(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 101652468 B
(45)授权公告日 2012.09.19

(21)申请号 200880011110.0
(22)申请日 2008.04.08
(30)优先权数据
102940/2007 2007.04.10 JP
(85)PCT申请进入国家阶段日
2009.09.30
(86)PCT申请的申请数据
PCT/JP2008/057229 2008.04.08
(87)PCT申请的公布数据
WO2008/126929 EN 2008.10.23
(73)专利权人 花王株式会社
地址 日本东京都
(72)发明人 刘生浩 远藤圭二 荒胜俊
(74)专利代理机构 北京尚诚知识产权代理有限公司 11322
代理人 龙淳
(51)Int.Cl.
C12N 1/20(2006.01)
C12R 1/125(2006.01)

(56)对比文件
CN 1875095 A,2006.12.06,
CN 1513057 A,2004.07.14,
WO 2006068148 A1,2006.06.29,
Nathalie Campo 等.subcellular sites for bacterial protein export.《molecular microbiology》.2004,第53卷(第6期),
沈卫锋等.枯草芽孢杆菌作为外源基因表达系统的研究进展.《浙江农业学报》.2005,第17卷(第4期),

审查员 罗洋

权利要求书 2 页 说明书 27 页
序列表 47 页 附图 4 页

(54)发明名称

重组微生物

(57)摘要

本发明提供了一种蛋白质或多肽生产率提高的重组微生物,以及使用该重组微生物产生蛋白质或多肽的方法。该重组微生物通过将编码所需的蛋白质或多肽的基因转染到微生物株中而获得,上述微生物株通过如下方法获得:遗传构建以过度表达枯草杆菌secY基因或与secY基因对应的基因,并使选自芽孢形成相关基因和基因组中与芽孢形成相关基因对应的基因的一种或多种基因缺失或灭活。

CN 101652468 B

1. 一种重组微生物，其通过将对所需的蛋白质或多肽进行编码的基因转染到微生物中而获得，所述微生物通过如下方法获得：以遗传方式构建成过度表达枯草杆菌 (Bacillus subtilis) 的 secY 基因，并使选自芽孢形成相关基因中的一种或多种基因从基因组中缺失或灭活，

其中，所述微生物是枯草杆菌，

所述芽孢形成相关基因选自 phrA、sigF 和 sigE 基因。

2. 一种重组微生物，其通过将对所需的蛋白质或多肽进行编码的基因转染到微生物株中而获得，所述微生物株通过如下方法获得：将在微生物中具有功能的转录起始控制区或转录起始控制区 - 核糖体结合的位点导入至枯草杆菌的 secY 基因的基因组的上游，或者导入至含有枯草杆菌的 secY 基因的基因组上的操纵子的先导基因的上游；或者所述微生物株通过如下方法获得：导入一基因片段，并使选自芽孢形成相关基因的一种或多种基因缺失或灭活，在所述基因片段中，在微生物中具有功能的转录起始控制区或转录起始控制区 - 核糖体结合的位点连接在枯草杆菌的 secY 基因的上游，

其中，所述微生物是枯草杆菌，

所述芽孢形成相关基因选自 phrA、sigF 和 sigE 基因。

3. 根据权利要求 2 所述的重组微生物，其中所述在微生物中具有功能的转录起始控制区或转录起始控制区 - 核糖体结合的位点来源于枯草杆菌的 spoVG 基因。

4. 根据权利要求 1-3 中任意一项所述的重组微生物，其中选自与基因的转录、翻译和分泌有关的控制区的任何一个或多个区连接在对所需的蛋白质或多肽进行编码的基因的上游，

其中，与基因的转录、翻译和分泌有关的控制区包括：含有启动子和转录起始点的转录起始控制区、含核糖体结合位点和起始密码子的翻译起始控制区、以及分泌信号肽区。

5. 根据权利要求 4 所述的重组微生物，其中连接有所述转录起始控制区、翻译起始控制区和分泌信号区这三个区。

6. 根据权利要求 4 所述的重组微生物，其中所述分泌信号区来自枯草杆菌的纤维素酶基因，并且所述转录起始控制区和所述翻译起始控制区来自所述纤维素酶基因上游的大小为 0.6-1kb 的区。

7. 根据权利要求 4 所述的重组微生物，其中所述转录起始控制区、翻译起始控制区和分泌信号区这三个区形成由一定碱基序列构成的 DNA 片段，其中所述一定碱基序列为由 SEQ ID NO：5 所示的碱基序列的纤维素酶基因的 1 号碱基至 659 号碱基的碱基序列。

8. 一种产生如权利要求 1 所述的重组微生物的方法，其包括：在微生物中，

将在微生物中具有功能的转录起始控制区或转录起始控制区 - 核糖体结合的位点，导入至枯草杆菌的 secY 基因的基因组的上游，或者导入至含有枯草杆菌的 secY 基因的基因组上的操纵子的先导基因的上游；或者导入一基因片段，在所述基因片段中，在微生物中具有功能的转录起始控制区或转录起始控制区 - 核糖体结合的位点连接在枯草杆菌的 secY 基因的上游；

使选自芽孢形成相关基因的一种或多种基因缺失或灭活；和

将对所需的蛋白质或多肽进行编码的基因转染到微生物株中，

所述芽孢形成相关基因选自 phrA、sigF 和 sigE 基因。

9. 一种使用根据权利要求 1-7 中任意一项所述的重组微生物产生所需的蛋白质或多肽的方法。

重组微生物

发明领域

[0001] 本发明涉及用于产生有用的蛋白质或多肽的重组微生物，以及产生蛋白质或多肽的方法。

背景技术

[0002] 使用微生物在工业上生产有用的物质被用于大范围的物质，其类型包括食品，例如酒精饮料、大豆酱和酱油，以及氨基酸、有机酸、核酸相关物质、抗生素物质、碳水化合物、脂质、蛋白质等。这些物质的应用还扩展至更宽的领域，包括食品、药品、洗涤剂、日用品例如化妆品，以及多种化学原料。

[0003] 关于这种通过微生物在工业上生产有用物质，一个重要的挑战是提高生产率，并且作为其措施，已经进行了通过遗传技术例如突变来培育生产性微生物。近年来，具体地，微生物遗传学和生物技术的进展已经使人们可以使用遗传重组技术和类似技术来更有效地培育生产性微生物。另外，近年来基因组分析技术的迅速进展使得能够尝试解读目标微生物的基因组数据，并且更积极地在工业上利用获得的信息。其基因组数据已公开的工业上有用的宿主微生物的实例包括枯草杆菌 (Bacillus subtilis)Marburg 168(非专利文献 1)、大肠埃希杆菌 (Escherichia coli)K-12 MG1655(非专利文献 2)、粒状棒状杆菌 (Corynebacterium glutamicum)ATCC132032 等等，并且已经使用这些基因组数据开发了进一步改良的微生物菌株。然而，尽管进行了这些努力，生产效率未必能令人满意。

[0004] 对于某些类型的微生物，近年来已经构建了其中与芽孢形成早期相关的基因已被缺失或灭活的菌株，从而正在获得提高蛋白质或多肽的生产率的效果。例如，已有报道称，通过使用其中枯草杆菌的 sigE 基因、sigF 基因、spoIIE 基因、spoIISB 基因或 sigG 基因、或者从 spoIVCB 基因延伸至 spoIIIC 基因的区中包括的一组基因已缺失的宿主株，增加纤维素酶和类似物的分泌的生产率（专利文献 1）。

[0005] 此外，在枯草杆菌中操纵蛋白质易位系统的功能 (Sec 路线）由 SecA 分担，其作用是作为将分泌的蛋白质排出至细胞外部的发动机，并且三种蛋白质，SecY、SecE 和 SecG 构成易位 (translocation)0 通道（分泌的蛋白质经过此通道）的主要部分，以及 SecDF 是易位通道的辅因子，等等。尤其是，已报道有能够过度表达编码 SecG 蛋白质的 secG 基因的表达载体（专利文献 2)，或者其中已通过改变 SecG 基因的核糖体结合位点来改变 secG 基因的表达的革兰氏阳性细菌（专利文献 3)。由此，显示出在产生外源蛋白质期间 SecG 基因被破坏的枯草杆菌种属的繁殖受到抑制。也有报道称，在大肠埃希杆菌中，secY 基因的变异抑制低温下的繁殖或蛋白质的分泌（非专利文献 3)。

[0006] 然而，目前为止，过度表达 secY 基因、并且芽孢形成相关基因缺失或灭活的微生物还是未知的。

[0007] [专利文献 1]JP-A-2003-47490

[0008] [专利文献 2]JW-A-2001-510046

[0009] [专利文献 3]US-A-2003/0157642

[0010] [非专利文献 1]Nature,390,249,1997

[0011] [非专利文献 2]Science,277,1453,1997

[0012] [非专利文献 3]Cell,32,789,1983。

发明内容

[0013] 本发明具有以下几个方面。

[0014] (1) 一种重组微生物,其通过将对所需的蛋白质或多肽进行编码的基因转染到微生物株中而获得,所述微生物株通过如下方法获得 :以遗传方式构建成过度表达枯草杆菌的 secY 基因或与所述 secY 基因对应的基因,并使选自芽孢形成相关基因和与所述芽孢形成相关基因对应的基因中的一种或多种基因从基因组中缺失或灭活。

[0015] (2) 一种产生如权利要求 1 所述的重组微生物的方法,其包括 :在微生物中,

[0016] 将在微生物中具有功能的转录起始控制区或转录起始控制区 - 核糖体结合的位点,导入至枯草杆菌的 secY 基因或与所述 secY 基因对应的基因的基因组的上游,或者导入至含有枯草杆菌的 secY 基因或对应基因的基因组上的操纵子的先导基因的上游 ;或者导入一基因片段,其中,在所述基因片段中,在微生物中具有功能的转录起始控制区或转录起始控制区 - 核糖体结合的位点连接在枯草杆菌的 secY 基因或与所述 secY 基因对应的基因的上游 ;

[0017] 使选自芽孢形成相关基因和与所述芽孢形成相关基因对应的基因的一种或多种基因缺失或灭活 ;和

[0018] 将对所需的蛋白质或多肽进行编码的基因转染到微生物株中。

[0019] (3) 一种使用重组微生物产生所需的蛋白质或多肽的方法。

附图说明

[0020] 图 1 是显示使用通过 SOE-PCR 法制备的结合核酸片段的基因转染的示意图。

[0021] 图 2 是说明在枯草杆菌中芽孢形成信号转导的图。

[0022] 图 3 是显示通过 SOE-PCR 制备用于产生 secY 表达增强的菌株的 DNA 片段的方法的示意图。

[0023] 图 4 是显示使用 SOE-PCR 片段通过双交换 (double crossover) 法使目标基因缺失的示意图。

具体实施方式

[0024] 最佳实施方式

[0025] 本发明涉及蛋白质或多肽生产率提高的微生物,以及使用该微生物产生所需的蛋白质或多肽的方法。

[0026] 本发明的发明人已经在微生物基因组上编码的多种基因当中搜索了影响有用的蛋白质或多肽的产生的基因,并发现,当对所需的蛋白质或多肽进行编码的基因被转染到其中枯草杆菌 secY 基因已被增强以过度表达枯草杆菌的 SecY、以及芽孢形成相关基因被调控的微生物株中时,与改变之前的生产率相比,所需蛋白质或多肽的生产率提高。

