggag ggcgtgtagg aaaggccgta ggtgatgttg 1260
 tctctgatgg tcccggacat caagaccggc tcctgctgga caacgccaaag atgttttctg 1320

```

agcgacttgg cgcttacctt ggatatgtcc tgaccgtcta tgaggatggg gccgctgttg 1380
gcacnngtaa aaccgtagga ggagggacgc aatggtggac ttgccccgcc cggagggccc 1440
gacgatgcag acgttggagc ccgggggttat cgtgaggcgg agtttcttca agatc 1495

```

```

<210> 6
<211> 36
<212> PRT
<213> 棉桃阿舒氏囊霉

```

```

<220>
<221> misc_feature
<223> Oligo 24

```

```

<400> 6

```

```

Ile Leu Lys Lys Leu Arg Leu Thr Ile Thr Pro Gly Ser Asn Val Cys
1          5          10          15
Ile Val Gly Pro Ser Gly Arg Gly Lys Ser Thr Ile Ala Ser Leu Leu
          20          25          30
Leu Arg Phe Tyr
          35

```

```

<210> 7
<211> 225
<212> PRT
<213> 棉桃阿舒氏囊霉

```

```

<220>
<221> misc_feature
<223> Oligo 24

```

```

<400> 7

```

```

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

```

```

Arg Ile Ala Ile Ala Arg Ala Leu Ile Lys Lys Pro Lys Ile Leu Ile
      115              120              125

Leu Asp Glu Ala Thr Ser Ala Leu Asp Val Glu Ser Glu Gly Ala Ile
      130              135              140

Asn Tyr Thr Leu Gly Arg Leu Met Arg Ser Lys Glu Ile Thr Ile Val
      145              150              155              160

Ser Ile Ala His Arg Leu Ser Thr Ile Arg Arg Ser Glu Asn Ile Val
              165              170              175

Val Leu Gly Asn Asp Gly Ser Val Ile Glu Ala Gly Lys Phe Ala Asp
      180              185              190

Leu Tyr Arg Asp Glu Asp Ser Glu Leu Phe Lys Leu Leu Asn Glu His
      195              200              205

Ala Ser Glu Pro Ala Lys Pro Glu Ala Ser Pro Ala Pro Ala Gln Glu
      210              215              220

```

Glu
225

<210> 8
<211> 4118
<212> DNA
<213> 棉桃阿舒氏囊霉

<220>
<221> CDS
<222> (820)..(3081)
<223>

<220>
<221> misc_feature
<223> Oligo 24

```

<400> 8
ctgtcacagc aagacctggc ttcccgtcac gctatggctc gcctgatcat ctcatctcgc      60

ctgcgcacatc agcgcagctc aaaattttta ccattacccc ggccagtaca catattacat      120

aatcaccagt gtccagggat taccgggaca tcaaacaatca cagagaacga acccaatgaa      180

atgcaggaaa agctgctcag cgccgggtgcc cggcggccat taagcgtgca cctacagcca      240

cttcaactcca ggccacacac tccacgcctc cccgcaagta aacacgatgt ctgcctggag      300

aaaagccggt ttgacctaca atagctacct ggccgttgcg gccaggaccg tgcgcgcggc      360

tctgaagaag gagctccaga gccctgccgt cctcaaccga tctgtgacgg aggcaaaggt      420

catcgactac gcctccaagg gctccgcccgc ggaggccgtg ccgctacgga agtgagcagc      480

acgcgtgctc ttctcgttga actgctggca cgcagcctgg gccctgttgg cggcaagcgg      540

gcggatggag acgcccagca caggcccttt gacgggctgt gctggtggca tgtgtgtagt      600

tcgactatat aaactctcgt tctctaataa taatgacaga aggtcgattg ggcgatttcg      660

```

gtgctgcgaa cggggacgtt gtgcaggggc aaaggcgggtt ttcgcgggca agcgaagagg	720
tgcgtttttcc gcaagcttgc ttgtgtcccc gcccaaggca agacacagtc tactcgcac	780
agttttggca gagcgtgttg ctacaggatt ccaggcgct atg tgg aga gcg ata	834
Met Trp Arg Ala Ile	
1 5	
agg ccc ttt ggg ggg ctt ggc cct gta gcg gtg tta cag agg cta ccg	882
Arg Pro Phe Gly Gly Leu Gly Pro Val Ala Val Leu Gln Arg Leu Pro	
10 15 20	
ggc ccg cgg cac aat gtc gcc gtg agg ctt cag gtg atg tcc agg gcc	930
Gly Pro Arg His Asn Val Ala Val Arg Leu Gln Val Met Ser Arg Ala	
25 30 35	
aga tgc ggc tgg cag tgt gtg cgt gcg aca tct cga ctg aca gtg ccg	978
Arg Cys Gly Trp Gln Cys Val Arg Ala Thr Ser Arg Leu Thr Val Pro	
40 45 50	
gtg tgc agg caa agc cgg gcg tac agc gct gcc cgt gtg ggc ggc gag	1026
Val Cys Arg Gln Ser Arg Ala Tyr Ser Ala Ala Arg Val Gly Gly Glu	
55 60 65	
gga acg cct tcg ccg cgg gac aag gag cat gag cag gag cgc aaa gtt	1074
Gly Thr Pro Ser Pro Arg Asp Lys Glu His Glu Gln Glu Arg Lys Val	
70 75 80 85	
gaa aag cga ccg gcg gag ccg tcg cgc gcg caa ccc gca aag cca gcc	1122
Glu Lys Arg Pro Ala Glu Pro Ser Arg Ala Gln Pro Ala Lys Pro Ala	
90 95 100	
ggc tgg cac gag gtg ttc cgg ctc ttg cga ttg gtg aag ccc gac tgg	1170
Gly Trp His Glu Val Phe Arg Leu Leu Arg Leu Val Lys Pro Asp Trp	
105 110 115	
aag ttg ttg ctt gcg gcg ttg ggg ctc ctc acg ata tct tgc tcg att	1218
Lys Leu Leu Leu Ala Ala Leu Gly Leu Leu Thr Ile Ser Cys Ser Ile	
120 125 130	
ggg atg gcg gtt ccg aag gtt att ggg ctt gtg cta gat gcc acg aaa	1266
Gly Met Ala Val Pro Lys Val Ile Gly Leu Val Leu Asp Ala Thr Lys	
135 140 145	
gac gca gtc gca cag gca gaa aag aac gaa acg ggc gcg ccg atg tct	1314
Asp Ala Val Ala Gln Ala Glu Lys Asn Glu Thr Gly Ala Pro Met Ser	
150 155 160 165	
ttg cta gac atg ccg cct atc atc ggg aac ttc acg cta gtc gac gtg	1362
Leu Leu Asp Met Pro Pro Ile Ile Gly Asn Phe Thr Leu Val Asp Val	
170 175 180	
ctt gct gcg ttt gcg ggc gcg ttg ctc gtg ggc tcc gcc gcg aac ttc	1410
Leu Ala Ala Phe Ala Gly Ala Leu Leu Val Gly Ser Ala Ala Asn Phe	
185 190 195	
ggg cgg atg ttc ctc ttg aaa atg ctt ggt gag cgt ttg gtt gcg cgg	1458
Gly Arg Met Phe Leu Leu Lys Met Leu Gly Glu Arg Leu Val Ala Arg	
200 205 210	
cta cgt gcg cag gtc atc aag aag acg ttg cat cac gac gcg gag ttt	1506

Leu	Arg	Ala	Gln	Val	Ile	Lys	Lys	Thr	Leu	His	His	Asp	Ala	Glu	Phe	
215						220				225						
ttc	gac	agg	aac	aag	gtc	ggc	gac	ttg	ata	tcg	cgg	ctt	ggg	tcc	gac	1554
Phe	Asp	Arg	Asn	Lys	Val	Gly	Asp	Leu	Ile	Ser	Arg	Leu	Gly	Ser	Asp	
230					235					240					245	
gcc	tac	att	gta	tcg	cgc	tcg	atg	acg	cag	aac	att	tcg	gac	ggg	ggt	1602
Ala	Tyr	Ile	Val	Ser	Arg	Ser	Met	Thr	Gln	Asn	Ile	Ser	Asp	Gly	Val	
				250					255					260		
aaa	gcc	gcg	ctc	tgc	gga	ggc	gtc	ggg	gtg	ggc	atg	atg	ttc	cac	att	1650
Lys	Ala	Ala	Leu	Cys	Gly	Gly	Val	Gly	Val	Gly	Met	Met	Phe	His	Ile	
			265					270					275			
tcg	agc	aca	ttg	act	tct	gca	atg	ata	atg	ttt	gcc	cct	ccc	ctc	ctc	1698
Ser	Ser	Thr	Leu	Thr	Ser	Ala	Met	Ile	Met	Phe	Ala	Pro	Pro	Leu	Leu	
		280					285					290				
atc	gcc	gcc	acc	atc	tac	ggc	aag	cgc	att	cgc	gcc	atc	tcg	cgc	gag	1746
Ile	Ala	Ala	Thr	Ile	Tyr	Gly	Lys	Arg	Ile	Arg	Ala	Ile	Ser	Arg	Glu	
	295					300					305					
ctc	cag	caa	tct	aca	ggg	aac	ttg	acg	cgg	gtg	tca	gag	gag	cag	ttc	1794
Leu	Gln	Gln	Ser	Thr	Gly	Asn	Leu	Thr	Arg	Val	Ser	Glu	Glu	Gln	Phe	
310					315					320					325	
aat	ggc	atc	aag	acc	atc	aag	tca	ttc	gtc	gcg	gag	gga	aag	gag	atg	1842
Asn	Gly	Ile	Lys	Thr	Ile	Lys	Ser	Phe	Val	Ala	Glu	Gly	Lys	Glu	Met	
				330					335					340		
cgg	aga	tac	aac	acg	gcc	gtg	cgg	agg	ctc	ttc	aac	gtt	gca	aag	agg	1890
Arg	Arg	Tyr	Asn	Thr	Ala	Val	Arg	Arg	Leu	Phe	Asn	Val	Ala	Lys	Arg	
			345					350					355			
gaa	ggg	ata	acg	agc	gcg	acg	ttt	ttc	agc	ggg	acg	aat	gtg	ttg	ggc	1938
Glu	Gly	Ile	Thr	Ser	Ala	Thr	Phe	Phe	Ser	Gly	Thr	Asn	Val	Leu	Gly	
		360					365					370				
gac	gtg	agt	ttt	ttg	ctt	gtt	ttg	gcg	tat	ggg	tcg	cac	ctc	ggt	ttg	1986
Asp	Val	Ser	Phe	Leu	Leu	Val	Leu	Ala	Tyr	Gly	Ser	His	Leu	Val	Leu	
	375					380					385					
cag	ggc	gaa	cta	tcc	ctt	ggg	agc	ttg	acg	gcg	ttc	atg	ctc	tac	act	2034
Gln	Gly	Glu	Leu	Ser	Leu	Gly	Ser	Leu	Thr	Ala	Phe	Met	Leu	Tyr	Thr	
390					395					400					405	
gag	tac	act	gga	agc	tcg	gtt	ttc	ggg	tta	tcg	aac	ttt	tac	tca	gag	2082
Glu	Tyr	Thr	Gly	Ser	Ser	Val	Phe	Gly	Leu	Ser	Asn	Phe	Tyr	Ser	Glu	
				410					415					420		
ttg	atg	aag	ggc	gct	ggg	gct	gct	tcg	cgc	ttg	ttc	gaa	ttg	aca	gac	2130
Leu	Met	Lys	Gly	Ala	Gly	Ala	Ala	Ser	Arg	Leu	Phe	Glu	Leu	Thr	Asp	
			425					430					435			
tac	aag	aac	tcc	gtg	ccg	tct	acc	gtt	ggc	cag	acg	ttt	gtg	ccg	cgc	2178
Tyr	Lys	Asn	Ser	Val	Pro	Ser	Thr	Val	Gly	Gln	Thr	Phe	Val	Pro	Arg	
		440					445					450				
gat	ggg	cgc	atc	gag	ttc	aag	gac	gtg	tcc	ttc	agc	tac	ccc	acg	cgg	2226
Asp	Gly	Arg	Ile	Glu	Phe	Lys	Asp	Val	Ser	Phe	Ser	Tyr	Pro	Thr	Arg	
	455					460					465					

ccg acg agc cag atc ttc aag aaa ctc cgc ctc acg ata acc ccg ggc Pro Thr Ser Gln Ile Phe Lys Lys Leu Arg Leu Thr Ile Thr Pro Gly 470 475 480 485	2274
tcc aac gtc tgc atc gtc ggg ccc tcc ggg cgg ggc aag tcc acc att Ser Asn Val Cys Ile Val Gly Pro Ser Gly Arg Gly Lys Ser Thr Ile 490 495 500	2322
gcg tcc ctc ctc cta cgg ttt tac gat gcc aac agc ggc acc atc ctc Ala Ser Leu Leu Leu Arg Phe Tyr Asp Ala Asn Ser Gly Thr Ile Leu 505 510 515	2370
ata gac ggt cag gac ata tcc aag gta agc gcc aag tcg ctc aga aaa Ile Asp Gly Gln Asp Ile Ser Lys Val Ser Ala Lys Ser Leu Arg Lys 520 525 530	2418
cat ctt ggc gtt gtc cag cag gag ccg gtc ttg atg tcc ggg acc atc His Leu Gly Val Val Gln Gln Glu Pro Val Leu Met Ser Gly Thr Ile 535 540 545	2466
aga gac aac atc acc tac ggc ctt tcc tac acg ccc tcc atg gat gag Arg Asp Asn Ile Thr Tyr Gly Leu Ser Tyr Thr Pro Ser Met Asp Glu 550 555 560 565	2514
ata cgc gca gtt gcc aag aag tgc ttc tgc cat aac ttc atc acc aag Ile Arg Ala Val Ala Lys Lys Cys Phe Cys His Asn Phe Ile Thr Lys 570 575 580	2562
ttc ccg gac ggc tat gac acc gtg att ggg cct aga ggc gcc cta ctc Phe Pro Asp Gly Tyr Asp Thr Val Ile Gly Pro Arg Gly Ala Leu Leu 585 590 595	2610
agc ggg ggc cag cgc cag cgc att gcc atc gcc cgc gcg ctc atc aag Ser Gly Gly Gln Arg Gln Arg Ile Ala Ile Ala Arg Ala Leu Ile Lys 600 605 610	2658
aaa ccg aaa atc cta atc ctc gac gag gcc acc tcg gcg ttg gac gtt Lys Pro Lys Ile Leu Ile Leu Asp Glu Ala Thr Ser Ala Leu Asp Val 615 620 625	2706
gaa agc gag ggc gcc atc aac tac acc ctc ggc agg cta atg cgc agc Glu Ser Glu Gly Ala Ile Asn Tyr Thr Leu Gly Arg Leu Met Arg Ser 630 635 640 645	2754
aag gag atc acc atc gtg agc atc gcc cac cgc ctc agc acc ata cgc Lys Glu Ile Thr Ile Val Ser Ile Ala His Arg Leu Ser Thr Ile Arg 650 655 660	2802
cgc tcc gag aac atc gtg gtt ctc ggc aat gac ggc tcc gtc atc gag Arg Ser Glu Asn Ile Val Val Leu Gly Asn Asp Gly Ser Val Ile Glu 665 670 675	2850
gcc ggc aag ttt gcc gac cta tac cgt gac gag gac agc gag ttg ttc Ala Gly Lys Phe Ala Asp Leu Tyr Arg Asp Glu Asp Ser Glu Leu Phe 680 685 690	2898
aag cta ttg aac gaa cac gcg tcg gag cct gcg aag ccc gag gcg tct Lys Leu Leu Asn Glu His Ala Ser Glu Pro Ala Lys Pro Glu Ala Ser 695 700 705	2946
cca gct cct gcc cag gaa gag gct tcc gct cca cag gcc gca gac att	2994

```

Pro Ala Pro Ala Gln Glu Glu Ala Ser Ala Pro Gln Ala Ala Asp Ile
710                               715                               720                               725

ccc cgc gat atc gag ggc gga aat ctt cgc cta aat gaa gag gtc ata      3042
Pro Arg Asp Ile Glu Gly Gly Asn Leu Arg Leu Asn Glu Glu Val Ile
                               730                               735                               740

aag cat gtg ctg gag gac acg gcc cgt gcc cca atg cct taagacatgc      3091
Lys His Val Leu Glu Asp Thr Ala Arg Ala Pro Met Pro
                               745                               750

cccagattca cccgtattat ctcccaataa ctacatatat atagaactga tatttgataa      3151
aatgttttag tttcatcgat tccatcgatt ccatcttttc ttaagtaagc acacggagcc      3211
ttgctacgag atgctgacgt ctcaataggg aaaaagcaga ttacacgcat caactacata      3271
aaatcagcgc aattgtatgg agacagcgtt gcttataaag ggcagctgat agcacagttg      3331
taatatttca aagtgaggtg gtgtccgcag catgatacag gtacctaacg cgtttatctc      3391
gccccctta gcgaccacac tagggcccgg agctgactca ctggaggatg tgatgctgcg      3451
aatcggaaac aagaacctga gcacggagga gatcaagcag gacctgttgc gattggtcaa      3511
cctgctgcag gatctgaagg tgtcgctgaa ggggacgttt tctagggacg tgggtggtgtc      3571
gcacatgcca cagtacttcg actggatcca ggacgggcgt ggctctttct acccgctgta      3631
caagtgcttt tcgcagatcg agcccgcgat gctgaagttc ctggaccgcg tggaggcgca      3691
gcagacgtcc ggcgcgctga acagggacca ggtggcgatc ttcaacctga tcgaccaggt      3751
gtccgagctc gtgctgagcc tgaagccgct gttccagtcg gtgaagaacc tattcgacac      3811
ggcgctggag ttcaatgaga tcttcaagga ccacatgaac tcgctactgg aggagatcga      3871
aggaaacatg aagaagtgcc tcgcgctgca ccaggactgc ttcgcgtcgc ccgtgagaca      3931
ccctccgtcc ttcacgttgg accagctcgt ggagcttata tcctcatcgt caaacaacca      3991
gcggctgcag atgccgacct tcaatcccct ggaaaagcgc atctaccagg gattactgcg      4051
agctaagaga atgctattgt gccgatccag acatcgctga agagacgtcc tgaaaacccg      4111
catccca                                                                4118

```

<210> 9

<211> 754

<212> PRT

<213> 棉桃阿舒氏囊霉

<220>

<221> misc_feature

<223> Oligo 24

<400> 9

```

Met Trp Arg Ala Ile Arg Pro Phe Gly Gly Leu Gly Pro Val Ala Val
1                               5                               10                               15

```

```

Leu Gln Arg Leu Pro Gly Pro Arg His Asn Val Ala Val Arg Leu Gln
      20                      25                      30

Val Met Ser Arg Ala Arg Cys Gly Trp Gln Cys Val Arg Ala Thr Ser
      35                      40                      45

Arg Leu Thr Val Pro Val Cys Arg Gln Ser Arg Ala Tyr Ser Ala Ala
      50                      55                      60

Arg Val Gly Gly Glu Gly Thr Pro Ser Pro Arg Asp Lys Glu His Glu
      65                      70                      75                      80

Gln Glu Arg Lys Val Glu Lys Arg Pro Ala Glu Pro Ser Arg Ala Gln
      85                      90                      95

Pro Ala Lys Pro Ala Gly Trp His Glu Val Phe Arg Leu Leu Arg Leu
      100                     105                     110

Val Lys Pro Asp Trp Lys Leu Leu Leu Ala Ala Leu Gly Leu Leu Thr
      115                     120                     125

Ile Ser Cys Ser Ile Gly Met Ala Val Pro Lys Val Ile Gly Leu Val
      130                     135                     140

Leu Asp Ala Thr Lys Asp Ala Val Ala Gln Ala Glu Lys Asn Glu Thr
      145                     150                     155                     160

Gly Ala Pro Met Ser Leu Leu Asp Met Pro Pro Ile Ile Gly Asn Phe
      165                     170                     175

Thr Leu Val Asp Val Leu Ala Ala Phe Ala Gly Ala Leu Leu Val Gly
      180                     185                     190

Ser Ala Ala Asn Phe Gly Arg Met Phe Leu Leu Lys Met Leu Gly Glu
      195                     200                     205

Arg Leu Val Ala Arg Leu Arg Ala Gln Val Ile Lys Lys Thr Leu His
      210                     215                     220

His Asp Ala Glu Phe Phe Asp Arg Asn Lys Val Gly Asp Leu Ile Ser
      225                     230                     235                     240

Arg Leu Gly Ser Asp Ala Tyr Ile Val Ser Arg Ser Met Thr Gln Asn
      245                     250                     255

Ile Ser Asp Gly Val Lys Ala Ala Leu Cys Gly Gly Val Gly Val Gly
      260                     265                     270

Met Met Phe His Ile Ser Ser Thr Leu Thr Ser Ala Met Ile Met Phe
      275                     280                     285

Ala Pro Pro Leu Leu Ile Ala Ala Thr Ile Tyr Gly Lys Arg Ile Arg
      290                     295                     300

Ala Ile Ser Arg Glu Leu Gln Gln Ser Thr Gly Asn Leu Thr Arg Val
      305                     310                     315                     320

Ser Glu Glu Gln Phe Asn Gly Ile Lys Thr Ile Lys Ser Phe Val Ala
      325                     330                     335

Glu Gly Lys Glu Met Arg Arg Tyr Asn Thr Ala Val Arg Arg Leu Phe
      340                     345                     350

```

```

Asn Val Ala Lys Arg Glu Gly Ile Thr Ser Ala Thr Phe Phe Ser Gly
    355                                360                                365

Thr Asn Val Leu Gly Asp Val Ser Phe Leu Leu Val Leu Ala Tyr Gly
    370                                375                                380

Ser His Leu Val Leu Gln Gly Glu Leu Ser Leu Gly Ser Leu Thr Ala
    385                                390                                395                                400

Phe Met Leu Tyr Thr Glu Tyr Thr Gly Ser Ser Val Phe Gly Leu Ser
    405                                410                                415

Asn Phe Tyr Ser Glu Leu Met Lys Gly Ala Gly Ala Ala Ser Arg Leu
    420                                425                                430

Phe Glu Leu Thr Asp Tyr Lys Asn Ser Val Pro Ser Thr Val Gly Gln
    435                                440                                445

Thr Phe Val Pro Arg Asp Gly Arg Ile Glu Phe Lys Asp Val Ser Phe
    450                                455                                460

Ser Tyr Pro Thr Arg Pro Thr Ser Gln Ile Phe Lys Lys Leu Arg Leu
    465                                470                                475                                480

Thr Ile Thr Pro Gly Ser Asn Val Cys Ile Val Gly Pro Ser Gly Arg
    485                                490                                495

Gly Lys Ser Thr Ile Ala Ser Leu Leu Leu Arg Phe Tyr Asp Ala Asn
    500                                505                                510

Ser Gly Thr Ile Leu Ile Asp Gly Gln Asp Ile Ser Lys Val Ser Ala
    515                                520                                525

Lys Ser Leu Arg Lys His Leu Gly Val Val Gln Gln Glu Pro Val Leu
    530                                535                                540

Met Ser Gly Thr Ile Arg Asp Asn Ile Thr Tyr Gly Leu Ser Tyr Thr
    545                                550                                555                                560

Pro Ser Met Asp Glu Ile Arg Ala Val Ala Lys Lys Cys Phe Cys His
    565                                570                                575

Asn Phe Ile Thr Lys Phe Pro Asp Gly Tyr Asp Thr Val Ile Gly Pro
    580                                585                                590

Arg Gly Ala Leu Leu Ser Gly Gly Gln Arg Gln Arg Ile Ala Ile Ala
    595                                600                                605

Arg Ala Leu Ile Lys Lys Pro Lys Ile Leu Ile Leu Asp Glu Ala Thr
    610                                615                                620

Ser Ala Leu Asp Val Glu Ser Glu Gly Ala Ile Asn Tyr Thr Leu Gly
    625                                630                                635                                640

Arg Leu Met Arg Ser Lys Glu Ile Thr Ile Val Ser Ile Ala His Arg
    645                                650                                655

Leu Ser Thr Ile Arg Arg Ser Glu Asn Ile Val Val Leu Gly Asn Asp
    660                                665                                670

Gly Ser Val Ile Glu Ala Gly Lys Phe Ala Asp Leu Tyr Arg Asp Glu

