

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公开说明书

[21] 申请号 200480031689.9

[51] Int. Cl.

C12N 15/75 (2006.01)

C12R 1/125 (2006.01)

C12P 21/02 (2006.01)

[43] 公开日 2006年12月6日

[11] 公开号 CN 1875106A

[22] 申请日 2004.11.5

[21] 申请号 200480031689.9

[30] 优先权

[32] 2003.11.7 [33] JP [31] 379167/2003

[86] 国际申请 PCT/JP2004/016891 2004.11.5

[87] 国际公布 WO2005/045013 英 2005.5.19

[85] 进入国家阶段日期 2006.4.27

[71] 申请人 花王株式会社

地址 日本东京

[72] 发明人 东畑正敏 泽田和久 尾崎克也

小林和夫 小笠原直毅

[74] 专利代理机构 北京纪凯知识产权代理有限公司

代理人 龙 淳

权利要求书 2 页 说明书 16 页 序列表 48 页
附图 1 页

[54] 发明名称

重组微生物

[57] 摘要

一种重组微生物，其通过将编码蛋白质或多肽的基因转移到能够以提高的生产率生产蛋白质或多肽的宿主微生物上而获得，以及使用该重组微生物生产蛋白质或多肽的方法。该重组微生物通过将编码异性蛋白质或多肽的基因转移到枯草杆菌基因 *comA*、*yopO*、*treR*、*yvbA*、*espB*、*yvaN*、*yttP*、*yurK*、*yozaA*、*licR*、*sigL*、*mntR*、*glcT*、*yvdE*、*ykvE*、*slr*、*rocR*、*ccpA*、*yaaT*、*yyaA*、*yycH*、*yacP*、*hprK*、*rsiX*、*yhdK* 以及 *ylbO* 的任意一种或一个或多个与这些基因的任意一种功能相当的基因缺失或被剔除的微生物的突变株上而制备。

1. 一种重组微生物，其通过将编码异性蛋白质或多肽的基因转移到枯草杆菌基因 *comA*、*yopO*、*treR*、*yvbA*、*cspB*、*yvaN*、*yttP*、*yurK*、*yoza*、*licR*、*sigL*、*mntR*、*glcT*、*yvdE*、*ykvE*、*slr*、*rocR*、*ccpA*、*yaaT*、*yaaA*、*yycH*、*yacP*、*hprK*、*rsiX*、*yhdK* 以及 *ylbO* 的任意一种或一个或多个与这些基因的任意一种功能相当的基因缺失或被剔除的微生物的突变株上而制备。

2. 如权利要求 1 所述的重组微生物，其中所述微生物是枯草杆菌或属于芽孢杆菌属的另一细菌。

3. 如权利要求 1 或 2 所述的重组微生物，其中一个或多个选自转录起始调节区域、转译起始调节区域和分泌信号区域的区域接合到编码异性蛋白质或多肽的基因的上游区域。

4. 如权利要求 3 所述的重组微生物，其中所述一个或多个区域是由转录起始调节区域、转译起始调节区域和分泌信号区域构成的三个区域。

5. 如权利要求 3 或 4 所述的重组微生物，其中所述分泌信号区域来自属于芽孢杆菌属的细菌的纤维素酶基因，而转录起始调节区域和转译起始调节区域各自来自该纤维素酶基因上游的 0.6 至 1kb 区域。

6. 如权利要求 4 所述的重组微生物，其中由所述转录起始调节区域、所述转译起始调节区域和所述分泌信号区域构成的三个区域是序列号：1 的纤维素酶基因的碱基数为 1 至 659 的核苷酸序列；序列号：3 的纤维素酶基因的碱基数为 1 至 696 的核苷酸序列；含有与上述核苷酸序列的任意一种具有 70%同源性的核苷酸序列的 DNA 片段；或含有缺失了上述核苷酸序列的任意一种的一部分的核苷酸序列的 DNA 片段。

7. 一种生产蛋白质或多肽的方法，其使用如权利要求 1 至 6 任意一项所述的重组微生物。

重组微生物

技术领域

本发明涉及可用于生产有用蛋白质或多肽的重组微生物，以及涉及该蛋白质和多肽。

背景技术

微生物广泛用于工业生产大量有用物质，包括酒精饮料、例如味噌和酱油之类的某些类型的食品、氨基酸、有机酸、核酸相关物质、抗生素、糖、脂质和蛋白质。这些物质还发现具有多种用途，包括食品、医药品、洗涤剂、例如化妆品之类的日用品和各种化工原料。

在利用微生物工业生产有用物质时，生产率的提高是一个主要关注点，因此一种方法是通过诱变或其它基因手段进行微生物育种。最近，特别是随着微生物遗传学和生物技术的进步，通过基因重组技术进行更有效的有用微生物育种，并与其相关地是正在开发用于获得重组基因的宿主微生物。例如，已经进一步改进了枯草杆菌 Marburg No. 168，其作为宿主微生物已经被证实是安全且具有优异特性的。

然而，微生物本身拥有各种基因使其能够应对自然界中的环境变化，因此，它们在只采用有限的生产媒介的工业生产中不一定表现出蛋白质和类似物质的高生产效率。

发明内容

本发明提供一种重组微生物，其通过将编码异性蛋白质或多肽的基因转移到枯草杆菌基因 *comA*、*yopO*、*treR*、*yvbA*、*cspB*、*yvaN*、*yttP*、*yurK*、*yoza*、*licR*、*sigL*、*mntR*、*glcT*、*yvdE*、*ykvE*、*slr*、*rocR*、*ccpA*、*yaaT*、*yyaA*、*yycH*、*yacP*、*hprK*、*rsiX*、*yhdK* 以及 *yIbO* 的任意一种或一个或多个与这些基因的任意一种功能相当的基因缺失或被剔除的微生物的突变株上而制备。

附图说明

图 1 示意性地显示了通过 SOE-PCR (SOE: 重叠延伸拼接法) (参见 Gene, 77, 61 (1989)) 缺失基因以制备 DNA 片段的方法, 和使用该 DNA 缺失目标基因的方法 (用抗药性基因取代目标基因)。

具体实施方式

本发明针对一种通过将编码蛋白质或多肽的基因转移到能够以提高的生产率生产蛋白质或多肽的宿主微生物上而获得的重组微生物, 以及针对使用该重组微生物生产蛋白质或多肽的方法。

本发明人在微生物基因组上编码的许多不同基因中对有用蛋白质或多肽的生产中不需要或有害的基因进行了广泛的研究, 并发现, 当特异性基因从微生物的基因组中缺失或剔除之后, 编码目标蛋白质或多肽的基因转移到例如枯草杆菌之类的微生物上, 与缺失或剔除之前的情况相比, 目标蛋白质或多肽的生产率提高。

在本发明的微生物中, 由于缺失或剔除了对目标蛋白质或多肽的生产不必要或有害的基因, 因此包括能量损耗、副产物生成和单位生产率降低在内的培养基的浪费显著减少, 并且此外, 蛋白质和多肽可以持续生产, 从而高效地生产目标产物。

在本发明中, 氨基酸序列之间的同源性和核酸序列之间的同源性都使用 Lipman-Pearson 法 (Science, 227, 1435 (1985)) 测定。具体而言, 将 ktup 参数 (相匹配的单元大小) 设定为 2, 使用 Genetyx-Win (Software Development Co., Ltd.) 的遗传信息处理软件开发的同源性分析程序 (Search Homology) 进行计算。

对构成本发明的微生物的母体微生物没有特别的限制, 只要其含有对于目标蛋白质或多肽的生产不是必需的基因。具体而言, 可以使用如表 1 所示的任何枯草杆菌基因或与其功能相当的基因, 其中基因可以是野生型或突变体。具体例子包括枯草杆菌和属于芽孢杆菌属的类似微生物、属于梭状芽孢杆菌属的微生物和酵母。其中, 优选属于芽孢杆菌属的微生物。特别是从已经获得该微生物的完全基因组信息、并由此确立了基因工程技术和基因组工程技术、并且该微生物具有细胞外分泌所得蛋白质的能力的角度考虑, 优选枯草杆菌。

使用本发明的微生物生产的目标蛋白质或多肽的例子包括酶、生

理活性物质，和其它发现在食品、医药品、化妆品、洗涤剂、纤维处理剂、临床化验剂等有用的蛋白质和多肽。

以已知在基因组上含有 4,106 个基因的枯草杆菌为例，要被缺失或剔除的一个或多个基因是表 1 所示的任一枯草杆菌基因，或者选自与其功能相当的基因。本发明人已经发现，这些基因不会直接参与目标蛋白质或多肽的生产并且对于微生物在普通工业生产培养基中的生长不是必需的。

此处所含表中的各个基因的名称、数量和功能与 Nature, 390, 249-256 (1997) 中报道并由 JAFAN (Japan Functional Analysis Network for *Bacillus subtilis*; BSORF DB) 在互联网上公开 (<http://bacillus.genome.ad.jp/>, 2003 年 6 月 17 日更新) 的枯草杆菌基因组数据相符。

表 1

基因名称	基因编号	该基因的功能或其它信息
<i>comA</i>	BG10381	两元组分反应调节因子
<i>yopO</i>	BG13648	推导 (deduced) 的转录调节因子, sp β 原噬菌体蛋白质
<i>treR</i>	BG11011	海藻糖操纵子转录阻遏物(GntR 族)
<i>yvbA</i>	BG14078	推导的转录调节因子(ArsR 族)
<i>cspB</i>	BG10824	冷休克-相关的主要因子
<i>yvaN</i>	BG14069	推导的转录调节因子
<i>yttP</i>	BG13927	推导的转录调节因子 (TetR 族)
<i>yurK</i>	BG13997	推导的转录调节因子 (GntR 族)
<i>yoza</i>	BG13748	推导的转录调节因子 (ArsR 族)
<i>licR</i>	BG11346	转录调节因子 (抗终止子), 地衣淀粉操控子 (LicBCAH) 调节
<i>sigL</i>	BG10748	RNA 聚合酶 σ 因子 ($\sigma 54$)
<i>mntR</i>	BG11702	锰转运调节因子
<i>glcT</i>	BG12593	对于 <i>ptsGHI</i> 操纵子的表达必不可少的转录调节因子 (BglG 族, 抗终止子)
<i>yvdE</i>	BG12414	推导的转录调节因子 (LacI 族)
<i>ykvE</i>	BG13310	推导的转录调节因子 (MarR 族)
<i>slr</i>	BG11858	感受态相关基因或孢子形成相关基因的转录激活剂

<i>rocR</i>	BG10723	精氨酸同化操纵子的转录激活剂 (NtrC 族)
<i>ccpA</i>	BG10376	碳源代谢阻遏相关的转录调节因子 (LacI 族)
<i>yaaT</i>	BG10096	II 型信号肽酶样蛋白质
<i>yyaA</i>	BG10057	DNA-结合蛋白 SpoOJ 样蛋白质
<i>yycH</i>	BG11462	功能未知 (同源基因在其它生物体中发现)
<i>yacP</i>	BG10158	功能未知 (同源基因在其它生物体中发现)
<i>hprK</i>	BG14125	Hpr 蛋白质 Ser 残基含磷酶 (phosphoenzyme) / 去磷酶 (dephosphoenzyme)
<i>rsiX</i>	BG10537	抗 σ_X 因子
<i>yhdK</i>	BG13017	功能未知, 与 σ_M 因子表达的阻遏有关
<i>ylbO</i>	BG13367	σ_E 相关的 metrocyein 中基因的表达调节因子

来源于其它微生物的基因, 优选来源于属于芽胞杆菌属的细菌的基因, 其具有与表 1 所示的任一枯草杆菌相同的功能, 或具有与表 1 所示任一基因的核苷酸序列 70%或更高的同源性, 优选 80%或更高的同源性, 更优选 90%或更高, 进一步优选 95%或更高, 仍然更优选 98%或更高的同源性的, 应该被认为与表 1 所示的基因功能相当, 并因此构成根据本发明要缺失或剔除的基因。在这一方面, 使用 Lipman-Pearson 法 (Science, 227, 1435, 1985) 用计算机计算核苷酸的同源性。

表 1 所示的编码枯草杆菌的多种基因是参与各种基因表达的活化或抑制的调节基因, 或被推导成这种调节基因的基因。基于此发现实现了本发明; 即, 在本发明中揭示了对蛋白质或多肽的生产不必要或有害的调节基因的存在。

值得注意地是下列事实已经引起了注意, 即所列的许多“不必要”或“有害”基因是参与糖摄取或代谢的调节基因, 以下列基因为例: *glcT* 基因, 其对葡萄糖 PTS 摄取操纵子起到抗终止子的作用; *LicT* 基因, 其对地衣淀粉水解操纵子起到抗终止子的作用; *treR* 基因, 其起到海藻糖摄取和代谢的阻遏物的作用; 和 *hprk* 基因和 *ccpA* 基因, 其关系到葡萄糖分解代谢物的阻遏。

而且, 除了参与糖摄取和代谢的调节基因外, 可以缺失或剔除参与激活精氨酸同化的 *rocR* 基因、和感受态相关的 *comA* 基因和 *slr* 基因, 它们也是调节基因, 从而提高蛋白质或多肽的生产率。

表 1 所示的基因包括 *yhdK* 基因, 以及编码抑制 ECF σ (σ) 因子表达的抗 ECF σ 因子, 即抗 σ X 因子的 *rsiX* 基因。有报道 *yhdK* 参与了 σ M 的抑制 (Mol. Microbiol., 32, 41, 1999)。编码 σ L 的 σ 基因也包括在表 1 的基因中。这表明, 在 σ X 或 σ M 调节下的基因表达有利于蛋白质生产, 相反, 在 σ L 调节下的一些基因表达是不利的。

通过缺失或剔除一个或多个选自上述基因的基因, 可以防止对蛋白质或多肽的生产不必要的或有害的表达, 从而使该蛋白质或多肽生产中的生产率提高。

要缺失或剔除的基因数为一个或更多, 优选两个或更多, 更优选三个或更多, 甚至更优选 5 个或更多。当构筑本发明的微生物时, 可以缺失或失活上述基因以外的一个或多个基因。在这种情况下, 预计会有更好的效果。另一种实现本发明的方法是通过将另一来源的 DNA 片段插入到目标基因中或在该基因的转录/转译起始区域引入突变体来使目标基因失活或剔除。然而, 优选物理缺失目标基因。