[0027] 本发明的重组微生物对所需的蛋白质或所需的多肽具有很高的生产率。因此,当

使用该重组微生物进行所需蛋白质或所需多肽的生产时，可以减少生产物质所需的时间或成本。

[0028] 本发明中的氨基酸序列和碱基序列的同一性通过 Lipman-Pearson 法 (Science, 227,1435(1985)) 计算。具体地，同一性通过使用遗传数据处理软件 Genetyx-Win(Software Development Co., Ltd.) 的同源性分析程序（同源性搜索）执行分析来计算，比较的单位大小 (ktup) 设定为 2。

[0029] 在本说明书中，转录起始控制区是包括启动子和转录起始点的区，核糖体结合位点是与 Shine-Dalgarno(SD) 序列对应的位点，上述 Shine-Dalgarno(SD) 序列与起始密码子一起形成翻译起始控制区 (Proc.Natl.Acad.Sci.USA,71,1342(1974))。

[0030] 根据本发明，术语基因的上游和下游不是指关于复制起始点的位置，而是上游表示接着目的基因或区的 5′ -端的区，而下游表示接着目的基因或区的 3′ -端的区。

[0031] 用于构建本发明的微生物的亲代微生物可以是具有枯草杆菌 secY 基因或与其对应的基因的任意微生物，并且这些微生物可以是野生型微生物以及突变的微生物。具体地，可以是芽孢杆菌属的细菌、梭状芽孢杆菌 (Clostridium) 属的细菌、酵母或类似微生物，并且其中优选芽孢杆菌属细菌。而且，从下列角度来看更优选枯草杆菌 :微生物的全部基因组信息已被披露，由此构建了遗传工程和基因组工程的相关技术，并且该微生物具有分泌和产生蛋白质至细菌细胞外部的能力。

[0032] 本说明书所述的枯草杆菌的各种基因和基因区的名称均基于在 Nature,390, 249-256(1997) 中报道、并且在日本站点 :Japan FunctionalAnalysis Network for Bacillus subtilis(BSORFDB)(2004 年 3 月 10 日更新的 http://bacillus.genome.ad.JP/) 上在互联网中公布的枯草杆菌基因组数据进行描述。

[0033] 本发明的枯草杆菌 secY 基因是指具有 SEQ ID NO :1 所示的碱基序列的基因。与枯草杆菌 secY 基因对应的基因是指功能与枯草杆菌 secY 基因基本相同的基因，例如，可以是主要通过基因组分析鉴定的各 secY 基因，以及在地衣芽孢杆菌 (Bacillus licheniformis)、炭疽芽孢杆菌 (Bacillus anthracis)、蜡状芽孢杆菌 (Bacillus cereus)、苏云金芽孢杆菌 (Bacillus thuringiensis)、Oceanobacillus iheyensis 等中编码 SecY 蛋白质的基因。此外，有许多情况下，在微生物诸如炭疽芽孢杆菌中鉴定出两种类型的相应基因。作为与枯草杆菌 secY 基因对应的基因，可以是下列 (1) 至 (4) 的基因的任意者。

[0034] (1) 包括如下 DNA 的基因 :与 SEQ ID NO :1 所示的碱基序列的同一性为至少 90％、优选至少 95％、且更优选至少 99％，并且编码与具有 SEQ ID NO :2 所示的氨基酸序列的蛋白质功能相当的蛋白质。

[0035] (2) 包括如下 DNA 的基因 :在严格条件下与包括与 SEQ ID NO :1 所示的碱基序列互补的碱基序列的 DNA 杂交，并且编码与具有 SEQID NO :2 所示的氨基酸序列的蛋白质功能相当的蛋白质。

[0036] 另外，对于本文所用的“严格条件”，可以是，例如，分子克隆 - 实验手册第三版 (Molecular Cloning-A LABORATORY MANUALTHIRD EDITION)[Joseph Sambrook, David W.Russell,Cold SpringHarbor Laboratory Press] 中所述的方法，例如，可以是如下杂交条件 :在含 6×SSC(1×SSC 的组成 :0.15M 氯化钠，0.015M 柠檬酸钠，pH 7.0)、0.5％ SDS、

5×Denhardts 和 1.00mg/mL 鲱精 (herring sperm)DNA 的溶液中，与探针一起在 65℃下保持恒温 8-16 小时。

[0037] (3) 包括如下 DNA 的基因：编码与 SEQ ID NO:2 所示的氨基酸序列的同一性为至少 90％、优选至少 95％、且更优选至少 99％的氨基酸序列，并且编码与具有 SEQ ID NO:2 所示的氨基酸序列的蛋白质功能相当的蛋白质。

[0038] (4) 包括如下 DNA 的基因：编码在 SEQ ID NO:2 所示的氨基酸序列中缺失、置换或添加一个或两个或更多个氨基酸的氨基酸序列，并且编码与具有 SEQ ID NO:2 所示的氨基酸序列的蛋白质功能相当的蛋白质。

[0039] 另外，如本文用，在 SEQ ID NO:2 所示的氨基酸序列中缺失、置换或添加一个或两个或更多个氨基酸的氨基酸序列包括缺失、置换或添加一个或数个、优选 1-10 个氨基酸的氨基酸序列，并且所述添加包括在氨基酸序列的两个末端上添加一个至数个氨基酸。

[0040] 另外，与具有 SEQ ID NO:2 所示的氨基酸序列的蛋白质功能相当的蛋白质是指功能与 secY 基因编码的蛋白质基本相同、并且能够构成分泌的蛋白质所通过的易位通道的重要部分的蛋白质。

[0041] 术语枯草杆菌 secY 基因或与 secY 基因对应的基因的过度表达表示观察到在枯草杆菌 secY 基因或与 secY 基因对应的基因的宿主中表达的量超过通常的表达量。用于过度表达枯草杆菌 secY 基因或与 secY 基因对应的基因的方式的实例包括：将在微生物中具有功能的转录起始控制区或转录起始控制区－核糖体结合的位点，导入至枯草杆菌的 secY 基因或与 secY 基因对应的基因在基因组中的上游，或者导入至含有枯草杆菌的 secY 基因或与 secY 基因对应的基因的基因组上的操纵子的先导基因的上游；或者导入一基因片段，在所述基因片段中，在微生物中具有功能的转录起始控制区或转录起始控制区－核糖体结合的位点连接在枯草杆菌的 secY 基因或与 secY 基因对应的基因的上游。

[0042] 在此，在微生物中具有功能的转录起始控制区或转录起始控制区－核糖体结合的位点不特别地限制，只要该区是在用作宿主的微生物中具有功能的转录起始控制区或转录起始控制区－核糖体结合位点即可，但优选，例如，位于枯草杆菌 spoVG 基因或 aprE 基因或与这些基因的任意者对应的基因的上游的转录起始控制区或转录起始控制区－核糖体结合位点，并且更优选位于枯草杆菌 spoVG 基因或与 spoVG 基因对应的基因的上游的转录起始控制区或转录起始控制区－核糖体结合位点。

[0043] 作为枯草杆菌 spoVG 基因的转录起始控制区，可以是用于控制 spoVG 基因转录的区，spoVG 基因是在 JAFAN:Japan FunctionalAnalysis Network for Bacillus subtilis(BSORF DB) 的互联网站点 (2004 年 3 月 10 日更新的 http://bacillus.genome.ad.jp/) 中作为 Gene No.BG10112 公开的基因。更具体地，可以是具有由 SEQ ID NO:9 所示的碱基序列的 38 号碱基至 210 号碱基的碱基序列的 DNA，或者包括与上述碱基序列同源的碱基序列、并且具有枯草杆菌 spoVG 基因的转录起始控制区的功能的 DNA。此外，作为枯草杆菌 spoVG 基因的转录起始控制区－核糖体结合位点，可以是包括具有由 SEQ ID NO:9 所示的碱基序列的 38 号碱基至 230 号碱基的碱基序列的 DNA，或者包括与上述碱基序列同源的碱基序列、并且具有枯草杆菌 spoVG 基因的转录起始控制区－核糖体结合位点的功能的 DNA。

[0044] 与由 SEQ ID NO:9 所示的碱基序列的 38 号碱基至 210 号碱基的碱基序列同源的

碱基序列或与由 SEQ ID NO :9 所示的碱基序列的 38 号碱基至 230 号碱基的碱基序列同源的碱基序列的实例包括 :(A) 包括与如下 DNA 在严格条件下杂交的 DNA 的碱基序列 :上述 DNA 包括与由 SEQ ID NO :9 所示的碱基序列的 38 号碱基至 210 号碱基或 38 号至 230 号碱基的碱基序列互补的碱基序列 ;(B) 与由 SEQ ID NO :9 所示的碱基序列的 38 号碱基至 210 号碱基或 38 号碱基至 230 号碱基的碱基序列的同源性为至少 90％、优选至少 95％、且更优选至少 99％的碱基序列 ;和类似序列。

[0045] 另外,在此使用的术语“严格条件”可以如上所述的相同条件作为例子。

[0046] 术语“具有枯草杆菌 spoVG 基因的转录起始控制区或转录起始控制区 - 核糖体结合位点的功能”表示当具有功能的 DNA 被导入至枯草杆菌 secY 基因或与 secY 基因对应的基因的上游,或导入至含有枯草杆菌 secY 基因或对应基因的基因组上的操纵子 (枯草杆菌 rpsJ 基因) 的先导基因的上游时, secY 基因或与 secY 基因对应的基因被过度表达,从而导致所需的蛋白质或多肽的生产率提高,并且提高的程度还与在下列情况下获得的提高相同 :其中枯草杆菌 spoVG 基因的转录起始控制区或转录起始控制区 - 核糖体结合位点被导入至枯草杆菌 secY 基因或与 secY 基因对应的基因的上游,或导入至含枯草杆菌 secY 基因或对应基因的基因组上的操纵子 (枯草杆菌 rpsJ 基因) 的先导基因的上游。

[0047] 转录起始控制区或转录起始控制区 - 核糖体结合位点导入至 secY 基因或与 secY 基因对应的基因在基因组上的上游,或导入至含有枯草杆菌 secY 基因或对应基因的基因组上的操纵子 (枯草杆菌 rpsJ 基因) 的先导基因的上游,包括部分或全部地置换枯草杆菌 secY 基因或与 secY 基因对应的基因或含有枯草杆菌 secY 基因或对应基因的基因组上的操纵子的初始转录起始控制区,以及插入并同时保留初始转录起始控制区或转录起始控制区 - 核糖体结合位点。