```

675	680	685
Asp Ser Glu Leu Phe Lys Leu Leu Asn Glu His Ala Ser Glu Pro Ala		
690	695	700
Lys Pro Glu Ala Ser Pro Ala Pro Ala Gln Glu Glu Ala Ser Ala Pro		
705	710	715
Gln Ala Ala Asp Ile Pro Arg Asp Ile Glu Gly Gly Asn Leu Arg Leu		
725	730	735
Asn Glu Glu Val Ile Lys His Val Leu Glu Asp Thr Ala Arg Ala Pro		
740	745	750

Met Pro

<210> 10
 <211> 1002
 <212> DNA
 <213> 棉桃阿舒氏囊霉

<220>
 <221> misc_feature
 <223> Oligo 109

<400> 10
 gatcggagcc gattaaaacg gcacaagctt gggagctgat cgaactactt tcacaaaatg 60
 atgtcgtgaa gtatggagat atcgttttca ggccgttggt taagtcttct ccagaggccg 120
 ggttggtgga gttagagaaa aatgggctga tcacaatcag ccgcaataga ggagtgttgc 180
 aagatatccg gcccgctaaa ccattgttta aagccgcatt cagctacctg ttacaggaca 240
 aggacctgtc catcgtcctc cgaacagggt actacttgcg actaatcgcg tttgaaactg 300
 ggcggatcaa aaagtgggag gaggaattgc gtctattggc aaaggtagt gaccaaagaa 360
 tatgcaagag tcggttgaat tacttagcaa gcaaaataga cgccagtagt ggggtaatta 420
 acagtttgtga ggacaaggta aaagaaatgt caaagcgtat ttagaaacgc tcgcatcaat 480
 aagaccccat tttgccacga tctggcaaga tgaactaaat agcgataatc aactagggtc 540
 tcattagttt agcattccag ctgcacgggc tttagtttag tacctcaatt ccaaaaacgg 600
 cactatcttg tccgagcgtt ggtgaatatc cttgccgtat gtgaaattat ccagggttga 660
 tgtagttggc cggagctacc agtaattggt ggccatgca cgggcccgtc ttttccgtaa 720
 cataatccag agatccgtta acaggcaaca tgaactgatt gattaatttt aatgaacttt 780
 tttttttggc tttgaaaggc ctaaaattgg atcatctaca ttccctgacg atataggatg 840
 gtaaaagccc taccaaccaa gaaattctga tccctttgac aaatccctta taacatagat 900
 attagaatat aatcaacccc attgtgtgac aactcttctc caaaaataat tgggtctttca 960
 cacaattaaa gattggcttc aaggtagctc cacggtgaga tc 1002

<210> 11
 <211> 147
 <212> PRT
 <213> 棉桃阿舒氏囊霉

<220>
 <221> misc_feature
 <223> Oligo 109

<400> 11

```
Lys Thr Ala Gln Ala Trp Glu Leu Ile Glu Leu Leu Ser Gln Asn Asp
1              5              10              15

Val Val Lys Tyr Gly Asp Ile Val Phe Arg Pro Leu Phe Lys Ser Ser
                20              25              30

Pro Glu Ala Gly Leu Leu Glu Leu Glu Lys Asn Gly Leu Ile Thr Ile
          35              40              45

Ser Arg Asn Arg Gly Val Leu Gln Asp Ile Arg Pro Ala Lys Pro Leu
          50              55              60

Phe Lys Ala Ala Phe Ser Tyr Leu Leu Gln Asp Lys Asp Leu Ser Ile
65              70              75              80

Val Leu Arg Thr Gly Tyr Tyr Leu Arg Leu Ile Ala Phe Glu Thr Gly
          85              90              95

Arg Ile Lys Lys Trp Glu Glu Glu Leu Arg Leu Leu Ala Lys Val Ser
          100             105             110

Asp Gln Arg Ile Cys Lys Ser Arg Leu Asn Tyr Leu Ala Ser Lys Ile
          115             120             125

Asp Ala Ser Ser Gly Val Ile Asn Ser Cys Glu Asp Lys Val Lys Glu
          130             135             140

Met Ser Lys
145
```

<210> 12
 <211> 4077
 <212> DNA
 <213> 棉桃阿舒氏囊霉

<220>
 <221> CDS
 <222> (502)..(2919)
 <223>

<220>
 <221> misc_feature
 <223> Oligo 109

```
<400> 12
cctgtaccgg gtgcggtgcc agatacatcg caatcccctt cgtgctgttg ccatgcggga      60
cacaagcagg tcccattccc atctgcccac ctgcaccgcg ggcgagtaat gcttctcccc      120
```

cccgtagttc gctccaccgc tcgatccgcc acgtgaaccg gccctcgtac tcggtctctt	180
ctccttctag cgggcgcaaa acctcctcga acgtcttgcc gagctccacc gcctccatgt	240
ttgtttcccg ggtttctggc acatccatta ctatcttgat atagttcgat gttgatactg	300
cctgggtcca gcctactatc caagtcctgg tttgttcatc gtgtgtagcc ctgtagagag	360
tgaatttcaa ggggcaggta acattgccgg acgcctacac ccgtggcttc tgaagaccat	420
cgcaaaaggt attcagaact ttgcaactag taccagcgat acaaggccgg tgacagcaga	480
agtaacgtcc cgagaaggcc a atg atg ttt ctg gag atg cag cga gcg ttc	531
Met Met Phe Leu Glu Met Gln Arg Ala Phe	
1 5 10	
atg cta cac ggg cgg cga gca gtg acg aga agt gcg gtc ggc gtg cgg	579
Met Leu His Gly Arg Arg Ala Val Thr Arg Ser Ala Val Gly Val Arg	
15 20 25	
tac atc tcg gag gac atc cag cag aag gac gca cag gcg ggc gag aag	627
Tyr Ile Ser Glu Asp Ile Gln Gln Lys Asp Ala Gln Ala Gly Glu Lys	
30 35 40	
gcg acg gcg acg gcg acg ggg gtt atc tac aag tcg gac gag gaa acg	675
Ala Thr Ala Thr Ala Thr Gly Val Ile Tyr Lys Ser Asp Glu Glu Thr	
45 50 55	
ctg atg tat ttc gac aac gta tac ccg cgg gcg acg tcg ttg tgg cgg	723
Leu Met Tyr Phe Asp Asn Val Tyr Pro Arg Ala Thr Ser Leu Trp Arg	
60 65 70	
ccg acg cag tgg tac aac att ctg ctg tcg aac cag tcg cgg gaa gcg	771
Pro Thr Gln Trp Tyr Asn Ile Leu Leu Ser Asn Gln Ser Arg Glu Ala	
75 80 85 90	
gtg agg gag aag atc atg cgg ttc gca agc ccg gca agc aac ccg gtg	819
Val Arg Glu Lys Ile Met Arg Phe Ala Ser Pro Ala Ser Asn Pro Val	
95 100 105	
cac gga ctg gag ctg cgg tcg acg atc ccg atc aag cgg gac ggc ggg	867
His Gly Leu Glu Leu Arg Ser Thr Ile Pro Ile Lys Arg Asp Gly Gly	
110 115 120	
gtg ttt gcg acg ttc cgg gtg ccg cgt gag tac acg agg gcg cag gtg	915
Val Phe Ala Thr Phe Arg Val Pro Arg Glu Tyr Thr Arg Ala Gln Val	
125 130 135	
aac gcg ttg atc cag gcg acc acg cag cag gag tcg tcg aag tcg ctg	963
Asn Ala Leu Ile Gln Ala Thr Thr Gln Gln Glu Ser Ser Lys Ser Leu	
140 145 150	
ctg gcg gcg ttt acg cgg gct gca gcg ttc ccg gtg aag ggc gtg ccg	1011
Leu Ala Ala Phe Thr Arg Ala Ala Ala Phe Pro Val Lys Gly Val Pro	
155 160 165 170	
tgg atc gag gac ctg aag cgg ctg ccg aac aac gtg gtg cgc gtg gag	1059
Trp Ile Glu Asp Leu Lys Arg Leu Pro Asn Asn Val Val Arg Val Glu	
175 180 185	
gtg gag ggc ccg gcg ctg tcg gag gag gag ttg tac tcg ctg ttc cgg	1107

Val	Glu	Gly	Pro	Ala	Leu	Ser	Glu	Glu	Glu	Leu	Tyr	Ser	Leu	Phe	Arg	
			190					195					200			
cgg	tac	ggc	acg	att	ctg	gac	atc	tac	ccc	gcg	ggc	aag	aac	ggc	tat	1155
Arg	Tyr	Gly	Thr	Ile	Leu	Asp	Ile	Tyr	Pro	Ala	Gly	Lys	Asn	Gly	Tyr	
		205					210					215				
gcg	acg	atc	cgc	tac	cgc	tcg	ttc	cgg	ggc	gcg	atc	tgc	gcg	aag	aac	1203
Ala	Thr	Ile	Arg	Tyr	Arg	Ser	Phe	Arg	Gly	Ala	Ile	Cys	Ala	Lys	Asn	
		220				225					230					
tgt	gtg	tcg	ggg	atc	gag	atc	aac	ggc	tcg	acg	ctg	cat	gtc	aag	ttt	1251
Cys	Val	Ser	Gly	Ile	Glu	Ile	Asn	Gly	Ser	Thr	Leu	His	Val	Lys	Phe	
235					240					245					250	
gag	cct	gta	gtc	cgc	gcg	cac	gcg	atc	cgc	gac	ttc	ttc	gtg	aac	cac	1299
Glu	Pro	Val	Val	Arg	Ala	His	Ala	Ile	Arg	Asp	Phe	Phe	Val	Asn	His	
				255					260					265		
ccg	cgc	att	gcc	att	ccc	ctg	ctg	att	gcc	ctg	tta	tcc	atc	tgt	gct	1347
Pro	Arg	Ile	Ala	Ile	Pro	Leu	Leu	Ile	Ala	Leu	Leu	Ser	Ile	Cys	Ala	
			270					275					280			
gtg	ctg	atc	ttc	gac	cct	atc	cgg	gag	ttt	tcc	att	gag	cag	aag	atc	1395
Val	Leu	Ile	Phe	Asp	Pro	Ile	Arg	Glu	Phe	Ser	Ile	Glu	Gln	Lys	Ile	
		285					290					295				
acg	cgc	atg	tac	aca	ctg	tcg	cgg	gac	aac	ttc	gtg	gtg	aaa	agc	atc	1443
Thr	Arg	Met	Tyr	Thr	Leu	Ser	Arg	Asp	Asn	Phe	Val	Val	Lys	Ser	Ile	
		300				305					310					
ctg	cgc	ctc	acc	tcc	tat	acg	gtc	tca	tct	gtg	aag	cat	ctc	tgg	ggc	1491
Leu	Arg	Leu	Thr	Ser	Tyr	Thr	Val	Ser	Ser	Val	Lys	His	Leu	Trp	Gly	
315					320					325					330	
tac	gac	gat	gac	cag	cca	gag	aaa	cga	cag	ttg	tgg	cag	gag	cgc	gtg	1539
Tyr	Asp	Asp	Asp	Gln	Pro	Glu	Lys	Arg	Gln	Leu	Trp	Gln	Glu	Arg	Val	
				335				340						345		
gag	aag	gtg	aac	gac	ctg	aag	atg	tgg	tta	gag	gaa	aac	aac	aac	act	1587
Glu	Lys	Val	Asn	Asp	Leu	Lys	Met	Trp	Leu	Glu	Glu	Asn	Asn	Asn	Thr	
			350					355					360			
ttc	gtc	gtg	gtg	acc	ggc	cct	cgc	ggc	tct	ggc	aag	cat	gag	ctc	gtg	1635
Phe	Val	Val	Val	Thr	Gly	Pro	Arg	Gly	Ser	Gly	Lys	His	Glu	Leu	Val	
		365					370					375				
atg	cag	cac	act	ctg	cat	gac	cgt	cct	aat	gtg	cta	tat	ctg	gat	tgt	1683
Met	Gln	His	Thr	Leu	His	Asp	Arg	Pro	Asn	Val	Leu	Tyr	Leu	Asp	Cys	
		380				385					390					
gat	act	ttg	atc	aag	tcg	cgt	act	gat	tcc	aag	ttt	ttg	agg	aat	gcc	1731
Asp	Thr	Leu	Ile	Lys	Ser	Arg	Thr	Asp	Ser	Lys	Phe	Leu	Arg	Asn	Ala	
395					400				405						410	
gcc	cac	cag	att	ggc	tat	ttc	cca	ata	ttc	ccg	tgg	ttg	aat	tcc	gtc	1779
Ala	His	Gln	Ile	Gly	Tyr	Phe	Pro	Ile	Phe	Pro	Trp	Leu	Asn	Ser	Val	
				415					420					425		
acg	act	ctg	gtt	gac	ctg	gct	gtt	cag	ggc	ctg	acc	ggc	cag	aag	tcc	1827
Thr	Thr	Leu	Val	Asp	Leu	Ala	Val	Gln	Gly	Leu	Thr	Gly	Gln	Lys	Ser	
			430					435					440			

ggg cta tct gag tca aag gaa acg caa ttc cgg aat atg ctg aac act	1875
Gly Leu Ser Glu Ser Lys Glu Thr Gln Phe Arg Asn Met Leu Asn Thr	
445 450 455	
gct atg atg tct att agg cac att gct ttg tcc ggt tat aag gcg act	1923
Ala Met Met Ser Ile Arg His Ile Ala Leu Ser Gly Tyr Lys Ala Thr	
460 465 470	
ctg cac tct ggc gat gat gtt acg act gtg aag gag gag gac tac ttg	1971
Leu His Ser Gly Asp Asp Val Thr Thr Val Lys Glu Glu Asp Tyr Leu	
475 480 485 490	
caa cag cat cct gag cgc aag ccg gtt att gta ata gac agg ttc agt	2019
Gln Gln His Pro Glu Arg Lys Pro Val Ile Val Ile Asp Arg Phe Ser	
495 500 505	
aac aag gct gag ata aac ggt ttt gtg tac aaa gaa ctc gca gat tgg	2067
Asn Lys Ala Glu Ile Asn Gly Phe Val Tyr Lys Glu Leu Ala Asp Trp	
510 515 520	
gcg tcc atg ctt gtt cag atg aac atc gcg cat gtt ata ttc ctc acc	2115
Ala Ser Met Leu Val Gln Met Asn Ile Ala His Val Ile Phe Leu Thr	
525 530 535	
gag tct gta tct cca aac cag ttg ttg gca gag gct ctt cca aat caa	2163
Glu Ser Val Ser Pro Asn Gln Leu Leu Ala Glu Ala Leu Pro Asn Gln	
540 545 550	
gtt ttc aag ttc ctg ttt ctc tca gat gcg tcg aag gac tca gcg agg	2211
Val Phe Lys Phe Leu Phe Leu Ser Asp Ala Ser Lys Asp Ser Ala Arg	
555 560 565 570	
agc tac gtg ctt tca cag tta tac cct tct agc cca gcg tat tct gaa	2259
Ser Tyr Val Leu Ser Gln Leu Tyr Pro Ser Ser Pro Ala Tyr Ser Glu	
575 580 585	
aaa atg cca gca gca gat gca gat gcc aat gag gag tat agg aag gag	2307
Lys Met Pro Ala Ala Asp Ala Asp Ala Asn Glu Glu Tyr Arg Lys Glu	
590 595 600	
ata gac cgt gca ttg gag cct atc ggc ggt aga atg ttg gac cta caa	2355
Ile Asp Arg Ala Leu Glu Pro Ile Gly Gly Arg Met Leu Asp Leu Gln	
605 610 615	
gcg ttt gtc cgc aga gtg aaa tca ggg gag gag ccg tct gaa gca ttg	2403
Ala Phe Val Arg Arg Val Lys Ser Gly Glu Glu Pro Ser Glu Ala Leu	
620 625 630	
gaa aag atg gtg gag caa gca tca gag cag atc act caa ata ttc tta	2451
Glu Lys Met Val Glu Gln Ala Ser Glu Gln Ile Thr Gln Ile Phe Leu	
635 640 645 650	
agt gag aga tcg gag ccg att aaa acg gca caa gct tgg gag ctg atc	2499
Ser Glu Arg Ser Glu Pro Ile Lys Thr Ala Gln Ala Trp Glu Leu Ile	
655 660 665	
gaa cta ctt tca caa aat gat gtc gtg aag tat gga gat atc gtt ttc	2547
Glu Leu Leu Ser Gln Asn Asp Val Val Lys Tyr Gly Asp Ile Val Phe	
670 675 680	
agg ccg ttg ttt aag tct tct cca gag gcc ggg ttg ttg gag tta gag	2595

Arg	Pro	Leu	Phe	Lys	Ser	Ser	Pro	Glu	Ala	Gly	Leu	Leu	Glu	Leu	Glu	
		685					690					695				
aaa	aat	ggg	ctg	atc	aca	atc	agc	cgc	aat	aga	gga	gtg	ttg	caa	gat	2643
Lys	Asn	Gly	Leu	Ile	Thr	Ile	Ser	Arg	Asn	Arg	Gly	Val	Leu	Gln	Asp	
	700					705					710					
atc	cgg	ccc	gct	aaa	cca	ttg	ttt	aaa	gcc	gca	ttc	agc	tac	ctg	tta	2691
Ile	Arg	Pro	Ala	Lys	Pro	Leu	Phe	Lys	Ala	Ala	Phe	Ser	Tyr	Leu	Leu	
715					720				725						730	
cag	gac	aag	gac	ctg	tcc	atc	gtc	ctc	cga	aca	ggg	tac	tac	ttg	cga	2739
Gln	Asp	Lys	Asp	Leu	Ser	Ile	Val	Leu	Arg	Thr	Gly	Tyr	Tyr	Leu	Arg	
				735					740					745		
cta	atc	gcg	ttt	gaa	act	ggg	cgg	atc	aaa	aag	tgg	gag	gag	gaa	ttg	2787
Leu	Ile	Ala	Phe	Glu	Thr	Gly	Arg	Ile	Lys	Lys	Trp	Glu	Glu	Glu	Leu	
		750					755					760				
cgt	cta	ttg	gca	aag	gtt	agt	gac	caa	aga	ata	tgc	aag	agt	cgg	ttg	2835
Arg	Leu	Leu	Ala	Lys	Val	Ser	Asp	Gln	Arg	Ile	Cys	Lys	Ser	Arg	Leu	
	765						770					775				
aat	tac	tta	gca	agc	aaa	ata	gac	gcc	agt	agt	ggg	gta	att	aac	agt	2883
Asn	Tyr	Leu	Ala	Ser	Lys	Ile	Asp	Ala	Ser	Ser	Gly	Val	Ile	Asn	Ser	
	780					785					790					
tgt	gag	gac	aag	gta	aaa	gaa	atg	tca	aag	cgt	att	tagaaa	cgc			2929
Cys	Glu	Asp	Lys	Val	Lys	Glu	Met	Ser	Lys	Arg	Ile					
795					800					805						
cgcatcaata	agacccatt	ttgccacgat	ctggcaagat	gaactaaata	gcgataatca											2989
cactaggtct	cattagttta	gcattccagc	tgcacgggct	ttagtttagt	acctcaattc											3049
caaaaacggc	actatcttgt	ccgagcgtg	gtgaatatcc	ttgccgtatg	tgaaattatc											3109
cagggttgat	gtagttggcc	ggagctacca	gtaattgttg	gcctatgcac	gggccgctat											3169
tttccgtaac	ataatccaga	gatccgttaa	caggcaacat	gaactgattg	attaatttta											3229
atgaactttt	ttttttggct	ttgaaaggcc	taaaattgga	tcattctacat	ttcctgacga											3289
tataggatgg	taaaagccct	accaaccaag	aaattctgat	ccctttgaca	aatcccttat											3349
aacatagata	ttagaatata	atcaacccca	ttgtgtgaca	actcttctcc	aaaaataatt											3409
ggcttttcac	acaattaaag	attggcttca	aggtagctcc	acggtgagat	ctgctaaaac											3469
agtatggaag	ctgctaccat	taatgaagag	caagtcattc	ctgaaagga	tttggaag											3529
ctgccagtgt	cgagaaatgt	cagtagagga	tgggcagttc	gtaagttggc	aatattatac											3589
tttctattag	cgtctttcgc	tatagggacg	gtgcacttat	tccagtctgg	ttacatgggt											3649
gctaatagta	aatctcctca	gatcggtcgc	tggcaactat	ttgcggagaa	ttctcagggc											3709
cttacaaaac	ggagtagatg	ggcgatgac	gctgaaaaac	accacgcgcc	cctcttagca											3769
gcggcaagtc	tatcaacggc	agtggagcct	aatgtttgcc	cgctgatgg	cagaagatgc											3829
atgggcgaac	taatcaccag	attattactt	caatcactcg	cgatagctca	ggacctctat											3889

```

ggtggtgagt ggaagattac tcagccacat ccaccataca acaacaacaa acgcagtctg 3949
tcgaactatg catttgaaac tgcacggaa ctttcaactc aaataaagag tagagcagac 4009
actttggcgg cgtaaataat caatctggtg ggacgcttcc gcagcaatga cccaaatggc 4069
cttgcgaa 4077

```

```

<210> 13
<211> 806
<212> PRT
<213> 棉桃阿舒氏囊霉

```

```

<220>
<221> misc_feature
<223> Oligo 109

```

```

<400> 13

```

```

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

```

210					215					220					
Ser 225	Phe	Arg	Gly	Ala	Ile 230	Cys	Ala	Lys	Asn	Cys 235	Val	Ser	Gly	Ile	Glu 240
Ile	Asn	Gly	Ser	Thr 245	Leu	His	Val	Lys	Phe 250	Glu	Pro	Val	Val	Arg	Ala
His	Ala	Ile	Arg 260	Asp	Phe	Phe	Val	Asn 265	His	Pro	Arg	Ile	Ala 270	Ile	Pro
Leu	Leu	Ile 275	Ala	Leu	Leu	Ser	Ile 280	Cys	Ala	Val	Leu	Ile 285	Phe	Asp	Pro
Ile	Arg 290	Glu	Phe	Ser	Ile	Glu 295	Gln	Lys	Ile	Thr	Arg 300	Met	Tyr	Thr	Leu
Ser 305	Arg	Asp	Asn	Phe	Val 310	Val	Lys	Ser	Ile	Leu 315	Arg	Leu	Thr	Ser	Tyr 320
Thr	Val	Ser	Ser	Val 325	Lys	His	Leu	Trp	Gly 330	Tyr	Asp	Asp	Asp	Gln 335	Pro
Glu	Lys	Arg	Gln 340	Leu	Trp	Gln	Glu	Arg 345	Val	Glu	Lys	Val	Asn 350	Asp	Leu
Lys	Met 355	Trp	Leu	Glu	Glu	Asn 360	Asn	Asn	Thr	Phe	Val 365	Val	Val	Thr	Gly
Pro 370	Arg	Gly	Ser	Gly	Lys	His 375	Glu	Leu	Val	Met	Gln 380	His	Thr	Leu	His
Asp 385	Arg	Pro	Asn	Val	Leu 390	Tyr	Leu	Asp	Cys	Asp 395	Thr	Leu	Ile	Lys	Ser 400
Arg	Thr	Asp	Ser	Lys 405	Phe	Leu	Arg	Asn	Ala 410	Ala	His	Gln	Ile	Gly 415	Tyr
Phe	Pro	Ile	Phe 420	Pro	Trp	Leu	Asn	Ser 425	Val	Thr	Thr	Leu	Val 430	Asp	Leu
Ala	Val	Gln 435	Gly	Leu	Thr	Gly	Gln 440	Lys	Ser	Gly	Leu	Ser 445	Glu	Ser	Lys
Glu 450	Thr	Gln	Phe	Arg	Asn 455	Met	Leu	Asn	Thr	Ala	Met 460	Met	Ser	Ile	Arg
His 465	Ile	Ala	Leu	Ser	Gly 470	Tyr	Lys	Ala	Thr	Leu 475	His	Ser	Gly	Asp	Asp 480
Val	Thr	Thr	Val	Lys 485	Glu	Glu	Asp	Tyr	Leu 490	Gln	Gln	His	Pro	Glu 495	Arg
Lys	Pro	Val	Ile 500	Val	Ile	Asp	Arg	Phe 505	Ser	Asn	Lys	Ala	Glu 510	Ile	Asn
Gly	Phe 515	Val	Tyr	Lys	Glu	Leu	Ala 520	Asp	Trp	Ala	Ser 525	Met	Leu	Val	Gln
Met	Asn 530	Ile	Ala	His	Val	Ile 535	Phe	Leu	Thr	Glu	Ser 540	Val	Ser	Pro	Asn