在缺失或剔除基因的示例性的方法中, 根据之前提出的方案缺失或剔除表 1 所示的任一目标基因。或者, 进行基因的随机缺失或通过剔除进行突变, 然后对蛋白质生产率进行评测并进行基因分析。

目标基因可以通过同源重组而缺失或剔除。也就是, 用合适的质粒载体克隆含有部分目标基因的 DNA 片段从而获得环形重组质粒, 并将所得质粒转移到母体微生物的细胞中。此后, 通过在目标基因的部分区域中进行同源重组, 使母体微生物基因组上的目标基因裂解, 从而完成目标基因的失活。或者, 通过碱基的取代或插入来使目标基因剔除, 或者通过 PCR 或类似方法构成含有目标基因序列外的区域但不含目标基因的线型 DNA 片段, 并将由此制成的基因或片段转移到母体微生物的细胞中。在母体微生物基因组的目标基因中突变点外部的两个位点上, 或在目标基因序列外的两个区域上, 发生双交换同源重组, 从而替换成基因组上的目标基因缺失或剔除的基因片段。

特别是当用于构筑本发明的微生物的母体微生物是枯草杆菌时, 由于一些文献已经描述了缺失或剔除目标基因的方法 (参见, 例如, Mol. Gen. Genet., 223, 268 1990), 可以根据这些方法的任何一种重复, 从而生产本发明的宿主微生物。

基因随机缺失或失活可以通过使用与上述通过使用随机克隆的 DNA 片段诱发同源重组的方法类似的方法，或通过用 γ 射线或类似射线辐射母体微生物而进行。

接下来将更详细地描述通过使用为缺失目的而设计的 DNA 片段采用双交换的缺失方法，该 DNA 片段是通过 SOE-PCR 制备的 (Gene, 77, 61, 1989)。然而，在本发明中，缺失基因的方法不仅仅限于下述方法。

用于缺失目的的 DNA 片段是如下构成的片段：抗药标记基因在 ca. 0.5 至 3kb 上游序列的侧面之间插入并且处于要缺失的基因的上游，以及抗药标记基因在 ca. 0.5 至 3kb 下游序列的侧面之间插入并且处于相同基因的下游。在 PCR 的第一次循环中，制备下列三个片段：要缺失的上游片段和下游片段，和抗药标记基因。例如，要在该步骤中使用的引物可以是专门设计的引物，以便将抗药性基因的上游 10-30 碱基对序列加入到上游片段的下端，以及将抗药性标记基因的下游 10-30 碱基对序列加入到下游片段的上端 (图 1)。

接下来，使用在第一次循环中制备的三个 PCR 片段作为模板，使用上游片段的上部引物和下游片段的下部引物进行 PCR 的第二次循环。该步骤导致与如上制成的抗药性基因序列中的抗药性标记基因片段退火接合，并通过 PCR 扩增，可以获得在上游片段和下游片段之间插有抗药性标记基因的 DNA 片段 (图 1)。

当采用抗氯霉素基因作为抗药性标记基因时，可以使用合适的 DNA 模板和如表 2 所示的引物组和传统的 PCR 酶试剂盒 (例如 Pyrobest DNA 聚合酶 (Takara Shuzo 的产品) 在文献中描述的典型条件下 (参见，例如，PCR Protocols. Current Methods and Applications, Edited by B. A. White, Humana Press, pp. 251 (1993), Gene, 77, 61, 1989) 通过 SOE-PCR 获得用于缺失基因的 DNA 片段。

当由此获得的用于实现基因缺失的 DNA 片段通过转化方法 (competent method) 或类似方法引入细胞中时，在要缺失的基因上游和下游存在的同源区域中发生细胞内基因重组。因此，其中的目标基因已经被抗药性基因取代的细胞可以通过使用抗药性标记物而选择性地分离 (图 1)。具体而言，当在细胞中引入通过表 2 所列的引物组而

制备的用于基因缺失的 DNA 片段时，分离出在含有氯霉素的琼脂培养基上生长的菌落，并通过例如使用基因组作模板的 PCR 之类的合适方法来证实通过被抗氯霉素基因取代而缺失了目标基因。

随后，当将编码目标蛋白质或多肽的基因转移到表 1 所示的任一枯草杆菌基因或一个或多个选自与其相对应的基因的基因从中缺失或剔除的宿主突变微生物菌株上时，可以获得本发明的微生物。

对编码目标蛋白质或多肽的基因没有特别的限制。蛋白质和多肽的例子包括用于工业用途的例如洗涤剂、食品、纤维、饲料、化学品、药品和诊断用药之类的生理活性多肽和酶。工业酶从功能上可以分成氧化还原酶、转移酶、水解酶、裂合酶、异构酶和连接酶/合成酶。优选地，可以使用例如纤维素酶、 α -淀粉酶和蛋白质酶之类的水解酶。具体例子包括属于水解酶分类中的族 5 的纤维素酶(Bioche M.J., 280, 309, 1991); 特别是源自微生物的纤维素酶，更特别是源自芽孢杆菌属的纤维素酶。这些类型的工业酶的其他具体例子包括源自芽孢杆菌属并含有氨基酸序列号：2 或 4 的碱性纤维素酶，以及含有与所述氨基酸序列具有 70%同源性，优选 80%同源性，更优选 90%同源性，进一步优选 95%同源性，又进一步优选 98%或更高同源性的另一氨基酸序列的纤维素酶。

α -淀粉酶的具体例子包括源自微生物的 α -淀粉酶，优选源自芽孢杆菌属的液化淀粉酶。更具体的例子包括源自芽孢杆菌属并含有序列号：6 的氨基酸序列的碱性淀粉酶，以及含有与所述氨基酸序列具有 70%同源性，优选 80%同源性，更优选 90%同源性，进一步优选 95%同源性，特别优选 98%或更高同源性的另一氨基酸序列的淀粉酶。氨基酸序列的同源性通过 Lipman-Pearson 法 (Science, 227, 1435 (1985)) 计算。蛋白酶的具体例子包括源自微生物，特别是源自属于芽孢杆菌属的微生物的丝氨酸蛋白酶和金属蛋白酶。

优选地，编码目标蛋白或多肽的基因在其上游区域具有一个或多个适当地接合到其上的与基因的转录、转译或分泌有关的调节区域(具体而言，一个或多个选自包含启动子和转录起始位点的转录起始调节区域；包含核蛋白体结合位点和起始密码子的转译起始区域；和分泌信号肽区域的区域)。优选将由转录起始调节区域、转译起始调节区域

和分泌信号区域构成的三个区域接合到目标基因上。更优选分泌信号肽区域是源自属于芽孢杆菌属的微生物的纤维素酶基因，且转录起始区域和转译起始区域是该纤维素酶基因上游的 0.6 至 1kb 区域。在一个优选例子中，在例如日本专利申请（公开）第 2000-210081 号和第 190793/1990 号中公开的源自属于芽孢杆菌属的微生物的纤维素酶基因；即，源自 KSM-S237 菌株 (FERM BP-7875) 或 KSM-64 菌株 (FERM BP-2886) 的纤维素酶基因的转录起始调节区域、转译起始区域、和分泌信号肽区域，适当地接合到目标蛋白或多肽的结构基因上。更具体地，要接合的优选 DNA 片段包括序列号：1 的碱基数量为 1 至 659 的核苷酸序列；序列号：3 的纤维素酶基因的碱基数量为 1 至 696 的核苷酸序列；含有与上述任意一种核苷酸序列具有 70% 同源性，优选 80% 同源性，更优选 90% 同源性，进一步优选 95% 同源性，甚至更优选 98% 或更高同源性的核苷酸序列的 DNA 片段；或含有缺失了上述任意一种核苷酸序列的一部分的核苷酸序列的 DNA 片段。优选将这些 DNA 片段的一种适当地接合到目标蛋白质或多肽的结构基因上。正如此处所用，含有缺失上述任意一种核苷酸序列的一部分的核苷酸序列的 DNA 片段是指不含上述任意一种核苷酸序列的一部分并具有与该基因的转录、转译和分泌有关的功能的 DNA 片段。

本发明的重组微生物可以通过传统的转化技术获得，其中将含有包括编码目标蛋白质或多肽的基因并接合到适当的质粒载体上的 DNA 片段的重组质粒转移到宿主微生物细胞中。或者，重组微生物可以利用通过将上述 DNA 片段接合到与宿主微生物基因组的某一部分同源的一个合适区域上而制备的 DNA 片段，并直接插入宿主微生物基因组中而获得。

使用本发明的重组微生物获得的目标蛋白质或多肽可以按照如下方式生产：将相应的细胞株植入含有可同化的碳源和氮源以及其它主要组分的培养基上；通过传统的微生物培养法培养细胞株；随后，收集并提纯蛋白质或多肽。

通过前述方法，可以生产宿主突变微生物株，其中表 1 所示的任一枯草杆菌基因或一个或多个选自与其功能相当的基因的基因已经缺失或剔除。此外，通过使用这种突变株，可以生产重组微生物。由此，

通过使用突变株或重组微生物可以有效地生产有用的蛋白质或多肽。

下面将以构筑属于枯草杆菌的已经从中缺失了枯草杆菌的 *ccpA* 基因 (BG10376) 的重组菌株的实施例为中心, 详细描述根据本发明构筑重组微生物的方法, 以及使用该重组微生物生产纤维素酶和 α -淀粉酶的方法。

实施例

实施例 1

使用从枯草杆菌 168 菌株中提取的用作模板的基因组 DNA 样品和表 1 和 2 所示的两个引物组 (*ccpA*-AF 和 *ccpA*-A/CmR; 和 *ccpA*-B/CmF 和 *ccpA*-BR) 制备在该基因组上的 *ccpA* 基因上游侧侧面的 0.6kb 片段 (A) 和在 *ccpA* 基因下游侧侧面的 0.6kb 片段 (B)。将质粒 pC194 的抗氯霉素基因 (J. Bacteriol. 150 (2), 815 (1982)) 插入质粒 pUC18 的 XbaI-BamHI 裂解位点, 由此制备重组质粒 pCBB 31。使用重组质粒 pCBB 和表 2 所示的由 CmF 和 CmR 构成的引物组制备含有抗氯霉素基因的 1kb 片段 (C)。随后, 使用表 2 所示的引物 *ccpA*-AF 和 *ccpA*-BR 进行 SOE-PCR, 并结合使用由此制成的三个片段 (A)、(B) 和 (C) 作为模板, 制备 2.2kb DNA 片段, 其中在该序列中接合了片段 (A)、(B) 和 (C) (参见图 1)。使用由此制成的 DNA 片段, 通过转化方法转化枯草杆菌 168 菌株。收集在含氯霉素的 LB 琼脂培养基中生长的菌落作为转化体。提取上述转化体的基因组, 在其上进行的 PCR 证实, *ccpA* 基因已经被缺失并被抗氯霉素基因取代。

表 2-1

引物	核苷酸序列	序列号:
comA-AF	AAGGATGATAATCCGTCCCCTG	7
comA-A/CmR	GTTATCCGCTCACAATTCGGATGGTCATCAATCACTAG	8
comA-B/CmF	CGTCGTGACTGGGAAAACCTGCCAAATCAGACGGTGTAC	9
comA-BR	CGTCGCCTATCCGGCGGGCAC	10
yopO-AF	ATGTATATAGGAGGTTGGTGGTATG	11
yopO-A/CmR	GTTATCCGCTCACAATTCGCTCTGACATGTCAACCTCC	12
yopO-B/CmF	CGTCGTGACTGGGAAAACAGATGAGAAAGGAGGAGAAG	13
yopO-BR	ATAACTGTTACTATAATAATGGCC	14
treR-AF	GCTGGGGATGACGAATCCGA	15
treR-A/CmR	GTTATCCGCTCACAATTCACCTTCATTATGGACCAC	16
treR-B/CmF	CGTCGTGACTGGGAAAACACCGTCTCGACAAATTCGG	17
treR-BR	GTTGCCAAGCCGATATAGG	18
yvbA-AF	TATACAGGGATTATCAGTATTGAGC	19
yvbA-A/CmR	GTTATCCGCTCACAATTCCTTTCTCCTTGTGGATCTG	20
yvbA-B/CmF	CGTCGTGACTGGGAAAACGGGATAACGATTTATGAAG	21
yvbA-BR	TTTTGTAATAATGATATGAAGCTAGTGTG	22
espB-AF	ATATCCAGCCCTGCCTCTC	23
espB-A/CmR	CTGTGTGAAATTGTTATCCGCTCACAATTCGAAATTCCTCCTAA AGCGATCATAACG	24
espB-B/CmF	GTCGTTTTACAACGTCGTTGACTGGGAAAACCCACAAGCTGCTAA CGTTAC	25
espB-BR	TCCGTGTTGGGCTCCTGTG	26
yvaN-AF	TGTTTATGTATGGCGCCTGCGGGAC	27
yvaN-A/CmR	GTTATCCGCTCACAATTCAGCTTCCATATATCTCACC	28
yvaN-B/CmF	CGTCGTGACTGGGAAAACACGGTCTGCTGATGACTGAC	29
yvaN-BR	GCGTTTACTTAAGATGTCCA	30
yttP-AF	TTTCTAGCGTTTCGGCAAATTGAGTTAAG	31
yttP-A/CmR	GTTATCCGCTCACAATTCCTTACTTTTCATACGGCTCAC	32
yttP-B/CmF	CGTCGTGACTGGGAAAACGAGACGTGGCGCTCACCAAC	33
yttP-BR	CGGATTAATAAAGAATATCGCGGACAGC	34
yurK-AF	TGCCGCTGCCCGCCGGAGAG	35