[0048] 转录起始控制区或转录起始控制区 - 核糖体结合位点的置换可以例如使用涉及同源重组的已知方法来进行。即,首先,通过已知的方法例如 SOE(重叠延伸拼接,splicing by overlap extension)-PCR 法 (Gene,77,61,1989),在含有该转录起始控制区或转录起始控制区 - 核糖体结合位点的 DNA 片段的上游,连接含有含 secY 基因的操纵子的初始转录起始控制区的上游区的 DNA 片段,以及药物抗性基因片段,而在前述基因片段的下游,连接含有部分或全部的 rpsJ 基因的转录起始控制区和结构基因区 (其为含有 secY 基因的操纵子的先导基因),或者连接含有部分或全部的 rpsJ 基因的结构基因区的 DNA 片段。以这种方式,获得 DNA 片段,其中含有含 secY 基因的操纵子的初始转录起始控制区的上游区的 DNA 片段、药物抗性基因片段、含目标转录起始控制区或目标转录起始控制区结合核糖体的位点的 DNA 片段、以及含有部分或全部的 rpsJ 基因的转录起始控制区和结构基因区或含有部分或全部的 rpsJ 基因的结构基因区的 DNA 片段以该顺序连接。

[0049] 接下来,当通过已知方法将该 DNA 片段转染到亲代微生物中时,在亲代微生物基因组的两点处发生双交换同源重组,例如在含 secY 基因的操纵子的初始转录起始控制区的上游区,以及包括部分或全部的 rpsJ 基因的转录起始控制区和结构基因区或包括部分或全部的 rpsJ 基因的结构基因区的区。结果,可以使用药物抗性基因作为标记物分离出如下转化体 :其中初始转录起始控制区或转录起始控制区 - 核糖体结合位点被目标转录起始控制区或目标转录起始控制区 - 核糖体结合位点置换。以这种方式,导入至基因组上含 secY 基因的操纵子的上游的转录起始控制区或转录起始控制区 - 核糖体结合位点可

以稳定地得以遗传保留。另外，作为用转染用 DNA 片段转染宿主微生物的已知方法，特别地可以是感受态细胞转化法 (J.Bacteriol.93,1925(1967))、原生质体转化法 (Mol.Gen.Genet.168,111(1979))、电穿孔法 (FEMSMicrobiol.Lett.55,135(1990)) 等，特别优选感受态细胞转化法。

[0050] 具体地，在使用枯草杆菌作为本发明的宿主微生物的情况下，可使用 Mol.Gen.Genet.,223,268(1990) 中所述的方法或类似方法，通过同源重组，进行从含 secY 基因的操纵子的初始转录起始控制区或转录起始控制区 - 核糖体结合位点到目标转录起始控制区或目标转录起始控制区 - 核糖体结合位点的置换。

[0051] 如果适当地选择要插入到需要插入的转录起始控制区或转录起始控制区 - 核糖体结合位点的两个末端的 DNA 片段的序列，可以通过如上述置换方法相同的方法，进行目标转录起始控制区或转录起始控制区 - 核糖体结合位点的插入。例如，在该转录起始控制区的上游，连接含有含 secY 基因的操纵子的初始转录起始控制区的上游区的 DNA 片段，以及药物抗性基因，而在该转录起始控制区的下游，连接含有部分或全部的初始转录起始控制区的 DNA 片段。以这种方式，获得 DNA 片段，其中含有含 secY 基因的操纵子的初始转录起始控制区的上游区的 DNA 片段、药物抗性基因片段、目标转录起始控制区、以及含有部分或全部的初始转录起始控制区的 DNA 片段以该顺序连接。随后，将该 DNA 片段插入至宿主微生物中，然后可以使用药物抗性基因作为标记物分离出转化体。在由此分离的转化体的基因组上，含 secY 基因的操纵子的初始转录起始控制区和目标转录起始控制区可以稳定地保留，同时二者相互邻接，其间没有任何空隙。另外可选地，当制备和使用其中含 secY 基因的上游区的 DNA 片段和药物抗性基因连接在目标转录起始控制区的上游、而含部分或全部的 secY 基因的 DNA 片段连接在目标转录起始控制区的下游的 DNA 片段时，目标转录起始控制区可以稳定地保留，同时立即被导入至 secY 基因的上游。

[0052] 根据本发明，导入有在微生物中具有功能的转录起始控制区或转录起始控制区 - 核糖体结合位点的基因组的上游包括 secY 基因。上游不受特别的限制，只要该区在 rpsJ 基因（其为操纵子先导基因）的起始密码子的上游侧，或 secY 基因的起始密码子的上游侧即可，但优选为包括 2000 个相邻碱基对的区，更优选包括 500 个碱基对的区，再更优选包括 100 个碱基对的区，再更优选包括 50 个碱基对的区。

[0053] 根据本发明，使用通过将目标转录起始控制区或转录起始控制区 - 核糖体结合位点的片段与 secY 基因或与 secY 基因对应的基因的片段连接获得的基因片段，可以进行其中目标转录起始控制区或转录起始控制区 - 核糖体结合位点连接在 secY 基因或与 secY 基因对应的基因的上游的基因片段的导入，上述片段已经通过已知的克隆方法获得，例如使用 PCR 法，使用枯草杆菌以外的微生物的基因组作为模板，通过已知方法例如限制性内切酶法或 SOE（重叠延伸拼接）-PCR 法 (Gene,77,61(1989)) 而获得。可以根据已知的转化方法，通过在导入至细胞中的核酸片段与染色体之间的同源重组，可以将这些片段导入染色体。

[0054] 要导入的 secY 基因或与 secY 基因对应的基因的碱基序列可以与微生物原始具有的 secY 基因或与该 secY 基因对应的基因的碱基序列不一致，只要它是 secY 基因或与 secY 基因对应的基因的碱基序列即可。此外，要导入的在微生物中具有功能的转录起始控制区或转录起始控制区 - 核糖体结合位点（例如枯草杆菌 spoVG 基因的转录起始控制区或转录

起始控制区－核糖体结合位点）的碱基序列，可以与微生物所具有的碱基序列不一致，只要它是目标碱基序列即可。将核酸片段导入至宿主中的方法的例子可以是感受态细胞转化法、原生质体转化法、电穿孔法等，并且特别优选感受态细胞转化法。

[0055] 此外，这些片段也可以通过载体例如质粒导入至细胞质中。另外，如稍后所述的实施例中所示，由于在通过质粒的方式导入时，通过每个细菌细胞导入一个拷贝，这些片段对于所需蛋白质或多肽的产生发挥充分的效应，因此即使在生产和培养期间一些质粒丢失也几乎不影响该片段。

[0056] 另外，至于允许导入在宿主中的染色体区，优选非必需基因的内在部分、或非必需基因上游的非基因区的内在部分。例如，可以是 aprE 基因、sacB 基因、nprE 基因、amyE 基因和 ybxG 基因的内在部分，或这些基因上游的非基因区的内在部分，但优选 amyE 基因的内在部分、或 ybxG 基因上游的非基因区的内在部分。

[0057] 如本文所用，术语“非必需基因”是指即使当该基因被破坏时，至少在特定条件下具有该基因的宿主仍能存活的基因。另外，即使导入伴随着非必需基因的缺失，或部分或全部的非必需基因上游的非基因区的缺失，仍不会由此产生任何问题。

[0058] 根据本发明，枯草杆菌的 secY 基因或与 secY 基因对应的基因的过度表达，以及与枯草杆菌的 Sec 途径相关的另一基因（例如，secE 基因等）的过度表达可以在不影响所需蛋白质或多肽的生产率的提高的范围内进行，并且一种或两种或多种基因的灭活或缺失也可以并行地实现。另外，基因的灭活或缺失包括基因中的部分或全部碱基的置换和缺失，以及将碱基插入至基因中。

[0059] 下文将更详细地说明使用根据 SOE（重叠延伸拼接）-PCR 法 (Gene，77，61(1989)) 制备的 DNA 片段，通过双交换的方式将其中 spoVG 基因的转录起始控制区或转录起始控制区－核糖体结合位点连接在 secY 基因上游的基因片段导入至宿主基因组的方法。然而，根据本发明的导入方法并无意于局限于以下方法。

[0060] 用于本发明方法中的导入用 DNA 片段是如下的 DNA 片段：其中在邻近宿主基因组上的导入位点上游的大小为大约 0.1-3kb、优选 0.4-3kb 的片段（下文中称为片段 (1)）和邻近导入位点下游的大小为大约 0.1-3kb、优选 0.4-3kb 的片段（下文中称为片段 (2)）之间，以如下顺序插入有含有 spoVG 基因的转录起始控制区或转录起始控制区－核糖体结合位点的片段（下文中称为片段 (3)）、secY 基因片段（下文中称为片段 (4)）、和药物抗性标记基因例如氯霉素抗性基因的片段（下文中称为片段 (5)）。首先，在第一轮 PCR 中制备片段 (1) 至片段 (5) 的 5 个片段。这时，使用如下设计的引物：例如，片段 (3) 上游侧上的 10-30 个碱基对的序列加入至片段 (1) 的下游端，片段 (3) 下游侧上的 10-30 个碱基对的序列加入至片段 (4) 的上游端，片段 (5) 上游侧上的 10-30 个碱基对的序列加入至片段 (4) 的下游端，并且片段 (5) 下游侧上的 10-30 个碱基对的序列加入至片段 (2) 的上游（图 1）。

[0061] 随后，使用在第一轮中制备的 5 种类型的 PCR 片段作为模板，并使用位于片段 (1) 上游的引物和位于片段 (2) 下游的引物进行第二轮 PCR。从而，片段 (3) 在加入至片段 (1) 下游端的片段 (3) 的序列中发生退火，片段 (3) 在加入至片段 (4) 上游端的片段 (3) 的序列中发生退火，片段 (5) 在加入至片段 (4) 下游端的片段 (5) 的序列中发生退火，并且片段 (5) 在加入至片段 (2) 上游的片段 (5) 的序列中发生退火。由此，作为 PCR 扩增的结果，可以得到其中片段 (1) 至片段 (5) 的五个片段以 (1)、(3)、(4)、(5) 和 (2) 的顺序连接的 DNA

片段（图 1）。

[0062] 在此进行的 PCR 反应可以在文献（PCR Protocols. Current methodsand Applications，B. A. White 主编，Humana Press，pp. 251，1993；Gene，77，61(1989)）中所述的常规条件下，使用表 1 所示的引物组，并使用常规 PCR 用酶试剂盒例如 Pyrobest DNA 聚合酶（Takara Shuzo Co.，Ltd.）而有利地进行。

[0063] 当由此获得的转染用 DNA 片段通过感受态法或类似方法转染到细胞中时，在细胞内，在基因组上导入位点的上游和下游存在同一性的同源区中发生基因重组，并且通过基于药物抗性标记物的选择，可以分离出用如下基因片段转染的细胞：其中 spoVG 基因的转录起始控制区或转录起始控制区 - 核糖体结合位点连接在 secY 基因的上游。基于药物抗性标记物的选择可以通过以下方法有利地进行：分离在含氯霉素的琼脂培养基上生长的集落，然后选择基因组上的导入通过 PCR 法使用该基因组作为模板而得到确认的细胞，等等。

[0064] 另外，药物抗性标记基因不受特别的限制，只要它可以用于使用通用的抗生素物质的选择即可，但除氯霉素抗性基因以外，还可以是例如红霉素抗性基因、新霉素抗性基因、壮观霉素（spectinomycin）抗性基因、四环素抗性基因和灭瘟素 S 抗性基因的药物抗性标记基因。

[0065] 在本发明的重组微生物中，除了由 secY 基因增强引起的 SecY 过度表达以外，还实现了从基因组中使一种或多种芽孢形成相关基因和与其对应的基因缺失或灭活。如实施例 2 和 3 所示，基因的缺失或灭活是抑制芽孢形成的变化。