Gln Leu Leu Ala Glu Ala Leu Pro Asn Gln Val Phe Lys Phe Leu Phe
 545 550 555 560
 Leu Ser Asp Ala Ser Lys Asp Ser Ala Arg Ser Tyr Val Leu Ser Gln
 565 570 575
 Leu Tyr Pro Ser Ser Pro Ala Tyr Ser Glu Lys Met Pro Ala Ala Asp
 580 585 590
 Ala Asp Ala Asn Glu Glu Tyr Arg Lys Glu Ile Asp Arg Ala Leu Glu
 595 600 605
 Pro Ile Gly Gly Arg Met Leu Asp Leu Gln Ala Phe Val Arg Arg Val
 610 615 620
 Lys Ser Gly Glu Glu Pro Ser Glu Ala Leu Glu Lys Met Val Glu Gln
 625 630 635 640
 Ala Ser Glu Gln Ile Thr Gln Ile Phe Leu Ser Glu Arg Ser Glu Pro
 645 650 655
 Ile Lys Thr Ala Gln Ala Trp Glu Leu Ile Glu Leu Leu Ser Gln Asn
 660 665 670
 Asp Val Val Lys Tyr Gly Asp Ile Val Phe Arg Pro Leu Phe Lys Ser
 675 680 685
 Ser Pro Glu Ala Gly Leu Leu Glu Leu Glu Lys Asn Gly Leu Ile Thr
 690 695 700
 Ile Ser Arg Asn Arg Gly Val Leu Gln Asp Ile Arg Pro Ala Lys Pro
 705 710 715 720
 Leu Phe Lys Ala Ala Phe Ser Tyr Leu Leu Gln Asp Lys Asp Leu Ser
 725 730 735
 Ile Val Leu Arg Thr Gly Tyr Tyr Leu Arg Leu Ile Ala Phe Glu Thr
 740 745 750
 Gly Arg Ile Lys Lys Trp Glu Glu Glu Leu Arg Leu Leu Ala Lys Val
 755 760 765
 Ser Asp Gln Arg Ile Cys Lys Ser Arg Leu Asn Tyr Leu Ala Ser Lys
 770 775 780
 Ile Asp Ala Ser Ser Gly Val Ile Asn Ser Cys Glu Asp Lys Val Lys
 785 790 795 800
 Glu Met Ser Lys Arg Ile
 805

<210> 14
 <211> 1337
 <212> DNA
 <213> 棉桃阿舒氏囊霉
 <220>
 <221> misc_feature
 <223> Oligo 163

```

<400> 14
gatccttgac ttgacacaga cagtggcagc ttacggctgg ttaccattca tcctttacat      60
gggctggtcg cacactgcc aactccccaa cctgctaaac cttctctccc cactcccaag      120
tgtctaattgt atttgcaacg gttccctcag tatgtgcttt tggtcacatg atcctcctag      180
ggcaaggccg ggacggaggg aacgatagcg cgaggctacc agcctcggtc ctaccacgtc      240
gagtgtgccg caattctagt gagcttgtat atacgtaatt tgaataaata agcgttcatt      300
gttaccttat ccgcctatga gttcgggaac tacctgttcg ctgcctttgt gaaaagtagc      360
tgacaatagc agcgggtgatg catcttccga aacagccata ggagagcaat gggccacgag      420
ctgacaagcc tgcagcgtga agtgatttcc ggactcacag caggaacaat cacaacaatc      480
gctagccacc cgctcgacct gcttaaactg cgactacaat tatcagcggg caacagggct      540
aacactacat atacaggact aatcagagac atatttgaac gccagcaatg gggacgagag      600
ctatatcgtg gtttaggcgt taacctgctt ggaaattcag tggcatgggc gctttacttt      660
ggctgttacc gatgcgcaaa ggacattgcg cttcgacact tgggcaatga gtctgctacg      720
ggatatcatgg accgccgcct gccagcgcac gcatacatgc ttgcagctgg tagcagcggg      780
attgcgacgg cagtccttac aaaccccata tgggtcatta agactcgcat aatggccact      840
tctcgtgctg gaccttacia gtgcagcttt gacgggggttt ataagttata tcaaactgaa      900
gggtgttctgg cattctggcg ggggtgttgt ccatcgctac tcggcgtctc gcaaggagct      960
atctactttg cgctgtatga tacattgaaa ttccattacc tgcactctag tactgacaag     1020
gccgagcgaa ggttgtcggt ttccgagatc atcggcataa catgtatctc taaaatgatt     1080
tccgtcacat cgggtctatcc atttcagctg ctgaaatcta agctgcagga ttttgggtgca     1140
ccatccggca tcacccaact tgttcagact gtttacagta gagaggggat cagaggcttc     1200
tacaggggccc taccctgcta atctcctgcg ggcagttcct gctacatgca taaccttttt     1260
cgtgtacgaa aatatcaaat atcgtctcta aatttgtgat ttggtgcagc gcgtagcaag     1320
tatacactga ctggatc                                     1337

```

```

<210> 15
<211> 267
<212> PRT
<213> 棉桃阿舒氏囊霉

```

```

<220>
<221> misc_feature
<223> Oligo 163

```

```

<400> 15

```

```

His Glu Leu Thr Ser Leu Gln Arg Glu Val Ile Ser Gly Leu Thr Ala
1           5           10           15

```

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

<210> 16

<211> 47

<212> PRT

<213> 棉桃阿舒氏囊霉

<220>

<221> misc_feature

<223> Oligo 163

<400> 16

Ala Ser Pro Asn Leu Phe Arg Leu Phe Thr Val Glu Arg Val Ser Glu
 1 5 10 15

Ala Ser Thr Gly Ala Tyr Pro Ala Asn Leu Leu Arg Ala Val Pro Ala
20 25 30

Thr Cys Ile Thr Phe Phe Val Tyr Glu Asn Ile Lys Tyr Arg Leu
35 40 45

<210> 17

<211> 1733

<212> DNA

<213> 棉桃阿舒氏囊霉

<220>

<221> CDS

<222> (329)..(1207)

<223>

<220>

<221> misc_feature

<223> Oligo 163

<400> 17

actcccccaa cctgataaac cttctctccc cactcccaag tgtctaattgt atttgcaacg 60

gttccctcag tatgtgcttt tgttcacatg atcctcctag ggcaaggccg ggacggaggg 120

aacgatagcg cgaggctacc agcctcggtc ctaccacgtc gagtgtgccg caattctagt 180

gagcttgtat atacagtaat ttgaataaat aagcgttcat tgttacctta tccgcctatg 240

agttcgggaa ctactgttcg ctgcctttgt gaaaagtagc tgacaatagc agcggtagtg 300

catcttccga aacagccata ggagagca atg ggc cac gag ctg aca agc ctg 352
Met Gly His Glu Leu Thr Ser Leu
1 5

cag cgt gaa gtg att tcc gga ctc aca gca gga aca atc aca aca atc 400
Gln Arg Glu Val Ile Ser Gly Leu Thr Ala Gly Thr Ile Thr Thr Ile
10 15 20

gct agc cac ccg ctc gac ctg ctt aaa ctg cga cta caa tta tca gcg 448
Ala Ser His Pro Leu Asp Leu Leu Lys Leu Arg Leu Gln Leu Ser Ala
25 30 35 40

ggc aac agg gct aac act aca tat aca gga cta atc aga gac ata ttt 496
Gly Asn Arg Ala Asn Thr Thr Tyr Thr Gly Leu Ile Arg Asp Ile Phe
45 50 55

gaa cgc cag caa tgg gga cga gag cta tat cgt ggt tta ggc gtt aac 544
Glu Arg Gln Gln Trp Gly Arg Glu Leu Tyr Arg Gly Leu Gly Val Asn
60 65 70

ctg ctt gga aat tca gtg gca tgg gcg ctt tac ttt ggc tgt tac cga 592
Leu Leu Gly Asn Ser Val Ala Trp Ala Leu Tyr Phe Gly Cys Tyr Arg
75 80 85

tgc gca aag gac att gcg ctt cga cac ttg ggc aat gag tct gct acg 640
Cys Ala Lys Asp Ile Ala Leu Arg His Leu Gly Asn Glu Ser Ala Thr
90 95 100

ggg atc atg gac cgc cgc ctg cca gcg cac gca tac atg ctt gca gct 688

Gly 105	Ile	Met	Asp	Arg	Arg 110	Leu	Pro	Ala	His	Ala 115	Tyr	Met	Leu	Ala	Ala 120	
ggt	agc	agc	ggg	att	gcg	acg	gca	gtc	ctt	aca	aac	ccc	ata	tgg	gtc	736
Gly	Ser	Ser	Gly	Ile 125	Ala	Thr	Ala	Val	Leu 130	Thr	Asn	Pro	Ile	Trp	Val 135	
att	aag	act	cgc	ata	atg	gcc	act	tct	cgt	gct	gga	cct	tac	aag	tcg	784
Ile	Lys	Thr	Arg 140	Ile	Met	Ala	Thr	Ser 145	Arg	Ala	Gly	Pro	Tyr 150	Lys	Ser	
acg	ttt	gac	ggg	gtt	tat	aag	tta	tat	caa	act	gaa	ggt	gtt	ctg	gca	832
Thr	Phe	Asp 155	Gly	Val	Tyr	Lys	Leu 160	Tyr	Gln	Thr	Glu	Gly 165	Val	Leu	Ala	
ttc	tgg	cgg	ggt	gtt	gtt	cca	tcg	cta	ctc	ggc	gtc	tcg	caa	gga	gct	880
Phe	Trp	Arg	Gly	Val	Val	Pro	Ser 175	Leu	Leu	Gly	Val 180	Ser	Gln	Gly	Ala	
atc	tac	ttt	gcg	ctg	tat	gat	aca	ttg	aaa	ttc	cat	tac	ctg	cac	tct	928
Ile	Tyr	Phe	Ala	Leu	Tyr 190	Asp	Thr	Leu	Lys	Phe 195	His	Tyr	Leu	His	Ser 200	
agt	act	gac	aag	gcc	gag	cga	agg	ttg	tcg	gtt	tcc	gag	atc	atc	ggc	976
Ser	Thr	Asp	Lys	Ala 205	Glu	Arg	Arg	Leu	Ser 210	Val	Ser	Glu	Ile	Ile	Gly 215	
ata	aca	tgt	atc	tct	aaa	atg	att	tcc	gtc	aca	tcg	gtc	tat	cca	ttt	1024
Ile	Thr	Cys	Ile 220	Ser	Lys	Met	Ile	Ser 225	Val	Thr	Ser	Val	Tyr 230	Pro	Phe	
cag	ctg	ctg	aaa	tct	aag	ctg	cag	gat	ttt	ggt	gca	cca	tcc	ggc	atc	1072
Gln	Leu	Leu	Lys	Ser	Lys	Leu	Gln	Asp	Phe	Gly	Ala	Pro	Ser	Gly	Ile	
acc	caa	ctt	gtt	cag	act	gtt	tac	agt	aga	gag	ggt	atc	aga	ggc	ttc	1120
Thr	Gln	Leu	Val	Gln	Thr	Val 255	Tyr	Ser	Arg	Glu	Gly 260	Ile	Arg	Gly	Phe	
tac	agg	ggc	cta	tcc	gct	aat	ctc	ctg	cgg	gca	gtt	cct	gct	aca	tgc	1168
Tyr	Arg	Gly	Leu	Ser	Ala	Asn 270	Leu	Leu	Arg	Ala 275	Val	Pro	Ala	Thr	Cys 280	
ata	acc	ttt	ttc	gtg	tac	gaa	aat	atc	aaa	tat	cgt	ctc	taaatttg	tg		1217
Ile	Thr	Phe	Phe	Val	Tyr	Glu	Asn	Ile	Lys	Tyr	Arg	Leu				
				285				290								
atttggtgca	gcgcgtagca	agtatatact	gactggatct	aaaagatgat	taatttgct											1277
taagctttac	aaaaaaacac	tttatatatt	gattgatatt	gggcaaaaga	ttcatgacct											1337
tatgatgaaa	tttagtaacc	gagcttgccc	aaatcctcag	cgatctcgga	cttagcatca											1397
tatttagcct	tgacagcagc	agccgttttg	ttgaaggccc	tcttcttgga	gtagtactca											1457
gcagatctga	cctttctctt	ctcttccaac	ttggcaacaa	cgttctcgta	cttccagcca											1517
acggtggaag	acaacttacc	caaggtgggtg	aactttctac	ctggcttcaa	tctcaaaact											1577
ctaaagcctg	tggacaaca	actctcttca	acttgctgta	tggaggtggg	acaccctcta											1637
agaccttta	tctctccata	gcagccttac	cacggggcag	tcttgtaggc	aaccatacct											1697

ctgacagcct tgtagaagaa tctggatggg gctctg

1733

<210> 18

<211> 293

<212> PRT

<213> 棉桃阿舒氏囊霉

<220>

<221> misc_feature

<223> Oligo 163

<400> 18

Met Gly His Glu Leu Thr Ser Leu Gln Arg Glu Val Ile Ser Gly Leu
1 5 10 15

Thr Ala Gly Thr Ile Thr Thr Ile Ala Ser His Pro Leu Asp Leu Leu
20 25 30

Lys Leu Arg Leu Gln Leu Ser Ala Gly Asn Arg Ala Asn Thr Thr Tyr
35 40 45

Thr Gly Leu Ile Arg Asp Ile Phe Glu Arg Gln Gln Trp Gly Arg Glu
50 55 60

Leu Tyr Arg Gly Leu Gly Val Asn Leu Leu Gly Asn Ser Val Ala Trp
65 70 75 80

Ala Leu Tyr Phe Gly Cys Tyr Arg Cys Ala Lys Asp Ile Ala Leu Arg
85 90 95

His Leu Gly Asn Glu Ser Ala Thr Gly Ile Met Asp Arg Arg Leu Pro
100 105 110

Ala His Ala Tyr Met Leu Ala Ala Gly Ser Ser Gly Ile Ala Thr Ala
115 120 125

Val Leu Thr Asn Pro Ile Trp Val Ile Lys Thr Arg Ile Met Ala Thr
130 135 140

Ser Arg Ala Gly Pro Tyr Lys Ser Thr Phe Asp Gly Val Tyr Lys Leu
145 150 155 160

Tyr Gln Thr Glu Gly Val Leu Ala Phe Trp Arg Gly Val Val Pro Ser
165 170 175

Leu Leu Gly Val Ser Gln Gly Ala Ile Tyr Phe Ala Leu Tyr Asp Thr
180 185 190

Leu Lys Phe His Tyr Leu His Ser Ser Thr Asp Lys Ala Glu Arg Arg
195 200 205

Leu Ser Val Ser Glu Ile Ile Gly Ile Thr Cys Ile Ser Lys Met Ile
210 215 220

Ser Val Thr Ser Val Tyr Pro Phe Gln Leu Leu Lys Ser Lys Leu Gln
225 230 235 240

Asp Phe Gly Ala Pro Ser Gly Ile Thr Gln Leu Val Gln Thr Val Tyr
245 250 255

Ser Arg Glu Gly Ile Arg Gly Phe Tyr Arg Gly Leu Ser Ala Asn Leu
 260 265 270

Leu Arg Ala Val Pro Ala Thr Cys Ile Thr Phe Phe Val Tyr Glu Asn
 275 280 285

Ile Lys Tyr Arg Leu
 290

<210> 19
 <211> 972
 <212> DNA
 <213> 棉桃阿舒氏囊霉

<220>
 <221> misc_feature
 <223> Oligo 31

<400> 19
 gatcgatatt gtcggcaact tcattcctgg attaatatatt atgcaatcaa tatttgggta 60
 cctatcgtgg gcaattatatt acaagtgggc aaaagattgg attaaagatg agctgcctgc 120
 gcctgggttta ttaaacaatgt tgattaacat gtttttatct cccggagttg tcgatgaaaa 180
 actatataca ggccagagct ttcttcaagt tatccttttg ctagctgctc tagtttgtgt 240
 tccgtggcta ctactttaca aaccgttgat gttgaagcgc cagaacgata tagctctcag 300
 caaaggattc agaagcctca gagaccaacg ggtgcatgag attcttcttg aagcacagga 360
 aaacgcaggc gaagatatgc tggttgcaga ttatgaaaat gaagatgaat cgtcggagga 420
 gttcaacttt ggtgacgtta tgatacatca gggtatccac accattgaat tttgtttgaa 480
 ttgtatatct catactgcat cgtatctcag attgtgggccc ttatctcttg cgcatgctca 540
 actctcgact gtggtgtggt ctatgaccat tcagaattcg ttctccgact ccaaccctgg 600
 ttcattttttc tctgtcacca aggtggtggt tttgtttgcc atgtgggttg tgttgactgt 660
 ttgcatttta gtcttaatgg aaggaacgtc ggctatgttg cactcggtta gattgcattg 720
 ggtggaagcg atgtccaaat tcttcgaggg cgaaggctat gcttacgaac cattttcctt 780
 caaggctatc aacagcgatg atgaatagta tcgtatatat taaaactaga cagatggagg 840
 tagtgtcaca tgtgtcaggt atagattgcc cgatgaataa gcataccttga aatttagaaa 900
 aatcataggc ctttagaatc gacatacaac gtctaaatat atttactatt cacttcatag 960
 ttcacgtgga tc 972

<210> 20
 <211> 263
 <212> PRT
 <213> 棉桃阿舒氏囊霉

```

<220>
<221> misc_feature
<223> Oligo 31

<400> 20

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

Ile Phe Gly Tyr Leu Ser Trp Ala Ile Ile Tyr Lys Trp Ser Lys Asp
          20          25          30

Trp Ile Lys Asp Glu Leu Pro Ala Pro Gly Leu Leu Asn Met Leu Ile
          35          40          45

Asn Met Phe Leu Ser Pro Gly Val Val Asp Glu Lys Leu Tyr Thr Gly
          50          55          60

Gln Ser Phe Leu Gln Val Ile Leu Leu Leu Ala Ala Leu Val Cys Val
          65          70          75          80

Pro Trp Leu Leu Leu Tyr Lys Pro Leu Met Leu Lys Arg Gln Asn Asp
          85          90          95

Ile Ala Leu Ser Lys Gly Phe Arg Ser Leu Arg Asp Gln Arg Val His
          100          105          110

Glu Ile Leu Leu Glu Ala Gln Glu Asn Ala Gly Glu Asp Met Leu Val
          115          120          125

Ala Asp Tyr Glu Asn Glu Asp Glu Ser Ser Glu Glu Phe Asn Phe Gly
          130          135          140

Asp Val Met Ile His Gln Val Ile His Thr Ile Glu Phe Cys Leu Asn
          145          150          155          160

Cys Ile Ser His Thr Ala Ser Tyr Leu Arg Leu Trp Ala Leu Ser Leu
          165          170          175

Ala His Ala Gln Leu Ser Thr Val Leu Trp Ser Met Thr Ile Gln Asn
          180          185          190

Ser Phe Ser Asp Ser Asn Pro Gly Ser Phe Phe Ser Val Thr Lys Val
          195          200          205

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

Leu Met Glu Gly Thr Ser Ala Met Leu His Ser Leu Arg Leu His Trp
          225          230          235          240

Val Glu Ala Met Ser Lys Phe Phe Glu Gly Glu Gly Tyr Ala Tyr Glu
          245          250          255

Pro Phe Ser Phe Lys Ala Ile
          260

<210> 21
<211> 4054
<212> DNA
<213> 棉桃阿舒氏囊霉

```


gat tcg gtg gac gag gtt gcg gag gaa atc atg cag ttc gag ggg cgc Asp Ser Val Asp Glu Val Ala Glu Glu Ile Met Gln Phe Glu Gly Arg 125 130 135	1036
gtg agg cag ctc gac cag agt ctt gta gcg atg cgc gag cga ctg tcg Val Arg Gln Leu Asp Gln Ser Leu Val Ala Met Arg Glu Arg Leu Ser 140 145 150	1084
aag ctt gtg cat gag cga cgt gtg atg ttc act tgc gaa cac ttt ctc Lys Leu Val His Glu Arg Arg Val Met Phe Thr Cys Glu His Phe Leu 155 160 165 170	1132
gaa gtg aac ccc ggg atc ggg gag cga ttg ccc aca cgc ccc gcg ggg Glu Val Asn Pro Gly Ile Gly Glu Arg Leu Pro Thr Arg Pro Ala Gly 175 180 185	1180
ttc gaa gct gat gag ttt gaa ctc acg cgt gtc ggg gag gag gac gag Phe Glu Ala Asp Glu Phe Glu Leu Thr Arg Val Gly Glu Glu Asp Glu 190 195 200	1228
gag acg gca agc cag ctg tct ttt gat att tcg gat gac gca gag aca Glu Thr Ala Ser Gln Leu Ser Phe Asp Ile Ser Asp Asp Ala Glu Thr 205 210 215	1276
cag ctg ccc ggg gac atg cgc acg tta ctc gaa ccg gtg tac cgg cat Gln Leu Pro Gly Asp Met Arg Thr Leu Leu Glu Pro Val Tyr Arg His 220 225 230	1324
cag tac cta ctt aca ggc tca att gag cgt gct aag gta gag gcg ctc Gln Tyr Leu Leu Thr Gly Ser Ile Glu Arg Ala Lys Val Glu Ala Leu 235 240 245 250	1372
aac aag atc ttg tgg cgc ctt ctc cgt ggg aat gtt ttc ttc cag aat Asn Lys Ile Leu Trp Arg Leu Leu Arg Gly Asn Val Phe Phe Gln Asn 255 260 265	1420
ttc cct gtc tca gtg tca ccg gtg gaa gaa gat gat acc gat ctc gag Phe Pro Val Ser Val Ser Pro Val Glu Glu Asp Asp Thr Asp Leu Glu 270 275 280	1468
act gac tgc ttt atc gtc ttt acg cat ggc gaa gtc ttg cta agc aag Thr Asp Cys Phe Ile Val Phe Thr His Gly Glu Val Leu Leu Ser Lys 285 290 295	1516
gcg aag aaa gtg ata gaa tcc cta aat ggc aca ata tat ccg ttt atg Ala Lys Lys Val Ile Glu Ser Leu Asn Gly Thr Ile Tyr Pro Phe Met 300 305 310	1564
caa gac ggg gcg aca gtg cag gag ttg aac gac aag ata gcg gac ctt Gln Asp Gly Ala Thr Val Gln Glu Leu Asn Asp Lys Ile Ala Asp Leu 315 320 325 330	1612
aag cag ata tgt tct acg aca gaa cag aca cta cat acg gaa ctt ttt Lys Gln Ile Cys Ser Thr Thr Glu Gln Thr Leu His Thr Glu Leu Phe 335 340 345	1660
ctt gtt gcc aac caa ttg ccc atg tgg aat gcc att att aag cgt gaa Leu Val Ala Asn Gln Leu Pro Met Trp Asn Ala Ile Ile Lys Arg Glu 350 355 360	1708
aag tac atc tac tct gcc ttg aat tta ttt agg cag gag tca cag ggc	1756