表 2-2

yurK-A/CmR	GTTATCCGCTCACAATTC AAGGTG TAGAACTTCCGTTG	36
yurK-B/CmF	CGTCGTGACTGGGAAAACACCATCAACAGCCCCTACAC	37
yurK-BR	TCAAATAAAGGCGGCATT CAGTCC	38
yoza-AF	ATAATGGTATCCAAATCCACGC	39
yoza-A/CmR	GTTATCCGCTCACAATTCATTCAGTCATATGTATCACC	40
yoza-B/CmF	CGTCGTGACTGGGAAAACGATCCATCATACACAGCATG	41
yoza-BR	CACTTCTCAACGGAGGGGATTT CACATC	42
licR-AF	TAATGGAGGAGAGAAGGCCG	43
licR-A/CmR	GTTATCCGCTCACAATTCAGTCGCCCCATGAAGCATGAG	44
licR-B/CmF	CGTCGTGACTGGGAAAACACCAAAAAATGCTGAGCTGACAGC	45
licR-BR	TTGCCAATGATGAGGAAAAAGGAACC	46
sigL-AF	CTGAACGTCTGAATAAAAAAGCAGG	47
sigL-A/CmR	GTTATCCGCTCACAATTCGCTGAAGTTTCATATCCATC	48
sigL-B/CmF	CGTCGTGACTGGGAAAACATTCGTCATCGGCAGCGAG	49
sigL-BR	AGCGGTTTACAAGTTGGAGG	50
mntR-AF	ATTT CAGAAGGCATACTTCAAG	51
mntR-A/CmR	GTTATCCGCTCACAATTCATACTTGGTGTGTCATCG	52
mntR-B/CmF	CGTCGTGACTGGGAAAACATAATCAGTAAAAAGCGCGTC	53
mntR-BR	TTCTGACCGCTCTGGCAACC	54
glcT-AF	ATAATGCCCGCTTCCCAACC	55
glcT-A/CmR	GTTATCCGCTCACAATTCGGATCCTCAGCTCCTTTGTC	56
glcT-B/CmF	CGTCGTGACTGGGAAAAC TCACTGATACCGATTAACC	57
glcT-BR	CAACTGAATCCGAAGGAATG	58
yvdE-AF	TCGGGGTCATGCCGAGCGGT	59
yvdE-A/CmR	GTTATCCGCTCACAATTC CAATGTTGCCATTTTCATCC	60
yvdE-B/CmF	CGTCGTGACTGGGAAAAC TGTACGAGAATCAACGCTG	61
yvdE-BR	CACGGCAATGCATTCTTCGG	62
ykvE-AF	AGATCTGTCCGCCAGGTTTAC	63
ykvE-A/CmR	GTTATCCGCTCACAATTC T GATTTTTCTGTCATGTCTC	64
ykvE-B/CmF	CGTCGTGACTGGGAAAACGGTAGAGATGTGCACCGAAA	65
ykvE-BR	GAGTCAGACGGCATCGATGA	66
slr-AF	TTCTGATTCATTTTCACTGCTGG	67
slr-A/CmR	GTTATCCGCTCACAATTC AACGGATAATTCTTCCAATC	68
slr-B/CmF	CGTCGTGACTGGGAAAAC TGTCCATGAAGTCAAATCC	69
slr-BR	CGCTGAAATATTCTCTCGCA	70
rocR-AF	CGCCGCTTTCACCGCGGATTC	71
rocR-A/CmR	GTTATCCGCTCACAATTC TTTGACCACTGTATGAACC	72

表 2-3

rocR-B/CmF	CGTCGTGACTGGGAAAACACTCGTCTAACGAATAATCC	73
rocR-BR	TGTCATCACGGAATTTGACG	74
ccpA-AF	CCAAATTATCCTTTGTGAGCGCGGAATCAG	75
ccpA-A/CmR	GTTATCCGCTCACAATTCGTTAGATCGTAATATTGCTC	76
ccpA-B/CmF	CGTCGTGACTGGGAAAACAGCTTAGAAAGTCAACCAAG	77
ccpA-BR	TTTGAGCATCAGCACAAAGCC	78
yaaT-AF	TGTAGCAGAAGCAGTCCAATT	79
yaaT-A/Cm2R	CTAATGGGTGCTTTAGTTGACAATTACGCAGCTGTCATGT	80
yaaT-B/Cm2F	CTGCCCGTTAGTTGAAGAACTGATAAACCGTGAAAAAGTG	81
yaaT-RV	CCTTTGAAAAAGGCTCCCGT	82
yaaA-AF	GTTTTCCAAGTCTGCCGATAAAAAATATGC	83
yaaA-A/CmR	GTTATCCGCTCACAATTCATGCTTCATGTACCTACACC	84
yaaA-B/CmF	CGTCGTGACTGGGAAAACCAATTAACGATTCGCATACC	85
yaaA-BR	AAAAAGAAGAAGTACAGTACAGAACGTGG	86
yycH-AF	ATTTTTCGCCATCTTGAATTTTC	87
yycH-A/Cm2R	CTAATGGGTGCTTTAGTTGGATGATCCTCTCGTTGAACTG	88
yycH-B/Cm2F	CTGCCCGTTAGTTGAAGGGATGAGCCTTCAGAAAAAGTT	89
yycH-BR	GCCGGACAGAGATCTGTATG	90
yacP-B/Cm4F	GAAGAAGGTTTTTATGTTGACGCTTTTTTGCCCAATACTGTATAA	91
yacP-B/Cm4R	CAAAAAAGCGTCAACATAAAAAACCTTCTTCAACTAACGGGGCAGG	92
yacP-BR	AAGACGAGTACTTTTCTCTCTAAATCACTT	93
yacP-AF	AACTCGATCAAATGGTGACAGGACAGCATC	94
yacP-A/Cm4F	GGAGAATAAAGACCCCTTCAACTAAAGCACCCATTAGTTCAACA	95
yacP-A/Cm4R	TGCTTTAGTTGAAGAGGGTCTTATTCTCCACAGGGTTTCGTTT	96
hprK-B/Cm4F	TTTTTATATTACAGCGAGTTGGCGTTAAATGAATGAAGCGATAGA	97
hprK-B/Cm4R	ATTTAACGCCAACTCGCTGTAATATAAAAAACCTTCTTCAACTAAC	98
hprK-BR	TTGATTGATGATAAATTCAGGCAGGTGCAG	99
hprK-AF	CAAAGCTTGAGAAATGTTCCCATGCTCTTG	100
hprK-A/Cm4F	CAGGAGGAACATATCTTCAACTAAAGCACCCATTAGTTCAACA	101
hprK-A/Cm4R	TGCTTTAGTTGAAGAGATATGTTCCCTCCTGTTCCGGGCTGCCCGG	102
rsiX-AF	ATCCAGTTACTCGTAATATAGTTG	103
rsiX-A/CmR	GTTATCCGCTCACAATTCACCTTCATCATCCATTAGCTC	104
rsiX-B/CmF	CGTCGTGACTGGGAAAACCTGCTCCAAATCCGATTTCC	105
rsiX-BR	GTCCTGCATTTTTCGAAGTCTGG	106
yhdK-AF	TACACATCCTTCAAACAAGTCTGAACAAAC	107

表 2-4

yhdK-A/Cm4R	TGCTTTAGTTGAAGATTACCAGTTCATAATTCCACCTCGCCGAC	108
yhdK-B/Cm4F	TTTTTATATTACAGCGTGTGTATAACCATTGTATCTGTAGATACGA	109
yhdK-BR	GCTATGATCATTGTAACGAAAGGAAAGGGG	110
yhdK-A/Cm4F	TTATGGAAGTGGTAATCTTCAACTAAAGCACCCATTAGTTCAACA	111
yhdK-B/Cm4R	CAATGGTATACACACGCTGTAATATAAAAAACCTTCTTCAACTAAC	112
ylbO-AF	AATCTGAACAAGAAAAAGGAGCTGCTCCTC	113
ylbO-A/Cm4R	TGCTTTAGTTGAAGAATCAATCTCCCTCCATGTCAGCTTATTTA	114
ylbO-B/Cm4F	TTTTTATATTACAGCAGAAACGCCTGAAATGAACCGGCCCTATAG	115
ylbO-BR	TGTTTGACAAAGGTAGAACGTCTGCTTATC	116
ylbO-A/Cm4F	GGAGGGAGATTGAATTCTTCAACTAAAGCACCCATTAGTTCAACA	117
ylbO-B/Cm4R	ATTCAGGCGTTTCTGCTGTAATATAAAAAACCTTCTTCAACTAAC	118
CmF	GAATTGTGAGCGGATAAC	119
CmR	GTTTTCCAGTCACGACG	120
Cm2F	CAACTAAAGCACCCATTAG	121
Cm2R	CTTCAACTAACGGGGCAG	122

实施例 2

按照与实施例 1 所述相似的方式，通过使用由选自表 2 所示的各种引物组；即，基因-AF、基因-A/CmR、基因-B/CmF、基因-BR、CmF 以及 CmR 的适当引物组而制成的用于实现缺失的 DNA 片段来分离孢子形成基因缺失了的菌株，在该菌株中已经通过取代下述缺失的基因而引入了抗氯霉素基因。从基因组中缺失的基因是 *comA*、*yopO*、*treR*、*yvbA*、*yvaN*、*yttP*、*yurK*、*yoza*、*licR*、*sigL*、*mntR*、*glcT*、*ykvE*、*slr*、*rocR*、*yvaA* 或 *rsiX*。

实施例 3

按照与实施例 2 所述相似的方式，通过使用选自表 2 所示的基因-AF、基因-A/Cm2R、基因-B/Cm2F、基因-BR、Cm2F 以及 Cm2R 的适当引物组来制备用于缺失的 DNA 片段。通过使用由此制成的 DNA 片段，分离孢子形成基因缺失了的菌株，在该菌株中已经通过取代下述缺失的基因而引入了抗氯霉素基因。从基因组中缺失的基因是 *cspB*、*yvdE*、*yaaT*、*yycH* 或 *ylbO*。

实施例 4

按照与实施例 2 所述相似的方式，由选自表 2 所示的基因-AF、基因-A/Cm4R、基因-B/Cm4F、基因-BR、Cm4F 以及 Cm4R 的适当引物组制备用于实现缺失的 DNA 片段。通过使用由此制成的 DNA 片段，分离孢子形成基因缺失了的菌株，在该菌株中已经通过取代缺失的基因 *yacP*、*hprK* 和 *yhdK* 而引入了抗氯霉素基因。

实施例 5

在实施例 1 到 4 获得的每种基因缺失的菌株中和在用作对照物的枯草杆菌 168 菌株中，通过原生质体转化法引入重组质粒 pHY-S237。重组质粒 pHY-S237 的制备如下：将源自芽孢杆菌 sp. KSM-S237 菌株的编码碱性纤维素酶的 DNA 片段（3.1kb）（序列号:1，日本专利申请（公开）第 2000-210081 号）插入穿梭载体 pHY300 PLK 的限制性内切酶 *Bam*HI 裂解位点。将由此制成的每种细胞株在 LB 培养基（5 毫升）中以 30℃ 震荡培养过夜。将培养液（0.03 毫升）注入 2×L-麦芽糖培养基（2%胰胨、1%酵母提取物、1%NaCl、7.5%麦芽糖、7.5ppm4-5 水合硫酸锰和 15ppm 四环素）中，然后在 30℃ 震荡培养三天。在培养完成后，通过离心去除细胞，并测定由培养物获得的上层清液的碱性纤维素酶活性，由此计算细胞在培养过程中分泌出的碱性纤维素酶的量；即，在细胞外生成的碱性纤维素酶的量。从表 3 中清楚看出，与对照物 168 菌株（野生型菌株）相比，在使用基因缺失的可形成孢子的菌株作为宿主的所有情况下，碱性纤维素酶的生产或分泌都被证实更为有效。

表 3

缺失基因的名称	基因编号	基因大小(bp)	缺失片段的大小(bp)	制得的(分泌的)碱性纤维素酶的量(相对值)
<i>comA</i>	BG10381	645	588	160
<i>yopO</i>	BG13648	213	169	154
<i>treR</i>	BG11011	717	656	139
<i>yvbA</i>	BG14078	273	210	137
<i>cspB</i>	BG10824	204	171	132
<i>yvaN</i>	BG14069	408	379	124
<i>yttP</i>	BG13927	624	590	121
<i>yurK</i>	BG13997	729	677	118
<i>yoza</i>	BG13748	324	289	117
<i>licR</i>	BG11346	1926	1889	116
<i>sigL</i>	BG10748	1311	1256	114
<i>mntR</i>	BG11702	429	399	114
<i>gltT</i>	BG12593	858	811	110
<i>yvdE</i>	BG12414	951	916	109
<i>ykvE</i>	BG13310	438	356	108
<i>slr</i>	BG11858	459	394	105
<i>rocR</i>	BG10723	1386	1359	128
<i>ccpA</i>	BG10376	1005	957	205
<i>yaaT</i>	BG10096	828	828	127
<i>yyaA</i>	BG10057	852	816	113
<i>yycH</i>	BG11462	1368	1368	146
<i>yacP</i>	BG10158	513	513	156
<i>hprK</i>	BG14125	933	933	196
<i>rsiX</i>	BG10537	1107	1068	125
<i>yhdK</i>	BG13017	291	228	114
<i>yIbO</i>	BG13367	582	582	136
None (Wild type)	—	—	—	100

实施例 6

在实施例 1 至 4 获得的每种基因缺失的菌株中和在用作对照物的枯草杆菌 168 菌株中，通过原生质体转化法引入重组质粒 pHSP-K38。重组质粒 pHSP-K38 的制备如下：在穿梭载体 pHY300 PLK 的限制性内切酶 *BagII*-*XbaI* 裂解位点插入如下制成的 2.1kb 片段(序列号：5)：将包含碱性纤维素酶基因的启动子区域和信号序列区域部分的上游 0.6kb 片段(序列号：3)与编码源自芽孢杆菌 *sp.* KSM-K38 菌株的碱性淀粉酶基因的成熟酶区域(Asp1-Gln480)的 DNA 片段(1.5kb)

的上游侧接合（日本专利申请（公开）第 2000-1884882 号, Eur. J. Biochem., 268, 2974 (2001)）。将由此获得的每种细胞株在 LB 培养基（5 毫升）中以 30℃ 震荡培养过夜。将培养液（0.03 毫升）注入 2×L-麦芽糖培养基（2%胰脲、1%酵母提取物、1%NaCl、7.5%麦芽糖、7.5ppm4-5 水合硫酸锰和 15ppm 四环素）中，然后在 30℃ 震荡培养三至六天。在培养完成后，通过离心去除细胞，并测定由培养物获得的上层清液的碱性淀粉酶的活性，由此计算在培养过程中细胞分泌出的碱性淀粉酶的量；即，在细胞外生成的碱性淀粉酶的量。从表 4 中清楚看出，与对照物 168 菌株（野生型菌株）相比，在使用基因缺失的菌株作为宿主的情况下，碱性淀粉酶的生产或分泌被证实更为有效。