[0066] 根据本发明，芽孢形成相关基因的实例包括加速芽孢形成的多组基因，由此这些基因的每一种的缺失或灭活导致芽孢形成的过程基本上被抑制，例如，编码对芽孢形成各期有特异性的 sigma 因子的一组基因，或者与 sigma 因子基因的表达和 sigma 因子的活化相关的一组基因。此外，还包括由相应的 sigma 因子转录并且参与加速芽孢形成的一组基因。

[0067] 在芽孢杆菌属细菌中，对于枯草杆菌，已经鉴定出 17 种 sigma 因子，并且已知存在有如下 sigma 因子：从 SigA 开始，SigA 是参与营养生长期中生长所必需的基因的转录的主要 sigma 因子（看家 sigma 因子），SigH、SigF、SigE、SigG 和 SigK 是控制芽孢形成过程的 sigma 因子，SigD 是控制鞭毛形成或细胞壁消化的 sigma 因子，SigL 是控制某些类型氨基酸或糖的代谢的 sigma 因子，SigB 是控制对环境变化的响应的 sigma 因子，称为 ECF sigma 因子的 sigma 因子等（Bacillussubtilis and Its Closest Relatives：From Genes to Cells，A. L. Sonenshein 主编，American Society for Microbiology，289 页（2002））。

[0068] 其中，已知控制芽孢形成过程的 sigma 因子根据芽孢形成过程的进程而相继地表达和活化，如图 2 所示。换句话说，当枯草杆菌进入营养饥饿状态时，首先经由涉及多种蛋白质的多级磷酸盐转导系统（称为磷酸中继系统）发生 Spo0A（其为芽孢形成起始控制因子）的磷酸化（Cell，64，545(1991)）。更具体地，由于营养饥饿导致细胞质中存在的 KinA 以及细胞膜中存在的 KinB 和 KinC 发生自磷酸化，并且磷酰基经由 Spo0F 和 Spo0B 转移至 Spo0A，从而产生磷酸化的 Spo0A（磷酸化 Spo0A）。此外，在涉及 KinB 的芽孢形成过程的活化过程中需要 KapB(Mol. Microbiol.，26，1097(1997))，同时 KipA 与 KipI（其为 KinA 的自磷酸化抑制剂）结合，以避免其芽孢形成抑制作用（GenesDev.，11，2569(1997)）。另外，PhrA 抑制 RapA 的功能，RapA 是磷酸化 Spo0F 的去磷酸化酶（Proc. Natl. Acad. Sci. USA，94，

8612(1997))。KinA、KinB、KinC、Spo0F、Spo0B、Spo0A、KapB、KipI、KipA、RapA 和 PhrA 分别由基因 kinA、kinB、kinC、spo0F、spo0B、spo0A、kapB、kipI、kipA、rapA 和 phrA 编码。

[0069] 伴随着磷酸化 Spo0A 浓度的增加，抑制 SigH 的结构基因 (sigH) 表达的阻抑物 AbrB 的诱导受到抑制，因此，sigH 的转录以依赖于 SigA 的方式被诱导 (J.Bacteriol.，173，521(1991))。另外，在磷酸化 Spo0A 转录调节功能被参与染色体分离的 Soj 抑制的同时，也参与染色体分离的 Spo0J 抑制 Soj 的该作用 (J.Bacteriol.，182，3446，2000)。Soj 和 Spo0J 分别由 soj 基因和 spo0J 基因编码。在 SigH 活化后，不对称隔膜的形成将枯草杆菌的细胞质分隔为母细胞侧和子细胞侧。随后，在子细胞侧，磷酸化 Spo0A 和 SigH 接合，以诱导含有 SigF 的结构基因 (sigF) 的操纵子 (spoIIAA-spoIIAB-sigF) 的表达 (Gene，101，113(1991))，并且在母细胞侧，磷酸化 Spo0A 和 SigA 接合，以诱导含 SigE 前体的结构基因 (sigE) 的操纵子 (spoIIGA-sigE) 的转录 (J.Bacteriol.，169，3329(1987))。存在有两级抑制，其中 SigF 被抗 sigma 因子 SpoIIAB 在功能上抑制，而抗 - 抗 sigma 因子 SpoIIAA 抑制 SpoIIAB 的作用。即，如 SigF 的功能缺失的情况一样，SpoIIAA 的功能缺失导致芽孢形成的抑制，已知 SpoIIAB 的功能缺失也能抑制芽孢形成 (Proc.Natl.Acad.Sci.USA，87，9221(1990)；J.Bacteriol.，173，6678(1991))。而且，活化由 SpoIIE 控制，SpoIIE 是 SpoIIAA 的去磷酸化酶 (Genes Cells，1，881(1996))，并且活化的 SigF 诱导 SpoIIR(其为信号转导蛋白质) 的结构基因的转录。据设想，从子细胞侧分泌的 SpoIIR 活化 SpoIIGA，SpoIIGA 是位于母细胞侧上的不对称隔膜中的 SigE 前体活化蛋白酶，并且由此发生 SigE 的活化 (Proc.Natl.Acad.Sci.USA，92，2012(1995))。SpoIIAA、SpoIIAB、SpoIIE、SpoIIR 和 SpoIIGA 分别由 spoIIAA、spoIIAB、spoIIE、spoIIR 和 SpoIIGA 基因编码。此外，在子细胞侧，SigF 诱导 SigG 的结构基因 (sigG) 的转录，并且在母细胞侧，SigE 诱导 SigK 的结构基因 (spoIIIC 基因和 spoIVCB gene) 的转录。然而，在母细胞侧上的 SigE 活化后，在子细胞侧上发生 SigG 的活化，并且此后在母细胞侧上发生 SigK 的活化 (Mol.Microbiol.，31，1285(1999))。

[0070] 从图 2 中很显然，在上述的基因中，当缺失或灭活时抑制芽孢形成过程的基因是 kinA、kinB、kinC、spo0F、spo0B、spo0A、kapB、kipA、phrA、spo0J、sigH、sigF、sigE、spoIIAA、spoIIAB、spoIIE、spoIIR、spoIIGA、sigG、spoIIIC 和 spoIVCB 基因。在本发明中要缺失或灭活的芽孢形成相关基因优选地选自属于枯草杆菌的这些基因，并且更优选地选自枯草杆菌的 sigF 基因、sigE 基因和 phrA 基因。

[0071] sigF 基因是编码 sigma 因子的基因，其负责在芽孢形成阶段期间从 II 期开始在枯草杆菌细胞中形成不对称隔膜处的子细胞侧中发生的基因表达，而 sigE 基因是编码 sigma 因子的基因，其负责在芽孢形成阶段期间从 II 期开始在枯草杆菌细胞中形成不对称隔膜处的母细胞侧中发生的基因表达。

[0072] 此外，phrA 基因是涉及感受外部生长环境中的变化并以各种方式对其响应所需要的细胞间信息传输的机制中的基因之一，并且基因产物暂时分泌至细胞外。据报道，在细胞外被加工后，该基因作为五肽被摄取至细胞中，并且与 RapA 蛋白质结合，RapA 控制用于传输芽孢形成起始信号的磷 - 中继系统中的 Spo0F 的磷酸化，从而参与芽孢形成起始信号的转导 (Proc.Natl.Acad.Sci.USA，94，8612(1997))。

[0073] 上述基因的每一种基因的基因编号和功能总结于表 6。

[0074] 与芽孢形成相关基因诸如 kinA、kinB、kinC、spo0F、spo0B、spo0A、kapB、kipA、phrA、spo0J、sigH、sigF、sigE、spoIIAA、spoIIAB、spoIIE、spoIIR、spoIIGA、sigG、spoIIIC 和 spoIVCB 对应的基因是功能与上述基因的任一种基本相同的基因，并且例如，可以是来自另一微生物的基因，优选来自芽孢杆菌属细菌的基因，其碱基序列与这些基因的任一种的碱基序列的同一性为至少 70％、优选至少 80％、更优选至少 90％、再更优选至少 95％、再更优选至少 98％。另外，碱基序列的同一性在此根据 Lipman-Pearson 法计算 (Science，227，1435(1985))。

[0075] 这些基因的缺失或灭活可以是前述各种基因当中单个基因的缺失或灭活，或可以是其两种或多种的组合的缺失或灭活。此外，目标基因之外的基因的增强或缺失或灭活也可以并行地进行。另外，基因的缺失或灭活包括基因中的部分或全部碱基的置换或缺失，以及将碱基插入至基因中。

[0076] 对于一组基因或单个基因的缺失或灭活的顺序，可以是有意地缺失或灭活目标基因（靶基因）的方法，以及通过缺失或灭活使随机基因突变、然后进行蛋白质生产率的评价并通过适当方法进行遗传分析的方法。

[0077] 为了使靶基因缺失或灭活，例如，可以使用涉及同源重组的方法。即，可以将通过将含有部分靶基因的 DNA 片段克隆到适当的质粒中获得的环状重组质粒，转染到亲代微生物的细胞中，使得亲代微生物的基因组上的靶基因被靶基因的某些区中的同源重组分开，从而使靶基因灭活。另外可选地，如图 3 所示，根据 PCR 法或类似方法，通过构建经由突变例如碱基置换或碱基插入而灭活的靶基因，或者构建含有靶基因的上游和下游区但不含有靶基因的线性 DNA 片段，并将所得物转染到亲代微生物细胞中，从而在亲代微生物的基因组上的靶基因内突变位点外的两个位点处，或者在靶基因的上游侧和下游侧上，引起双交换同源重组，也可以用缺失或灭活的基因片段置换基因组上的靶基因。

[0078] 具体地，在使用枯草杆菌作为构建本发明的微生物的亲代微生物的情况下，已经有几个关于通过同源重组缺失或灭活靶基因的方法的报道 (Mol.Gen.Genet.，223，268(1990) 等)，由此通过重复这些方法可以获得本发明的宿主微生物。

[0079] 随机基因的缺失或灭活也可以通过诱导同源重组的方法进行，例如使用随机克隆 DNA 片段的上述方法，或者通过用放射照射亲代微生物的方法进行。

[0080] 下文将更具体地解释使用根据 SOE（重叠延伸拼接）-PCR 法 (Gene，77，61(1989)) 制备的缺失用 DNA 片段通过双交换进行缺失的方法，但本发明中的基因缺失方法并不限于以下方法。

[0081] 本发明中使用的缺失用 DNA 片段是通过将药物抗性标记基因片段插入在邻接于要被缺失基因上游的大小为大约 0.1-3kb、优选 0.4-3kb 的片段和邻接于要被缺失基因下游的大小为大约 0.1-3kb、优选 0.4-3kb 的片段之间而制备的片段。首先，通过第一轮 PCR 制备三个片段：要被缺失基因的上游片段和下游片段，以及药物抗性标记基因片段，这时，使用如下设计的引物：例如，药物抗性标记基因上游侧上的 10-30 个碱基对的序列加入至上游片段的下游端，反过来，药物抗性标记基因下游侧上的 10-30 个碱基对的序列加入至下游片段的上游端（图 4)。

[0082] 随后，使用在第一轮中制备的三种类型的 PCR 片段作为模板，并使用上游片段的上游侧引物和下游片段的下游引物进行第二轮 PCR。从而，药物抗性标记基因片段在加入