Lys	Tyr	Ile	Tyr	Ser	Ala	Leu	Asn	Leu	Phe	Arg	Gln	Glu	Ser	Gln	Gly	
		365					370					375				
ctc	ggt	gca	gag	gga	tgg	ctt	ccc	acg	tac	gat	cta	cca	gga	ggt	cag	1804
Leu	Val	Ala	Glu	Gly	Trp	Leu	Pro	Thr	Tyr	Asp	Leu	Pro	Gly	Val	Gln	
	380					385					390					
gcg	gca	cta	aag	gac	tat	ggg	gag	agc	gta	gga	tcc	gca	aat	tcg	gcc	1852
Ala	Ala	Leu	Lys	Asp	Tyr	Gly	Glu	Ser	Val	Gly	Ser	Ala	Asn	Ser	Ala	
395					400					405					410	
ggt	ttg	aat	gta	att	tcg	aca	acg	agg	aca	ccg	cca	act	ttc	cac	agg	1900
Val	Leu	Asn	Val	Ile	Ser	Thr	Thr	Arg	Thr	Pro	Pro	Thr	Phe	His	Arg	
				415					420					425		
act	aac	aag	ttt	acg	cag	gca	ttc	cag	tcc	att	ggt	gat	gcc	tat	ggc	1948
Thr	Asn	Lys	Phe	Thr	Gln	Ala	Phe	Gln	Ser	Ile	Val	Asp	Ala	Tyr	Gly	
			430					435					440			
att	gca	aca	tat	aaa	gaa	gtg	aac	ccc	ggg	ttg	gca	act	atc	gtc	aca	1996
Ile	Ala	Thr	Tyr	Lys	Glu	Val	Asn	Pro	Gly	Leu	Ala	Thr	Ile	Val	Thr	
		445					450					455				
ttt	ccc	ttt	atg	ttt	gcc	ggt	atg	ttt	ggc	gat	gcc	gga	cac	ggc	gca	2044
Phe	Pro	Phe	Met	Phe	Ala	Val	Met	Phe	Gly	Asp	Ala	Gly	His	Gly	Ala	
	460					465					470					
ttg	atg	ctc	ata	gcc	gcg	ctc	tat	ttg	gtg	tta	aat	gaa	aag	aag	ttg	2092
Leu	Met	Leu	Ile	Ala	Ala	Leu	Tyr	Leu	Val	Leu	Asn	Glu	Lys	Lys	Leu	
475					480					485					490	
gga	gca	atg	aaa	cga	ggc	gag	att	ttc	gac	atg	gca	tac	act	gga	aga	2140
Gly	Ala	Met	Lys	Arg	Gly	Glu	Ile	Phe	Asp	Met	Ala	Tyr	Thr	Gly	Arg	
				495					500					505		
tat	ggt	att	cta	ctg	atg	gga	atc	ttt	tct	atc	tat	acc	ggc	ata	atg	2188
Tyr	Val	Ile	Leu	Leu	Met	Gly	Ile	Phe	Ser	Ile	Tyr	Thr	Gly	Ile	Met	
			510					515					520			
tac	aat	gat	att	ttt	tcc	aag	tcc	atg	cat	ttg	ttt	tcc	act	ggc	tggt	2236
Tyr	Asn	Asp	Ile	Phe	Ser	Lys	Ser	Met	His	Leu	Phe	Ser	Thr	Gly	Trp	
		525					530					535				
aaa	tgg	cct	tca	aac	ttt	caa	gag	ggc	gag	atg	att	gag	gct	caa	aaa	2284
Lys	Trp	Pro	Ser	Asn	Phe	Gln	Glu	Gly	Glu	Met	Ile	Glu	Ala	Gln	Lys	
	540					545					550					
ggt	ggc	ggt	tac	cca	ttt	gga	ttg	gac	tat	gcc	tggt	cac	ggc	tcg	gac	2332
Val	Gly	Val	Tyr	Pro	Phe	Gly	Leu	Asp	Tyr	Ala	Trp	His	Gly	Ser	Asp	
555					560					565					570	
aat	agt	ttg	tta	ttc	acc	aat	tca	tat	aaa	atg	aag	tta	tca	atc	ctt	2380
Asn	Ser	Leu	Leu	Phe	Thr	Asn	Ser	Tyr	Lys	Met	Lys	Leu	Ser	Ile	Leu	
				575					580					585		
cta	ggc	ttt	atc	cac	atg	tcg	tat	tca	tat	att	ttc	tca	tat	ctc	aac	2428
Leu	Gly	Phe	Ile	His	Met	Ser	Tyr	Ser	Tyr	Ile	Phe	Ser	Tyr	Leu	Asn	
			590					595					600			
tac	cac	tat	aaa	ggc	tcg	agg	atc	gat	att	gtc	ggc	aac	ttc	att	cct	2476
Tyr	His	Tyr	Lys	Gly	Ser	Arg	Ile	Asp	Ile	Val	Gly	Asn	Phe	Ile	Pro	
		605					610					615				

gga tta ata ttt atg caa tca ata ttt ggg tac cta tcg tgg gca att Gly Leu Ile Phe Met Gln Ser Ile Phe Gly Tyr Leu Ser Trp Ala Ile 620 625 630	2524
att tac aag tgg tca aaa gat tgg att aaa gat gag ctg cct gcg cct Ile Tyr Lys Trp Ser Lys Asp Trp Ile Lys Asp Glu Leu Pro Ala Pro 635 640 645 650	2572
ggg tta tta aac atg ttg att aac atg ttt tta tct ccc gga gtt gtc Gly Leu Leu Asn Met Leu Ile Asn Met Phe Leu Ser Pro Gly Val Val 655 660 665	2620
gat gaa aaa cta tat aca ggc cag agc ttt ctt caa gtt atc ctt ttg Asp Glu Lys Leu Tyr Thr Gly Gln Ser Phe Leu Gln Val Ile Leu Leu 670 675 680	2668
cta gct gct cta gtt tgt gtt ccg tgg cta cta ctt tac aaa ccg ttg Leu Ala Ala Leu Val Cys Val Pro Trp Leu Leu Leu Tyr Lys Pro Leu 685 690 695	2716
atg ttg aag cgc cag aac gat ata gct ctc agc aaa gga ttc aga agc Met Leu Lys Arg Gln Asn Asp Ile Ala Leu Ser Lys Gly Phe Arg Ser 700 705 710	2764
ctc aga gac caa cgg gtg cat gag att ctt ctg gaa gca cag gaa aac Leu Arg Asp Gln Arg Val His Glu Ile Leu Leu Glu Ala Gln Glu Asn 715 720 725 730	2812
gca ggc gaa gat atg ctg gtt gca gat tat gaa aat gaa gat gaa tcg Ala Gly Glu Asp Met Leu Val Ala Asp Tyr Glu Asn Glu Asp Glu Ser 735 740 745	2860
tcg gag gag ttc aac ttt ggt gac gtt atg ata cat cag gtt atc cac Ser Glu Glu Phe Asn Phe Gly Asp Val Met Ile His Gln Val Ile His 750 755 760	2908
acc att gaa ttt tgt ttg aat tgt ata tct cat act gca tcg tat ctc Thr Ile Glu Phe Cys Leu Asn Cys Ile Ser His Thr Ala Ser Tyr Leu 765 770 775	2956
aga ttg tgg gcc tta tct ctt gcg cat gct caa ctc tcg act gtg ttg Arg Leu Trp Ala Leu Ser Leu Ala His Ala Gln Leu Ser Thr Val Leu 780 785 790	3004
tgg tct atg acc att cag aat tcg ttc tcc gac tcc aac cct ggt tca Trp Ser Met Thr Ile Gln Asn Ser Phe Ser Asp Ser Asn Pro Gly Ser 795 800 805 810	3052
ttt ttc tct gtc acc aag gtg gtg gtt ttg ttt gcc atg tgg ttt gtg Phe Phe Ser Val Thr Lys Val Val Val Leu Phe Ala Met Trp Phe Val 815 820 825	3100
ttg act gtt tgc att tta gtc tta atg gaa gga acg tcg gct atg ttg Leu Thr Val Cys Ile Leu Val Leu Met Glu Gly Thr Ser Ala Met Leu 830 835 840	3148
cac tcg tta aga ttg cat tgg gtg gaa gcg atg tcc aaa ttc ttc gag His Ser Leu Arg Leu His Trp Val Glu Ala Met Ser Lys Phe Phe Glu 845 850 855	3196
ggc gaa ggc tat gct tac gaa cca ttt tcc ttc aag gct atc aac agc	3244

Gly Glu Gly Tyr Ala Tyr Glu Pro Phe Ser Phe Lys Ala Ile Asn Ser
 860 865 870

gat gat gaa tagtatcgta tatattaaaa ctagacagat ggaggtagtg 3293
 Asp Asp Glu
 875

tcacatgtgt caggtataga ttgcccgatg aataagcatc cttgaaattt agaaaaatca 3353
 taggccttta gaatcgacat acaacgtcta aatatattta ctattcactt catagttcat 3413
 cgtgatcgat gtccgaagtg gaactatctg cgctagaggt ggatgcatct gctggttggt 3473
 gcgtagaggt cgcttttgct tcactcgca agtcagtagt agatgaagag gctgcttctt 3533
 tgcttttctc ttctttctc tttgccgctc ttagttttct tgtgtataat gagttagct 3593
 cacgcaagcg gtatgcatgt gctctttcgc ggaatttttg catttctcgc atggaattta 3653
 gatgaccttc ggatttcatg aacaatttta ctaactgttc gcggtcaatt gtatcaaag 3713
 catctggagt taatatgggc tcgagttcgc tcagtatctt accgagtgcg gccaattctg 3773
 actcagcaaa ggttgaattg tactgtaaat tctttggaac cttcagttta gtatattctc 3833
 tagtgacatt taggccaaga gagttaaagt taccctcaag ttcattagct tgtgccactg 3893
 cagtatcatc aaatgcgtta aaggtagaaa tggcttcaga aacatttttg cgtagaagct 3953
 caaatactgg gagatccaaa accttagcgt ctatggaaag gtaatggcta agcaacttat 4013
 acagactcat ggtttcctgg aatctgtggt ggatatcgcc a 4054

<210> 22
 <211> 877
 <212> PRT
 <213> 棉桃阿舒氏囊霉

<220>
 <221> misc_feature
 <223> Oligo 31

<400> 22

Met Phe Leu Ser Met Arg Ala Glu Glu Ala Met Phe Arg Ser Ala Asp
 1 5 10 15

Met Thr Tyr Ile Glu Leu Tyr Ile Pro Leu Glu Ile Ala Arg Glu Val
 20 25 30

Val Cys Val Leu Gly Asn Leu Gly Ser Val Met Leu Lys Asp Met Asn
 35 40 45

Lys Asp Leu Ser Thr Phe Gln Arg Gly Tyr Val Asn Gln Val Arg Arg
 50 55 60

Phe Asp Glu Val Glu Arg Gln Val Gly Tyr Met Glu Gly Val Val Arg
 65 70 75 80

Arg His Lys Asn Glu Thr Trp Arg Tyr Leu Tyr Arg His Leu Gln Arg
 85 90 95

Glu Glu Gln Gln Glu Tyr Pro Gly Arg Glu His Pro Thr Leu Ala Gln
 100 105 110
 Leu Ile Gly Ser Met His Thr His Ser Ile Asp Ser Val Asp Glu Val
 115 120 125
 Ala Glu Glu Ile Met Gln Phe Glu Gly Arg Val Arg Gln Leu Asp Gln
 130 135 140
 Ser Leu Val Ala Met Arg Glu Arg Leu Ser Lys Leu Val His Glu Arg
 145 150 155 160
 Arg Val Met Phe Thr Cys Glu His Phe Leu Glu Val Asn Pro Gly Ile
 165 170 175
 Gly Glu Arg Leu Pro Thr Arg Pro Ala Gly Phe Glu Ala Asp Glu Phe
 180 185 190
 Glu Leu Thr Arg Val Gly Glu Glu Asp Glu Glu Thr Ala Ser Gln Leu
 195 200 205
 Ser Phe Asp Ile Ser Asp Asp Ala Glu Thr Gln Leu Pro Gly Asp Met
 210 215 220
 Arg Thr Leu Leu Glu Pro Val Tyr Arg His Gln Tyr Leu Leu Thr Gly
 225 230 235 240
 Ser Ile Glu Arg Ala Lys Val Glu Ala Leu Asn Lys Ile Leu Trp Arg
 245 250 255
 Leu Leu Arg Gly Asn Val Phe Phe Gln Asn Phe Pro Val Ser Val Ser
 260 265 270
 Pro Val Glu Glu Asp Asp Thr Asp Leu Glu Thr Asp Cys Phe Ile Val
 275 280 285
 Phe Thr His Gly Glu Val Leu Leu Ser Lys Ala Lys Lys Val Ile Glu
 290 295 300
 Ser Leu Asn Gly Thr Ile Tyr Pro Phe Met Gln Asp Gly Ala Thr Val
 305 310 315 320
 Gln Glu Leu Asn Asp Lys Ile Ala Asp Leu Lys Gln Ile Cys Ser Thr
 325 330 335
 Thr Glu Gln Thr Leu His Thr Glu Leu Phe Leu Val Ala Asn Gln Leu
 340 345 350
 Pro Met Trp Asn Ala Ile Ile Lys Arg Glu Lys Tyr Ile Tyr Ser Ala
 355 360 365
 Leu Asn Leu Phe Arg Gln Glu Ser Gln Gly Leu Val Ala Glu Gly Trp
 370 375 380
 Leu Pro Thr Tyr Asp Leu Pro Gly Val Gln Ala Ala Leu Lys Asp Tyr
 385 390 395 400
 Gly Glu Ser Val Gly Ser Ala Asn Ser Ala Val Leu Asn Val Ile Ser
 405 410 415
 Thr Thr Arg Thr Pro Pro Thr Phe His Arg Thr Asn Lys Phe Thr Gln

420					425					430					
Ala	Phe	Gln	Ser	Ile	Val	Asp	Ala	Tyr	Gly	Ile	Ala	Thr	Tyr	Lys	Glu
		435					440					445			
Val	Asn	Pro	Gly	Leu	Ala	Thr	Ile	Val	Thr	Phe	Pro	Phe	Met	Phe	Ala
	450					455					460				
Val	Met	Phe	Gly	Asp	Ala	Gly	His	Gly	Ala	Leu	Met	Leu	Ile	Ala	Ala
465					470					475					480
Leu	Tyr	Leu	Val	Leu	Asn	Glu	Lys	Lys	Leu	Gly	Ala	Met	Lys	Arg	Gly
				485					490					495	
Glu	Ile	Phe	Asp	Met	Ala	Tyr	Thr	Gly	Arg	Tyr	Val	Ile	Leu	Leu	Met
			500					505					510		
Gly	Ile	Phe	Ser	Ile	Tyr	Thr	Gly	Ile	Met	Tyr	Asn	Asp	Ile	Phe	Ser
		515					520					525			
Lys	Ser	Met	His	Leu	Phe	Ser	Thr	Gly	Trp	Lys	Trp	Pro	Ser	Asn	Phe
	530					535					540				
Gln	Glu	Gly	Glu	Met	Ile	Glu	Ala	Gln	Lys	Val	Gly	Val	Tyr	Pro	Phe
545					550					555					560
Gly	Leu	Asp	Tyr	Ala	Trp	His	Gly	Ser	Asp	Asn	Ser	Leu	Leu	Phe	Thr
				565					570					575	
Asn	Ser	Tyr	Lys	Met	Lys	Leu	Ser	Ile	Leu	Leu	Gly	Phe	Ile	His	Met
			580					585					590		
Ser	Tyr	Ser	Tyr	Ile	Phe	Ser	Tyr	Leu	Asn	Tyr	His	Tyr	Lys	Gly	Ser
		595					600					605			
Arg	Ile	Asp	Ile	Val	Gly	Asn	Phe	Ile	Pro	Gly	Leu	Ile	Phe	Met	Gln
	610					615					620				
Ser	Ile	Phe	Gly	Tyr	Leu	Ser	Trp	Ala	Ile	Ile	Tyr	Lys	Trp	Ser	Lys
625					630					635					640
Asp	Trp	Ile	Lys	Asp	Glu	Leu	Pro	Ala	Pro	Gly	Leu	Leu	Asn	Met	Leu
			645						650					655	
Ile	Asn	Met	Phe	Leu	Ser	Pro	Gly	Val	Val	Asp	Glu	Lys	Leu	Tyr	Thr
			660					665					670		
Gly	Gln	Ser	Phe	Leu	Gln	Val	Ile	Leu	Leu	Leu	Ala	Ala	Leu	Val	Cys
		675					680					685			
Val	Pro	Trp	Leu	Leu	Leu	Tyr	Lys	Pro	Leu	Met	Leu	Lys	Arg	Gln	Asn
	690					695					700				
Asp	Ile	Ala	Leu	Ser	Lys	Gly	Phe	Arg	Ser	Leu	Arg	Asp	Gln	Arg	Val
705					710					715					720
His	Glu	Ile	Leu	Leu	Glu	Ala	Gln	Glu	Asn	Ala	Gly	Glu	Asp	Met	Leu
				725					730					735	
Val	Ala	Asp	Tyr	Glu	Asn	Glu	Asp	Glu	Ser	Ser	Glu	Glu	Phe	Asn	Phe

740					745					750					
Gly	Asp	Val	Met	Ile	His	Gln	Val	Ile	His	Thr	Ile	Glu	Phe	Cys	Leu
	755					760					765				
Asn	Cys	Ile	Ser	His	Thr	Ala	Ser	Tyr	Leu	Arg	Leu	Trp	Ala	Leu	Ser
	770					775					780				
Leu	Ala	His	Ala	Gln	Leu	Ser	Thr	Val	Leu	Trp	Ser	Met	Thr	Ile	Gln
785						790					795				800
Asn	Ser	Phe	Ser	Asp	Ser	Asn	Pro	Gly	Ser	Phe	Phe	Ser	Val	Thr	Lys
			805						810					815	
Val	Val	Val	Leu	Phe	Ala	Met	Trp	Phe	Val	Leu	Thr	Val	Cys	Ile	Leu
			820					825					830		
Val	Leu	Met	Glu	Gly	Thr	Ser	Ala	Met	Leu	His	Ser	Leu	Arg	Leu	His
		835					840						845		
Trp	Val	Glu	Ala	Met	Ser	Lys	Phe	Phe	Glu	Gly	Glu	Gly	Tyr	Ala	Tyr
	850					855					860				
Glu	Pro	Phe	Ser	Phe	Lys	Ala	Ile	Asn	Ser	Asp	Asp	Glu			
865						870					875				

<210> 23

<211> 872

<212> DNA

<213> 棉桃阿舒氏囊霉

<220>

<221> misc_feature

<223> Oligo 4

<400> 23

```

gatctcatgg aggattgatt atataactat ttcttgctga gaggaagtta tggctataat      60
cgccgtgtac acatgtat    ccaagacgct cactttttct tcatagaggg tgctccatat      120
aacctgatag agcttaagcg ggtctttgaa accaaagata aacttcttaa agtttattcc      180
atattactcg taaaatggaa ctctaaagtt ttacagcctc aagggtgagg cagtgacgac      240
ttttttcatc cattccggca ccttgagaat atttgatggt aagtgctctg ttgttagtag      300
atagctttca tggctatctc gaatgtttga ctctgtattg gctccgcctg agtcaagctt      360
ttgaagctaa gaaaggacca acttgtaaaa cttcagacaa attttatcta ggaacatggc      420
ctgattcagc tcgaggtatc ctctttcaac tcaaatacaca ttaatgcact aaaactgagt      480
gtaagcaaga cctgcaaata ataaaatttg ggcacatttt cattaaattt ctcccatata      540
agatagtatc ttctcatgta aagcatatta atgaaacca cacaagacc agaagtatta      600
tactgaaaag tataggaaga aacagcaata atcatcccat tattaaataa ttctactaaa      660
ccggtcagga taaggtatat ttcagacatg agactgagct ttttccatag tccgaagaac      720
aatccttcac acgccacaat taaccagcac catttcatta taggatagaa atgttcagaa      780

```

attctcttttg gatcaataga atctcacatg aaagaagtat tgacgtatatt tctgattcct 840
 gggcatgtaa acttaaaccg ctgttaagga tc 872

<210> 24
 <211> 114
 <212> PRT
 <213> 棉桃阿舒氏囊霉

<220>
 <221> misc_feature
 <223> Oligo 4

<400> 24

Pro Gln Arg Phe Lys Phe Thr Cys Pro Gly Ile Arg Lys Tyr Val Asn
 1 5 10 15

Thr Ser Phe Met Asp Ser Ile Asp Pro Lys Arg Ile Ser Glu His Phe
 20 25 30

Tyr Pro Ile Met Lys Trp Cys Trp Leu Ile Val Ala Cys Glu Gly Leu
 35 40 45

Phe Phe Gly Leu Trp Lys Lys Leu Ser Leu Met Ser Glu Ile Tyr Leu
 50 55 60

Ile Leu Thr Gly Leu Val Glu Leu Phe Asn Asn Gly Met Ile Ile Ala
 65 70 75 80

Val Ser Ser Tyr Thr Phe Gln Tyr Asn Thr Ser Gly Leu Cys Val Gly
 85 90 95

Phe Ile Asn Met Leu Tyr Met Arg Arg Tyr Tyr Leu Ile Trp Glu Lys
 100 105 110

Phe Asn

<210> 25
 <211> 2093
 <212> DNA
 <213> 棉桃阿舒氏囊霉

<220>
 <221> misc_feature
 <223> Oligo 4

<400> 25

ccatgaatga gcgataatat atataacaag agtgggtgtag catggcatga gcoctgcttgc 60

ggatggataa aggaggagtc ttcaggatag gaataatttc ccaagaatct attcatgaat 120

ggtatccaca ataggtcttt aatgtgagtt acgtagagtc tcttatgaga tttttacaga 180

aatttgatatt taaaaatggc catttagcat gtttagtcta tgtcgccagg ggaatgggtgt 240

ttaaaaattt tctcatttgt tgtgaattca agatatgcta tgatacacia ataaacgata 300

tacggactct accacatgca aattatagac taaaaagata cggttgataa aagcattttt	360
aaggagcaag attaacccgtc cagttgcgta atattatact taatttggtt tatagagggtg	420
tttttactcg tcttcatgat gtttatcggt attggataaa catgatggca ccactctcta	480
tcatatcgct cgcagaatag ccgatataat cttttcagta gcgaaaaaat attttgcgtt	540
atcttaaata aatttaagct tagttaccat tttttttatt tagagaactg tctactgtga	600
tctcatggag gattgattat ataactatct cttgctgaga ggaagttatg gctataatcg	660
ccgtgtacac atgtatttcc aagacgctca ctttttcttc atagaggggtg ctccatataa	720
cctgatagag ctttaagcggg tctttgaaac caaagataaa cttcttaaag tttattccat	780
attactcgta aaatggaact cttaaagttt acagcctcaa gggtgaggca gtgacgactt	840
ttttcatcca ttccggcacc ttgagaatat ttgatgttaa gtgctctggt gttagtagat	900
agctttcatg gctatctcga atgtttgact ctgtattggc tccgcctgag tcaagctttt	960
gaagctaaga aaggaccaac ttgtaaaact tcagacaaat tttatctagg aacatggcct	1020
gattcagctc gaggtatcct ctttcaactc aaatcacatt aatgcactaa aactgagtgt	1080
aagcaagacc tgcaaataat aaaatttggg cacattttca ttaaatttct cccatataag	1140
atagtatctt ctcatgtaaa gcatattaat gaaaccaca caaagaccag aagtattata	1200
ctgaaaagta taggaagaaa cagcaataat catcccatta ttaaataatt ctactaaacc	1260
ggtcaggata aggtatatct cagacatgag actgagcttt ttccatagtc cgaagaacaa	1320
tccttcacac gccacaatta accagcacca tttcattata ggatagaaat gttcagaaat	1380
tctctttgga tcaatagaat ctacatgaa agaagtattg acgtattttc tgattcctgg	1440
gcatgtaaac ttaaaccgct gttaaggatc gcaatagttc tcaatattat gaagttgcca	1500
attaataacg ccaagattca caaagatctg tataagaatc attaaaagct gaccccgaaa	1560
tattggaacg ggtggtatct ttgcgtaatg tgccaactta atatcgccaa gatagttatc	1620
agtttgata tcaatattat acgcgaatga cttgataaaa gttagcgcac aatgatttct	1680
cggcactttt caaaccactt attacctgca acagcagcat tgttgctata gtaatactat	1740
ctgtgacttg tagcaacgcg acggaaagtg atagatcgga gttctaactc atggaaaacc	1800
atattgctca gatcaattcg acttccaagg ctatactacc agagatattg caaaggttgt	1860
gaagaacggt ccctacagtc actcctgctg ccattgggtc ttaagagata ccatatctta	1920
aatacccctt gtcgggcaat gatgaggatt agcaacattc attgatagtt ggtatccttt	1980
tcttagtacc aatgcggtgt gcaggtgtaa tggtttacag gcggaacaat ggaggatatg	2040
cagtttctgg acaaggataa tcaaggaatg aaaccattag gagtgggcca ata	2093