表 4

缺失基因的名称	基因编号	基因大小 (bp)	缺失片段的大小(bp)	制得的(分泌的)碱性淀粉酶的量(相对值)
培养 3 天				
<i>slr</i>	BG11858	459	394	178
<i>treR</i>	BG11011	717	656	124
<i>yopO</i>	BG13648	213	169	364
<i>yvaN</i>	BG14069	408	379	148
<i>yvbA</i>	BG14078	273	210	171
无(野生型)	—	—	—	100
培养 5 天(野生型)				
<i>cspB</i>	BG10824	204	171	195
<i>rocR</i>	BG10723	1386	1359	215
<i>sigL</i>	BG10748	1311	1256	204
<i>glcT</i>	BG12593	858	811	132
<i>yvdE</i>	BG12414	951	916	127
<i>yacP</i>	BG10158	513	513	110
无(野生型)	—	—	—	100
培养 6 天				
<i>yycH</i>	BG11462	1368	1368	120
<i>licR</i>	BG11346	1926	1889	122
无(野生型)	—	—	—	100

<110> 花王株式会社

<120> 重组微生物

<130> KS0795

<150> JP 2003-379167

<151> 2003. 11. 7

<160> 122

<170> PatentIn Ver. 2.1

<210> 1

<211> 3150

<212> DNA

<213> 芽孢杆菌KSM-S237

<220>

<221> CDS

<222> (573).. (3044)

<223>

<220>

<221> sig_peptide

<222> (573).. (659)

<223>

<220>

<221> mat_peptide

<222> (660).. ()

<223>

<400> 1

gatttgccga tgcaacaggc ttatatttag aggaaatttc ttttaaatt gaatacggaa 60

taaaatcagg taaacaggtc ctgattttat tttttgagt ttttagaga actgaagatt 120

gaaataaaag tagaagacaa aggacataag aaaattgcat tagttttaat tatagaaac 180

gcctttttat aattatttat acctagaacg aaaactgtg ttcgaaagcg gtttactata 240

aaaccttata ttcggctct tttttaaac aggggtaaa aattcactct agtattctaa 300

ttcaacatg ctataataaa tttgtaagac gcaatatgca tctcttttt tacgatatat 360

gtaagcggtt aaccttgtgc tatatgccga ttaggaagg ggggtagatt gactcaagta 420

gtaataatat agataactta taagttgttg agaagcagga gagcatctgg gttactcaca	480
agttttttta aaactttaac gaaagcactt tcgtaatgc ttaatgaattt agctatttga	540
ttcaattact ttaaaaaatat ttaggaggta at atg atg tta aga aag aaa aca	593
Met Met Leu Arg Lys Lys Thr	
-25	
aag cag ttg att tct tcc att ctt att tta gtt tta ctt cta tct tta	641
Lys Gln Leu Ile Ser Ser Ile Leu Ile Leu Val Leu Leu Leu Ser Leu	
-20 -15 -10	
ttt ccg gca gct ctt gca gca gaa gga aac act cgt gaa gac aat ttt	689
Phe Pro Ala Ala Leu Ala Ala Glu Gly Asn Thr Arg Glu Asp Asn Phe	
-5 -1 1 5 10	
aaa cat tta tta ggt aat gac aat gtt aaa cgc cct tct gag gct ggc	737
Lys His Leu Leu Gly Asn Asp Asn Val Lys Arg Pro Ser Glu Ala Gly	
15 20 25	
gca tta caa tta caa gaa gtc gat gga caa atg aca tta gta gat caa	785
Ala Leu Gln Leu Gln Glu Val Asp Gly Gln Met Thr Leu Val Asp Gln	
30 35 40	
cat gga gaa aaa att caa tta cgt gga atg agt aca cac gga tta cag	833
His Gly Glu Lys Ile Gln Leu Arg Gly Met Ser Thr His Gly Leu Gln	
45 50 55	
tgg ttt cct gag atc ttg aat gat aac gca tac aaa gct ctt tct aac	881
Trp Phe Pro Glu Ile Leu Asn Asp Asn Ala Tyr Lys Ala Leu Ser Asn	
60 65 70	
gat tgg gat tcc aat atg att cgt ctt gct atg tat gta ggt gaa aat	929
Asp Trp Asp Ser Asn Met Ile Arg Leu Ala Met Tyr Val Gly Glu Asn	
75 80 85 90	
ggg tac gct aca aac cct gag tta atc aaa caa aga gtg att gat gga	977
Gly Tyr Ala Thr Asn Pro Glu Leu Ile Lys Gln Arg Val Ile Asp Gly	
95 100 105	
att gag tta gcg att gaa aat gac atg tat gtt att gtt gac tgg cat	1025
Ile Glu Leu Ala Ile Glu Asn Asp Met Tyr Val Ile Val Asp Trp His	
110 115 120	
gtt cat gcg cca ggt gat cct aga gat cct gtt tat gca ggt gct aaa	1073
Val His Ala Pro Gly Asp Pro Arg Asp Pro Val Tyr Ala Gly Ala Lys	
125 130 135	

gat ttc ttt aga gaa att gca gct tta tac cct aat aat cca cac att Asp Phe Phe Arg Glu Ile Ala Ala Leu Tyr Pro Asn Asn Pro His Ile 140 145 150	1121
att tat gag tta gcg aat gag ccg agt agt aat aat aat ggt gga gca Ile Tyr Glu Leu Ala Asn Glu Pro Ser Ser Asn Asn Asn Gly Gly Ala 155 160 165 170	1169
ggg att ccg aat aac gaa gaa ggt tgg aaa gcg gta aaa gaa tat gct Gly Ile Pro Asn Asn Glu Glu Gly Trp Lys Ala Val Lys Glu Tyr Ala 175 180 185	1217
gat cca att gta gaa atg tta cgt aaa agc ggt aat gca gat gac aac Asp Pro Ile Val Glu Met Leu Arg Lys Ser Gly Asn Ala Asp Asp Asn 190 195 200	1265
att atc att gtt ggt agt cca aac tgg agt cag cgt ccg gac tta gca Ile Ile Ile Val Gly Ser Pro Asn Trp Ser Gln Arg Pro Asp Leu Ala 205 210 215	1313
gct gat aat cca att gat gat cac cat aca atg tat act gtt cac ttc Ala Asp Asn Pro Ile Asp Asp His His Thr Met Tyr Thr Val His Phe 220 225 230	1361
tac act ggt tca cat gct gct tca act gaa agc tat ccg tct gaa act Tyr Thr Gly Ser His Ala Ala Ser Thr Glu Ser Tyr Pro Ser Glu Thr 235 240 245 250	1409
cct aac tct gaa aga gga aac gta atg agt aac act cgt tat gcg tta Pro Asn Ser Glu Arg Gly Asn Val Met Ser Asn Thr Arg Tyr Ala Leu 255 260 265	1457
gaa aac gga gta gcg gta ttt gca aca gag tgg gga acg agt caa gct Glu Asn Gly Val Ala Val Phe Ala Thr Glu Trp Gly Thr Ser Gln Ala 270 275 280	1505
agt gga gac ggt ggt cct tac ttt gat gaa gca gat gta tgg att gaa Ser Gly Asp Gly Gly Pro Tyr Phe Asp Glu Ala Asp Val Trp Ile Glu 285 290 295	1553
ttt tta aat gaa aac aac att agc tgg get aac tgg tct tta acg aat Phe Leu Asn Glu Asn Asn Ile Ser Trp Ala Asn Trp Ser Leu Thr Asn 300 305 310	1601
aaa aat gaa gta tct ggt gca ttt aca cca ttc gag tta ggt aag tct Lys Asn Glu Val Ser Gly Ala Phe Thr Pro Phe Glu Leu Gly Lys Ser 315 320 325 330	1649
aac gca acc aat ctt gac cca ggt cca gat cat gtg tgg gca cca gaa	1697

Asn Ala Thr Asn Leu Asp Pro Gly Pro Asp His Val Trp Ala Pro Glu	
335 340 345	
gaa tta agt ctt tct gga gaa tat gta cgt gct cgt att aaa ggt gtg	1745
Glu Leu Ser Leu Ser Gly Glu Tyr Val Arg Ala Arg Ile Lys Gly Val	
350 355 360	
aac tat gag cca atc gac cgt aca aaa tac acg aaa gta ctt tgg gac	1793
Asn Tyr Glu Pro Ile Asp Arg Thr Lys Tyr Thr Lys Val Leu Trp Asp	
365 370 375	
ttt aat gat gga acg aag caa gga ttt gga gtg aat tcg gat tct cca	1841
Phe Asn Asp Gly Thr Lys Gln Gly Phe Gly Val Asn Ser Asp Ser Pro	
380 385 390	
aat aaa gaa ctt att gca gtt gat aat gaa aac aac act ttg aaa gtt	1889
Asn Lys Glu Leu Ile Ala Val Asp Asn Glu Asn Asn Thr Leu Lys Val	
395 400 405 410	
tcg gga tta gat gta agt aac gat gtt tca gat ggc aac ttc tgg gct	1937
Ser Gly Leu Asp Val Ser Asn Asp Val Ser Asp Gly Asn Phe Trp Ala	
415 420 425	
aat gct cgt ctt tct gcc aac ggt tgg gga aaa agt gtt gat att tta	1985
Asn Ala Arg Leu Ser Ala Asn Gly Trp Gly Lys Ser Val Asp Ile Leu	
430 435 440	
ggt gct gag aag ctt aca atg gat gtt att gtt gat gaa cca acg acg	2033
Gly Ala Glu Lys Leu Thr Met Asp Val Ile Val Asp Glu Pro Thr Thr	
445 450 455	
gta gct att gcg gcg att cca caa agt agt aaa agt gga tgg gca aat	2081
Val Ala Ile Ala Ala Ile Pro Gln Ser Ser Lys Ser Gly Trp Ala Asn	
460 465 470	
cca gag cgt gct gtt cga gtg aac gcg gaa gat ttt gtc cag caa acg	2129
Pro Glu Arg Ala Val Arg Val Asn Ala Glu Asp Phe Val Gln Gln Thr	
475 480 485 490	
gac ggt aag tat aaa gct gga tta aca att aca gga gaa gat gct cct	2177
Asp Gly Lys Tyr Lys Ala Gly Leu Thr Ile Thr Gly Glu Asp Ala Pro	
495 500 505	
aac cta aaa aat atc gct ttt cat gaa gaa gat aac aat atg aac aac	2225
Asn Leu Lys Asn Ile Ala Phe His Glu Glu Asp Asn Asn Met Asn Asn	
510 515 520	
atc att ctg ttc gtg gga act gat gca gct gac gtt att tac tta gat	2273
Ile Ile Leu Phe Val Gly Thr Asp Ala Ala Asp Val Ile Tyr Leu Asp	

525	530	535	
aac att aaa gta att gga aca gaa gtt gaa att cca gtt gtt cat gat			2321
Asn Ile Lys Val Ile Gly Thr Glu Val Glu Ile Pro Val Val His Asp			
540	545	550	
cca aaa gga gaa gct gtt ctt cct tct gtt ttt gaa gac ggt aca cgt			2369
Pro Lys Gly Glu Ala Val Leu Pro Ser Val Phe Glu Asp Gly Thr Arg			
555	560	565	570
caa ggt tgg gac tgg gct gga gag tct ggt gtg aaa aca gct tta aca			2417
Gln Gly Trp Asp Trp Ala Gly Glu Ser Gly Val Lys Thr Ala Leu Thr			
	575	580	585
att gaa gaa gca aac ggt tct aac gcg tta tca tgg gaa ttt gga tat			2465
Ile Glu Glu Ala Asn Gly Ser Asn Ala Leu Ser Trp Glu Phe Gly Tyr			
	590	595	600
cca gaa gta aaa cct agt gat aac tgg gca aca gct cca cgt tta gat			2513
Pro Glu Val Lys Pro Ser Asp Asn Trp Ala Thr Ala Pro Arg Leu Asp			
	605	610	615
ttc tgg aaa tct gac ttg gtt cgc ggt gag aat gat tat gta gct ttt			2561
Phe Trp Lys Ser Asp Leu Val Arg Gly Glu Asn Asp Tyr Val Ala Phe			
	620	625	630
gat ttc tat cta gat cca gtt cgt gca aca gaa ggc gca atg aat atc			2609
Asp Phe Tyr Leu Asp Pro Val Arg Ala Thr Glu Gly Ala Met Asn Ile			
	635	640	645
aat tta gta ttc cag cca cct act aac ggg tat tgg gta caa gca cca			2657
Asn Leu Val Phe Gln Pro Pro Thr Asn Gly Tyr Trp Val Gln Ala Pro			
	655	660	665
aaa acg tat acg att aac ttt gat gaa tta gag gaa gcg aat caa gta			2705
Lys Thr Tyr Thr Ile Asn Phe Asp Glu Leu Glu Glu Ala Asn Gln Val			
	670	675	680
aat ggt tta tat cac tat gaa gtg aaa att aac gta aga gat att aca			2753
Asn Gly Leu Tyr His Tyr Glu Val Lys Ile Asn Val Arg Asp Ile Thr			
	685	690	695
aac att caa gat gac acg tta cta cgt aac atg atg atc att ttt gca			2801
Asn Ile Gln Asp Asp Thr Leu Leu Arg Asn Met Met Ile Ile Phe Ala			
	700	705	710
gat gta gaa agt gac ttt gca ggg aga gtc ttt gta gat aat gtt cgt			2849
Asp Val Glu Ser Asp Phe Ala Gly Arg Val Phe Val Asp Asn Val Arg			
	715	720	725
			730

ttt gag ggg gct gct act act gag ccg gtt gaa cca gag cca gtt gat 2897
 Phe Glu Gly Ala Ala Thr Thr Glu Pro Val Glu Pro Glu Pro Val Asp
 735 740 745

cct ggc gaa gag acg cca cct gtc gat gag aag gaa gcg aaa aaa gaa 2945
 Pro Gly Glu Glu Thr Pro Pro Val Asp Glu Lys Glu Ala Lys Lys Glu
 750 755 760

caa aaa gaa gca gag aaa gaa gag aaa gaa gca gta aaa gaa gaa aag 2993
 Gln Lys Glu Ala Glu Lys Glu Lys Glu Ala Val Lys Glu Glu Lys
 765 770 775

aaa gaa gct aaa gaa gaa aag aaa gca gtc aaa aat gag gct aag aaa 3041
 Lys Glu Ala Lys Glu Glu Lys Lys Ala Val Lys Asn Glu Ala Lys Lys
 780 785 790