至上游片段的下游端和下游片段的上游端的药物抗性标记基因序列中发生退火，并且作为PCR扩增的结果，可以得到药物抗性标记基因插入在上游侧片段和下游侧片段之间的DNA片段（图4）。

[0083] 在使用氯霉素抗性基因作为药物抗性标记基因的情况下，通过在文献（PCR Protocols.Current methods and Applications,B.A.White主编,Humana Press,pp.251,1993;Gene,77,61(1989)）中所述的常规条件下，使用表1所示的引物组，并使用常规PCR用酶试剂盒例如Pyrobest DNA聚合酶（Takara Shuzo Co.,Ltd.）来进行SOE-PCR，获得用于缺失各种基因的DNA片段。

[0084] 当由此获得的缺失用DNA片段通过感受态法或类似方法转染到细胞中时，在细胞内，在要缺失的基因的上游和下游的存在同一性的同源区中发生基因重组，并且通过基于药物抗性标记物的选择，可以分离出其中所需基因已经被药物抗性基因置换的细胞。即，当使用表1所示的引物组制备的缺失用DNA片段被转染时，可以分离在含氯霉素的琼脂培养基上生长的集落，并通过使用基因组作为模板的PCR法或类似方法，确认基因组上的所需基因被氯霉素抗性基因置换。

[0085] 本发明的微生物可以通过将编码所需蛋白质或所需多肽的基因转染到由此产生的微生物中而获得。在此，术语“所需蛋白质或多肽”是指目的之一是生产或纯化的蛋白质或多肽。另外，关于“具有编码所需蛋白质或多肽的基因的微生物”，其中的基因意在包括微生物原本具有的基因，以及该微生物原本不具有的基因，即外源基因。

[0086] 所需蛋白质或所需多肽不受特别的限制，其实例包括在洗涤剂、食品、纤维、饲料、化学品、药物、诊断等中使用的各种工业酶或生理活性肽，并且优选工业酶。此外，就工业酶的功能而言，包括氧化还原酶、转移酶、水解酶、裂解酶、异构酶、连接酶/合成酶等，并且可以优选为水解酶，例如纤维素酶、γ-淀粉酶和蛋白酶。

[0087] 对于蛋白酶，可以是来自微生物的蛋白酶，优选来自芽孢杆菌属细菌的蛋白酶，更优选来自克劳氏芽孢杆菌(Bacillus clausii)KSM-K16株(FERM BP-3376)的蛋白酶。来自克劳氏芽孢杆菌KSM-K16株的碱性蛋白酶的更具体的实例包括来自芽孢杆菌属细菌的碱性蛋白酶，其包括SEQ ID NO:4所示的氨基酸序列的1号氨基酸至380号氨基酸的氨基酸序列，或者是包括与上述氨基酸序列的同一性为至少70%、优选至少80%、更优选至少90%、再更优选至少95%、再更优选至少98%的氨基酸序列的蛋白酶。

[0088] 对于纤维素酶，可以是属于多糖水解酶类别中第5家族的纤维素酶(Biochem.J.,280,309(1991))，并且其中，可以是来自微生物的纤维素酶，尤其是来自芽孢杆菌属细菌的纤维素酶。例如，可以是来自芽孢杆菌属KSM-S237株(FERM BP-7875)和芽孢杆菌属KSM-64株(FERM BP-2886)的纤维素酶，并且其合适的实例包括来自芽孢杆菌属细菌的碱性纤维素酶，其包括SEQ ID NO:6所示的氨基酸序列的1号氨基酸至795号氨基酸的氨基酸序列，或者是包括SEQ ID NO:8所示的氨基酸序列的1号氨基酸至793号氨基酸的氨基酸序列的碱性纤维素酶，或者是包括与上述氨基酸序列的同一性为至少70%、优选至少80%、更优选至少90%、再更优选95%、再更优选至少98%的氨基酸序列的纤维素酶。

[0089] 此外，对于γ-淀粉酶，可以是来自微生物的γ-淀粉酶，优选来自芽孢杆菌属细菌的γ-淀粉酶，并且更优选来自芽孢杆菌属KSM-K38株的γ-淀粉酶。

[0090] 合意地，要转染到本发明的微生物中的所需蛋白质或所需多肽的基因与上游的一

个或多个区以适当的顺序连接，该一个或多个区选自以下区：与基因的转录、翻译和分泌有关的控制区，即，含有启动子和转录起始点的转录起始控制区；含核糖体结合位点和起始密码子的翻译起始控制区；以及分泌信号肽区。具体地，优选具有三个结合区，包括转录起始控制区、翻译起始控制区和分泌信号肽区的基因，另外，合意地，分泌信号肽区来自芽孢杆菌属细菌的纤维素酶基因，而转录起始控制区和翻译起始控制区是纤维素酶基因上游的大小为 0.6-1kb 的区，并且这些区适当地与所需蛋白质或所需多肽的基因连接。例如，合意地，来自芽孢杆菌属细菌（即 KSM-S237 株 (FERM BP-7875) 或 KSM-64 株 (FERM BP-2886)）的纤维素酶基因以及该纤维素酶基因的转录起始控制区、翻译起始控制区和分泌信号肽区适当地与所需蛋白质或所需多肽的结构基因连接。更具体地，合意地，包括 SEQ ID NO:5 所示的碱基序列的 1 号碱基至 659 号碱基的碱基序列的 DNA 片段，或包括 SEQ ID NO:7 所示的碱基序列的 1 号碱基至 696 号碱基的碱基序列的 DNA 片段，或者包括与上述碱基序列的同一性为至少 70％、优选至少 80％、更优选至少 90％、再更优选至少 95％、再更优选至少 98％的碱基序列的 DNA 片段，或包括由任意上述碱基序列的一部分的缺失、置换或添加得到的碱基序列的 DNA 片段，适当地与所需蛋白质或所需多肽的结构基因连接。另外，如本文所用，包括由任意上述碱基序列的一部分的缺失、置换或添加得到的碱基序列的 DNA 片段是指其中任意上述碱基序列的一部分被缺失、置换或添加，但保留了与基因的转录、翻译和分泌相关的功能的 DNA 片段。

[0091] 编码所需蛋白质或所需多肽的这些基因的导入可以通过如下方法进行：例如，(1) 通过载体导入，或 (2) 插入到基因组中。在 (1) 通过载体导入的情况下，可以通过适当的转化法，例如感受态细胞转化法、原生质体转化法或电穿孔法，导入含有以下基因的载体：该基因编码所需蛋白质或所需多肽，并且在上游与选自与基因的转录、翻译和分泌相关的控制区（即含启动子和转录起始点的转录起始控制区）、含核糖体结合位点和起始密码子的翻译起始控制区和分泌信号肽区的一个或多个区以适当的形式连接。在此，载体不受特别的限制，只要它是用于将所需基因转染到宿主中以便增殖并表达该基因的合适的载体核酸分子即可，并且，载体可以是质粒，也可以是，例如，人工染色体，例如 YAC 和 BAC，使用转座子的载体，黏端质粒等等。质粒的实例包括 pUB110 和 pHY300PLK。

[0092] 此外，(2) 插入至基因组中可以使用例如涉及同源重组的方法进行。即，具有用于诱导与编码所需蛋白质或所需多肽的基因连接的导入的染色体区的一部分的 DNA 片段，可以通过将该 DNA 片段转染到微生物的细胞中、并在染色体区的一些部分中诱导同源重组而结合至基因组中。在此，用于诱导导入的染色体区不受特别的限制，但优选非必需基因区或非必需基因区上游的非基因区。

[0093] 使用本发明的重组微生物产生所需的蛋白质或多肽可以通过如下方法进行：将细菌菌株接种至含有可同化碳源、氮源和其他必需组分的培养基中，通过常规微生物培养方法培养菌株，并且培养完成后，收集和纯化蛋白质或多肽。如在后述的实施例中所述，与使用基因未被改变的微生物的情况相比，所需蛋白质或多肽的生产率得到提高。

[0094] 下文将详细地描述构建本发明的重组微生物的方法，以及使用重组微生物产生纤维素酶和淀粉酶的方法。

[0095] 实施例

[0096] 在以下实施例中用于扩增 DNA 片段的聚合酶链式反应 (PCR) 中，使用 GeneAmp PCR

系统(Applied Biosystems,Inc.)并使用PyrobestDNA聚合酶(Takara Bio,Inc.)和辅助试剂来进行DNA扩增。通过加入1μL适当稀释的模板DNA、20pmol各正义引物和反义引物、以及2.5U Pyrobest DNA聚合酶,并将反应溶液的总量调节至50μL,制备PCR反应溶液的组合物。在下列反应条件下进行PCR:重复30轮在98℃下10秒、在55℃下30秒和在72℃下1-5分钟(根据所需的扩增产物调节。大致标准是1分钟/1kb)的三阶段温度变化,然后反应在72℃下继续进行5分钟。

[0097] 此外,在下列实施例中,基因的上游和下游不是指从复制起始点开始的位置,而是,上游表示在各种操作和过程中紧接着基因或目的区域的5′-端的区域,而下游表示在各种操作和过程中紧接着基因或目的区域的3′-端的区域。

[0098] 此外,在下列实施例中各种基因和基因区的名称基于在Nature,390,249-256(1997)中报道、并在日本站点Japan Functional Analysis Networkfor Bacillus subtilis(BSORF DB)(http://bacillus.genome.ad.JP/,2004年3月10日更新)上在互联网上公布的枯草杆菌基因组数据进行描述。

[0099] 枯草杆菌的转化以下列方式进行。具体地,将枯草杆菌株在SPI培养基(0.20%硫酸铵、1.40%磷酸氢二钾、0.60%磷酸二氢钾、0.10%柠檬酸三钠二水合物、0.50%葡萄糖、0.02%酪蛋白氨基酸(DifcoLaboratories,Inc.)、5mM硫酸镁、0.25μM氯化锰和50μg/mL色氨酸)中在37℃下振荡培养,直至生长度值(OD600)达到大约1。在振荡培养后,将一些培养溶液接种至9倍量的SPII培养基(0.20%硫酸铵、1.40%磷酸氢二钾、0.60%磷酸二氢钾、0.10%柠檬酸三钠二水合物、0.50%葡萄糖、0.01%酪蛋白氨基酸(Difco Laboratories,Inc.)、5mM硫酸镁、0.40μM氯化锰和5μg/mL色氨酸)中,并将细胞进一步振荡培养,直至生长度值(OD600)达到大约0.4。由此,制备枯草杆菌的感受态细胞。

[0100] 随后,向100μL由此制备的感受态细胞悬液(SPII培养基中的培养溶液)中加入5μL含各种DNA片段的溶液(SOE-PCR反应溶液等),将混合物在37℃下振荡孵育1小时,并且将全部的量涂在含适当药物的LB琼脂培养基(1%色氨酸、0.5%酵母提取物、1% NaCl和1.5%琼脂)上。在37℃下静止培养后,分离出生长的集落作为转化体。提取所获得的转化体的基因组,并使用该基因组作为模板通过PCR确认实现了所需的基因组结构的改变。