<210> 26
 <211> 2093
 <212> DNA
 <213> 棉桃阿舒氏囊霉

 <220>
 <221> CDS
 <222> (738)..(1037)
 <223>

 <220>
 <221> misc_feature
 <223> Oligo 4

<400> 26
 tattggccca ctctaattgg ttctattcct tgattatcct tgtccagaaa ctgcatatcc 60

 tccattgttc cgctgtataa ccattacacc tgcacaccgc attggtacta agaaaaggat 120
 accaactatc aatgaatggt gctaattctc atcattgccc gacaaggggt atttaagata 180
 tggatatctt taagaaccaa tggcagcagg agtgactgta gggaccgttc ttcacaacct 240
 ttgcaatatc tctggtagta tagccttgga agtcgaattg atctgagcaa tatggttttc 300
 catagattag aactccgac taccactttc cgtcgcgttg ctacaagtca cagatagtat 360
 tactatagca acaatgctgc tgttgccagg aataagtggg ttgaaaagtg ccgagaaatc 420
 attatgcgct aacttttatc aagtcattcg cgtataatat tgatatccaa actgataact 480
 atcttggcga tattaagttg gcacattacg caaagatacc acccggtcca atatttcggg 540
 gtcagctttt aatgattctt atacagatct ttgtgaatct tggcggtatt aattggcaac 600
 ttcataatat tgagaactat tgcgatcctt aacagcgggt taagtttaca tgcccaggaa 660
 tcagaaaata cgtcaatact tctttcatgt gagattctat tgatccaaag agaattttctg 720
 aacattttcta tctata atg aaa tgg tgc tgg tta att gtg gcg tgt gaa 770
 Met Lys Trp Cys Trp Leu Ile Val Ala Cys Glu
 1 5 10

 gga ttg ttc ttc gga cta tgg aaa aag ctc agt ctc atg tct gaa ata 818
 Gly Leu Phe Phe Gly Leu Trp Lys Lys Leu Ser Leu Met Ser Glu Ile
 15 20 25

 tac ctt atc ctg acc ggt tta gta gaa tta ttt aat aat ggg atg att 866
 Tyr Leu Ile Leu Thr Gly Leu Val Glu Leu Phe Asn Asn Gly Met Ile
 30 35 40

 att gct gtt tct tcc tat act ttt cag tat aat act tct ggt ctt tgt 914
 Ile Ala Val Ser Ser Tyr Thr Phe Gln Tyr Asn Thr Ser Gly Leu Cys
 45 50 55

 gtg ggt ttc att aat atg ctt tac atg aga aga tac tat ctt ata tgg 962
 Val Gly Phe Ile Asn Met Leu Tyr Met Arg Arg Tyr Tyr Leu Ile Trp
 60 65 70 75

 gag aaa ttt aat gaa aat gtg ccc aaa ttt tat tat ttg cag gtc ttg 1010
 Glu Lys Phe Asn Glu Asn Val Pro Lys Phe Tyr Tyr Leu Gln Val Leu

80	85	90	
ctt aca ctc agt ttt agt gca tta atg tgatttgagt tgaaagagga			1057
Leu Thr Leu Ser Phe Ser Ala Leu Met			
95	100		
tacctcgagc tgaatcaggc catgttccta gataaaattht gtctgaagtt ttacaagttg			1117
gtccttttctt agcttcaaaa gcttgactca ggcggagcca atacagagtc aaacattcga			1177
gatagccatg aaagctatct actaacaaca gagcacttaa catcaaatat tctcaagggtg			1237
ccggaatgga tgaaaaaagt cgtcactgcc tcacccttga ggctgtaaaa ctttagagtt			1297
ccatttttacg agtaatatgg aataaacttht aagaagttta tctttggtht caaagacccg			1357
cttaagctct atcagggttat atggagcacc ctctatgaag aaaaagtgag cgtcttgga			1417
atacatgtgt acacggcgat tatagccata acttcctctc agcaagaaat agttatataa			1477
tcaatcctcc atgagatcac agtagacagt tctctaaata aaaaaaatgg taactaagct			1537
taaattttatt taagataacg caaaatatttht tttcgctact gaaaagatta tatcggctat			1597
tctgcgagcg atatgataga gagtggtgcc atcatgttta tccaataccg ataaacatca			1657
tgaagacgag taaaaacacc tctataaacc aaattaagta taatattacg caactggacg			1717
gttaatcttg ctcttataaaa atgcttttht caaccgtatc tttttagtht ataatttgca			1777
tgtggtagag tccgtatatc gtttatttht gtatcatagc atatcttgaa ttcacaacaa			1837
atgagaaaaat ttttaaacac cattcccttg gcgacataga ctaaacatgc taaatgacca			1897
tttttaaaat caaattthtctg taaaaatctc ataagagact ctacgtaact cacattaaag			1957
acctattgtg gataccattc atgaatagat tcttgggaaa ttattcctat cctgaagact			2017
cctccttht atccatccgcaa gcaggctcat gccatgctac accactcttg ttatatatat			2077
tatcgctcat tcatgg			2093

<210> 27

<211> 100

<212> PRT

<213> 棉桃阿舒氏囊霉

<220>

<221> misc_feature

<223> Oligo 4

<400> 27

Met	Lys	Trp	Cys	Trp	Leu	Ile	Val	Ala	Cys	Glu	Gly	Leu	Phe	Phe	Gly
1				5					10					15	

Leu	Trp	Lys	Lys	Leu	Ser	Leu	Met	Ser	Glu	Ile	Tyr	Leu	Ile	Leu	Thr
			20					25					30		

Gly	Leu	Val	Glu	Leu	Phe	Asn	Asn	Gly	Met	Ile	Ile	Ala	Val	Ser	Ser
		35					40					45			

```

Tyr Thr Phe Gln Tyr Asn Thr Ser Gly Leu Cys Val Gly Phe Ile Asn
  50                      55                      60

Met Leu Tyr Met Arg Arg Tyr Tyr Leu Ile Trp Glu Lys Phe Asn Glu
  65                      70                      75                      80

Asn Val Pro Lys Phe Tyr Tyr Leu Gln Val Leu Leu Thr Leu Ser Phe
          85                      90                      95

Ser Ala Leu Met
          100

```

```

<210> 28
<211> 853
<212> DNA
<213> 棉桃阿舒氏囊霉

```

```

<220>
<221> misc_feature
<223> Oligo 6

```

```

<400> 28
gatcatattc cccttgtagg tctagccgca cgggtaaacg gcaatgtatg tcgcagaaca      60
agtccccaca cctccggcag ttctggattc ctgtcctcgg tgtgagcttt cgcccgcatc      120
tcgagcattg tgacttcctt ggcacaagag gtttgtaatg tgacctgttt ctgccggatg      180
tctcccgact ggacgccaga cccaccgaag ttggtgtagt acttgcaatg ttcacsggaa      240
ccgctgtctg tggatagcga tcgttcatct ttaaaattgt gttcaacatc cagatgcata      300
ttcaactggt ccaagttttc aatatttcga ttgcatacgg ggcaactcgac cgattccgag      360
gcgtcagagg cagcagcatt actgtactct gtaggtatgt gttcatctgc ctccgggggt      420
ggcattactg ggccacctag ttggaactga catttatgat ataacgcaca gcactaatgt      480
gtgaaacgca aaagttcacc aactgtcaa aaaccaccac acagtcaact cgagatagac      540
acaataatga tggaagagct gtttatgcta ccagttattg gtatcacggc gttggttggtg      600
tataaatacg cttacgatat cctttggtac aagttacaga atctatttga actgattcct      660
tcatcaagct catctcacc ctactcttcg gctatcgcta gcattaataa gtctcagaat      720
gggtttctgt acaagctata cacggaatat tcagtttctt ctaacaatgt gctgcgagtg      780
attcggactg ctagtctccg ggactgttgg ctgctgtgcg tctgtggctg tggagatagt      840
gctatggcag atc                                             853

```

```

<210> 29
<211> 38
<212> PRT
<213> 棉桃阿舒氏囊霉

```

```

<220>

```

<221> misc_feature

<223> Oligo 6

<400> 29

Glu Ser Val Glu Cys Pro Val Cys Asn Arg Asn Ile Glu Asn Leu Glu
1 5 10 15

Gln Leu Asn Met His Leu Asp Val Glu His Asn Phe Lys Asp Glu Arg
20 25 30

Ser Leu Ser Thr Asp Ser
35

<210> 30

<211> 55

<212> PRT

<213> 棉桃阿舒氏囊霉

<220>

<221> misc_feature

<223> Oligo 6

<400> 30

Arg Ser His Tyr Lys Pro Leu Val Pro Gly Lys Ser Gln Cys Ser Arg
1 5 10 15

Cys Gly Arg Lys Leu Thr Pro Arg Thr Gly Ile Gln Asn Cys Arg Arg
20 25 30

Cys Gly Asp Leu Phe Cys Asp Ile His Cys Arg Leu Pro Val Arg Leu
35 40 45

Asp Leu Gln Gly Glu Tyr Asp
50 55

<210> 31

<211> 2389

<212> DNA

<213> 棉桃阿舒氏囊霉

<220>

<221> CDS

<222> (428)..(1993)

<223>

<220>

<221> misc_feature

<223> Oligo 6

<400> 31

gatctgccat agcactatct ccacagccac agacgcacag cagccaacag tcccggagac 60

tagcgtccga atcactcgca gcacattggt agaagaaact gaatattccg tgtatagctt 120

gtacagaaac ccattctgag acttattaat gctagcgata gccgaagagt aggggtgaga 180

tgagcttgat gaaggaatca gttcaaatag attctgtaac ttgtaccaaa ggatatcgta	240
agcgtatttta tacaccaaca acgccgtgat accaataact ggtagcataa acagctcttc	300
catcattatt gtgtctatct cgagttgact gtgtggtggt ttttgacagt gtggtgaact	360
tttgcgtttc acacattagt gctgtgcggt atatcataaa tgtcagttcc aactaggtgg	420
cccagta atg cca acc ccg gag gca gat gaa caa ata cct aca gag tac Met Pro Thr Pro Glu Ala Asp Glu Gln Ile Pro Thr Glu Tyr 1 5 10	469
agt aat gct gct gcc tct gac gcc tcg gaa tcg gtc gag tgc ccc gta Ser Asn Ala Ala Ala Ser Asp Ala Ser Glu Ser Val Glu Cys Pro Val 15 20 25 30	517
tgc aat cga aat att gaa aac ttg gaa cag ttg aat atg cat ctg gat Cys Asn Arg Asn Ile Glu Asn Leu Glu Gln Leu Asn Met His Leu Asp 35 40 45	565
gtt gaa cac aat ttt aaa gat gaa cga tcg cta tcc aca gac agc ggt Val Glu His Asn Phe Lys Asp Glu Arg Ser Leu Ser Thr Asp Ser Gly 50 55 60	613
tcc gtg aac att gca agt act aca cca act tcg gtg ggt ctg gcg tcc Ser Val Asn Ile Ala Ser Thr Thr Pro Thr Ser Val Gly Leu Ala Ser 65 70 75	661
agt cgg gag aca tcc ggc aga aac agg tca cat tac aaa cct ctt gtg Ser Arg Glu Thr Ser Gly Arg Asn Arg Ser His Tyr Lys Pro Leu Val 80 85 90	709
cca ggg aag tca caa tgc tcg aga tgc ggg cga aag ctc aca ccg agg Pro Gly Lys Ser Gln Cys Ser Arg Cys Gly Arg Lys Leu Thr Pro Arg 95 100 105 110	757
aca gga atc cag aac tgc cgg agg tgt ggg gac ttg ttc tgc gac ata Thr Gly Ile Gln Asn Cys Arg Arg Cys Gly Asp Leu Phe Cys Asp Ile 115 120 125	805
cat tgc cgt tta ccc gtg cgg cta gac cta caa ggg gaa tat gat ccg His Cys Arg Leu Pro Val Arg Leu Asp Leu Gln Gly Glu Tyr Asp Pro 130 135 140	853
aga aac ggc gac tgg tgc aag tgc tgc cat gcc tgc atg gcg ggg cgc Arg Asn Gly Asp Trp Cys Lys Cys Cys His Ala Cys Met Ala Gly Arg 145 150 155	901
cct gga tac aac aaa ctg ggc tta tcg gtg gac cgg acc gac gag ttt Pro Gly Tyr Asn Lys Leu Gly Leu Ser Val Asp Arg Thr Asp Glu Phe 160 165 170	949
ata cgc cat cgt acc tct aaa aac gag gat aaa cag ctt cgc att ctg Ile Arg His Arg Thr Ser Lys Asn Glu Asp Lys Gln Leu Arg Ile Leu 175 180 185 190	997
cag ctt gag aat cgg ctt gtg cgt ctt gtt gac ggc atc gct act att Gln Leu Glu Asn Arg Leu Val Arg Leu Val Asp Gly Ile Ala Thr Ile 195 200 205	1045
gtt cgc gcg cac aac cag tct ttg ttt tat gga agt gga atg tat agg Val Arg Ala His Asn Gln Ser Leu Phe Tyr Gly Ser Gly Met Tyr Arg	1093

210					215					220						
gaa	atc	act	gct	cta	cag	aag	tcc	gtt	acg	cca	tgg	aaa	gaa	aat	tca	1141
Glu	Ile	Thr	Ala	Leu	Gln	Lys	Ser	Val	Thr	Pro	Trp	Lys	Glu	Asn	Ser	
		225					230					235				
cag	gca	tca	agc	tgc	tat	ctc	tgt	tct	cgg	ccg	ttc	aat	ctt	ctt	ctt	1189
Gln	Ala	Ser	Ser	Cys	Tyr	Leu	Cys	Ser	Arg	Pro	Phe	Asn	Leu	Leu	Leu	
	240					245					250					
cgc	aag	cac	cat	tgc	aag	ttg	tgt	ggt	ctg	ata	gtc	tgc	gaa	aac	aac	1237
Arg	Lys	His	His	Cys	Lys	Leu	Cys	Gly	Leu	Ile	Val	Cys	Glu	Asn	Asn	
255					260					265					270	
ttt	acc	aac	tgt	tcc	aag	gag	ttc	ccg	atc	gcg	cag	ctg	gtg	agt	gct	1285
Phe	Thr	Asn	Cys	Ser	Lys	Glu	Phe	Pro	Ile	Ala	Gln	Leu	Val	Ser	Ala	
				275					280					285		
gcc	aca	gac	cta	cca	ttc	agg	agc	aat	ccc	cag	gaa	tta	gcc	gca	ttg	1333
Ala	Thr	Asp	Leu	Pro	Phe	Arg	Ser	Asn	Pro	Gln	Glu	Leu	Ala	Ala	Leu	
			290					295					300			
ccg	gtt	cga	cta	cgc	gtc	tgc	gtg	gtt	tgc	ata	aga	tcc	gtg	ttt	ctg	1381
Pro	Val	Arg	Leu	Arg	Val	Cys	Val	Val	Cys	Ile	Arg	Ser	Val	Phe	Leu	
		305				310						315				
cgc	gca	agg	ttg	cag	gac	aat	ctc	gca	aac	gat	gct	tct	cag	cta	ttt	1429
Arg	Ala	Arg	Leu	Gln	Asp	Asn	Leu	Ala	Asn	Asp	Ala	Ser	Gln	Leu	Phe	
	320					325					330					
tcc	aag	tac	act	gag	ttg	cag	cgt	gta	tca	aga	gcc	atc	ttg	cgt	ata	1477
Ser	Lys	Tyr	Thr	Glu	Leu	Gln	Arg	Val	Ser	Arg	Ala	Ile	Leu	Arg	Ile	
335					340					345					350	
atg	cct	cgt	ttc	gag	caa	ctg	ctc	ggg	gac	ctc	aac	gca	ccc	gac	gcg	1525
Met	Pro	Arg	Phe	Glu	Gln	Leu	Leu	Gly	Asp	Leu	Asn	Ala	Pro	Asp	Ala	
				355					360					365		
agc	ccg	aat	cga	agc	gag	cta	gac	gag	ctg	gcg	cac	ctg	cga	agg	aaa	1573
Ser	Pro	Asn	Arg	Ser	Glu	Leu	Asp	Glu	Leu	Ala	His	Leu	Arg	Arg	Lys	
			370					375					380			
ctc	ttg	gag	act	ttc	aag	ctg	tac	gac	acc	ata	gcc	aag	cag	ata	ttt	1621
Leu	Leu	Glu	Thr	Phe	Lys	Leu	Tyr	Asp	Thr	Ile	Ala	Lys	Gln	Ile	Phe	
		385				390						395				
gcc	atc	gac	ccg	gcc	aat	acc	gca	gaa	ttg	aag	atc	cag	cag	gct	ata	1669
Ala	Ile	Asp	Pro	Ala	Asn	Thr	Ala	Glu	Leu	Lys	Ile	Gln	Gln	Ala	Ile	
	400					405					410					
aag	gcc	aaa	tcg	atg	tcc	ttc	ata	cag	gat	aaa	atg	cta	ccg	cta	aag	1717
Lys	Ala	Lys	Ser	Met	Ser	Phe	Ile	Gln	Asp	Lys	Met	Leu	Pro	Leu	Lys	
415					420					425					430	
aat	atc	ccc	ggc	ctg	ctg	aag	ccg	aag	gat	gcc	gaa	cct	gag	atc	aat	1765
Asn	Ile	Pro	Gly	Leu	Leu	Lys	Pro	Lys	Asp	Ala	Glu	Pro	Glu	Ile	Asn	
				435					440					445		
tac	act	act	agc	aat	cta	ctt	ttc	aac	aat	ctg	act	gtc	cgc	gag	gtt	1813
Tyr	Thr	Thr	Ser	Asn	Leu	Leu	Phe	Asn	Asn	Leu	Thr	Val	Arg	Glu	Val	
			450					455					460			

```

aaa ctc tac cgt gaa cag ctg atg gta ctc aaa gag cag agg ttt ata      1861
Lys Leu Tyr Arg Glu Gln Leu Met Val Leu Lys Glu Gln Arg Phe Ile
      465                      470                      475

gtg gag ggc atg ctc gag aac gcc aag aaa cag cgg cgt ttt gaa gag      1909
Val Glu Gly Met Leu Glu Asn Ala Lys Lys Gln Arg Arg Phe Glu Glu
      480                      485                      490

gtt aat acg tta aag gaa aat acc aaa gag cta gac aat cag ata gcc      1957
Val Asn Thr Leu Lys Glu Asn Thr Lys Glu Leu Asp Asn Gln Ile Ala
      495                      500                      505                      510

cag ctc gaa gaa acc cta ggc gac cag ggt ttt gtt tagtatctag      2003
Gln Leu Glu Glu Thr Leu Gly Asp Gln Gly Phe Val
      515                      520

catggagtttt tttgcttaac tataattact gtgtagatgc cgcagatagc atgtcgtagc      2063
ataattgcga attttcacca acatgaaaaa gtgtatgtgt ataaggcatc cagtgaactc      2123
ctaacatgct gatgaggtttt taagtaaaga tatcactagc aatgaacgta agtgcagttt      2183
ttgagcttta tgtcctctgt agaacataat attaacgaca gggggatagg atgaaagaag      2243
acagcagtta tttgagctga acagtgaagc ctgggtctgga attgatgcgt tcccgaataa      2303
aaccagcaag cttgactcaa gcatcaagag aaacacaggg tttatcaaaa agctgaaaca      2363
gggtatcacg aaagactcga aagatc      2389

```

```

<210> 32
<211> 522
<212> PRT
<213> 棉桃阿舒氏囊霉

```

```

<220>
<221> misc_feature
<223> Oligo 6

```

```

<400> 32

```

```

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

Ala Ala Ala Ser Asp Ala Ser Glu Ser Val Glu Cys Pro Val Cys Asn
20          25          30

Arg Asn Ile Glu Asn Leu Glu Gln Leu Asn Met His Leu Asp Val Glu
35          40          45

His Asn Phe Lys Asp Glu Arg Ser Leu Ser Thr Asp Ser Gly Ser Val
50          55          60

Asn Ile Ala Ser Thr Thr Pro Thr Ser Val Gly Leu Ala Ser Ser Arg
65          70          75          80

Glu Thr Ser Gly Arg Asn Arg Ser His Tyr Lys Pro Leu Val Pro Gly
85          90          95

Lys Ser Gln Cys Ser Arg Cys Gly Arg Lys Leu Thr Pro Arg Thr Gly
100         105         110

```

```

Ile Gln Asn Cys Arg Arg Cys Gly Asp Leu Phe Cys Asp Ile His Cys
    115                      120                      125

Arg Leu Pro Val Arg Leu Asp Leu Gln Gly Glu Tyr Asp Pro Arg Asn
    130                      135                      140

Gly Asp Trp Cys Lys Cys Cys His Ala Cys Met Ala Gly Arg Pro Gly
    145                      150                      155                      160

Tyr Asn Lys Leu Gly Leu Ser Val Asp Arg Thr Asp Glu Phe Ile Arg
    165                      170                      175

His Arg Thr Ser Lys Asn Glu Asp Lys Gln Leu Arg Ile Leu Gln Leu
    180                      185                      190

Glu Asn Arg Leu Val Arg Leu Val Asp Gly Ile Ala Thr Ile Val Arg
    195                      200                      205

Ala His Asn Gln Ser Leu Phe Tyr Gly Ser Gly Met Tyr Arg Glu Ile
    210                      215                      220

Thr Ala Leu Gln Lys Ser Val Thr Pro Trp Lys Glu Asn Ser Gln Ala
    225                      230                      235                      240

Ser Ser Cys Tyr Leu Cys Ser Arg Pro Phe Asn Leu Leu Leu Arg Lys
    245                      250                      255

His His Cys Lys Leu Cys Gly Leu Ile Val Cys Glu Asn Asn Phe Thr
    260                      265                      270

Asn Cys Ser Lys Glu Phe Pro Ile Ala Gln Leu Val Ser Ala Ala Thr
    275                      280                      285

Asp Leu Pro Phe Arg Ser Asn Pro Gln Glu Leu Ala Ala Leu Pro Val
    290                      295                      300

Arg Leu Arg Val Cys Val Val Cys Ile Arg Ser Val Phe Leu Arg Ala
    305                      310                      315                      320

Arg Leu Gln Asp Asn Leu Ala Asn Asp Ala Ser Gln Leu Phe Ser Lys
    325                      330                      335

Tyr Thr Glu Leu Gln Arg Val Ser Arg Ala Ile Leu Arg Ile Met Pro
    340                      345                      350

Arg Phe Glu Gln Leu Leu Gly Asp Leu Asn Ala Pro Asp Ala Ser Pro
    355                      360                      365

Asn Arg Ser Glu Leu Asp Glu Leu Ala His Leu Arg Arg Lys Leu Leu
    370                      375                      380

Glu Thr Phe Lys Leu Tyr Asp Thr Ile Ala Lys Gln Ile Phe Ala Ile
    385                      390                      395                      400

Asp Pro Ala Asn Thr Ala Glu Leu Lys Ile Gln Gln Ala Ile Lys Ala
    405                      410                      415

Lys Ser Met Ser Phe Ile Gln Asp Lys Met Leu Pro Leu Lys Asn Ile
    420                      425                      430

Pro Gly Leu Leu Lys Pro Lys Asp Ala Glu Pro Glu Ile Asn Tyr Thr

```

435	440	445
Thr Ser Asn Leu Leu Phe	Asn Asn Leu Thr Val	Arg Glu Val Lys Leu
450	455	460
Tyr Arg Glu Gln Leu Met Val	Leu Lys Glu Gln Arg Phe Ile Val	Glu
465	470	475
Gly Met Leu Glu Asn Ala Lys Lys	Gln Arg Arg Phe Glu Glu Val	Asn
485	490	495
Thr Leu Lys Glu Asn Thr Lys Glu	Leu Asp Asn Gln Ile Ala Gln	Leu
500	505	510
Glu Glu Thr Leu Gly Asp Gln Gly	Phe Val	
515	520	