aaa taatctatta aactagtat agggttatct aaaggtctga ttagatctt 3094
 Lys
 795

ttagataacc tttttcttgc ataactggac acagagttgt tattaagaa agtaag 3150

<210> 2
 <211> 824
 <212> PRT
 <213> 芽孢杆菌KSM-S237

<400> 2

Met Met Leu Arg Lys Lys Thr Lys Gln Leu Ile Ser Ser Ile Leu Ile
 -25 -20 -15

Leu Val Leu Leu Leu Ser Leu Phe Pro Ala Ala Leu Ala Ala Glu Gly
 -10 -5 -1 1

Asn Thr Arg Glu Asp Asn Phe Lys His Leu Leu Gly Asn Asp Asn Val
 5 10 15

Lys Arg Pro Ser Glu Ala Gly Ala Leu Gln Leu Gln Glu Val Asp Gly
 20 25 30 35

Gln Met Thr Leu Val Asp Gln His Gly Glu Lys Ile Gln Leu Arg Gly
 40 45 50

Met Ser Thr His Gly Leu Gln Trp Phe Pro Glu Ile Leu Asn Asp Asn
55 60 65

Ala Tyr Lys Ala Leu Ser Asn Asp Trp Asp Ser Asn Met Ile Arg Leu
70 75 80

Ala Met Tyr Val Gly Glu Asn Gly Tyr Ala Thr Asn Pro Glu Leu Ile
85 90 95

Lys Gln Arg Val Ile Asp Gly Ile Glu Leu Ala Ile Glu Asn Asp Met
100 105 110 115

Tyr Val Ile Val Asp Trp His Val His Ala Pro Gly Asp Pro Arg Asp
120 125 130

Pro Val Tyr Ala Gly Ala Lys Asp Phe Phe Arg Glu Ile Ala Ala Leu
135 140 145

Tyr Pro Asn Asn Pro His Ile Ile Tyr Glu Leu Ala Asn Glu Pro Ser
150 155 160

Ser Asn Asn Asn Gly Gly Ala Gly Ile Pro Asn Asn Glu Glu Gly Trp
165 170 175

Lys Ala Val Lys Glu Tyr Ala Asp Pro Ile Val Glu Met Leu Arg Lys
180 185 190 195

Ser Gly Asn Ala Asp Asp Asn Ile Ile Ile Val Gly Ser Pro Asn Trp
200 205 210

Ser Gln Arg Pro Asp Leu Ala Ala Asp Asn Pro Ile Asp Asp His His
215 220 225

Thr Met Tyr Thr Val His Phe Tyr Thr Gly Ser His Ala Ala Ser Thr
230 235 240

Glu Ser Tyr Pro Ser Glu Thr Pro Asn Ser Glu Arg Gly Asn Val Met
 245 250 255

Ser Asn Thr Arg Tyr Ala Leu Glu Asn Gly Val Ala Val Phe Ala Thr
 260 265 270 275

Glu Trp Gly Thr Ser Gln Ala Ser Gly Asp Gly Gly Pro Tyr Phe Asp
 280 285 290

Glu Ala Asp Val Trp Ile Glu Phe Leu Asn Glu Asn Asn Ile Ser Trp
 295 300 305

Ala Asn Trp Ser Leu Thr Asn Lys Asn Glu Val Ser Gly Ala Phe Thr
 310 315 320

Pro Phe Glu Leu Gly Lys Ser Asn Ala Thr Asn Leu Asp Pro Gly Pro
 325 330 335

Asp His Val Trp Ala Pro Glu Glu Leu Ser Leu Ser Gly Glu Tyr Val
 340 345 350 355

Arg Ala Arg Ile Lys Gly Val Asn Tyr Glu Pro Ile Asp Arg Thr Lys
 360 365 370

Tyr Thr Lys Val Leu Trp Asp Phe Asn Asp Gly Thr Lys Gln Gly Phe
 375 380 385

Gly Val Asn Ser Asp Ser Pro Asn Lys Glu Leu Ile Ala Val Asp Asn
 390 395 400

Glu Asn Asn Thr Leu Lys Val Ser Gly Leu Asp Val Ser Asn Asp Val
 405 410 415

Ser Asp Gly Asn Phe Trp Ala Asn Ala Arg Leu Ser Ala Asn Gly Trp
 420 425 430 435

Gly Lys Ser Val Asp Ile Leu Gly Ala Glu Lys Leu Thr Met Asp Val
 440 445 450

Ile Val Asp Glu Pro Thr Thr Val Ala Ile Ala Ala Ile Pro Gln Ser
 455 460 465

Ser Lys Ser Gly Trp Ala Asn Pro Glu Arg Ala Val Arg Val Asn Ala
 470 475 480

Glu Asp Phe Val Gln Gln Thr Asp Gly Lys Tyr Lys Ala Gly Leu Thr
 485 490 495

Ile Thr Gly Glu Asp Ala Pro Asn Leu Lys Asn Ile Ala Phe His Glu
 500 505 510 515

Glu Asp Asn Asn Met Asn Asn Ile Ile Leu Phe Val Gly Thr Asp Ala
 520 525 530

Ala Asp Val Ile Tyr Leu Asp Asn Ile Lys Val Ile Gly Thr Glu Val
 535 540 545

Glu Ile Pro Val Val His Asp Pro Lys Gly Glu Ala Val Leu Pro Ser
 550 555 560

Val Phe Glu Asp Gly Thr Arg Gln Gly Trp Asp Trp Ala Gly Glu Ser
 565 570 575

Gly Val Lys Thr Ala Leu Thr Ile Glu Glu Ala Asn Gly Ser Asn Ala
 580 585 590 595

Leu Ser Trp Glu Phe Gly Tyr Pro Glu Val Lys Pro Ser Asp Asn Trp
 600 605 610

Ala Thr Ala Pro Arg Leu Asp Phe Trp Lys Ser Asp Leu Val Arg Gly
 615 620 625

Glu Asn Asp Tyr Val Ala Phe Asp Phe Tyr Leu Asp Pro Val Arg Ala

	630		635		640	
Thr	Glu Gly Ala Met Asn	Ile Asn Leu Val Phe	Gln Pro Pro Thr Asn			
	645		650		655	
Gly Tyr Trp Val Gln Ala Pro Lys Thr Tyr Thr Ile Asn Phe Asp Glu						
660		665		670		675
Leu Glu Glu Ala Asn Gln Val Asn Gly Leu Tyr His Tyr Glu Val Lys						
	680		685			690
Ile Asn Val Arg Asp Ile Thr Asn Ile Gln Asp Asp Thr Leu Leu Arg						
	695		700			705
Asn Met Met Ile Ile Phe Ala Asp Val Glu Ser Asp Phe Ala Gly Arg						
	710		715			720
Val Phe Val Asp Asn Val Arg Phe Glu Gly Ala Ala Thr Thr Glu Pro						
	725		730			735
Val Glu Pro Glu Pro Val Asp Pro Gly Glu Glu Thr Pro Pro Val Asp						
740		745		750		755
Glu Lys Glu Ala Lys Lys Glu Gln Lys Glu Ala Glu Lys Glu Glu Lys						
	760		765			770
Glu Ala Val Lys Glu Glu Lys Lys Glu Ala Lys Glu Glu Lys Lys Ala						
	775		780			785
Val Lys Asn Glu Ala Lys Lys Lys						
	790		795			

<210> 3
 <211> 3332
 <212> DNA
 <213> 芽孢杆菌KSM-64
 <220>

```

<221> CDS
<222> (610)..(3075)
<223>

<220>
<221> sig_peptide
<222> (610)..(696)
<223>

<220>
<221> mat_peptide
<222> (697)..()
<223>

<400> 3
agtacttacc atttttagagt caaaagatag aagccaagca ggatttgccg atgcaaccgg      60
cttatattta gagggaattt ctttttaaat tgaatacggg ataaaatcag gtaaacaggt      120
cctgatttta tttttttgaa tttttttgag aactaaagat tgaaatagaa gtagaagaca      180
acggacataa gaaaattgta ttagttttaa ttatagaaaa cgcttttcta taattattta      240
tacctagaac gaaaatactg tttcgaaagc ggtttactat aaaaccttat attccggctc      300
tttttttaaa caggggggga aaattcactc tagtattceta atttcaacat gctataataa      360
atttgaaga cgcaatatac atcttttttt tatgatattt gtaageggtt aaccttgtgc      420
tatatgccga tttaggaagg gggtagattg agtcaagtag tcataattta gataacttat      480
aagttgttga gaagcaggag agaatctggg ttactcacia gttttttaa acattatcga      540
aagcactttc ggttatgctt atgaatttag ctatttgatt caattacttt aataatttta      600
ggaggtaat atg atg tta aga aag aaa aca aag cag ttg att tct tcc att      651
      Met Met Leu Arg Lys Lys Thr Lys Gln Leu Ile Ser Ser Ile
                    -25                               -20

ctt att tta gtt tta ctt cta tct tta ttt ccg aca gct ctt gca gca      699
Leu Ile Leu Val Leu Leu Leu Ser Leu Phe Pro Thr Ala Leu Ala Ala
-15                -10                -5                -1  1

gaa gga aac act cgt gaa gac aat ttt aaa cat tta tta ggt aat gac      747
Glu Gly Asn Thr Arg Glu Asp Asn Phe Lys His Leu Leu Gly Asn Asp
          5                10                15

aat gtt aaa cgc cct tct gag gct ggc gca tta caa tta caa gaa gtc      795
Asn Val Lys Arg Pro Ser Glu Ala Gly Ala Leu Gln Leu Gln Glu Val
    
```

20	25	30	
gat gga caa atg aca tta gta gat caa cat gga gaa aaa att caa tta			843
Asp Gly Gln Met Thr Leu Val Asp Gln His Gly Glu Lys Ile Gln Leu			
35	40	45	
cgt gga atg agt aca cac gga tta caa tgg ttt cct gag atc ttg aat			891
Arg Gly Met Ser Thr His Gly Leu Gln Trp Phe Pro Glu Ile Leu Asn			
50	55	60	65
gat aac gca tac aaa gct ctt gct aac gat tgg gaa tca aat atg att			939
Asp Asn Ala Tyr Lys Ala Leu Ala Asn Asp Trp Glu Ser Asn Met Ile			
70	75	80	
cgt cta gct atg tat gtc ggt gaa aat ggc tat gct tca aat cca gag			987
Arg Leu Ala Met Tyr Val Gly Glu Asn Gly Tyr Ala Ser Asn Pro Glu			
85	90	95	
tta att aaa agc aga gtc att aaa gga ata gat ctt gct att gaa aat			1035
Leu Ile Lys Ser Arg Val Ile Lys Gly Ile Asp Leu Ala Ile Glu Asn			
100	105	110	
gac atg tat gtc atc gtt gat tgg cat gta cat gca cct ggt gat cct			1083
Asp Met Tyr Val Ile Val Asp Trp His Val His Ala Pro Gly Asp Pro			
115	120	125	
aga gat ccc gtt tac gct gga gca gaa gat ttc ttt aga gat att gca			1131
Arg Asp Pro Val Tyr Ala Gly Ala Glu Asp Phe Phe Arg Asp Ile Ala			
130	135	140	145
gca tta tat cct aac aat cca cac att att tat gag tta gcg aat gag			1179
Ala Leu Tyr Pro Asn Asn Pro His Ile Ile Tyr Glu Leu Ala Asn Glu			
150	155	160	
cca agt agt aac aat aat ggt gga gct ggg att cca aat aat gaa gaa			1227
Pro Ser Ser Asn Asn Asn Gly Gly Ala Gly Ile Pro Asn Asn Glu Glu			
165	170	175	
ggt tgg aat gcg gta aaa gaa tac gct gat cca att gta gaa atg tta			1275
Gly Trp Asn Ala Val Lys Glu Tyr Ala Asp Pro Ile Val Glu Met Leu			
180	185	190	
cgt gat agc ggg aac gca gat gac aat att atc att gtg ggt agt cca			1323
Arg Asp Ser Gly Asn Ala Asp Asp Asn Ile Ile Ile Val Gly Ser Pro			
195	200	205	
aac tgg agt cag cgt cct gac tta gca gct gat aat cca att gat gat			1371
Asn Trp Ser Gln Arg Pro Asp Leu Ala Ala Asp Asn Pro Ile Asp Asp			
210	215	220	225

cac cat aca atg tat act gtt cac ttc tac act ggt tca cat gct gct His His Thr Met Tyr Thr Val His Phe Tyr Thr Gly Ser His Ala Ala 230 235 240	1419
tca act gaa agc tat ccg cct gaa act cct aac tct gaa aga gga aac Ser Thr Glu Ser Tyr Pro Pro Glu Thr Pro Asn Ser Glu Arg Gly Asn 245 250 255	1467
gta atg agt aac act cgt tat gcg tta gaa aac gga gta gca gta ttt Val Met Ser Asn Thr Arg Tyr Ala Leu Glu Asn Gly Val Ala Val Phe 260 265 270	1515
gca aca gag tgg gga act agc caa gca aat gga gat ggt ggt cct tac Ala Thr Glu Trp Gly Thr Ser Gln Ala Asn Gly Asp Gly Gly Pro Tyr 275 280 285	1563
ttt gat gaa gca gat gta tgg att gag ttt tta aat gaa aac aac att Phe Asp Glu Ala Asp Val Trp Ile Glu Phe Leu Asn Glu Asn Asn Ile 290 295 300 305	1611
agc tgg gct aac tgg tct tta acg aat aaa aat gaa gta tct ggt gca Ser Trp Ala Asn Trp Ser Leu Thr Asn Lys Asn Glu Val Ser Gly Ala 310 315 320	1659
ttt aca cca ttc gag tta ggt aag tct aac gca aca agt ctt gac cca Phe Thr Pro Phe Glu Leu Gly Lys Ser Asn Ala Thr Ser Leu Asp Pro 325 330 335	1707
ggg cca gac caa gta tgg gta cca gaa gag tta agt ctt tct gga gaa Gly Pro Asp Gln Val Trp Val Pro Glu Glu Leu Ser Leu Ser Gly Glu 340 345 350	1755
tat gta cgt gct cgt att aaa ggt gtg aac tat gag cca atc gac cgt Tyr Val Arg Ala Arg Ile Lys Gly Val Asn Tyr Glu Pro Ile Asp Arg 355 360 365	1803
aca aaa tac acg aaa gta ctt tgg gac ttt aat gat gga acg aag caa Thr Lys Tyr Thr Lys Val Leu Trp Asp Phe Asn Asp Gly Thr Lys Gln 370 375 380 385	1851
gga ttt gga gtg aat gga gat tct cca gtt gaa gat gta gtt att gag Gly Phe Gly Val Asn Gly Asp Ser Pro Val Glu Asp Val Val Ile Glu 390 395 400	1899
aat gaa gcg ggc gct tta aaa ctt tca gga tta gat gca agt aat gat Asn Glu Ala Gly Ala Leu Lys Leu Ser Gly Leu Asp Ala Ser Asn Asp 405 410 415	1947