[0101] 将编码所需蛋白质或多肽的基因转染到宿主微生物中遵循下列的任一种方法进行:感受态细胞转化法(J.Bacteriol.,93,1925(1967))、电穿孔法(FEMS Microbiol. Lett.,55,135(1990))和原生质体转化法(Mol.Gen.Genet.,168,111(1979))。

[0102] 对于用于通过重组微生物产生蛋白质的培养,使用LB培养基(1%色氨酸、0.5%酵母提取物和1% NaCl)、2×YT培养基(1.6%色氨酸、1%酵母提取物和0.5% NaCl)、2×L-麦芽糖培养基(2%色氨酸、1%酵母提取物、1% NaCl、7.5%麦芽糖和7.5ppm硫酸锰四水合物或五水合物)或CSL发酵培养基(2%酵母提取物、0.5%玉米浸出液(CSL)、0.05%氯化镁七水合物、0.6%尿素、0.2% L-色氨酸、10%葡萄糖、0.15%磷酸二氢钠和0.35%磷酸氢二钠,pH 7.2)。

[0103] 实施例1构建过度表达secY基因的菌株

[0104] 如下进行过度表达secY基因的变异体的构建(见图3)。使用从枯草杆菌168株中提取的基因组DNA作为模板,并使用PVG-FW和PVG-R,以及secY/PVG-F和secY/Cm-R引物组,通过PCR扩增含spoVG基因的转录起始控制区和核糖体结合位点的0.2kb片段(A)和

含 spoVG 基因的 1.3kb 片段 (B)。此外，使用质粒 pC194(J.Bacteriol.，150(2)，815(1982)) 作为模板，并使用表 1 所示的 catf 和 catr 引物组，通过 PCR 扩增含氯霉素 (Cm) 抗性基因的 0.9kb 片段 (C)。

[0105] 接下来，通过使用获得的三个片段 (A)、(B) 和 (C) 的混合物作为模板，并使用表 1 所示的 PVG-FW2 和 catr2 引物组进行 SOE-PCR，获得大小为 2.2kb 的 DNA 片段 (D)，其中三个片段 (A)、(B) 和 (C) 以该顺序连接，spoVG 基因的转录起始控制区和核糖体结合位点连接在 secY 基因的上游（连接成 secY 基因的起始密码子位于 spoVG 基因的起始密码子的位置处），并且 Cm 抗性基因结合在其下游。随后，使用从枯草杆菌 168 株提取的基因组数据作为模板，并使用表 1 所示的 amyEfw2 和 amyE/PVG2-R 以及 amyE/Cm2-F 和 amyErv2 引物组，通过 PCR 扩增含 amyE 基因的 5′ -端上的区的 1.0kb 片段 (E)，和含 amyE 基因的 3′ -端上的区的 1.0kb 片段 (F)。

[0106] 随后，通过使用获得的三个片段 (E)、(F) 和 (D) 的混合物作为模板，并使用表 1 所示的 amyEfw1 和 amyErv1 引物组进行 SOE-PCR，获得总碱基长度为 4.2kb 的 DNA 片段 (G)，其中三个片段 (E)、(D) 和 (F) 以该顺序连接，secY 基因连接在 spoVG 基因的转录起始控制区和核糖体结合位点的下游，并且在其下游连接有氯霉素抗性基因的大小为 2.2kb 的 DNA 片段插入至 amyE 基因的中央。

[0107] 使用获得的 4.2kb 的 DNA 片段 (G)，通过感受态细胞法将枯草杆菌 168 株转化，并分离出在含 (10μg/mL) 的 LB 琼脂培养基上生长的集落作为转化体。通过使用从获得的转化体提取的基因组 DNA 作为模板，并使用表 1 所示的 amyEfw2 和 secY/Cm-R 以及 secY/PVG-F 和 amyErv2 引物组，通过 PCR 确认大小分别为 2.5kb 和 3.1kb 的 DNA 片段的扩增，并且确认其中 secY 基因连接在 spoVG 基因的转录起始控制区和核糖体结合位点的下游的 DNA 片段在枯草杆菌 168 株的基因组上的 amyE 基因位点处被插入。由此获得的菌株称为 secY-K 株。

[0108] [表 1-1]

[0109]

引物名称	序列 (5′ -3′)	SEQ ID NO.
PVG-FW	GTTAGTCGAGATCGAAGTTA	10
PVG-R	AGTAGTTCACCACCTTTTCC	11
secY/PVG-F	GGAAAAGGTGGTGAACTACTATGTTGTTTAAAACAA TCTCCAA	12
secY/Cm-R	ATGGGTGCTTTAGTTGAAGACTAGTTTTTCATAAATC CAC	13
catf	CAACTAAAGCACCCATTAG	14
catr	CTTCAACTAACGGGGCAG	15
PVG-FW2	TAAGAAAAGTGATTCTGGGA	16

catr2	CTCATATTATAAAAGCCAGTC	17
amyEfw2	GGAGTGTCAAGAATGTTTGC	18
amyE/PVG2-R	TCCCAGAATCACTTTTCTTAATCATCGCTCATCCATGTCG	19
amyE/Cm2-F	GACTGGCTTTTATAATATGAGGTTTAGGCTGGGCGGTGATA	20
amyErv2	TCAATGGGGAAGAGAACC	21
amyErv1	TCAAAACCTCTTTACTGCCG	22
amyErv1	CACGTAATCAAAGCCAGGCT	23
spf	ATCGATTTTCGTTCGTG	24
spr	CATATGCAAGGGTTTATTG	25
sigF-FW	GAAGAAAGCCGGGTTTATCA	26
sigF/Sp-R	CACGAACGAAAATCGATCTGAGCGTTTTTGCCGTTTT	27
sigF/Sp-F	CAATAAACCCTTGCATATGTCTGCAGTGCAGGCTAGCTT	28
sigF-RV	CCCGACGAACAAACCTGCCA	29
sigF-FW2	CGAATGACCACTAGTTTTGT	30
sigF-RV2	TGAAGCGTCTCCCATCCCCC	31
sigE-FW	AGTCAGATGTGAAAATCTATT	32
sigE/Sp-R	CACGAACGAAAATCGATCTTCCTCTCCCTTCTAAATG	33
sigE/Sp-F	CAATAAACCCTTGCATATGAAAATTTTATGGTTAGAACCC	34
sigE-RV	CCTTACTTTTTCCAAAACGT	35

[0110] [表 1-2]

[0111]

引物名称	序列 (5′ -3′)	SEQ ID NO.

sigE-FW2	CTCACGGCATTTATTTTAAAA	36
sigE-RV2	GCTTTTCATTATTGATGAATAT	37
phrA-FW	AGAAGACCAAGATTTGCTGC	38
phrA/Sp-R	CACGAACGAAAATCGATATGAAATGTTTTCCCTTCT G	39
phrA/Sp-F	CAATAAACCCTTGCATATGGGTTCATGCAGGTGAAA C	40
phrA-RV	ACTGGCCCCGTGTGATGCGG	41
phrA-FW2	GAGTTTTCAGAATTGTTAGAA	42
phrA-RV2	GAAGAGACTGCAGCTTTTT	43
S237pKAPpp -F	ACTTTAAAAATATTTAGGAGGTAATATGAAGAAACC GTTGGGGAAA	44
KAPter-R(Bg lII)	GGGAGATCTTCAGCGATCTATTTCTCTTTTTC	45
S237ppp-F2(BamHI)	CCCGGATCCAACAGGCTTATATTTA	46
S237pKAPpp -R	TTTCCCCAACGGTTTCTTCATATTACCTCCTAAATAT TTTTAAAGT	47
237UB1	TTGCGGATCCAACAGGCTTATATTTAGAGGAAATTT C	48
237DB1	TTGCGGATCCAACAACTCTGTGTCCAGTTATGCAAG	49
rsiX-FW	ATTCCAGTTACTCGTAATATAGTTG	50
rsiX/Cm-R	CTAATGGGTGCTTTAGTTGACTTCATCATCCATTAGC TC	51
rsiX/Cm-F	CTGCCCCGTTAGTTGAAGCTGCTCCAAATCCGATTT CC	52
rsiX-RV	GTCCTGCATTTTTCGAAGTCTGG	53
rsiX-FW2	ACTCCGGGTCTGGCATACCG	54

rsiX-RV2	ACATCTGGAAGATAAAATTGT	55
yacP-FW	CAGGCTGAGATCCTATTTTT	56
yacP/Cm-R	CTAATGGGTGCTTTAGTTGGGGTCTTTATTCTCCCA CAG	57
yacP/Cm-F	CTGCCCCGTTAGTTGAAGGTTGACGCTTTTTTGCCC AA	58
yacP-RV	ACGCATGTAAAAGACCTCCA	59
yacP-FW2	GAGGCAGAAATGCCAAGTCA	60

[0112] [表 1-3]

[0113]

引物名称	序列(5′ -3′)	SEQ ID NO.
yacP-RV2	TTGCAAGTACTGCAGTATTT	61
yvdE-FW	CTTCCTCCATTAAAAAGCCG	62
yvdE/Cm-R	CTAATGGGTGCTTTAGTTGTTTCATCCCCTCCTTATC TG	63
yvdE/Cm-F	CTGCCCCGTTAGTTGAAGGCGCCTTATTCTGTTATC GG	64
yvdE-RV	CGGCATATCAGCTGTAAAAG	65
yvdE-FW2	TTTCATCCATTTTTCTGCATC	66
yvdE-RV2	CAGTCCTTATAGCGGGATTG	67
yurK-FW	CTTCAGCCGCTTTGCTTTTT	68
vurK/Cm-R	CTAATGGGTGCTTTAGTTGAGGGTAGCCTCCTTTTA ACC	69
vurK/Cm-F	CTGCCCCGTTAGTTGAAGCAGGCATAAAAAACGAG ACA	70
yurK-RV	GTCCTGCTGGCGGGGTTAAC	71
yurK-FW2	TGCTGCTGTTCTATGATGCC	72

yurK-RV2	TTGTCCGCGGGATTGCAAGC	73
yhdQ-FW	TCACAAATCCAAGCGTTCGA	74
yhdQ/Cm-R	CTAATGGGTGCTTTAGTTGCACGTTATAGTTATGAG AATA	75
yhdQ/Cm-F	CTGCCCCGTTAGTTGAAGAACCATTTTATCTAACAG GAG	76
yhdQ-RV	TGTGGACCCTCTCTTTTTGC	77
yhdQ-FW2	GTCCAATCCGATATACCCGA	78
yhdQ-RV2	AGGGTTGACGAATTGAGAAA	79
glcT-FW	AAGCCGGTGTCTCTGTTACA	80
glcT/Cm-R	CTAATGGGTGCTTTAGTTGTCAATACCTCATATCGTA CA	81
glcT/Cm-F	CTGCCCCGTTAGTTGAAGAATTTCATAAATTCAGTT TATCC	82
glcT-RV	CTTATAGCTGAAGAATTCATA	83
glcT-FW2	AAAAAGAGTGTTTGAGGCAA	84
glcT-RV2	GTTCAATCACCCCGAAGATA	85