<210> 33

<211> 975

<212> DNA

<213> 棉桃阿舒氏囊霉

<220>

<221> misc_feature

<223> Oligo 146

<400> 33

```

gatccggacg acagggctgc ggcacaggag tcgtaccgta ggggcggggc agtccgaggc      60
cagcacgaat ggggttccggc cgttcacgga cgagggcatg ttcttccgag gggccggggg      120
ggctccggag gacctgtttg actttttctt tcggggcggg ggccccggcg gcccgttcgg      180
gatggctgat ccctatgatt cgtttggggc ctttgggggc gctacgacgt ttacgtttgg      240
cgggcccccg ggcttcaagg tgtacagcgg aggggtctgg gggcagttcc gccgcggacc      300
gtttggcatg gcacaggctg ccaccgaagc gcaacgcaa cgtgcgaacg gccagccgga      360
ggcgcaagat ccgctacaac acgtagtgtt cgttctcctg atagtgttct tcttcctgct      420
actgccaagt ctgggcttct gacttggagc agtatccgtt gtaccataca gggccggcgc      480
cgacctgacc aaggcttccc atggctggcc ttccatgatg gcatttcgta aatacctatt      540
tttatagtca acaagtataa ataaataata atatacaaat aatttacaaa tctgcagcac      600
gtcgcgtgac gctgactgcg acataaaaat gttcaagata gtttgattag aatgcacaaa      660
caaacattgc aacacttaca taaaccagat gaagctacag ccgctattcc aggttttgca      720
ttctgcgttc tattatttcg tcaatatcct cggctgctgg cttaagcttt agcttcctgc      780
ctaatttggc ggcaaatcca tcgtttcccg agccgccaga attgtcaggt actagatttt      840
ttctcttgat aagtttgcca ctcggtaact gctcatattc ctctgtgtta tatgtggcac      900
tgctgtgtat agtcactttg ctaggcgttc caggaggtaa ctcggtcgca ccgggtagtg      960
tattcttcct ggatc                                          975

```

<210> 34
 <211> 66
 <212> PRT
 <213> 棉桃阿舒氏囊霉

<220>
 <221> misc_feature
 <223> Oligo 146

<400> 34

```
Ser Ser Ala Thr Tyr Asn Thr Glu Glu Tyr Glu Gln Leu Pro Ser Gly
1          5          10          15

Lys Leu Ile Lys Arg Lys Asn Leu Val Pro Asp Asn Ser Gly Gly Ser
          20          25          30

Gly Asn Asp Gly Phe Ala Ala Lys Leu Gly Arg Lys Leu Lys Leu Lys
          35          40          45

Pro Ala Ala Glu Asp Ile Asp Glu Ile Ile Glu Arg Arg Met Gln Asn
          50          55          60

Leu Glu
65
```

<210> 35
 <211> 2509
 <212> DNA
 <213> 棉桃阿舒氏囊霉

<220>
 <221> CDS
 <222> (882)..(1208)
 <223>

<220>
 <221> misc_feature
 <223> Oligo 146

```
<400> 35
ctgcagtttc tgaaggtgct ggtggcagat ccctacggtt tgaacaagat ccgcaccgcg      60
ttcgagaagg gctacatcac gccgaatgac acatggttcc aggggtggcac gacgacatac      120
gacaacattg cctatacgct ggcctgatg atcgactgcg gcattgttgg catgtcgtgg      180
atcacactac caaagggcaa gtatgccatg atcccaaaga acaaaaaaat atctacgtgc      240
cagctggaag tgtcaataaa ctataaggac ctgatttcgc ggccggcaga cagtgcactgg      300
tcgcacagtg ctccattgcg aatcctctct ttcgacatag aatgtgctgg ccgcgtgggc      360
gtgtttccag aaccggaggt tgatcctgtc attcagatag cgaatgttgt cagtattgcg      420
ggagagtcaa aaccatttat tcgcaatgtc ttcacagttg acacctgtgc gccattact      480
ggctcgcaaa tttttgagca tgagacagaa gaagcgatgc ttaaacactg gcgtgacttc      540
```

cttggtgagg ttgaccccgga cggtatcatt ggttacaata cgacaaatTT tgatcttcct	600
tacctgatca acagagctgc agccttgaac gtgtcttcat ttccctactt tggccggttg	660
gttaactcca agcaggtcat caaagaaact gtatttttcat ccaaggctca tggaacgcgg	720
gagtcgaagt cgatcaatat cgagggccgc cttcagttgg atatttttca gtttgtaaga	780
cgagaatatc agttgagatc gtacactttg aattcgggtgc gtttaatgag ttgaagcaag	840
gatggacatg gcaacgtgat cctccaacaa ttaagcgctt g atg aac aag gcg att	896
	Met Asn Lys Ala Ile
	1 5
gga aat aaa agc gat act ttt gat ggc gtt atc aac cat acg cga tcc	944
Gly Asn Lys Ser Asp Thr Phe Asp Gly Val Ile Asn His Thr Arg Ser	
	10 15 20
agg aag aat aca cta ccc ggt gcg acc gag tta cct cct gga acg cct	992
Arg Lys Asn Thr Leu Pro Gly Ala Thr Glu Leu Pro Pro Gly Thr Pro	
	25 30 35
agc aaa gtg act ata cac agc agt gcc aca tat aac aca gag gaa tat	1040
Ser Lys Val Thr Ile His Ser Ser Ala Thr Tyr Asn Thr Glu Glu Tyr	
	40 45 50
gag cag tta ccg agt ggc aaa ctt atc aag aga aaa aat cta gta cct	1088
Glu Gln Leu Pro Ser Gly Lys Leu Ile Lys Arg Lys Asn Leu Val Pro	
	55 60 65
gac aat tct ggc ggc tcg gga aac gat gga ttt gcc gcc aaa tta ggc	1136
Asp Asn Ser Gly Gly Ser Gly Asn Asp Gly Phe Ala Ala Lys Leu Gly	
	70 75 80 85
agg aag cta aag ctt aag cca gca gcc gag gat att gac gaa ata ata	1184
Arg Lys Leu Lys Leu Lys Pro Ala Ala Glu Asp Ile Asp Glu Ile Ile	
	90 95 100
gaa cgc aga atg caa aac ctg gaa tagcggctgt agcttcatct gggttatgta	1238
Glu Arg Arg Met Gln Asn Leu Glu	
	105
agtgttgcaa tgtttggttg tgcattctaa tcaaactatc ttgaacatTT ttatgtcgca	1298
gtcagcgtea cgcgacgtgc tgcagatttg taaattatTT gtatattatt atttatttat	1358
acttgttgac tataaaaata ggtattttacg aaatgccatc atggaaggcc agccatggga	1418
agccttggtc aggtcggcgc cggccctgta tgggtacaacg gatactgctc caagtcagaa	1478
gcccagactt ggcagtagca ggaagagaag cactatcagg agaacgaaca ctacgtggtg	1538
tagcggatct tgcgcctccg gctggccggtt cgcacgttgg cggtgcgctt cggtggcagc	1598
ctgtgccatg ccaaacggtc cgcggcggaa ctgcccccca gacccaccgc tgtacacctt	1658
gaagcccgcg ggcccggcaa acgtaaacgt cgtagcgcgc ccaaagggcc caaacgaatc	1718
atagggatca gccatcccga acgggcccgc ggggcccccg ccccgaaaga aaaagtcaaa	1778
caggtcctcc ggagcccccc cggccccctc gaagaacatg ccctcgtccg tgaacggccg	1838

```

gaacccattc gtgctggcct cggactgccc cgcccctacg gtacgactcc tgtgccgcag 1898
ccctgtcgtc cggatcgtag ccagctgggt cgtacacgcg gcgcttcttc tcgtcagaca 1958
aaacctcgaa cgcacgattt accttcttga acgcttccgc cgccctcggg tgccgatttt 2018
tgtccggatg tagtttgatt gccatcttgc ggtacgcctt ctttatgtcc ccatcgctcg 2078
ccttctcgtc cacctgtagc agtcataga acgaatgctt gtccttatcc acaatcagca 2138
gcgtcagctt ttcttgctct tctgtgtact ctttctccgc catcttgata gctatcttag 2198
ctctccgcca atctgccaga ccttgttctt cttgccaccg agtacaatt cctcgagcct 2258
ctcaaccaaa ctacottata acagctcctt cagatcacac gctgctggaa tcgatgactg 2318
gtccccggtc tgtcactatt catgcacaaa gcaactatag aacactcctc ttttcttaaa 2378
tgtaaaagtt gaagccaccg tgcagcgccg tttacgggaa gttccccaag gtagcacaac 2438
agaaggtaaa gatgaaacga ctacgcaccc gcataccggc ctagggccaa cggcgacatt 2498
cggcgtacag c 2509

```

<210> 36
 <211> 109
 <212> PRT
 <213> 棉桃阿舒氏囊霉

<220>
 <221> misc_feature
 <223> Oligo 146

<400> 36

```

Met Asn Lys Ala Ile Gly Asn Lys Ser Asp Thr Phe Asp Gly Val Ile
1           5           10           15

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

Pro Pro Gly Thr Pro Ser Lys Val Thr Ile His Ser Ser Ala Thr Tyr
          35           40           45

Asn Thr Glu Glu Tyr Glu Gln Leu Pro Ser Gly Lys Leu Ile Lys Arg
50           55           60

Lys Asn Leu Val Pro Asp Asn Ser Gly Gly Ser Gly Asn Asp Gly Phe
65           70           75           80

Ala Ala Lys Leu Gly Arg Lys Leu Lys Leu Lys Pro Ala Ala Glu Asp
          85           90           95

Ile Asp Glu Ile Ile Glu Arg Arg Met Gln Asn Leu Glu
          100          105

```

<210> 37
 <211> 1031
 <212> DNA

<213> 棉桃阿舒氏囊霉

<220>

<221> misc_feature

<223> Oligo 56

<400> 37

```

gatccactaa taattggacg atacgcgcgt acccttcagt cgctgccgca aagattgggg      60
tccagccaaa aagcttctca ccaagttcca tatttgcacc ctgtttaact agatattcgg      120
ccgcctcata gatgtctaac ttgcaagcaa cataaagagc agtctcgaga ttttccggct      180
cctgatagtc aatatcaaag cccttagcag acaataaaga atttaacaaa gacggcgagt      240
ttagtcttgt cgcgaggtgt agcaaacgag ggctgttaag agatttctct ggggtcataa      300
acgacaatag cgtgctgaca gtaagagggt gtgtcccgat aacagccaaa tgcaatgggtg      360
tgaggttctc cgaatctccc cacacgtcaa tatcatcgat tgcaacgtct gcattccaag      420
catcccactc actcagagcy tgcagaatta tcctgggtamc ctcacataga ccatattggg      480
cagagtaatg tagcggcggt cgcttgtagt tatctctttg taacagcgac ggtcttaaatt      540
gggctggcaa ctggtgtaat atgtgtgaca ggggagctgg agagtcgtca gagttaacgc      600
catccggacc aaaagcacca acaagtctag tattgggttc cgggggtatc gctgcctcta      660
gatcgagaga gtcgctcagt agcgagttta tggtatcacg ctcttcgaga ctcttggtcc      720
gctttttgcc cagtgcaatt acatgatgat ggaagaagtt tctgccattt atatcggatg      780
gatctccgag cgtaggtata ataccaagga tttcatccac acagctgaag gaccgcgata      840
aagcggactt attcaatagc gtcaccagcg tccgcgtcgg tatcaggaca acagaactgg      900
tacatagaaa tgagctcggt tatcaagccc gccccgtcat ccttggtaac tagctgggac      960
gccacatcaa ttggcgacga cgagccttgg agcagacgac ggtcttcgcc gcgtagttta     1020
tcctgtagat c                                     1031

```

<210> 38

<211> 17

<212> PRT

<213> 棉桃阿舒氏囊霉

<220>

<221> misc_feature

<223> Oligo 56

<400> 38

```

Val Thr Lys Asp Asp Gly Ala Gly Leu Ile Asn Glu Leu Ile Ser Met
1           5           10           15

```

Tyr

<210> 39
 <211> 298
 <212> PRT
 <213> 棉桃阿舒氏囊霉

<220>
 <221> misc_feature
 <223> Oligo 56

<220>
 <221> misc_feature
 <222> (147)..(147)
 <223> 未知氨基酸

<400> 39

```

Ser Ser Val Val Leu Ile Pro Thr Arg Thr Leu Val Thr Leu Leu Asn
1          5          10          15

Lys Ser Ala Leu Ser Arg Ser Phe Ser Cys Val Asp Glu Ile Leu Gly
          20          25          30

Ile Ile Pro Thr Leu Gly Asp Pro Ser Asp Ile Asn Gly Arg Asn Phe
          35          40          45

Phe His His His Val Ile Ala Leu Gly Lys Lys Arg Thr Lys Ser Leu
          50          55          60

Glu Glu Arg Asp Asn Ile Asn Ser Leu Leu Ser Asp Ser Leu Asp Leu
65          70          75          80

Glu Ala Ala Ile Pro Pro Glu Pro Asn Thr Arg Leu Val Gly Ala Phe
          85          90          95

Gly Pro Asp Gly Val Asn Ser Asp Asp Ser Pro Ala Pro Leu Ser His
          100          105          110

Ile Leu Gln Gln Leu Pro Ala His Leu Arg Pro Ser Leu Leu Gln Arg
          115          120          125

Asp Asn Tyr Lys Arg Thr Pro Leu His Tyr Ser Ala Gln Tyr Gly Leu
          130          135          140

Cys Glu Xaa Thr Arg Ile Ile Leu Gln Ala Leu Ser Glu Trp Asp Ala
145          150          155          160

Trp Asn Ala Asp Val Ala Ile Asp Asp Ile Asp Val Trp Gly Asp Ser
          165          170          175

Glu Asn Leu Thr Pro Leu His Leu Ala Val Ile Gly Thr His Pro Leu
          180          185          190

Thr Val Ser Thr Leu Leu Ser Phe Met Asn Pro Glu Lys Ser Leu Asn
          195          200          205

Ser Pro Arg Leu Leu His Leu Ala Thr Arg Leu Asn Ser Pro Ser Leu
          210          215          220

Leu Asn Ser Leu Leu Ser Ala Lys Gly Phe Asp Ile Asp Tyr Gln Glu
225          230          235          240

```

Pro Glu Asn Leu Glu Thr Ala Leu Tyr Val Ala Cys Lys Leu Asp Ile
 245 250 255

Tyr Glu Ala Ala Glu Tyr Leu Val Lys Gln Gly Ala Asn Met Glu Leu
 260 265 270

Gly Glu Lys Leu Phe Gly Trp Thr Pro Ile Phe Ala Ala Ala Thr Glu
 275 280 285

Gly Tyr Ala Arg Ile Val Gln Leu Leu Val
 290 295

<210> 40
 <211> 4925
 <212> DNA
 <213> 棉桃阿舒氏囊霉

<220>
 <221> CDS
 <222> (426) .. (4388)
 <223>

<220>
 <221> misc_feature
 <223> Oligo 56

<400> 40
 ctcaagggttt ggctgggtct ggaactacgc cagctcatga cgctatccgt gcgtatcatc 60
 aatgacctcg ttaaaacgga tcagctatgt cctagttact gaaggccaaa aagaaatata 120
 taaaagtaac gctgggtgcct ttccacctta tccacgtgcg gaagctcggg cgtaaagggtt 180
 acgagctacg cttgcagatt gtgtttgagg aaggccaggg tgctcgtgcc tagcattcat 240
 tccgcgcaca tgtgtacata tatactagct ttataggcca cacaacaaac agatatgcgc 300
 acggctggag atatctcggc atacaccgcg tggcacagtg ttttctcttg acttctgcat 360
 attagagggtt tgtatagggg atactagctt cagctgggga gcacacagac taggctagga 420
 caagt atg aag ttc ggc aag aca ttt ccc aac cat cag gtg ccg gaa tgg 470
 Met Lys Phe Gly Lys Thr Phe Pro Asn His Gln Val Pro Glu Trp
 1 5 10 15

gca cac aag tac gtg aat tac aag ggt ctg aag aag cag atc aag gaa 518
 Ala His Lys Tyr Val Asn Tyr Lys Gly Leu Lys Lys Gln Ile Lys Glu
 20 25 30

atc acg ttg gtg cag gat gcg ctg ttc cgc cag gag caa ggc gca gct 566
 Ile Thr Leu Val Gln Asp Ala Leu Phe Arg Gln Glu Gln Gly Ala Ala
 35 40 45

tcg cag gac gga ccc gct cgg cgg cgg gga aga gag agc aag gag cag 614
 Ser Gln Asp Gly Pro Ala Arg Arg Arg Gly Arg Glu Ser Lys Glu Gln
 50 55 60

tat ctt ggc cat cca gag gtg aag aag ctg ctt gca gca ttt ttc ttt 662
 Tyr Leu Gly His Pro Glu Val Lys Lys Leu Leu Ala Ala Phe Phe Phe
 65 70 75

gcc ctg gac cgg gat atc gag aag gtg gac ggt ttc tac aac atg cag	710
Ala Leu Asp Arg Asp Ile Glu Lys Val Asp Gly Phe Tyr Asn Met Gln	
80 85 90 95	
ttt atg gag tat gac cgg cgg ctg agg aag ctt cta tca agc gcg cag	758
Phe Met Glu Tyr Asp Arg Arg Leu Arg Lys Leu Leu Ser Ser Ala Gln	
100 105 110	
ctg gca gac atc acg tcg gta cag cgc ggc gct acc ggc tac ctg cac	806
Leu Ala Asp Ile Thr Ser Val Gln Arg Gly Ala Thr Gly Tyr Leu His	
115 120 125	
gcg cca ctt ccg cag tac ata gca tac ggg gag cgc gaa cgg gat gga	854
Ala Pro Leu Pro Gln Tyr Ile Ala Tyr Gly Glu Arg Glu Arg Asp Gly	
130 135 140	
ttg cca gag cgc tat gta ccg ccg cac gcc act gac atg tcg gag gac	902
Leu Pro Glu Arg Tyr Val Pro Pro His Ala Thr Asp Met Ser Glu Asp	
145 150 155	
cta gcg gag gtg ctg acg att ctg ctg gag ctg cgg tcg cac ttc cgc	950
Leu Ala Glu Val Leu Thr Ile Leu Leu Glu Leu Arg Ser His Phe Arg	
160 165 170 175	
aac ctg aag tgg tac ggt gag ctc aac aag cgg gca ttc acg aaa atc	998
Asn Leu Lys Trp Tyr Gly Glu Leu Asn Lys Arg Ala Phe Thr Lys Ile	
180 185 190	
atg aag aaa ctg gat aag aag gtt ggc aca aac cag cag cac tcc tac	1046
Met Lys Lys Leu Asp Lys Lys Val Gly Thr Asn Gln Gln His Ser Tyr	
195 200 205	
ttc cag gcc cgc att aag ccc tta gaa ttt gct gac gat aca ccg atc	1094
Phe Gln Ala Arg Ile Lys Pro Leu Glu Phe Ala Asp Asp Thr Pro Ile	
210 215 220	
gtc aag gcg cta gct acc att aat gag atc cta gat cgc atc tcg ccc	1142
Val Lys Ala Leu Ala Thr Ile Asn Glu Ile Leu Asp Arg Ile Ser Pro	
225 230 235	
tgc gtg aag gat cta cag gat aaa cta cgc ggc gaa gac cgt cgt ctg	1190
Cys Val Lys Asp Leu Gln Asp Lys Leu Arg Gly Glu Asp Arg Arg Leu	
240 245 250 255	
ctc caa ggc tcg tcg tcg cca att gat gtg gcg tcc cag cta gtt acc	1238
Leu Gln Gly Ser Ser Ser Pro Ile Asp Val Ala Ser Gln Leu Val Thr	
260 265 270	
aag gat gac ggg gcg ggc ttg ata aac gag ctc att tct atg tac cgt	1286
Lys Asp Asp Gly Ala Gly Leu Ile Asn Glu Leu Ile Ser Met Tyr Arg	
275 280 285	
tct gtt gtc ctg ata ccg acg cgg acg ctg gtg acg cta ttg aat aag	1334
Ser Val Val Leu Ile Pro Thr Arg Thr Leu Val Thr Leu Leu Asn Lys	
290 295 300	
tcc gct tta tcg cgg tcc ttc agc tgt gtg gat gaa atc ctt ggt att	1382
Ser Ala Leu Ser Arg Ser Phe Ser Cys Val Asp Glu Ile Leu Gly Ile	
305 310 315	
ata cct acg ctc gga gat cca tcc gat ata aat ggc aga aac ttc ttc	1430

Ile 320	Pro	Thr	Leu	Gly	Asp 325	Pro	Ser	Asp	Ile	Asn 330	Gly	Arg	Asn	Phe	Phe 335	
cat	cat	cat	gta	att	gca	ctg	ggc	aaa	aag	cgg	acc	aag	agt	ctc	gaa	1478
His	His	His	Val	Ile 340	Ala	Leu	Gly	Lys	Lys 345	Arg	Thr	Lys	Ser	Leu	Glu 350	
gag	cgt	gat	aac	ata	aac	tcg	cta	ctg	agc	gac	tct	ctc	gat	cta	gag	1526
Glu	Arg	Asp	Asn 355	Ile	Asn	Ser	Leu	Leu 360	Ser	Asp	Ser	Leu	Asp 365	Leu	Glu	
gca	gcg	ata	ccc	ccg	gaa	ccc	aat	act	aga	ctt	gtt	ggc	gct	ttt	ggc	1574
Ala	Ala	Ile 370	Pro	Pro	Glu	Pro	Asn 375	Thr	Arg	Leu	Val	Gly 380	Ala	Phe	Gly	
ccg	gat	ggc	gtt	aac	tct	gac	gac	tct	cca	gct	ccc	ctg	tca	cac	ata	1622
Pro	Asp 385	Gly	Val	Asn	Ser	Asp 390	Asp	Ser	Pro	Ala	Pro 395	Leu	Ser	His	Ile	
tta	caa	cag	ttg	cca	gcc	cat	tta	aga	ccg	tcg	ctg	tta	caa	aga	gat	1670
Leu 400	Gln	Gln	Leu	Pro	Ala 405	His	Leu	Arg	Pro	Ser	Leu 410	Leu	Gln	Arg	Asp 415	
aac	tac	aag	cga	acg	ccg	cta	cat	tac	tct	gcc	caa	tat	ggc	cta	tgt	1718
Asn	Tyr	Lys	Arg	Thr 420	Pro	Leu	His	Tyr	Ser 425	Ala	Gln	Tyr	Gly	Leu 430	Cys	
gag	gtt	acc	agg	ata	att	ctg	cag	gct	ctg	agt	gag	tgg	gat	gct	tgg	1766
Glu	Val	Thr	Arg 435	Ile	Ile	Leu	Gln	Ala 440	Leu	Ser	Glu	Trp	Asp 445	Ala	Trp	
aat	gca	gac	gtt	gca	atc	gat	gat	att	gac	gtg	tgg	gga	gat	tcg	gag	1814
Asn	Ala 450	Asp	Val	Ala	Ile	Asp	Asp 455	Ile	Asp	Val	Trp	Gly 460	Asp	Ser	Glu	
aac	ctc	aca	cca	ttg	cat	ttg	gct	gtt	atc	ggg	aca	cac	cct	ctt	act	1862
Asn 465	Leu	Thr	Pro	Leu	His	Leu 470	Ala	Val	Ile	Gly	Thr 475	His	Pro	Leu	Thr	
gtc	agc	acg	cta	ttg	tcg	ttt	atg	aac	cca	gag	aaa	tct	ctt	aac	agc	1910
Val 480	Ser	Thr	Leu	Leu	Ser 485	Phe	Met	Asn	Pro	Glu 490	Lys	Ser	Leu	Asn 495	Ser	
cct	cgt	ttg	cta	cac	ctc	gcg	aca	aga	cta	aac	tcg	ccg	tct	ttg	tta	1958
Pro	Arg	Leu	Leu 500	His	Leu	Ala	Thr	Arg	Leu 505	Asn	Ser	Pro	Ser	Leu 510	Leu	
aat	tct	tta	ttg	tct	gct	aag	ggc	ttt	gat	att	gac	tat	cag	gag	ccg	2006
Asn	Ser	Leu	Leu 515	Ser	Ala	Lys	Gly	Phe 520	Asp	Ile	Asp	Tyr	Gln 525	Glu	Pro	
gaa	aat	ctc	gag	act	gct	ctt	tat	gtt	gct	tgc	aag	tta	gac	atc	tat	2054
Glu	Asn 530	Leu	Glu	Thr	Ala	Leu	Tyr 535	Val	Ala	Cys	Lys	Leu 540	Asp	Ile	Tyr	
gag	gcg	gcc	gaa	tat	cta	gtt	aaa	cag	ggc	gca	aat	atg	gaa	ctt	ggc	2102
Glu 545	Ala	Ala	Glu	Tyr	Leu	Val 550	Lys	Gln	Gly	Ala	Asn 555	Met	Glu	Leu	Gly	
gag	aag	ctt	ttt	ggc	tgg	acc	cca	atc	ttt	gcg	gca	gcg	act	gaa	ggg	2150
Glu 560	Lys	Leu	Phe	Gly	Trp 565	Thr	Pro	Ile	Phe	Ala 570	Ala	Ala	Thr	Glu	Gly 575	