ggt tct gaa ggt aat tac tgg gct aat gct cgt ctt tct gcc gac ggt Val Ser Glu Gly Asn Tyr Trp Ala Asn Ala Arg Leu Ser Ala Asp Gly 420 425 430	1995
tgg gga aaa agt gtt gat att tta ggt gct gaa aaa ctt act atg gat Trp Gly Lys Ser Val Asp Ile Leu Gly Ala Glu Lys Leu Thr Met Asp 435 440 445	2043
gtg att gtt gat gag ccg acc acg gta tca att gct gca att cca caa Val Ile Val Asp Glu Pro Thr Thr Val Ser Ile Ala Ala Ile Pro Gln 450 455 460 465	2091
ggg cca tca gcc aat tgg gtt aat cca aat cgt gca att aag gtt gag Gly Pro Ser Ala Asn Trp Val Asn Pro Asn Arg Ala Ile Lys Val Glu 470 475 480	2139
cca act aat ttc gta ccg tta gga gat aag ttt aaa gcg gaa tta act Pro Thr Asn Phe Val Pro Leu Gly Asp Lys Phe Lys Ala Glu Leu Thr 485 490 495	2187
ata act tca gct gac tct cca teg tta gaa gct att gcg atg cat gct Ile Thr Ser Ala Asp Ser Pro Ser Leu Glu Ala Ile Ala Met His Ala 500 505 510	2235
gaa aat aac aac atc aac aac atc att ctt ttt gta gga act gaa ggt Glu Asn Asn Asn Ile Asn Asn Ile Ile Leu Phe Val Gly Thr Glu Gly 515 520 525	2283
gct gat gtt atc tat tta gat aac att aaa gta att gga aca gaa gtt Ala Asp Val Ile Tyr Leu Asp Asn Ile Lys Val Ile Gly Thr Glu Val 530 535 540 545	2331
gaa att cca gtt gtt cat gat cca aaa gga gaa gct gtt ctt cct tct Glu Ile Pro Val Val His Asp Pro Lys Gly Glu Ala Val Leu Pro Ser 550 555 560	2379
gtt ttt gaa gac ggt aca cgt caa ggt tgg gac tgg gct gga gag tct Val Phe Glu Asp Gly Thr Arg Gln Gly Trp Asp Trp Ala Gly Glu Ser 565 570 575	2427
ggt gtg aaa aca gct tta aca att gaa gaa gca aac ggt tct aac gcg Gly Val Lys Thr Ala Leu Thr Ile Glu Glu Ala Asn Gly Ser Asn Ala 580 585 590	2475
tta tca tgg gaa ttt gga tac cca gaa gta aaa cct agt gat aac tgg Leu Ser Trp Glu Phe Gly Tyr Pro Glu Val Lys Pro Ser Asp Asn Trp 595 600 605	2523
gca aca gct cca cgt tta gat ttc tgg aaa tct gac ttg gtt cgc ggt	2571

Ala Thr Ala Pro Arg Leu Asp Phe Trp Lys Ser Asp Leu Val Arg Gly	
610	615 620 625
gaa aat gat tat gta act ttt gat ttc tat cta gat cca gtt cgt gca 2619	
Glu Asn Asp Tyr Val Thr Phe Asp Phe Tyr Leu Asp Pro Val Arg Ala	
630	635 640
aca gaa ggc gca atg aat atc aat tta gta ttc cag cca cct act aac 2667	
Thr Glu Gly Ala Met Asn Ile Asn Leu Val Phe Gln Pro Pro Thr Asn	
645	650 655
ggg tat tgg gta caa gca cca aaa acg tat acg att aac ttt gat gaa 2715	
Gly Tyr Trp Val Gln Ala Pro Lys Thr Tyr Thr Ile Asn Phe Asp Glu	
660	665 670
tta gag gaa gcg aat caa gta aat ggt tta tat cac tat gaa gtg aaa 2763	
Leu Glu Glu Ala Asn Gln Val Asn Gly Leu Tyr His Tyr Glu Val Lys	
675	680 685
att aac gta aga gat att aca aac att caa gat gac acg tta cta cgt 2811	
Ile Asn Val Arg Asp Ile Thr Asn Ile Gln Asp Asp Thr Leu Leu Arg	
690	695 700 705
aac atg atg atc att ttt gca gat gta gaa agt gac ttt gca ggg aga 2859	
Asn Met Met Ile Ile Phe Ala Asp Val Glu Ser Asp Phe Ala Gly Arg	
710	715 720
gtc ttt gta gat aat gtt cgt ttt gag ggg gct gct act act gag ccg 2907	
Val Phe Val Asp Asn Val Arg Phe Glu Gly Ala Ala Thr Thr Glu Pro	
725	730 735
gtt gaa cca gag cca gtt gat cct ggc gaa gag acg ccg cct gtc gat 2955	
Val Glu Pro Glu Pro Val Asp Pro Gly Glu Glu Thr Pro Pro Val Asp	
740	745 750
gag aag gaa gcg aaa aaa gaa caa aaa gaa gca gag aaa gaa gag aaa 3003	
Glu Lys Glu Ala Lys Lys Glu Gln Lys Glu Ala Glu Lys Glu Glu Lys	
755	760 765
gaa gca gta aaa gaa gaa aag aaa gaa gct aaa gaa gaa aag aaa gca 3051	
Glu Ala Val Lys Glu Glu Lys Lys Glu Ala Lys Glu Glu Lys Lys Ala	
770	775 780 785
atc aaa aat gag gct acg aaa aaa taatctaata aactagttat agggttatct 3105	
Ile Lys Asn Glu Ala Thr Lys Lys	
790	
aaaggtctga tgcagatctt ttagataacc tttttttgca taactggaca tagaatggtt 3165	

attaaagaaa gcaaggtggtt tatacgatat taaaaggta gcgattttaa attgaaacct 3225
 ttaataatgt cttgtgatag aatgatgaag taatttaaga gggggaaacg aagtgaaaac 3285
 ggaaatttct agtagaagaa aaacagacca agaataactg caagctt 3332

<210> 4
 <211> 822
 <212> PRT
 <213> 芽孢杆菌KSM-64

<400> 4

Met Met Leu Arg Lys Lys Thr Lys Gln Leu Ile Ser Ser Ile Leu Ile
 -25 -20 -15

Leu Val Leu Leu Leu Ser Leu Phe Pro Thr Ala Leu Ala Ala Glu Gly
 -10 -5 -1 1

Asn Thr Arg Glu Asp Asn Phe Lys His Leu Leu Gly Asn Asp Asn Val
 5 10 15

Lys Arg Pro Ser Glu Ala Gly Ala Leu Gln Leu Gln Glu Val Asp Gly
 20 25 30 35

Gln Met Thr Leu Val Asp Gln His Gly Glu Lys Ile Gln Leu Arg Gly
 40 45 50

Met Ser Thr His Gly Leu Gln Trp Phe Pro Glu Ile Leu Asn Asp Asn
 55 60 65

Ala Tyr Lys Ala Leu Ala Asn Asp Trp Glu Ser Asn Met Ile Arg Leu
 70 75 80

Ala Met Tyr Val Gly Glu Asn Gly Tyr Ala Ser Asn Pro Glu Leu Ile
 85 90 95

Lys Ser Arg Val Ile Lys Gly Ile Asp Leu Ala Ile Glu Asn Asp Met
 100 105 110 115

Tyr Val Ile Val Asp Trp His Val His Ala Pro Gly Asp Pro Arg Asp
 120 125 130

Pro Val Tyr Ala Gly Ala Glu Asp Phe Phe Arg Asp Ile Ala Ala Leu
 135 140 145

Tyr Pro Asn Asn Pro His Ile Ile Tyr Glu Leu Ala Asn Glu Pro Ser
 150 155 160

Ser Asn Asn Asn Gly Gly Ala Gly Ile Pro Asn Asn Glu Glu Gly Trp
 165 170 175

Asn Ala Val Lys Glu Tyr Ala Asp Pro Ile Val Glu Met Leu Arg Asp
 180 185 190 195

Ser Gly Asn Ala Asp Asp Asn Ile Ile Ile Val Gly Ser Pro Asn Trp
 200 205 210

Ser Gln Arg Pro Asp Leu Ala Ala Asp Asn Pro Ile Asp Asp His His
 215 220 225

Thr Met Tyr Thr Val His Phe Tyr Thr Gly Ser His Ala Ala Ser Thr
 230 235 240

Glu Ser Tyr Pro Pro Glu Thr Pro Asn Ser Glu Arg Gly Asn Val Met
 245 250 255

Ser Asn Thr Arg Tyr Ala Leu Glu Asn Gly Val Ala Val Phe Ala Thr
 260 265 270 275

Glu Trp Gly Thr Ser Gln Ala Asn Gly Asp Gly Gly Pro Tyr Phe Asp
 280 285 290

Glu Ala Asp Val Trp Ile Glu Phe Leu Asn Glu Asn Asn Ile Ser Trp
 295 300 305

Ala Asn Trp Ser Leu Thr Asn Lys Asn Glu Val Ser Gly Ala Phe Thr
310 315 320

Pro Phe Glu Leu Gly Lys Ser Asn Ala Thr Ser Leu Asp Pro Gly Pro
325 330 335

Asp Gln Val Trp Val Pro Glu Glu Leu Ser Leu Ser Gly Glu Tyr Val
340 345 350 355

Arg Ala Arg Ile Lys Gly Val Asn Tyr Glu Pro Ile Asp Arg Thr Lys
360 365 370

Tyr Thr Lys Val Leu Trp Asp Phe Asn Asp Gly Thr Lys Gln Gly Phe
375 380 385

Gly Val Asn Gly Asp Ser Pro Val Glu Asp Val Val Ile Glu Asn Glu
390 395 400

Ala Gly Ala Leu Lys Leu Ser Gly Leu Asp Ala Ser Asn Asp Val Ser
405 410 415

Glu Gly Asn Tyr Trp Ala Asn Ala Arg Leu Ser Ala Asp Gly Trp Gly
420 425 430 435

Lys Ser Val Asp Ile Leu Gly Ala Glu Lys Leu Thr Met Asp Val Ile
440 445 450

Val Asp Glu Pro Thr Thr Val Ser Ile Ala Ala Ile Pro Gln Gly Pro
455 460 465

Ser Ala Asn Trp Val Asn Pro Asn Arg Ala Ile Lys Val Glu Pro Thr
470 475 480

Asn Phe Val Pro Leu Gly Asp Lys Phe Lys Ala Glu Leu Thr Ile Thr
485 490 495

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

500	505	510	515
Asn Asn Ile	Asn Asn Ile Ile Leu Phe	Val Gly Thr Glu Gly Ala Asp	
	520	525	530
Val Ile Tyr	Leu Asp Asn Ile Lys Val	Ile Gly Thr Glu Val Glu Ile	
	535	540	545
Pro Val Val	His Asp Pro Lys Gly Glu Ala Val	Leu Pro Ser Val Phe	
	550	555	560
Glu Asp Gly	Thr Arg Gln Gly Trp Asp Trp	Ala Gly Glu Ser Gly Val	
	565	570	575
Lys Thr Ala	Leu Thr Ile Glu Glu Ala Asn Gly	Ser Asn Ala Leu Ser	
580	585	590	595
Trp Glu Phe	Gly Tyr Pro Glu Val Lys Pro	Ser Asp Asn Trp Ala Thr	
	600	605	610
Ala Pro Arg	Leu Asp Phe Trp Lys Ser Asp	Leu Val Arg Gly Glu Asn	
	615	620	625
Asp Tyr Val	Thr Phe Asp Phe Tyr Leu Asp	Pro Val Arg Ala Thr Glu	
	630	635	640
Gly Ala Met	Asn Ile Asn Leu Val Phe Gln	Pro Pro Thr Asn Gly Tyr	
	645	650	655
Trp Val Gln	Ala Pro Lys Thr Tyr Thr Ile	Asn Phe Asp Glu Leu Glu	
660	665	670	675
Glu Ala Asn	Gln Val Asn Gly Leu Tyr His Tyr	Glu Val Lys Ile Asn	
	680	685	690
Val Arg Asp	Ile Thr Asn Ile Gln Asp Asp	Thr Leu Leu Arg Asn Met	
	695	700	705