[0114] 实施例 2 用药物抗性基因置换基因组中的 sigF 基因

[0115] 用药物抗性基因置换基因组中的 sigE 基因的方法将基于图 4 进行说明。

[0116] 使用从枯草杆菌 168 株提取的基因组 DNA 作为模板，使用表 1 所示的 sigF-FW 和 sigF/Sp-R 引物组，通过 PCR 扩增邻近基因组中 sigE 基因上游的 1.0kb 片段 (A)。另外，使用上述基因组 DNA 作为模板，使用 sigF/sp-F 和 sigF-RV 引物组，通过 PCR 扩增邻近基因组中 sigF 基因下游的 1.0kb 片段 (B)。

[0117] 此外，使用质粒 pDG1727(Gene,167,335(1995))DNA 作为模板，并使用表 1 所示的 spf 和 spr 引物组，通过 PCR 制备大小为 1.2kb 的壮观霉素 (Sp) 抗性基因区 (C)。

[0118] 随后，如图 4 所示，使用获得的 1.0kb 片段 (A)、1.0kb 片段 (B) 和 Sp 抗性基因区 (C) 三个片段的混合物作为模板，并使用表 1 所示的 sigF-FW2 和 sigF-RV2 引物组，根据 SOE-PCR 法获得其中以 1.0kb 片段 (A)、Sp 抗性基因区 (C) 和 1.0kb 片段 (B) 的顺序含有以上三个片段的 2.8kb 的 DNA 片段 (D)。

[0119] 然后，使用获得的 DNA 片段 (D)，根据感受态细胞转化法进行 168 株的转化。转化后，分离出在含壮观霉素 (100μg/mL) 的 LB 琼脂培养基上生长的集落作为转化株。

[0120] 提取获得的转化体的基因组 DNA，并通过 PCR 确认 sigF 基因被 Sp 抗性基因置换。这样，构建缺失 sigF 基因的菌株（Δ sigF 株）。此外，使用在实施例 1 中构建的 secY-K 株通过在转化中用此株更换枯草杆菌 168 株，构建其中在 secY-K 株的基因组中 sigF 基因被 Sp 抗性基因置换的菌株（secYK Δ sigF 株）。

[0121] 实施例 3 用药物抗性基因置换基因组中的 sigE 基因和 phrA 基因

[0122] 以在实施例 2 中所示的用药物抗性基因置换 sigE 基因相同的方式，进行用壮观霉素抗性基因置换 168 株基因组中的 sigE 基因和 phrA 基因，从而构建缺失 sigE 基因的菌株（Δ sigE 株）和缺失 phrA 基因的菌株（Δ phrA 株）。对于各菌株的构建，使用表 1 所示的引物，并且各引物与构建 Δ sigF 株中使用的引物的对应关系显示于表 2。

[0123]

[表 2]

	用于缺失 sigF 基因	用于缺失 sigE 基因	用于缺失 phrA 基因	用于缺失 rsiX 基因	用于缺失 yacP 基因	用于缺失 yvdE 基因	用于缺失 yurK 基因	用于缺失 yhdQ 基因	用于缺失 glcT 基因
片段（A）的扩增	sigF-FW	sigE-FW	phrA-FW	rsiX-FW	yacP-FW	yvdE-FW	yurK-FW	yhdQ-FW	glcT-FW
	sigF/Cm-R	sigE/Cm-R	phrA/Cm-R	rsiX/Cm-R	yacP/Cm-R	yvdE/Cm-R	yurK/Cm-R	yhdQ/Cm-R	glcT/Cm-R
片段（B）的扩增	sigF/Cm-F	sigE/Cm-F	phrA/Cm-F	rsiX/Cm-F	yacP/Cm-F	yvdE/Cm-F	yurK/Cm-F	yhdQ/Cm-F	glcT/Cm-F
	sigF-RV	sigE-RV	phrA-RV	rsiX-RV	yacP-RV	yvdE-RV	yurK-RV	yhdQ-RV	glcT-RV
片段（C）的扩增	spf	spf	spf	catf	catf	catf	catf	catf	catf
	spr	spr	spr	catr	catr	catr	catr	catr	catr
片段（D）的扩增	sigF-FW2	sigE-FW2	phrA-FW2	rsiX-FW2	yacP-FW2	yvdE-FW2	yurK-FW2	yhdQ-FW2	glcT-FW2
	sigF-RV2	sigE-RV2	phrA-RV2	rsiX-RV2	yacP-RV2	yvdE-RV2	yurK-RV2	yhdQ-RV2	glcT-RV2

[0124] 通过用壮观霉素抗性基因置换实施例 1 中构建的 secY-K 株的基因组上的 sigE 基因或 phrA 基因，构建 secYK Δ sigE 株和 secYK Δ phrA 株。

[0125] 实施例 4 评价碱性蛋白酶的分泌和产生

[0126] 实施例 1-3 中获得的 secY-K 株、secYKΔsigF 株、secYKΔsigE 株和 secYKΔphrA 株的异源蛋白质生产率的评价如下进行：使用来自芽孢杆菌属细菌的碱性蛋白酶的生产率作为指标，上述蛋白酶包括 SEQID NO：4 所示的氨基酸序列。作为对照，对枯草杆菌 168 株、ΔsigF 株、ΔsigE 株和 ΔphrA 株也进行评价。即，使用从克劳氏芽孢杆菌 KSM-K16 株 (FERM BP-3376) 中提取的基因组 DNA 作为模板，并使用表 1 所示的 S237pKAPpp-F 和 KAPter-R(BglH) 引物组，通过 PCR 扩增大小为 1.3kb 的编码具有 SEQ ID NO：3 所示的氨基酸序列的碱性蛋白酶的 DNA 片段 (W)(Appl.Microbiol.Biotechnol.，43，473(1995))。另外，使用从枯草杆菌 KSM-S237 株 (FERM BP-7875) 中提取的基因组 DNA 作为模板，并使用表 1 所示的 S237ppp-F2(BamHI) 和 S237pKAPpp-R 引物组，通过 PCR 扩增大小为 0.6kb 的含碱性纤维素酶基因的启动子区 (JP-A 第 2000-210081 号) 的 DNA 片段 (X)。

[0127] 随后，使用获得的两个片段 (W) 和 (X) 的混合物作为模板，并使用表 1 所示的 S237ppp-F2(BamHI) 和 KAPter-R(BglII) 等引物组进行 SOE-PCR，获得大小为 1.8kb 的 DNA 片段 (Y)，其中碱性蛋白酶基因连接在碱性纤维素酶基因的启动子区下游。将得到的 1.8kb 的 DNA 片段 (Y) 插入在穿梭载体 pHY300PLK(Yakult Honsha Co.，Ltd.) 的 BamHI-BglII 限制性酶切位点处，以构建用于评价碱性蛋白酶产量的质粒 pHYKAP(S237p)。

[0128] 将根据原生质体转化法构建的质粒 pHYKAP(S237p) 转染到各菌株中。获得的各重组株在 37℃下在 10mL LB 培养基中振荡培养过夜，将 0.05mL 该培养溶液接种至 50mL 2×L- 麦芽糖培养基 (2％胰蛋白质胨、1％酵母提取物、1％ NaCl、7.5％麦芽糖、7.5ppm 硫酸锰四水合物或五水合物和 15ppm 四环素) 中，在 30℃下振荡培养 3 天。培养后，测定通过离心去除细菌细胞的培养溶液的上清液的碱性蛋白酶活性，以确定培养期间分泌和产生至细菌细胞外的碱性蛋白酶的量。培养上清液中蛋白酶活性的测定如下进行。具体地，向 50μL 用 2mM $CaCl_2$ 溶液适当稀释的培养上清液中，加入含 7.5mM 琥珀酰 -L- 丙氨酰 -L- 丙氨酰 -L- 丙氨酸对硝基苯胺 (STANA，Peptide Institute，Inc.) 作为底物的 100μL 75mM 硼酸 -KCl 缓冲溶液 (pH 10.5) 并混合。当反应在 30℃下进行时，脱离的对硝基苯胺量的定量通过测量 420nm 处的吸光度 (OD420nm) 的变化来进行。在 1 分钟内使 1μ-mol 对硝基苯胺脱离的酶量作为 1U。

[0129] 如表 3 所示，作为碱性蛋白酶活性的测定结果，当使用 secY-K 株作为宿主时，碱性蛋白酶的生产率等于对照 168 株 (野生型) 的生产率，并且特别地，观察不到生产率的提高。另一方面，观察到 secYKΔsigF 株、secYKΔsigE 株和 secYKΔphrA 株的生产率显著提高，在这些菌株中结合了芽孢形成相关基因的缺失。对于没有进行 secY 基因表达增强的导入的 ΔsigF 株、ΔsigE 株和 ΔphrA 株也观察到生产率的提高超过了 168 株的提高；然而，与 secY 基因表达的增强相结合，由这些基因的缺失所引起的生产率提高作用的表现显然是增加的。换句话说，设想当首先增加 SecY 蛋白质 (其为分泌机构) 的量，然后使芽孢形成相关基因例如 sigF、sigE 和 phrA 缺失时，在异源蛋白质生产率的提高方面获得了协同作用。

[0130] [表 3]

[0131]

	蛋白酶活性（相对值，％）	缺失的作用（％）
168	100	-
ΔsigF	219	119
ΔsigE	253	153
ΔphrA	154	54
secY-K	100	-
SecYKΔsigF	233	133
SecYKΔsigE	278	178
SecYKΔphrA	209	109

[0132] 实施例 5 评价碱性纤维素酶的分泌和产生

[0133] 其他异源蛋白质生产率的评价如下进行：使用来自芽孢杆菌属细菌的包括SEQ ID NO：6 所示氨基酸序列的碱性纤维素酶的生产率作为指标。具体地，使用表 1 所示的 237UB1 和 237DB1 引物组，扩增来自芽孢杆菌属 KSM-S237 株 (FERM BP-7875) 的碱性纤维素酶基因 (JP-A 第 2000-210081 号）的片段 (3.1kb)，然后用 BamHI 限制性内切酶处理，以将该片段插入在穿梭载体 pHY300PLK 的 BamHI 限制性酶切位点处。根据原生质体转化法将由此获得的重组质粒 pHY-S237 转染到各菌株中。由此获得的各重组株在与实施例 4 相同的条件下振荡培养 3 天。测定离心去除细菌细胞的培养溶液的上清液的碱性纤维素酶活性，以确定培养期间分泌和产生至细菌细胞外的碱性纤维素酶的量。

[0134] 对于纤维素酶活性的测定，向 50μL 用 1/7.5M 磷酸盐缓冲液 (pH7.4，Wako Pure Chemical Industries，Ltd.）适当稀释的样品溶液中加入 50μL 0.4mM 对硝基苯基 -β-D- 纤维三糖苷 (Seikagaku Corporation) 并混合，并且通过测量 420nm 处吸光度 (OD420nm) 的变化，当反应在 30℃下进行时进行脱离的对硝基苯酚量的定量。在 1 分钟内使 1μ-mol 对硝基苯酚脱离的酶量作为 1U。