tac gcg cgt atc gtc caa tta tta gtg gat cat ggg gcc aaa tat gat	2198
Tyr Ala Arg Ile Val Gln Leu Leu Val Asp His Gly Ala Lys Tyr Asp	
580 585 590	
ctt ttt gat gag agc ggt tgg aca cca atg gag cac gcg gct ctg cgg	2246
Leu Phe Asp Glu Ser Gly Trp Thr Pro Met Glu His Ala Ala Leu Arg	
595 600 605	
gga cat ttg gat atc tct caa ctg att cgc ata aca gat aat aaa gcg	2294
Gly His Leu Asp Ile Ser Gln Leu Ile Arg Ile Thr Asp Asn Lys Ala	
610 615 620	
atc act cgg ccg aag ttc gcc acc gac tgg aac aag tct acc aga cca	2342
Ile Thr Arg Pro Lys Phe Ala Thr Asp Trp Asn Lys Ser Thr Arg Pro	
625 630 635	
aca gaa acc acg aat ggt ttg tta tcc gca tta aca cca tca gaa tct	2390
Thr Glu Thr Thr Asn Gly Leu Leu Ser Ala Leu Thr Pro Ser Glu Ser	
640 645 650 655	
ggg tct aca acc acg ggc tcc gaa aac aag agc tct tct ttg aca cct	2438
Gly Ser Thr Thr Thr Gly Ser Glu Asn Lys Ser Ser Ser Leu Thr Pro	
660 665 670	
agc acc agc aat gaa atg tat gct ctg cca gca cgc tct tca acg tca	2486
Ser Thr Ser Asn Glu Met Tyr Ala Leu Pro Ala Arg Ser Thr Ser	
675 680 685	
ata gac aag ata tct gaa cca aac aaa gga aac cac agg aag gtt ttg	2534
Ile Asp Lys Ile Ser Glu Pro Asn Lys Gly Asn His Arg Lys Val Leu	
690 695 700	
aag tcc cag cta agc cac ggt aag gta caa act att aag gac aca caa	2582
Lys Ser Gln Leu Ser His Gly Lys Val Gln Thr Ile Lys Asp Thr Gln	
705 710 715	
ttg ccg cag cag ccg att aag tct ttc ggt cat agc ttc ctg cag aaa	2630
Leu Pro Gln Gln Pro Ile Lys Ser Phe Gly His Ser Phe Leu Gln Lys	
720 725 730 735	
gac gag tct gtt atc ctt ttg act ttg ggc act aat gat aat cgc agt	2678
Asp Glu Ser Val Ile Leu Leu Thr Leu Gly Thr Asn Asp Asn Arg Ser	
740 745 750	
acc ata cct gct gtt tct ctg aat aag gtt cct gtt gct aag gct tcc	2726
Thr Ile Pro Ala Val Ser Leu Asn Lys Val Pro Val Ala Lys Ala Ser	
755 760 765	
tca acg gag cta gat aca gcg tta tct ttg ttg gtt aca tgc atg gat	2774
Ser Thr Glu Leu Asp Thr Ala Leu Ser Leu Leu Val Thr Cys Met Asp	
770 775 780	
aat ttg gat gca gaa cct gtt atg ctg gac ctt cca ttg cat gag aac	2822
Asn Leu Asp Ala Glu Pro Val Met Leu Asp Leu Pro Leu His Glu Asn	
785 790 795	
ttg gat tcg gtc act ttt aaa gtc cca tac aag aaa gat tcc tct tat	2870
Leu Asp Ser Val Thr Phe Lys Val Pro Tyr Lys Lys Asp Ser Ser Tyr	
800 805 810 815	
act ata ttt ttt gat atc gtt ccc acc tat ggc tat tca atg gca aac	2918

Thr	Ile	Phe	Phe	Asp	Ile	Val	Pro	Thr	Tyr	Gly	Tyr	Ser	Met	Ala	Asn	
				820					825					830		
atg	aac	cgc	gaa	aat	tcc	tct	ggg	atg	cat	tcg	aat	gtt	ggg	aat	agt	2966
Met	Asn	Arg	Glu	Asn	Ser	Ser	Gly	Met	His	Ser	Asn	Val	Gly	Asn	Ser	
			835					840					845			
act	ggc	ccc	gcg	tac	cta	gac	gca	caa	gtg	ggc	cag	tgc	ggc	tct	cgt	3014
Thr	Gly	Pro	Ala	Tyr	Leu	Asp	Ala	Gln	Val	Gly	Gln	Cys	Gly	Ser	Arg	
		850					855					860				
tta	cac	tat	gac	cag	ctt	ggg	aga	gat	aca	cca	aac	act	tac	gat	cag	3062
Leu	His	Tyr	Asp	Gln	Leu	Gly	Arg	Asp	Thr	Pro	Asn	Thr	Tyr	Asp	Gln	
	865					870					875					
cgt	tcg	cgg	cac	caa	gca	tcc	caa	caa	aag	gaa	caa	att	gcc	acg	aag	3110
Arg	Ser	Arg	His	Gln	Ala	Ser	Gln	Gln	Lys	Glu	Gln	Ile	Ala	Thr	Lys	
880					885					890					895	
aag	caa	agt	aag	ata	tta	ggg	agg	gca	gtt	gct	tta	ttg	gac	tcc	gcg	3158
Lys	Gln	Ser	Lys	Ile	Leu	Gly	Arg	Ala	Val	Ala	Leu	Leu	Asp	Ser	Ala	
				900					905					910		
cca	act	tcg	gtg	ggc	ccc	aac	agg	cgc	tct	att	gcg	gag	gct	atc	act	3206
Pro	Thr	Ser	Val	Gly	Pro	Asn	Arg	Arg	Ser	Ile	Ala	Glu	Ala	Ile	Thr	
			915					920					925			
ata	cct	atc	att	ggg	agc	gac	aca	ctt	gaa	gtc	ctc	ggg	att	att	cga	3254
Ile	Pro	Ile	Ile	Gly	Ser	Asp	Thr	Leu	Glu	Val	Leu	Gly	Ile	Ile	Arg	
		930					935					940				
ttt	gac	ttc	ctc	gta	gtt	act	ccg	ttt	gtc	cac	aag	aat	tta	tca	gtt	3302
Phe	Asp	Phe	Leu	Val	Val	Thr	Pro	Phe	Val	His	Lys	Asn	Leu	Ser	Val	
	945					950					955					
gga	cct	gct	gaa	acg	tat	tgg	aaa	tcg	ctg	gtt	tcg	acc	cgg	gtg	atc	3350
Gly	Pro	Ala	Glu	Thr	Tyr	Trp	Lys	Ser	Leu	Val	Ser	Thr	Arg	Val	Ile	
960					965				970						975	
gga	cac	aga	ggc	ttg	ggc	aaa	aac	atg	aac	acg	aac	aag	tcg	ttg	cag	3398
Gly	His	Arg	Gly	Leu	Gly	Lys	Asn	Met	Asn	Thr	Asn	Lys	Ser	Leu	Gln	
			980						985					990		
ctt	ggc	gag	aat	acg	gtg	gag	tcc	ttt	atc	gca	gct	gcc	tca	ttg	ggg	3446
Leu	Gly	Glu	Asn	Thr	Val	Glu	Ser	Phe	Ile	Ala	Ala	Ala	Ser	Leu	Gly	
			995					1000					1005			
gcc	tcg	tat	gtc	gaa	ttc	gat	gtc	cag	ttg	aca	aag	gat	aac	att		3491
Ala	Ser	Tyr	Val	Glu	Phe	Asp	Val	Gln	Leu	Thr	Lys	Asp	Asn	Ile		
		1010					1015					1020				
ccc	gtt	gtg	tac	cac	gat	ttt	ctt	gtt	gca	gaa	tcg	ggg	gtc	gat		3536
Pro	Val	Val	Tyr	His	Asp	Phe	Leu	Val	Ala	Glu	Ser	Gly	Val	Asp		
		1025					1030					1035				
att	ccc	atg	cat	gag	ctc	act	ttg	gaa	cag	ttt	ctc	gat	ctg	aac		3581
Ile	Pro	Met	His	Glu	Leu	Thr	Leu	Glu	Gln	Phe	Leu	Asp	Leu	Asn		
		1040					1045					1050				
ggg	gag	cgt	caa	agg	cat	cag	gac	gcg	aga	gag	gcc	cac	aga	aac		3626
Gly	Glu	Arg	Gln	Arg	His	Gln	Asp	Ala	Arg	Glu	Ala	His	Arg	Asn		
		1055					1060					1065				

cac agg agc	ccg aac ggc	aga cgc	ttg tcc atg	gat gac	agc tct	3671
His Arg Ser	Pro Asn Gly	Arg Arg	Leu Ser Met	Asp Asp	Ser Ser	
1070		1075		1080		
gct gag cta	att aag cgg	tcc ctg	atg atg cgg	ggc gat	gaa gac	3716
Ala Glu Leu	Ile Lys Arg	Ser Leu	Met Met Arg	Gly Asp	Glu Asp	
1085		1090		1095		
cgc aca gct	aga gac ctt	aac aca	ata tac ggc	gac cgt	atg cgg	3761
Arg Thr Ala	Arg Asp Leu	Asn Thr	Ile Tyr Gly	Asp Arg	Met Arg	
1100		1105		1110		
ctg acc aga	acg ttc aag	aag aat	gcc ttc aaa	gcc aac	tcc agg	3806
Leu Thr Arg	Thr Phe Lys	Lys Asn	Ala Phe Lys	Ala Asn	Ser Arg	
1115		1120		1125		
ggt cat gcc	att gca tct	agt ttt	gtt acg cta	aaa gag	ctg ttc	3851
Gly His Ala	Ile Ala Ser	Ser Phe	Val Thr Leu	Lys Glu	Leu Phe	
1130		1135		1140		
aag aag atc	ccc cag aat	gtt ggc	ttc aac atc	gag tgc	aag tat	3896
Lys Lys Ile	Pro Gln Asn	Val Gly	Phe Asn Ile	Glu Cys	Lys Tyr	
1145		1150		1155		
cca atg gta	gac gag gcc	gag gag	gaa gac atc	ggc ccg	atc gcc	3941
Pro Met Val	Asp Glu Ala	Glu Glu	Glu Asp Ile	Gly Pro	Ile Ala	
1160		1165		1170		
gtg gaa atg	aac cat tgg	atc gat	acc gtg ctg	gag gtt	gtc tac	3986
Val Glu Met	Asn His Trp	Ile Asp	Thr Val Leu	Glu Val	Val Tyr	
1175		1180		1185		
gac aac gtc	gag ggc cgt	gac gtc	atc ttt tcg	tcg ttt	cag cca	4031
Asp Asn Val	Glu Gly Arg	Asp Val	Ile Phe Ser	Ser Phe	Gln Pro	
1190		1195		1200		
gac gtg tgc	ctc atg ctc	tcc cta	aag cag ccc	tcc ttc	ccg atc	4076
Asp Val Cys	Leu Met Leu	Ser Leu	Lys Gln Pro	Ser Phe	Pro Ile	
1205		1210		1215		
ctg ttc cta	acg gaa ggt	ggg acc	gcg aag cgc	tgc gac	atc cgc	4121
Leu Phe Leu	Thr Glu Gly	Gly Thr	Ala Lys Arg	Cys Asp	Ile Arg	
1220		1225		1230		
gcg gcg tcg	ctg cag aat	gcc atc	gcg ttc gcg	cac cgc	tgg aac	4166
Ala Ala Ser	Leu Gln Asn	Ala Ile	Arg Phe Ala	His Arg	Trp Asn	
1235		1240		1245		
ctg ctg ggc	atc gtc tcg	gcc gcg	gcg cca atc	gtc atc	gcg ccc	4211
Leu Leu Gly	Ile Val Ser	Ala Ala	Ala Pro Ile	Val Ile	Ala Pro	
1250		1255		1260		
cgc ctg gcc	cag atc gtc	aag tcc	agt ggc ctt	gtg tgc	gtg acg	4256
Arg Leu Ala	Gln Ile Val	Lys Ser	Ser Gly Leu	Val Cys	Val Thr	
1265		1270		1275		
tac ggc gtc	gag aac aac	gac ccc	gag atc gcc	cgc gtc	gag atg	4301
Tyr Gly Val	Glu Asn Asn	Asp Pro	Glu Ile Ala	Arg Val	Glu Met	
1280		1285		1290		
gac gcc ggc	gtc gac gcg	gtc atc	gtc gac agc	gtg ctc	gcg gtc	4346

Asp Ala Gly Val Asp Ala Val Ile Val Asp Ser Val Leu Ala Val
 1295 1300 1305

cgc aag ggc ctc acc cgc gag gca cag gac gcc gac acg ctc 4388
 Arg Lys Gly Leu Thr Arg Glu Ala Gln Asp Ala Asp Thr Leu
 1310 1315 1320

taagaacaaa tcgtgattca ctaagcagaa gatatatcgt tatgtagagc ctgccgggctc 4448
 ccaggctaata ccacccattg ccgtcacaaa atcgtaaccg agatgcgccg cccgtgcccc 4508
 cgcgcgccgc cctgcagact cgcgcccgcg ctctgcatcg tcagcttctc caagctgact 4568
 aacttgctcg tcacgtgatc atccttgaac atttccccga tctgggtgctt gccgcgtca 4628
 tccgtgaaca cctgcaacgt cttcgtctgc gccgcctcgt tgatgtgggtg gtactcgtcc 4688
 gggaagggct ggatgtgctc gaccgtcagt tcgttcaatt ggatgaactt catcactgtg 4748
 aggttcgaca cctcgactgg aatattcgaa actatcacac tgtgcgccat ggggactctg 4808
 tttgtactgc cgcccttggt gcgccactga tgctgccacc gccacgtcca gcttatatac 4868
 ccgcgtggaa agttgtcgaa ttgtgaaata gccgacgcta tcggtaccat ttacacc 4925

<210> 41
 <211> 1321
 <212> PRT
 <213> 棉桃阿舒氏囊霉

<220>
 <221> misc_feature
 <223> Oligo 56

<400> 41

Met Lys Phe Gly Lys Thr Phe Pro Asn His Gln Val Pro Glu Trp Ala
 1 5 10 15

His Lys Tyr Val Asn Tyr Lys Gly Leu Lys Lys Gln Ile Lys Glu Ile
 20 25 30

Thr Leu Val Gln Asp Ala Leu Phe Arg Gln Glu Gln Gly Ala Ala Ser
 35 40 45

Gln Asp Gly Pro Ala Arg Arg Arg Gly Arg Glu Ser Lys Glu Gln Tyr
 50 55 60

Leu Gly His Pro Glu Val Lys Lys Leu Leu Ala Ala Phe Phe Phe Ala
 65 70 75 80

Leu Asp Arg Asp Ile Glu Lys Val Asp Gly Phe Tyr Asn Met Gln Phe
 85 90 95

Met Glu Tyr Asp Arg Arg Leu Arg Lys Leu Leu Ser Ser Ala Gln Leu
 100 105 110

Ala Asp Ile Thr Ser Val Gln Arg Gly Ala Thr Gly Tyr Leu His Ala
 115 120 125

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

130	135	140
Pro Glu Arg Tyr Val 145	Pro Pro His Ala Thr 150	Asp Met Ser Glu Asp Leu 155 160
Ala Glu Val Leu Thr 165	Ile Leu Leu Glu Leu 170	Arg Ser His Phe Arg Asn 175
Leu Lys Trp Tyr Gly 180	Glu Leu Asn Lys Arg 185	Ala Phe Thr Lys Ile Met 190
Lys Lys Leu Asp Lys 195	Lys Val Gly Thr Asn 200	Gln Gln His Ser Tyr Phe 205
Gln Ala Arg Ile Lys 210	Pro Leu Glu Phe Ala 215	Asp Asp Thr Pro Ile Val 220
Lys Ala Leu Ala Thr 225	Ile Asn Glu Ile Leu 230	Asp Arg Ile Ser Pro Cys 235 240
Val Lys Asp Leu Gln 245	Asp Lys Leu Arg Gly 250	Glu Asp Arg Arg Leu Leu 255
Gln Gly Ser Ser Ser 260	Pro Ile Asp Val Ala 265	Ser Gln Leu Val Thr Lys 270
Asp Asp Gly Ala Gly 275	Leu Ile Asn Glu Leu 280	Ile Ser Met Tyr Arg Ser 285
Val Val Leu Ile Pro 290	Thr Arg Thr Leu Val 295	Thr Leu Leu Asn Lys Ser 300
Ala Leu Ser Arg Ser 305	Phe Ser Cys Val Asp 310	Glu Ile Leu Gly Ile Ile 315 320
Pro Thr Leu Gly Asp 325	Pro Ser Asp Ile Asn 330	Gly Arg Asn Phe Phe His 335
His His Val Ile Ala 340	Leu Gly Lys Lys Arg 345	Thr Lys Ser Leu Glu Glu 350
Arg Asp Asn Ile Asn 355	Ser Leu Leu Ser Asp 360	Ser Leu Asp Leu Glu Ala 365
Ala Ile Pro Pro Glu 370	Pro Asn Thr Arg Leu 375	Val Gly Ala Phe Gly Pro 380
Asp Gly Val Asn Ser 385	Asp Asp Ser Pro Ala 390	Pro Leu Ser His Ile Leu 395 400
Gln Gln Leu Pro Ala 405	His Leu Arg Pro Ser 410	Leu Leu Gln Arg Asp Asn 415
Tyr Lys Arg Thr Pro 420	Leu His Tyr Ser Ala 425	Gln Tyr Gly Leu Cys Glu 430
Val Thr Arg Ile Ile 435	Leu Gln Ala Leu Ser 440	Glu Trp Asp Ala Trp Asn 445
Ala Asp Val Ala Ile 450	Asp Asp Ile Asp Val 455	Trp Gly Asp Ser Glu Asn 460

```

Leu Thr Pro Leu His Leu Ala Val Ile Gly Thr His Pro Leu Thr Val
465                               470                     475           480

Ser Thr Leu Leu Ser Phe Met Asn Pro Glu Lys Ser Leu Asn Ser Pro
                               485                     490           495

Arg Leu Leu His Leu Ala Thr Arg Leu Asn Ser Pro Ser Leu Leu Asn
                               500                     505           510

Ser Leu Leu Ser Ala Lys Gly Phe Asp Ile Asp Tyr Gln Glu Pro Glu
                               515                     520           525

Asn Leu Glu Thr Ala Leu Tyr Val Ala Cys Lys Leu Asp Ile Tyr Glu
530                               535                     540

Ala Ala Glu Tyr Leu Val Lys Gln Gly Ala Asn Met Glu Leu Gly Glu
545                               550                     555           560

Lys Leu Phe Gly Trp Thr Pro Ile Phe Ala Ala Ala Thr Glu Gly Tyr
                               565                     570           575

Ala Arg Ile Val Gln Leu Leu Val Asp His Gly Ala Lys Tyr Asp Leu
                               580                     585           590

Phe Asp Glu Ser Gly Trp Thr Pro Met Glu His Ala Ala Leu Arg Gly
                               595                     600           605

His Leu Asp Ile Ser Gln Leu Ile Arg Ile Thr Asp Asn Lys Ala Ile
610                               615                     620

Thr Arg Pro Lys Phe Ala Thr Asp Trp Asn Lys Ser Thr Arg Pro Thr
625                               630                     635           640

Glu Thr Thr Asn Gly Leu Leu Ser Ala Leu Thr Pro Ser Glu Ser Gly
                               645                     650           655

Ser Thr Thr Thr Gly Ser Glu Asn Lys Ser Ser Ser Leu Thr Pro Ser
                               660                     665           670

Thr Ser Asn Glu Met Tyr Ala Leu Pro Ala Arg Ser Ser Thr Ser Ile
                               675                     680           685

Asp Lys Ile Ser Glu Pro Asn Lys Gly Asn His Arg Lys Val Leu Lys
690                               695                     700

Ser Gln Leu Ser His Gly Lys Val Gln Thr Ile Lys Asp Thr Gln Leu
705                               710                     715           720

Pro Gln Gln Pro Ile Lys Ser Phe Gly His Ser Phe Leu Gln Lys Asp
                               725                     730           735

Glu Ser Val Ile Leu Leu Thr Leu Gly Thr Asn Asp Asn Arg Ser Thr
                               740                     745           750

Ile Pro Ala Val Ser Leu Asn Lys Val Pro Val Ala Lys Ala Ser Ser
                               755                     760           765

Thr Glu Leu Asp Thr Ala Leu Ser Leu Leu Val Thr Cys Met Asp Asn
770                               775                     780

```

```

Leu Asp Ala Glu Pro Val Met Leu Asp Leu Pro Leu His Glu Asn Leu
785                               790                               795                               800

Asp Ser Val Thr Phe Lys Val Pro Tyr Lys Lys Asp Ser Ser Tyr Thr
                               805                               810                               815

Ile Phe Phe Asp Ile Val Pro Thr Tyr Gly Tyr Ser Met Ala Asn Met
                               820                               825                               830

Asn Arg Glu Asn Ser Ser Gly Met His Ser Asn Val Gly Asn Ser Thr
                               835                               840                               845

Gly Pro Ala Tyr Leu Asp Ala Gln Val Gly Gln Cys Gly Ser Arg Leu
850                               855                               860

His Tyr Asp Gln Leu Gly Arg Asp Thr Pro Asn Thr Tyr Asp Gln Arg
865                               870                               875                               880

Ser Arg His Gln Ala Ser Gln Gln Lys Glu Gln Ile Ala Thr Lys Lys
                               885                               890                               895

Gln Ser Lys Ile Leu Gly Arg Ala Val Ala Leu Leu Asp Ser Ala Pro
                               900                               905                               910

Thr Ser Val Gly Pro Asn Arg Arg Ser Ile Ala Glu Ala Ile Thr Ile
                               915                               920                               925

Pro Ile Ile Gly Ser Asp Thr Leu Glu Val Leu Gly Ile Ile Arg Phe
930                               935                               940

Asp Phe Leu Val Val Thr Pro Phe Val His Lys Asn Leu Ser Val Gly
945                               950                               955                               960

Pro Ala Glu Thr Tyr Trp Lys Ser Leu Val Ser Thr Arg Val Ile Gly
                               965                               970                               975

His Arg Gly Leu Gly Lys Asn Met Asn Thr Asn Lys Ser Leu Gln Leu
                               980                               985                               990

Gly Glu Asn Thr Val Glu Ser Phe Ile Ala Ala Ala Ser Leu Gly Ala
                               995                               1000                               1005

Ser Tyr Val Glu Phe Asp Val Gln Leu Thr Lys Asp Asn Ile Pro
1010                               1015                               1020

Val Val Tyr His Asp Phe Leu Val Ala Glu Ser Gly Val Asp Ile
1025                               1030                               1035

Pro Met His Glu Leu Thr Leu Glu Gln Phe Leu Asp Leu Asn Gly
1040                               1045                               1050

Glu Arg Gln Arg His Gln Asp Ala Arg Glu Ala His Arg Asn His
1055                               1060                               1065

Arg Ser Pro Asn Gly Arg Arg Leu Ser Met Asp Asp Ser Ser Ala
1070                               1075                               1080

Glu Leu Ile Lys Arg Ser Leu Met Met Arg Gly Asp Glu Asp Arg
1085                               1090                               1095

Thr Ala Arg Asp Leu Asn Thr Ile Tyr Gly Asp Arg Met Arg Leu
1100                               1105                               1110