Met Ile Ile Phe Ala Asp Val Glu Ser Asp Phe Ala Gly Arg Val Phe
 710 715 720

Val Asp Asn Val Arg Phe Glu Gly Ala Ala Thr Thr Glu Pro Val Glu
 725 730 735

Pro Glu Pro Val Asp Pro Gly Glu Glu Thr Pro Pro Val Asp Glu Lys
 740 745 750 755

Glu Ala Lys Lys Glu Gln Lys Glu Ala Glu Lys Glu Glu Lys Glu Ala
 760 765 770

Val Lys Glu Glu Lys Lys Glu Ala Lys Glu Glu Lys Lys Ala Ile Lys
 775 780 785

Asn Glu Ala Thr Lys Lys
 790

<210> 5
 <211> 2343
 <212> DNA
 <213> 芽孢杆菌pHSP-K38

<220>
 <221> CDS
 <222> (580)..(2067)
 <223>

<220>
 <221> sig_peptide
 <222> (580)..(627)
 <223>

<220>
 <221> mat_peptide
 <222> (628)..()
 <223>

<400> 5
 agatctagca ggatttgccg atgcaaccgg cttatattta gagggaattt ctttttaaat 60

tgaatacga ataaatcag gtaaacaggt cctgatttta ttttttgaa ttttttgag	120
aactaaagat tgaatagaa gtagaagaca acggacataa gaaaattgta ttagtttta	180
ttatagaaaa cgcttttcta taattattta tacctagaac gaaaatactg tttcgaaagc	240
ggtttactat aaaaccttat attccggctc tttttttaa caggggggta aaattcactc	300
tagtattcta atttcaacat gctataataa atttgtaaga cgcaatatac atctttttt	360
tatgatattt gtaagcggtt aaccttgtgc tatatgccga tttaggaagg gggtagattg	420
agtcaagtag tcataattta gataacttat aagtgttga gaagcaggag agaatctggg	480
ttactcacia gttttttaa acattatcga aagcactttc ggttatgctt atgaatttag	540
ctatttgatt caattacttt aataatttta ggagtaat atg atg tta aga aag	594
Met Met Leu Arg Lys	
-15	
aaa aca aag cag ttg ggt cga cca gca caa gcc gat gga ttg aac ggt	642
Lys Thr Lys Gln Leu Gly Arg Pro Ala Gln Ala Asp Gly Leu Asn Gly	
-10 -5 -1 1 5	
acg atg atg cag tat tat gag tgg cat ttg gaa aac gac ggg cag cat	690
Thr Met Met Gln Tyr Tyr Glu Trp His Leu Glu Asn Asp Gly Gln His	
10 15 20	
tgg aat cgg ttg cac gat gat gcc gca gct ttg agt gat gct ggt att	738
Trp Asn Arg Leu His Asp Asp Ala Ala Ala Leu Ser Asp Ala Gly Ile	
25 30 35	
aca gct att tgg att ccg cca gcc tac aaa ggt aat agt cag gcg gat	786
Thr Ala Ile Trp Ile Pro Pro Ala Tyr Lys Gly Asn Ser Gln Ala Asp	
40 45 50	
gtt ggg tac ggt gca tac gat ctt tat gat tta gga gag ttc aat caa	834
Val Gly Tyr Gly Ala Tyr Asp Leu Tyr Asp Leu Gly Glu Phe Asn Gln	
55 60 65	
aag ggt act gtt cga acg aaa tac gga act aag gca cag ctt gaa cga	882
Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys Ala Gln Leu Glu Arg	
70 75 80 85	
gct att ggg tcc ctt aaa tct aat gat atc aat gta tac gga gat gtc	930
Ala Ile Gly Ser Leu Lys Ser Asn Asp Ile Asn Val Tyr Gly Asp Val	
90 95 100	
gtg atg aat cat aaa atg gga gct gat ttt acg gag gca gtg caa gct	978

Val Met Asn His Lys Met Gly Ala Asp Phe Thr Glu Ala Val Gln Ala	
105	110
115	
gtt caa gta aat cca acg aat cgt tgg cag gat att tca ggt gcc tac	1026
Val Gln Val Asn Pro Thr Asn Arg Trp Gln Asp Ile Ser Gly Ala Tyr	
120	125
130	
acg att gat gcg tgg acg ggt ttc gac ttt tca ggg cgt aac aac gcc	1074
Thr Ile Asp Ala Trp Thr Gly Phe Asp Phe Ser Gly Arg Asn Asn Ala	
135	140
145	
tat tca gat ttt aag tgg aga tgg ttc cat ttt aat ggt gtt gac tgg	1122
Tyr Ser Asp Phe Lys Trp Arg Trp Phe His Phe Asn Gly Val Asp Trp	
150	155
160	165
gat cag cgc tat caa gaa aat cat att ttc cgc ttt gca aat acg aac	1170
Asp Gln Arg Tyr Gln Glu Asn His Ile Phe Arg Phe Ala Asn Thr Asn	
170	175
180	
tgg aac tgg cga gtg gat gaa gag aac ggt aat tat gat tac ctg tta	1218
Trp Asn Trp Arg Val Asp Glu Glu Asn Gly Asn Tyr Asp Tyr Leu Leu	
185	190
195	
gga tcg aat atc gac ttt agt cat cca gaa gta caa gat gag ttg aag	1266
Gly Ser Asn Ile Asp Phe Ser His Pro Glu Val Gln Asp Glu Leu Lys	
200	205
210	
gat tgg ggt agc tgg ttt acc gat gag tta gat ttg gat ggt tat cgt	1314
Asp Trp Gly Ser Trp Phe Thr Asp Glu Leu Asp Leu Asp Gly Tyr Arg	
215	220
225	
tta gat gct att aaa cat att cca ttc tgg tat aca tct gat tgg gtt	1362
Leu Asp Ala Ile Lys His Ile Pro Phe Trp Tyr Thr Ser Asp Trp Val	
230	235
240	245
cgg cat cag cgc aac gaa gca gat caa gat tta ttt gtc gta ggg gaa	1410
Arg His Gln Arg Asn Glu Ala Asp Gln Asp Leu Phe Val Val Gly Glu	
250	255
260	
tat tgg aag gat gac gta ggt gct ctc gaa ttt tat tta gat gaa atg	1458
Tyr Trp Lys Asp Asp Val Gly Ala Leu Glu Phe Tyr Leu Asp Glu Met	
265	270
275	
aat tgg gag atg tct cta ttc gat gtt cca ctt aat tat aat ttt tac	1506
Asn Trp Glu Met Ser Leu Phe Asp Val Pro Leu Asn Tyr Asn Phe Tyr	
280	285
290	
cgg gct tca caa caa ggt gga agc tat gat atg cgt aat att tta cga	1554
Arg Ala Ser Gln Gln Gly Gly Ser Tyr Asp Met Arg Asn Ile Leu Arg	

295	300	305	
gga tct tta gta gaa gcg cat ccg atg cat gca gtt acg ttt gtt gat			1602
Gly Ser Leu Val Glu Ala His Pro Met His Ala Val Thr Phe Val Asp			
310	315	320	325
aat cat gat act cag cca ggg gag tca tta gag tca tgg gtt gct gat			1650
Asn His Asp Thr Gln Pro Gly Glu Ser Leu Glu Ser Trp Val Ala Asp			
	330	335	340
tgg ttt aag cca ctt gct tat gcg aca att ttg acg cgt gaa ggt ggt			1698
Trp Phe Lys Pro Leu Ala Tyr Ala Thr Ile Leu Thr Arg Glu Gly Gly			
	345	350	355
tat cca aat gta ttt tac ggt gat tac tat ggg att cct aac gat aac			1746
Tyr Pro Asn Val Phe Tyr Gly Asp Tyr Tyr Gly Ile Pro Asn Asp Asn			
	360	365	370
att tca gct aaa aaa gat atg att gat gag ctg ctt gat gca cgt caa			1794
Ile Ser Ala Lys Lys Asp Met Ile Asp Glu Leu Leu Asp Ala Arg Gln			
	375	380	385
aat tac gca tat ggc acg cag cat gac tat ttt gat cat tgg gat gtt			1842
Asn Tyr Ala Tyr Gly Thr Gln His Asp Tyr Phe Asp His Trp Asp Val			
390	395	400	405
gta gga tgg act agg gaa gga tct tcc tcc aga cct aat tca ggc ctt			1890
Val Gly Trp Thr Arg Glu Gly Ser Ser Ser Arg Pro Asn Ser Gly Leu			
	410	415	420
gcg act att atg tcg aat gga cct ggt ggt tcc aag tgg atg tat gta			1938
Ala Thr Ile Met Ser Asn Gly Pro Gly Gly Ser Lys Trp Met Tyr Val			
	425	430	435
gga cgt cag aat gca gga caa aca tgg aca gat tta act ggt aat aac			1986
Gly Arg Gln Asn Ala Gly Gln Thr Trp Thr Asp Leu Thr Gly Asn Asn			
	440	445	450
gga gcg tcc gtt aca att aat ggc gat gga tgg ggc gaa ttc ttt acg			2034
Gly Ala Ser Val Thr Ile Asn Gly Asp Gly Trp Gly Glu Phe Phe Thr			
	455	460	465
aat gga gga tct gta tcc gtg tac gtg aac caa taacaaaaag ccttgagaag			2087
Asn Gly Gly Ser Val Ser Val Tyr Val Asn Gln			
470	475	480	
ggattcctcc ctaactcaag gctttcttta tgtegettag ctttacgctt ctaegacttt			2147
gaagcttggg gatccgtcga gacaaggtaa aggataaaac agcacaattc caagaaaaac			2207

acgatttaga acctaaaaag aacgaatttg aactaactca taaccgagag gtaaaaaaag 2267
aacgaagtcg agatcaggga atgagtttat aaaataaaaa aagcacctga aaaggtgtct 2327
ttttttgatg tctaga 2343

<210> 6
<211> 496
<212> PRT
<213> 芽孢杆菌pHSP-K38

<400> 6

Met Met Leu Arg Lys Lys Thr Lys Gln Leu Gly Arg Pro Ala Gln Ala
-15 -10 -5 -1

Asp Gly Leu Asn Gly Thr Met Met Gln Tyr Tyr Glu Trp His Leu Glu
1 5 10 15

Asn Asp Gly Gln His Trp Asn Arg Leu His Asp Asp Ala Ala Ala Leu
20 25 30

Ser Asp Ala Gly Ile Thr Ala Ile Trp Ile Pro Pro Ala Tyr Lys Gly
35 40 45

Asn Ser Gln Ala Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr Asp Leu
50 55 60

Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys
65 70 75 80

Ala Gln Leu Glu Arg Ala Ile Gly Ser Leu Lys Ser Asn Asp Ile Asn
85 90 95

Val Tyr Gly Asp Val Val Met Asn His Lys Met Gly Ala Asp Phe Thr
100 105 110

Glu Ala Val Gln Ala Val Gln Val Asn Pro Thr Asn Arg Trp Gln Asp
115 120 125

Ile Ser Gly Ala Tyr Thr Ile Asp Ala Trp Thr Gly Phe Asp Phe Ser
 130 135 140

Gly Arg Asn Asn Ala Tyr Ser Asp Phe Lys Trp Arg Trp Phe His Phe
 145 150 155 160

Asn Gly Val Asp Trp Asp Gln Arg Tyr Gln Glu Asn His Ile Phe Arg
 165 170 175

Phe Ala Asn Thr Asn Trp Asn Trp Arg Val Asp Glu Glu Asn Gly Asn
 180 185 190

Tyr Asp Tyr Leu Leu Gly Ser Asn Ile Asp Phe Ser His Pro Glu Val
 195 200 205

Gln Asp Glu Leu Lys Asp Trp Gly Ser Trp Phe Thr Asp Glu Leu Asp
 210 215 220

Leu Asp Gly Tyr Arg Leu Asp Ala Ile Lys His Ile Pro Phe Trp Tyr
 225 230 235 240

Thr Ser Asp Trp Val Arg His Gln Arg Asn Glu Ala Asp Gln Asp Leu
 245 250 255

Phe Val Val Gly Glu Tyr Trp Lys Asp Asp Val Gly Ala Leu Glu Phe
 260 265 270

Tyr Leu Asp Glu Met Asn Trp Glu Met Ser Leu Phe Asp Val Pro Leu
 275 280 285

Asn Tyr Asn Phe Tyr Arg Ala Ser Gln Gln Gly Gly Ser Tyr Asp Met
 290 295 300

Arg Asn Ile Leu Arg Gly Ser Leu Val Glu Ala His Pro Met His Ala
 305 310 315 320

Val Thr Phe Val Asp Asn His Asp Thr Gln Pro Gly Glu Ser Leu Glu
 325 330 335

Ser Trp Val Ala Asp Trp Phe Lys Pro Leu Ala Tyr Ala Thr Ile Leu
 340 345 350

Thr Arg Glu Gly Gly Tyr Pro Asn Val Phe Tyr Gly Asp Tyr Tyr Gly
 355 360 365

Ile Pro Asn Asp Asn Ile Ser Ala Lys Lys Asp Met Ile Asp Glu Leu
 370 375 380

Leu Asp Ala Arg Gln Asn Tyr Ala Tyr Gly Thr Gln His Asp Tyr Phe
 385 390 395 400

Asp His Trp Asp Val Val Gly Trp Thr Arg Glu Gly Ser Ser Ser Arg
 405 410 415

Pro Asn Ser Gly Leu Ala Thr Ile Met Ser Asn Gly Pro Gly Gly Ser
 420 425 430

Lys Trp Met Tyr Val Gly Arg Gln Asn Ala Gly Gln Thr Trp Thr Asp
 435 440 445

Leu Thr Gly Asn Asn Gly Ala Ser Val Thr Ile Asn Gly Asp Gly Trp
 450 455 460

Gly Glu Phe Phe Thr Asn Gly Gly Ser Val Ser Val Tyr Val Asn Gln
 465 470 475 480

<210> 7

<211> 22

<212> DNA

<213> 人工序列

<400> 7

aaggatgata atccgtcccg tg 22

<210> 8
<211> 38
<212> DNA
<213> 人工序列

<400> 8
gttatccgct cacaattcgg atggatca atcaactag 38

<210> 9
<211> 38
<212> DNA
<213> 人工序列

<400> 9
cgtcgtgact gggaaaactg cgaatcaga cgggtgtac 38

<210> 10
<211> 20
<212> DNA
<213> 人工序列

<400> 10
cgteccctat cggcgggcac 20

<210> 11
<211> 25
<212> DNA
<213> 人工序列

<400> 11
atgtatatag gagttggtg gtatg 25

<210> 12
<211> 38
<212> DNA
<213> 人工序列

<400> 12
gttatccgct cacaattegc tetgacatgt caacctcc 38

<210> 13
<211> 38
<212> DNA
<213> 人工序列

<400> 13
cgtcgtgact gggaaaacag atgagaaagg aggagaag 38

<210> 14
<211> 23
<212> DNA
<213> 人工序列

<400> 14
ataactgtta ctatataatg gcc 23

<210> 15
<211> 20
<212> DNA
<213> 人工序列

<400> 15
gctggggatg acgaatccga 20

<210> 16
<211> 38
<212> DNA
<213> 人工序列

<400> 16
gttatccgct cacaattctc accttcatta tggaccac 38

<210> 17
<211> 38
<212> DNA
<213> 人工序列

<400> 17
cgtcgtgact gggaaaacca cgtctcgaca aattccg 38

<210> 18
<211> 20
<212> DNA
<213> 人工序列

<400> 18
gttgccaagc gcgatatagg 20

<210> 19
<211> 25
<212> DNA
<213> 人工序列

<400> 19
tatacagga ttatcagtat tgagc 25

<210> 20
<211> 38
<212> DNA
<213> 人工序列

<400> 20
gttatccgct cacaattctt ttctccttgt tggatctg 38

<210> 21
<211> 38
<212> DNA
<213> 人工序列

<400> 21
cgtcgtgact gggaaaacgg ggataacgat ttatgaag 38

<210> 22
<211> 30
<212> DNA
<213> 人工序列

<400> 22
tttgtaata atgatatgaa gctagtgttg 30

<210> 23
<211> 20
<212> DNA
<213> 人工序列

<400> 23
atatecagcc ctgcctcttc 20

<210> 24
<211> 58
<212> DNA

<213> 人工序列

<400> 24

ctgtgtgaaa ttgttatccg ctcacaattc gaaatttcct cctaaagcga tcataacg 58

<210> 25

<211> 51

<212> DNA

<213> 人工序列

<400> 25

gtcgtttttac aacgtcgttg actgggaaaa cccacaagct gctaacgtta c 51

<210> 26

<211> 20

<212> DNA

<213> 人工序列

<400> 26

tctgttttg gctcctgttg 20

<210> 27

<211> 26

<212> DNA

<213> 人工序列

<400> 27

tgtttatgta tggcggcctg cgggac 26

<210> 28

<211> 38

<212> DNA

<213> 人工序列

<400> 28

gltatccgct cacaattcag ctttccatat atctcacc 38

<210> 29

<211> 38

<212> DNA

<213> 人工序列

<400> 29

cgtegtgact gggaaaacac ggtctgctga tgactgac 38

<210> 30
<211> 20
<212> DNA
<213> 人工序列

<400> 30
gcgtttactt aagatgtcga 20

<210> 31
<211> 39
<212> DNA
<213> 人工序列

<400> 31
tttctagcgt ttcggcaaat tgagttaag 39

<210> 32
<211> 38
<212> DNA
<213> 人工序列

<400> 32
gttatccgct cacaattcct tactttcata cggetcac 38

<210> 33
<211> 38
<212> DNA
<213> 人工序列

<400> 33
cgtcgtgact gggaaaacga gacgtggcgc tcaccaac 38

<210> 34
<211> 29
<212> DNA
<213> 人工序列

<400> 34
cggattaataa aaagaatatc gcggacagc 29

<210> 35
<211> 20

<212> DNA
<213> 人工序列

<400> 35
tgccgctgcc cgccggagag 20

<210> 36
<211> 38
<212> DNA
<213> 人工序列

<400> 36
gttatccgct cacaattcaa ggtgtagaac ttccgttg 38

<210> 37
<211> 38
<212> DNA
<213> 人工序列

<400> 37
cgtcgtgact gggaaaacac catcaacagc ccctacac 38

<210> 38
<211> 24
<212> DNA
<213> 人工序列

<400> 38
tcaaataaag gcggcattca gtcc 24

<210> 39
<211> 22
<212> DNA
<213> 人工序列

<400> 39
ataatggtat ccaaatccac gc 22

<210> 40
<211> 38
<212> DNA
<213> 人工序列

<400> 40

gttatccgct cacaattcat tcagtcatat gtatcacc 38

<210> 41
<211> 38
<212> DNA
<213> 人工序列

<400> 41
cgtcgtgact gggaaaacga tccatcatac acagcatg 38

<210> 42
<211> 28
<212> DNA
<213> 人工序列

<400> 42
cacttctcaa cggaggggat ttcacatc 28

<210> 43
<211> 20
<212> DNA
<213> 人工序列

<400> 43
taatggagga gagaaggccg 20

<210> 44
<211> 38
<212> DNA
<213> 人工序列

<400> 44
gttatccgct cacaattcag tcgcccatac agcatgag 38

<210> 45
<211> 42
<212> DNA
<213> 人工序列

<400> 45
cgtcgtgact gggaaaacac caaaaaatgc tgagctgaca gc 42

<210> 46

<211> 26
<212> DNA
<213> 人工序列

<400> 46
ttgccaatga tgaggaaaa ggaacc 26

<210> 47
<211> 26
<212> DNA
<213> 人工序列

<400> 47
ctgaacgtcttgaataaaaaagcagg 26

<210> 48
<211> 38
<212> DNA
<213> 人工序列

<400> 48
gttatccgct cacaattcgc tgaagtttca tatccatc 38

<210> 49
<211> 38
<212> DNA
<213> 人工序列

<400> 49
cgctcgtgact gggaaaacat tccgtcatcg gcagcgag 38

<210> 50
<211> 20
<212> DNA
<213> 人工序列

<400> 50
agcggtttac aagttggagg 20

<210> 51
<211> 22
<212> DNA
<213> 人工序列

<400> 51
atttcagaag gcataactca ag 22

<210> 52
<211> 38
<212> DNA
<213> 人工序列

<400> 52
gttatccgct cacaattcca tacttgggtg tgcatcgc 38

<210> 53
<211> 40
<212> DNA
<213> 人工序列

<400> 53
cgtcgtgact gggaaaacca taatcagtaa aaaggcggtc 40

<210> 54
<211> 20
<212> DNA
<213> 人工序列

<400> 54
ttctgaccgc tctggcaacc 20

<210> 55
<211> 20
<212> DNA
<213> 人工序列

<400> 55
ataatgcccg ctteccaacc 20

<210> 56
<211> 38
<212> DNA
<213> 人工序列

<400> 56
gttatccgct cacaattccg atcctcagct cctttgtc 38

<210> 57
<211> 38
<212> DNA
<213> 人工序列

<400> 57
cgtcgtgact gggaaaactc atctgatacc gattaacc 38

<210> 58
<211> 20
<212> DNA
<213> 人工序列

<400> 58
caactgaatc cgaaggaatg 20

<210> 59
<211> 20
<212> DNA
<213> 人工序列

<400> 59
tcggggtcat gccgagcgg 20

<210> 60
<211> 38
<212> DNA
<213> 人工序列

<400> 60
gttatccgct cacaattcca atggtgcat tttcatcc 38

<210> 61
<211> 38
<212> DNA
<213> 人工序列

<400> 61
cgtcgtgact gggaaaactt gtacgagaat caacgctg 38

<210> 62
<211> 20
<212> DNA
<213> 人工序列

<400> 62
cacggcaatg cattcttcgg 20

<210> 63
<211> 21
<212> DNA
<213> 人工序列

<400> 63
agatctgtcg gccaggttta c 20

<210> 64
<211> 38
<212> DNA
<213> 人工序列

<400> 64
gttatccgct cacaattctg atttttctgt catgtctc 61

<210> 65
<211> 38
<212> DNA
<213> 人工序列

<400> 65
cgtcgtgact gggaaaacgg tagagatgtg caccgaaa 38

<210> 66
<211> 20
<212> DNA
<213> 人工序列

<400> 66
gagtcagacg gcatcgatga 20

<210> 67
<211> 23
<212> DNA
<213> 人工序列

<400> 67
ttctgattca ttttcaactgc tgg 23

<210> 68
<211> 38
<212> DNA
<213> 人工序列

<400> 68
gttatccgct cacaattcaa cggataattc ttccaatc 38

<210> 69
<211> 37
<212> DNA
<213> 人工序列

<400> 69
cgctgtgact gggaaaactg tccatgaagt caaatcc 37

<210> 70
<211> 20
<212> DNA
<213> 人工序列

<400> 70
cgctgaaata ttctetcgca 20

<210> 71
<211> 21
<212> DNA
<213> 人工序列

<400> 71
cgccgctttc accgaggatt c 21

<210> 72
<211> 38
<212> DNA
<213> 人工序列

<400> 72
gttatccgct cacaattcct ttgaccactg tatgaacc 38

<210> 73
<211> 38
<212> DNA

<213> 人工序列

<400> 73

cgtcgtgact gggaaaacac tcgtctaacg aataatcc 38

<210> 74

<211> 20

<212> DNA

<213> 人工序列

<400> 74

tgtcatcacg gaatttgacg 20

<210> 75

<211> 30

<212> DNA

<213> 人工序列

<400> 75

ccaaattatc ctttgtgagc gcggaatcag 30

<210> 76

<211> 38

<212> DNA

<213> 人工序列

<400> 76

gttatccgct cacaattccg tagatcgtaa tattgctc 38

<210> 77

<211> 38

<212> DNA

<213> 人工序列

<400> 77

cgtcgtgact gggaaaacag cttagaaagt caaccaag 38

<210> 78

<211> 20

<212> DNA

<213> 人工序列

<400> 78

tttgagcatc agcacaagcc 20

<210> 79
<211> 21
<212> DNA
<213> 人工序列

<400> 79
tgtagcagaa gcagtcgaat t 21

<210> 80
<211> 40
<212> DNA
<213> 人工序列

<400> 80
ctaattgggtg ctttagttga caattacgca gctgtcatgt 40

<210> 81
<211> 41
<212> DNA
<213> 人工序列

<400> 81
ctgccccgtt agttgaagaa ctgataaacc gtgaaaaagt g 41

<210> 82
<211> 20
<212> DNA
<213> 人工序列

<400> 82
cctttgaaaa aggctcccgt 20

<210> 83
<211> 29
<212> DNA
<213> 人工序列

<400> 83
gttttccaag tctgccgata aaaatatgc 29

<210> 84
<211> 38

<212> DNA

<213> 人工序列

<400> 84

gttatccgct cacaattcat gcttcatgta cctacacc 38

<210> 85

<211> 38

<212> DNA

<213> 人工序列

<400> 85

cgtcgtgact gggaaaacca attaacgatt cgcatacc 38

<210> 86

<211> 30

<212> DNA

<213> 人工序列

<400> 86

aaaaagaaga agtcacagta cagaacgtgg 30

<210> 87

<211> 23

<212> DNA

<213> 人工序列

<400> 87

atTTTTcgcc atcttgaatt ttc 23

<210> 88

<211> 40

<212> DNA

<213> 人工序列

<400> 88

ctaattgggtg ctttagttgg atgacacctt cgttgaactg 40

<210> 89

<211> 39

<212> DNA

<213> 人工序列

<400> 89

ctgccccggtt agttgaagg atgagccttc agaaaagtt 39

<210> 90
<211> 20
<212> DNA
<213> 人工序列

<400> 90
gccggacaga gatctgtatg 20

<210> 91
<211> 45
<212> DNA
<213> 人工序列

<400> 91
gaagaagggtt tttatggtga cgcttttttg cccaatactg tataa 45

<210> 92
<211> 45
<212> DNA
<213> 人工序列

<400> 92
caaaaaagcg tcaacataaa aaccttcttc aactaacggg gcagg 45

<210> 93
<211> 30
<212> DNA
<213> 人工序列

<400> 93
aagacgagta cttttctctc taaatcactt 30

<210> 94
<211> 30
<212> DNA
<213> 人工序列

<400> 94
aactcgatca aatggtgaca ggacagcatc 30

<210> 95

<211> 45
<212> DNA
<213> 人工序列

<400> 95
ggagaataaaa gaccctcttc aactaaagca cccattagtt caaca 45

<210> 96
<211> 45
<212> DNA
<213> 人工序列

<400> 96
tgctttagtt gaagagggtc tttattctcc cacagggttt cgttt 45

<210> 97
<211> 45
<212> DNA
<213> 人工序列

<400> 97
ttttatatt acagcgagtt ggcgttaaat gaatgaagcg ataga 45

<210> 98
<211> 45
<212> DNA
<213> 人工序列

<400> 98
atttaacgcc aactcgctgt aatataaaaa ccttcttcaa ctaac 45

<210> 99
<211> 30
<212> DNA
<213> 人工序列

<400> 99
ttgattgatg ataaattcag gcaggtgcag 30

<210> 100
<211> 30
<212> DNA
<213> 人工序列

<400> 100
caaagcttga gaaatgttcc catgctcttg 30

<210> 101
<211> 45
<212> DNA
<213> 人工序列

<400> 101
caggaggaac atatctcttc aactaaagca cccattagtt caaca 45

<210> 102
<211> 45
<212> DNA
<213> 人工序列

<400> 102
tgcttttagtt gaagagatat gttcctcctg ttccgggctg ccccg 45

<210> 103
<211> 25
<212> DNA
<213> 人工序列

<400> 103
attccagtta ctcgtaatat agttg 25

<210> 104
<211> 38
<212> DNA
<213> 人工序列

<400> 104
gttatccgct cacaattcac ttcatcatcc attagctc 38

<210> 105
<211> 38
<212> DNA
<213> 人工序列

<400> 105
cgtcgtgact gggaaaacct gctccaaatc cgatttcc 38

<210> 106
<211> 23
<212> DNA
<213> 人工序列

<400> 106
gtcctgcatt tttcgaagtc tgg 23

<210> 107
<211> 30
<212> DNA
<213> 人工序列

<400> 107
tacacatcct tcaaacaagt ctgaacaaac 30

<210> 108
<211> 45
<212> DNA
<213> 人工序列

<400> 108
tgctttagtt gaagattacc agttccataa ttccaccteg cggac 45

<210> 109
<211> 45
<212> DNA
<213> 人工序列

<400> 109
ttttatatt acagcgtgtg tataaccattg tatctgtaga tacga 45

<210> 110
<211> 30
<212> DNA
<213> 人工序列

<400> 110
gctatgatca ttgtaacgaa aggaaagggg 30

<210> 111
<211> 45
<212> DNA
<213> 人工序列

<400> 111
ttatggaact ggtaatcttc aactaaagca cccattagtt caaca 45

<210> 112
<211> 45
<212> DNA
<213> 人工序列

<400> 112
caatggtata cacacgctgt aatataaaaa ccttcttcaa ctaac 45

<210> 113
<211> 30
<212> DNA
<213> 人工序列

<400> 113
aatctgaaca agaaaaagga gctgctctc 30

<210> 114
<211> 45
<212> DNA
<213> 人工序列

<400> 114
tgctttagtt gaagaattca atctccctcc atgtcagctt attta 45

<210> 115
<211> 45
<212> DNA
<213> 人工序列

<400> 115
tttttatatt acagcagaaa cgctgaaat gaaccggccc tatag 45

<210> 116
<211> 30
<212> DNA
<213> 人工序列

<400> 116
tgtttgacaa aggtagaacg tctgcttate 30

<210> 117
<211> 45
<212> DNA
<213> 人工序列

<400> 117
ggaggagat tgaattcttc aactaaagca cccattagtt caaca 45

<210> 118
<211> 45
<212> DNA
<213> 人工序列

<400> 118
atttcaggcg tttctgctgt aatataaaaa ccttcttcaa ctaac 45

<210> 119
<211> 18
<212> DNA
<213> 人工序列

<400> 119
gaattgtgag cggataac 18

<210> 120
<211> 18
<212> DNA
<213> 人工序列

<400> 120
gttttcccag tcacgacg 18

<210> 121
<211> 19
<212> DNA
<213> 人工序列

<400> 121
caactaaagc acccattag 19

<210> 122
<211> 18
<212> DNA

<213> 人工序列

<400> 122

cttcaactaa cggggcag 18