```

```

Thr Arg  Thr Phe Lys Lys Asn  Ala Phe Lys Ala Asn  Ser Arg Gly
1115                      1120                      1125

His Ala  Ile Ala Ser Ser  Phe  Val Thr Leu Lys Glu  Leu Phe Lys
1130                      1135                      1140

Lys Ile  Pro Gln Asn Val  Gly  Phe Asn Ile Glu Cys  Lys Tyr Pro
1145                      1150                      1155

Met Val  Asp Glu Ala Glu Glu  Glu Asp Ile Gly Pro  Ile Ala Val
1160                      1165                      1170

Glu Met  Asn His Trp Ile  Asp  Thr Val Leu Glu Val  Val Tyr Asp
1175                      1180                      1185

Asn Val  Glu Gly Arg Asp Val  Ile Phe Ser Ser  Phe  Gln Pro Asp
1190                      1195                      1200

Val Cys  Leu Met Leu Ser  Leu  Lys Gln Pro Ser  Phe  Pro Ile Leu
1205                      1210                      1215

Phe Leu  Thr Glu Gly Gly Thr  Ala Lys Arg Cys Asp  Ile Arg Ala
1220                      1225                      1230

Ala Ser  Leu Gln Asn Ala  Ile  Arg Phe Ala His Arg  Trp Asn Leu
1235                      1240                      1245

Leu Gly  Ile Val Ser Ala  Ala  Ala Pro Ile Val  Ile  Ala Pro Arg
1250                      1255                      1260

Leu Ala  Gln Ile Val Lys Ser  Ser Gly Leu Val  Cys  Val Thr Tyr
1265                      1270                      1275

Gly Val  Glu Asn Asn Asp Pro  Glu Ile Ala Arg Val  Glu Met Asp
1280                      1285                      1290

Ala Gly  Val Asp Ala Val  Ile  Val Asp Ser Val  Leu  Ala Val Arg
1295                      1300                      1305

Lys Gly  Leu Thr Arg Glu  Ala  Gln Asp Ala Asp Thr  Leu
1310                      1315                      1320

```

```

<210>  42
<211>  1078
<212>  DNA
<213>  棉桃阿舒氏囊霉

```

```

<220>
<221>  misc_feature
<223>  Oligo 167

```

```

<220>
<221>  misc_feature
<222>  (436)..(436)
<223>  未知核苷酸

```

```

<220>
<221>  misc_feature
<222>  (584)..(584)

```

<223> 未知核苷酸

<400> 42

```

gatccttcga gcgtagacga agaaaaggtc aaacgcctga aatcgacctt taccttttag      60
ctctcaacta atccggtcgt tatcaatcgt gtaattctat accttaaata gactacctaa      120
ccagcacgca gtgtcccata cagggttgct agtatggtgc ctaggtctca ttcgttggtat      180
tcttttgctc tgcagctttc ctttggggccc atgcagcgta agccctgccg ttcgcaatca      240
acggattatc atgcttgaaa tagtatgggc cagagagcgc tgaatattgt ctcttggtcta      300
agttatacac tctttccatt tcttccttat agacgtcctc gataatttta acatctccct      360
catctatctt cgtgttggtg gaggcggtgt ataaatctgc atcgaatggc cgaatgcaat      420
tgccgacata gaaganggtg gatgtgtgcc cctcacccaa ttcaaactcg gatttgggggt      480
tacgcamccg ccctagatgt tccataagcc gccgtgctag ctgcggatgg tctggctgct      540
gcttttgctg atgtggcctg tcactctctg tttcctctcc acctgagac gccatagatg      600
gccaacaga gtttcgtggc acgtcaatcc ctccgtctca tggctgggta tgcgtcgcag      660
tccggaatgg ggacagttat atacgcagca aactgcagat cgaccacgc ggcctgtgcc      720
cacgacacac attaccgtag tcgcaggcgc ttatgctcac aaaagaaagc cctgtatata      780
aggtcctaata tatatagagc tcttacaaca ctttctgttt catgaagaag gtgcgcaggt      840
ggttgacctg ccagaagcat gtgccgcaa gcacaaacaa cgtcatgacc gcacaccata      900
caaccacga attggcgctc tcggacgcat ctgcggaatt gggcctccct ctcaactgcac      960
caacctgctg ctctgttgct ggatcgctt cacttctctg ccgagctgta gcacgcgcga     1020
atgcaggtac tgcagcttgt ctttcttctt gggatccagc agtaggtccg atccgatc     1078

```

<210> 43

<211> 42

<212> PRT

<213> 棉桃阿舒氏囊霉

<220>

<221> misc_feature

<223> Oligo 167

<400> 43

```

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

```

```

<210> 44
<211> 2549
<212> DNA
<213> 棉桃阿舒氏囊霉

<220>
<221> CDS
<222> (563)..(1216)
<223>

<220>
<221> misc_feature
<223> Oligo 167

<400> 44
tatgccccgc ggctgcgggc ccgccagcac ccggcacatc caaaccggcc catcaaacat      60
gtactcctgg ttctgccgcc ccaccgccgc gaggtccac atcttcactc tccaatcccg      120
cgagctgctc cataggtacc gtccacattg cgaccacgcc accgactgca ccgctcgtac      180
atggccgccc ccatggtttc ccagcaccga aataggcttc ttcgtatcca tgtcgtagat      240
caccagagac ccgttggagc atccaactgc caggtagtca ccgcctgggc tgaactccag      300
acaatcgcat tgcagcggga tttcaaacgt atatgttagc gtctcagggt attccttcag      360
tacactgaat ggatcctgca acaaccgatt tgccatcgcc ccttccccgc ctccaagcct      420
ccaacctgat gctttcctcc aggctgtacc tggaaagtaa aaaaaattc gaacagcgaa      480
ccagatatgg acaccgaatt gaagatcttg ttccccctgg aatctcagta tatcgatcga      540
tcaggcgctg aaggcttcca gg atg aac ttg aag aca ttt tta cag gta ctg      592
                        Met Asn Leu Lys Thr Phe Leu Gln Val Leu
                        1             5             10

gtg ctc ggc atg ttg ccg ttg cag gtc agt ggg ttc tac ttc tat gtc      640
Val Leu Gly Met Leu Pro Leu Gln Val Ser Gly Phe Tyr Phe Tyr Val
                        15             20             25

aat ggt gga gac cgt aag tgc ttc cac agt gaa tta atg aaa gat gcg      688
Asn Gly Gly Asp Arg Lys Cys Phe His Ser Glu Leu Met Lys Asp Ala
                        30             35             40

gtt ttg aac ctg aag tac aac gtg caa tcg tac gac tcc cag tct gga      736
Val Leu Asn Leu Lys Tyr Asn Val Gln Ser Tyr Asp Ser Gln Ser Gly
                        45             50             55

acg tac cgg gac atg cgg gac ggt gaa cta atg atg atg atc gac ata      784
Thr Tyr Arg Asp Met Arg Asp Gly Glu Leu Met Met Met Ile Asp Ile
                        60             65             70

gag gag gtg ttt gac gac aac cac cgt gtt atg cac cag aaa ttg ccg      832
Glu Glu Val Phe Asp Asp Asn His Arg Val Met His Gln Lys Leu Pro
                        75             80             85             90

gcg tac ggt aca act acg ttc aac gcg gtt gac tct ggc gag cac aag      880
Ala Tyr Gly Thr Thr Thr Phe Asn Ala Val Asp Ser Gly Glu His Lys
                        95             100             105

gtt tgc gtt cag ccg cag ttc cag ggg tgg atg gcg cgc acc aag aca      928

```

Val	Cys	Val	Gln	Pro	Gln	Phe	Gln	Gly	Trp	Met	Ala	Arg	Thr	Lys	Thr	
			110					115					120			
aag	atc	acg	ctt	gac	ttt	gag	atc	gga	tcg	gac	cta	ctg	ctg	gat	ccc	976
Lys	Ile	Thr	Leu	Asp	Phe	Glu	Ile	Gly	Ser	Asp	Leu	Leu	Leu	Asp	Pro	
		125					130					135				
aag	aag	aag	gac	aag	ctg	cag	tac	ctg	cat	tcg	cgc	gtg	cta	cag	ctc	1024
Lys	Lys	Lys	Asp	Lys	Leu	Gln	Tyr	Leu	His	Ser	Arg	Val	Leu	Gln	Leu	
	140					145					150					
ggc	gag	aag	gtg	aag	gcg	atc	cgc	aac	gag	cag	cgg	ttg	gtg	cgt	gag	1072
Gly	Glu	Lys	Val	Lys	Ala	Ile	Arg	Asn	Glu	Gln	Arg	Leu	Val	Arg	Glu	
155					160				165						170	
agg	gag	gcc	caa	ttc	cgc	gat	gcg	tcc	gag	agc	gcc	aat	tcg	tgg	gtt	1120
Arg	Glu	Ala	Gln	Phe	Arg	Asp	Ala	Ser	Glu	Ser	Ala	Asn	Ser	Trp	Val	
				175				180						185		
gta	tgg	tgt	gcg	gtc	atg	acg	ttg	ttt	gtg	ctt	ggc	ggc	aca	tgc	ttc	1168
Val	Trp	Cys	Ala	Val	Met	Thr	Leu	Phe	Val	Leu	Gly	Gly	Thr	Cys	Phe	
			190					195					200			
tgg	cag	gtc	aac	cac	ctg	cgc	acc	ttc	ttc	atg	aaa	cag	aaa	gtg	ttg	1216
Trp	Gln	Val	Asn	His	Leu	Arg	Thr	Phe	Phe	Met	Lys	Gln	Lys	Val	Leu	
	205						210					215				
taagagctct atataattag gaccttatat acagggcttt cttttgtgag cataagcgcc																1276
tgcgactacg gtaatgtgtg tcgtgggcac aggccgcgtg ggtcgatctg cagtttgctg																1336
cgtatataac tgtccccatt ccggactgcg acgcataacc agccatgaga cggagggatt																1396
gacgtgccac gaaactctgt tgggccatct atggcgtctc aggggtggaga ggaaacagag																1456
agtgacaggc cacatcagca aaagcagcag ccagaccatc cgcagctagc acggcggctt																1516
atggaacatc tagggcgggt gcgtaacccc aaatccgagt ttgaattggg tgaggggcac																1576
acatccaccc tcttctatgt cggcaattgc attcggccat tcgatgcaga tttatacacc																1636
gcctccaaca acacgaagat agatgaggga gatgttaaaa ttatcgagga cgtctataag																1696
gaagaaatgg aaagagtgtg taacttagac aagagacaat attcagcgct ctctggccca																1756
tactattttca agcatgataa tccgttgatt gcgaacggca gggcttacgc tgcattggcc																1816
caaaggaaag ctgcagagca aaagaatata acgaatgaga cctagcacca tacgcaagcc																1876
ctgtatggga cactgcgtgc tggttaggta gtctatttaa ggtatagaat tacacgattg																1936
ataacgaccg gattagttga gagctaaaag gtaaaggctc atttcaggcg tttgaccttt																1996
tcttcgtcta cgctcgaagg atcaatccca gatacaggta aggttccgaa gtcaaagtgg																2056
aacggaaccg cttcggcaga tgctgcagtg ttctctgcaa atgaggagtc atccttgggt																2116
aagtcagggt tcgagaagga aaatgcaaag ggggtgcggg atgtcgtatt gtgcttatct																2176
tcacttttat gaacttggga gactttctga atgccattag aaaaattgaa cgacccattc																2236
gactgcaatg agctccccga tatgttatcc tccatgggcc cgtaaagagt agctccgtgg																2296

```

gtattagatt cctctacttt ggccagaggt tgaaacgaaa atgtgctggt tgcacaggt 2356
atagtgtgtg gcttctcaga caaagtatcc acattgtggt ctttgggctt ttggaaagaa 2416
aagttaaatt tagagctctg tggaacaggc tgccgaggcg acttttcagg cgacttttca 2476
acagtcacct caacgcgcca gcgatgaagg cctggaaggt ttaaaccggg acactgggtg 2536
tntagcatga gga 2549

```

```

<210> 45
<211> 218
<212> PRT
<213> 棉桃阿舒氏囊霉

```

```

<220>
<221> misc_feature
<223> Oligo 167

```

```

<400> 45

```

```

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

```

Arg	Thr	Phe	Phe	Met	Lys	Gln	Lys	Val	Leu
210						215			

图1

Oligo 19

609 KEPGKKVTLKXRMVSASAGSLVTSFLTPLDVVRVRLQQQEMLPSCCTCTGQLSKPAG--- 439
 + +TLK RM+SA AGS++TSL LTP+DVVR+RLQQQ+M+P C+C G P
 4 RNTSNSLTCLKERMLSAGAGSVLTSLLITPMDVVRIRLQQQQMIPDCSCDGAAEVFNNAVSS 63

 438 -----KVFWDQECFANVGCREPAARLQGTLEGLRKIAQLEGLPTLWR 313
 K+FW+ CF + C+ + + GTLE KIA +EG+ +LWR
 64 GSKMKTFTNVGGQNLNNAKIPWESACFQELHCKNSSLKFNGTLEAFTKIASVEGITSCLR 123

 312 GLGITLVMAMPANVVYFSGYEALEDNSPLASRLPVANPLVCGAFARILAATTIAPLELLR 133
 G+ +TL+MA+PAN+VYFSGYE +RD SP+AS P NPL CGA AR+ AAT+IAPLEL++
 124 GISLTLMLAIPANMVYFSGYBYIRDVSPIASTYPTLNPLFCGAIARVFAATSIAPLELVK 183

 132 TRLQSVPRARDTERTIYLGDLLREMRHEVSVMG-YRALFKGLEI 1 棉桃阿舒氏囊霉
 T+LQS+PR+ + +T ++ DLL E R E+ ++G RALFKGLEI
 184 TKLQSIPIRSSKSTKTWMMVKDLLNETRQEMKMGVPSRALFKGLEI 228 酿酒酵母

图2

Oligo 24

1494 ILKILRLTITPGSNVCIVGPSGRGKSTIASLLLRFY 1387 棉桃阿舒氏囊霉

I K L I PGS+VCIVGPSGRGKSTIA LLLR+Y

510 IFKILNFKIAPGSSVCIVGPSGRGKSTIALLLRYY 545 酿酒酵母

图3

15 KTAQAWELIELLSQNDVVKYGDIVFRPLFKSSPEAGLLELEKNGLITISRNRGVLQDIRP 194
 K+AQAWELIELLS N V+ + +IV +PLFK++PE G++ELE NGLIT+SR+RGVLQ+IRP
 704 KSAQAWELIELLSANPVIPFHEIVNKPLFKAAPETGIMELENNGLITVSRDRGVLQEIRP 763

 195 AKPLFKAAFSYLLQDKDLSIVLRTGYLRLIAFETGRIKKWEEELRLAKVSDQRIKSR 374
 AKPL++AAF+YL+ D +L+ VL+T Y L+++ FETGRIKKWEEEL+ L KV DQ++ K+R
 764 AKPLYRAAFTYLINDPELAKVLKTRYLLKVVGFTGRIKKWEEELKPLGKVPDQKLFKTR 823

 375 LNYLASKIDASSGVINSCEDKVKEMSK 455 棉桃阿舒氏囊霉
 L+YL+ KI+AS+ VI CE+++K +SK
 824 LDYLSGKINASNAVITKCEBBIKNLSK 850 酿酒酵母

图4

Oligo163

415 HELTSLQREVISGLTAGTITTTIASHPLDLLKLRLQLSAGNRANTTYTG--LIRDIFERQ-
 H+ T LQ+EVISGL+AG++TT+ HPLDLLK+RLQLSA + Y +I++I
 4 HQWTPQLKEVISGLSAGSVTTLVVHPLDLLKVRQLSATSQAQKHYGPFMVIKEIRSSA 63

 QWGR----ELYRGLGVNLLGNSVAWALYFGCYRCAKDIALRHLGNESATGIM----DRRL
 GR ELYRGL +NL GN++AW +YFG Y K++ + + T + D ++
 64 NSGRSVTNELYRGLSINLFGNAIANGVYFGLYGVTKELIYKSVAKPGETQLKGVGNDRHM 123

 PAHAYMLAAGSSGIATAVLTNPIWIKTRIMATSRA--GPYKSTFDGVYKLYQTEGVLA
 + Y+ A SSG+ TA+LTNPIWIKTRIM+TS+ G Y S ++GV +L +T+G
 124 NSLIYLSAGASSGLMTAILTNPIWIKTRIMSTSKGAQAYTSMYNGVQQLLRDGFQGL 183

 WRGVVPSLLGVSQGAIFYFALYDTLKFHYLHSSTDKA-ERRLSVSEIIGITCISKMISVTS
 W+G+VP+L GVSQGA+YFA+YDTLK L + + L+ E I IT + KM+SVT
 184 WKGLVPALFGVSQGALYFAVYDTLKQRKLRKRENGLDIHLTNLETIETSLGKMVSRTL 243

 VYFPQLLKSKLQDFGAPS---GITQLVQTVYSREGIRGFYRGLS 1215 棉桃阿舒氏囊霉
 VYFPQLLKS LQ F A + L++ + + +G G Y+GLS
 244 VYFPQLLKS NLQSFRANEQKFRLEPLIKLIANDGFVGLYKGLS 酿酒酵母

图5

Oligo 31

2 IDIVGNFIPGLIFMQSIFGYLSWAIYKWSKDWDKDELPAPGLLNMLINMFLSPGVVDEK 181
 +DI+GNFIPGL+FMQSIFGYLSWAI+YKWSKDWDK+ PAPGLLNMLINMFL+PG +D++
 612 VDIIGNFIPGLVFMQSIFGYLSWAIYKWSKDWDKDDKPAPGLLNMLINMFLAPGTIDDQ 671

 182 LYTGQSFLQVILLALLVCVPWLLLYKPLMLKRQNDIA---LSKGFRSLRDQRVHEILLE 352
 LY+GQ+ LQV+LLLAALVCVPWLLLYKPL L+R N G++S+ + E + +
 672 LYSGQAKLQVLLALLVCVPWLLLYKPLTLRRLNKNKGGRPHGYQSVGNIEHEEQIAQ 731

 353 AQENAG--EDMLVAXXXXXXXXXXXXXXG-----DVMIHQVIHTIEFCLNCISHTA 499
 + +A + M+++ G DVMIHQVIHTIEFCLNCISHTA
 732 QRHSABGFQGMII SDVASVADSINESVGGGEQGFNF GDVMIHQVIHTIEFCLNCISHTA 791

 500 SYLRLWALSLAHAQLSTVLNSMTIQNSFSDSNPGSFFSVTKVVVLFAMWFVLTVCILVLM 679
 SYLRLWALSLAHAQLS+VLW MFI N+FS N GS +V KVV LFAMWFVLTVCILV M
 792 SYLRLWALSLAHAQLSSVLNDMTISNAFSSKNSGSPLAVMKVVFLFAMWFVLTVCILVFM 851

 680 EGTSAMLHSLRLHWVEAMSKFFE GEGYAYEPFSFKAI 790 棉桃阿舒氏囊霉
 EGTSAMLH+LRLHWVEAMSKFFE GEGYAYEPFSF+AI
 852 EGTSAMLHALRLHWVEAMSKFFE GEGYAYEPFSFRAI 888 酿酒酵母

图6

Oligo 4

869 P*QRFKFTCPGIRKYVNTSFM*DSIDPKRISEHFYPIMKWCNLIVACEGLFFGLWKLSL 690

P Q KFTCP Y N S + +I PKRI + YPI KWCNLI AC G+FFG+WK+

698 FHQNAKFTCPDAVTYYNASVVWGAIGPKRIFNYIYPIFKWCNLIGACIGIFFGVNKRNGK 757

689 MSEIYLILTGLVELFNNGMIIAVSSYTPQYNTSGLCVGFINMLYMRRYYL-IWEKFN 522 (i)

Y LF GM+ Y Y TSG+ V +I+ YM+R++L +WEK+N

758 FYPRYFD----PMLFVGGLNMSPPYNLMYYTSGMIVSYISQYYMKRHHLLNLWEKYN 810 (ii)

(i) 棉桃阿舒氏囊霉

(ii) 酿酒酵母

图7A

Oligo 6

356 ESVECPVCNRNIENLEQLNMHLDVEHNFKD-ERSL-STDS 棉桃阿舒氏囊霉
E+V CP+C R +NL+ LN HLDVEH F D E SL ++
4 ENVSPICLRKFDNLQALNAHLDVEHGFNDNEDSLGSNDS 酿酒酵母

图7B

166 RSHYKPLVPKKSQCSRCGRKLTFRGTGIONCRRCGDLFCDIHCRLPVRLDLQGEYD 2 棉桃阿舒氏囊霉
RSH++ GKS C CGR L G NCR+CG L+C H ++L+L +YD
65 RSHWEKFKKGKSCCHTCGRITLNNNIGAINCRKCGKLYCRRHLPNMIKLNLQAQYD 119 酿酒酵母

图8

Oligo 146

904 SSATYNTTEEYEQLP SGKLIK RKNLV PDNSGGSGNDGFAAKLGRK LKL---KPAAEDIDEI 734
+S+ ++E YE LPSGKLIK R +L +S S KL +KLKL ED+++I
1588 NSSKKSSERYEVLPSGKLIK RPSLKTQSSKDSIGGNITTKLT KKLKLPSRNV EDEDVNQI 1647

733 IEREMQNLE 707 棉桃阿舒氏囊霉
I+ R+++LE
1648 IQARLKOLE 1656 酿酒酵母

图9

Oligo 56

898 SSVVLIPTRTLVTLLNKSALSRSFSCVDEILGIIPTLGDPDINGRNFFHHHVIALGK-- 725
 S + LIPTR ++++LNK+ALS+S +C+D IL +IP+L D DIN RNFFHHH+IA+GK
 292 SELSLIPTRIMISVLNKAALSKSLACIDAILKVIPSLNDSSEINRRNFFHHHIIAIGKLI 351

 724 KRTKSLEERDNINXXXXXXXXLEAAIPPEP-NTRLVGAFGPDGVNSDD--SPAPLSHILQ 554
 ++ + L + E +T L D + ++ S LS+IL+
 352 RKQEILSRKKKSQPSKYTNSEGEIVTDLRTLHTTLSAPAESDSITEEEKSSACTLSYILE 411

 553 QLPAPHLRPSLLQORDNYKRTPLHYSAQYGLCEXTRIILQALSEWDANNADVAIDDDIDVWGD 374
 +LP HLRP L Q DNYKRTPLHYS QYGL E T++I++ + EW+ WN ++ IDD+ +GD
 412 ELPIHLRPLCLFQHDNYKRTPLHYSCQYGLSEVTKLI IKLMKEWNIWN-EIPIDDVSAFGD 470

 373 SENLTPLHLAVIGTHPLTVSTLLSFMNPEKSLNSPRLLHLATRXXXXXXXXXXXSAGGFD 194
 +E+LTPLHL V+G HP T LL ++P L S LLHLAT S+K PD
 471 AESLTPLHLCVLGAHPKTTTEVLLQSLDPNVKLKSSSLHLATEWNNYPPLLHVLLSSKRFD 530

 193 IDYQEPENLETALYVACKLDIYEAAYLVKQGANMELGEKLFGWTPIFAAATEGYARIVQ 14
 I+YQ+ E ET LY+AC+L+ +EAA L+ GA++E+ EKLPGWTF IF AA EG+ IV+
 531 INYQDNELHETPLYLACRLNFFFAAVCLLYNGADLEIREKLPGWTAIFVAAABGFTDIVK 590

 13 LLV 5 棉桃阿舒氏囊霉
 LL+
 591 LLI 593 酿酒酵母

图10

Oligo 167

931 DASESANSWVWCAVMTL FVLGGTCFWQVNHRLRTFFMKQKVL 806 棉桃阿舒氏囊霉

DASE+ NS +W V+ L VL TC WQ+ HL FF+KQK+L

178 DASEAVNSRAMWWIVIQLIVLAVTCGWQMKHLGKFFVKQKIL 219 酿酒酵母