[0135] 如表 4 所示，作为碱性纤维素酶活性的测定结果，当使用 secY-K 株作为宿主时，与对照 168 株（野生型）相比，观察到较高的碱性纤维素酶的分泌和产生。另外，观察到 secYKΔsigF 株、secYKΔsigE 株和 secYKΔphrA 株的生产率进一步显著提高，在这些菌株中结合了芽孢形成相关基因的缺失。对于没有进行 secY 基因表达增强的导入的 ΔsigF 株、ΔsigE 株和 ΔphrA 株，也观察到生产率的提高超过了 168 株的提高；然而，与 secY 基因表达的增强相结合，由这些基因的缺失所引起的生产率提高作用的表现显然是增加的。换句话说，设想当首先增加 SecY 蛋白质（其为分泌机构）的量，然后使芽孢形成相关基因例如 sigF、sigE 和 phrA 缺失时，在生产率的提高方面获得了协同作用。

[0136] [表 4]

[0137]

	纤维素酶活性（相对值，％）	缺失的作用（％）
168	100	-
ΔsigF	142	42
ΔsigE	145	45
ΔphrA	130	30
secY-K	115	-
SecYKΔsigF	162	62
SecYKΔsigE	167	67
SecYKΔphrA	155	55

[0138] 比较例 1 用药物抗性基因置换 rsiX 基因

[0139] 以与实施例 2 所示的用壮观霉素抗性基因置换 sigF 基因相同的方式，用氯霉素抗性基因置换枯草杆菌 168 株的基因组上的 rsiX 基因，从而构建 ΔrsiX 株。使用的引物显示于表 1，并且使用的各引物与构建 ΔsigF 株使用的引物的对应关系显示于表 2。另外，以相同的方式用氯霉素抗性基因置换实施例 2 中构建的 ΔsigF 株的基因组上的 rsiX 基因，从而构建 ΔsigFΔrsiX 株。此外，rsiX 基因是编码抑制 SigX 功能的抗 sigma 因子（抗 -SigX）的基因，SigX 是属于枯草杆菌的 ECF（胞质外功能）家族的 sigma 因子之一。SigX 在细胞周围环境在热应激或类似情况下发生变化时受到活化，并且具有通过诱导具有识别该活化作用的启动子的基因的转录或操纵子的转录来应付环境变化的功能 (J.Bacteriol.，179，2915(1997))。

[0140] 比较例 2 评价碱性蛋白酶的分泌和产生

[0141] 比较例 1 中构建的 ΔsigFΔrsiX 株的碱性蛋白酶的分泌生产率的评价以与实施例 4 相同的方式进行。作为对照，对枯草杆菌 168 株、实施例 1 中构建的 ΔrsiX 株和 ΔsigF 株进行相同的评价。

[0142] 结果，如表 5 所示，尽管 rsiX 基因缺失的 ΔrsiX 株的蛋白酶生产率高于野生株的生产率，但其中结合有 sigF 基因的缺失的 ΔsigFΔrsiX 株的生产率则低于 ΔsigF 株的生产率。换句话说，确认对于 rsiX 基因的缺失，没有观察到表现出当与芽孢形成相关基因的缺失结合时的产生异源蛋白质的协同作用。

[0143] [表 5]

[0144]

	蛋白酶活性（相对值，％）
168	100

ΔsigF	188
ΔrsiX	123
ΔrsiXΔsigF	178

[0145] 比较例 3 用药物抗性基因置换 yacP 基因、yvdE 基因、yurK 基因、yhdQ 基因和 glcT 基因

[0146] 通过实施例 2 所示的用壮观霉素抗性基因置换 sigF 基因的相同方式，用氯霉素抗性基因置换枯草杆菌 168 株的基因组上的 yacP 基因、yvdE 基因、yurK 基因、yhdQ 基因和 glcT 基因，从而分别构建 ΔyacP 株、ΔyvdE 株、ΔyurK 株、ΔyhdQ 株和 ΔglcT 株。使用的引物显示于表 1，并且使用的各引物与构建 ΔsigF 株使用的引物的对应关系显示于表 2。另外，通过用氯霉素抗性基因置换实施例 2 中构建的 ΔsigF 株的基因组上的 yacP 基因、yvdE 基因、yurK 基因、yhdQ 基因和 glcT 基因，分别构建 ΔsigFΔyacP 株、ΔsigFΔyvdE 株、ΔsigFΔyurK 株、ΔsigFΔyhdQ 株和 ΔsigFΔglcT 株。

[0147] 关于与本发明相关的基因，基因编号及其功能在表 6 中给出。

[0148] [表 6]

[0149]

基因名称	基因编号	功能
secY	BG10445	前蛋白质移位酶 SecY 亚单位
spoVG	BG10112	V 期芽孢形成蛋白质 G（孢子皮质的合成）
amyE	BG10473	α－淀粉酶
sigA	BG10314	RNA 聚合酶主要 sigma 因子
kinA	BG10204	参与芽孢形成起始的双组分调控感受器组氨酸激酶 A
kinB	BG10745	参与芽孢形成起始的双组分调控感受器组氨酸激酶 B
kinC	BG10989	参与芽孢形成起始的双组分调控感受器组氨酸激酶 C
spo0F	BG10411	参与芽孢形成起始的双组分调控响应调节物
spo0B	BG10336	引发芽孢形成的磷酸转移酶
spo0A	BG10765	导致芽孢形成起始的双组分调控相应调节物
kapB	BG10746	芽孢形成起始中 KinB 的活化
kipI	BG11231	抑制 KinA 的功能

kipA	BG11214	kip 操纵子转录的控制因子
rapA	BG10652	通过磷酸化 Spo0F 的去磷酸化抑制磷酸中继系统的天冬氨酸磷酸酶 A
phrA	BG10653	天冬氨酸磷酸酶 A(RapA) 活性的抑制因子
soj	BG10055	参与抑制细胞周期和复制起始的 Soj 蛋白质
spo0J	BG10054	0 期芽孢形成蛋白质 J。抑制 Soj 的功能
sigH	BG10159	对数期和静止期早期的 RNA 聚合酶 sigma H 因子
sigF	BG10298	RNA 聚合酶芽孢形成 - 特异性 sigma F 因子
sigE	BG10235	RNA 聚合酶芽孢形成 - 特异性 sigma E 因子
spoIIAA	BG10296	抗 - 抗 -sigma 因子 (SpoIIAB 的拮抗剂)
spoIIAB	BG10297	抗 -sigma 因子 (抑制 sigma F)
spoIIE	BG10127	参与不对称隔膜形成的 PP2C 丝氨酸磷酸酶 (活化 sigma F)
spoIIR	BG10937	II 期芽孢形成蛋白质 R(活化 SpoIIGA)
spoIIGA	BG10234	通过加工 Prosigma E 来参与活性 sigma E 的生成的蛋白酶
sigG	BG10236	RNA 聚合酶芽孢形成 - 特异性 sigma G 因子
spoIIIC	BG10919	RNA 聚合酶芽孢形成 - 特异性 sigma K 因子 (C- 端侧)
spoIVCB	BG10459	RNA 聚合酶芽孢形成 - 特异性 sigma K 因子 (N- 端侧)
rsiX	BG10537	抗 -sigma X 蛋白质 (抑制 sigma X 因子)
yacP	BG10158	功能未知的基因
yvdE	BG12414	疑似转录因子 (LacI 家族)
yurK	BG13997	疑似转录因子 (GntR 家族)
yhdQ	BG13023	活化铜离子转运系统操纵子 (copZA) 的转录的因子
glcT	BG12593	表达 ptsGHI 操纵子必需的抗转录终止因子

[0150] 比较例 4 评价碱性纤维素酶的分泌和产生

[0151] 在比较例 3 中构建的 ΔyacP 株、ΔyvdE 株、ΔyurK 株、ΔyhdQ 株和 ΔglcT 株的碱性纤维素酶的分泌生产率的评价通过与实施例 5 相同的方式进行。作为对照，对枯草杆菌 168 株也进行评价。结果，如表 7 所示，观察到 ΔyacP 株、ΔyvdE 株、ΔyurK 株和 ΔglcT 株的生产率高于野生株的生产率，而 ΔyhdQ 株的生产率则稍有下降。而且，在比较例 3 中构建的 ΔsigFΔyacP 株、ΔsigFΔyvdE 株、ΔsigFΔyurK 株、ΔsigFΔyhdQ 株和 ΔsigFΔglcT 株的碱性纤维素酶的分泌生产率的评价通过与实施例 5 相同的方式进行。作为对照，对枯草杆菌 158 株和实施例 1 中构建的 ΔsigF 株也进行评价。结果，如表 8 所示，任何构建的菌株的纤维素酶生产率均低于 ΔsigF 株的生产率。换句话说，有力地表明，当与芽孢形成相关基因的缺失结合时表现出产生异源蛋白质的协同作用，是与增强 secY 基因的表达结合的特征。

[0152] [表 7]

[0153]

	纤维素酶活性（相对值，%）
168	100
ΔyacP	156
ΔyvdE	109
ΔyurK	118
ΔyhdQ	97
ΔglcT	110

[0154] [表 8]

[0155]

	纤维素酶活性（相对值，%）
168	100
ΔsigF	161
ΔsigFΔyacP	154
ΔsigFΔyvdE	147
ΔsigFΔyurk	158
ΔsigFΔyhdQ	154

ΔsigFΔglcT	151

序列表

<110> 花王株式会社

<120> 重组微生物

<130>

<150>JP2007-102940
<151>2007-04-10

<160>85

<170>PatentIn version 3.1

<210>1
<211>1296
<212>DNA
<213> 枯草杆菌 (Bacillus subtilis)

<220>
<221>CDS
<222>(1)..(1296)

<400>1

```
ttg ttt aaa aca atc tcc aac ttt atg cgt gtg agt gat atc agg aat     48
Met Phe Lys Thr Ile Ser Asn Phe Met Arg Val Ser Asp Ile Arg Asn
1               5                  10                  15
aaa atc ata ttc act tta ctc atg ctt atc gtc ttt cgc ata ggt gcg     96
Lys Ile Ile Phe Thr Leu Leu Met Leu Ile Val Phe Arg Ile Gly Ala
            20                  25                  30
ttt att cct gtg cct tac gtt aac gct gaa gcg tta cag gca cag tct    144
Phe Ile Pro Val Pro Tyr Val Asn Ala Glu Ala Leu Gln Ala Gln Ser
        35                  40                  45
caa atg ggt gtt ttt gat ctc ctt aat aca ttt ggc ggc ggt gcg ctt    192
Gln Met Gly Val Phe Asp Leu Leu Asn Thr Phe Gly Gly Gly Ala Leu
    50                  55                  60
tac caa ttt tcc att ttc gca atg gga att act cct tat atc acg gct    240
Tyr Gln Phe Ser Ile Phe Ala Met Gly Ile Thr Pro Tyr Ile Thr Ala
```

65 70 75 80

tcg atc atc att cag ctg ctt cag atg gat gtg gta ccg aag ttt acc 288
Ser Ile Ile Ile Gln Leu Leu Gln Met Asp Val Val Pro Lys Phe Thr
85 90 95

gag tgg tct aag caa ggt gaa gtt ggc cgc cgt aaa tta gct cag ttc 336
Glu Trp Ser Lys Gln Gly Glu Val Gly Arg Arg Lys Leu Ala Gln Phe
100 105 110

aca agg tac ttt acg att gtg ctt ggt ttc atc caa gcg tta ggt atg 384
Thr Arg Tyr Phe Thr Ile Val Leu Gly Phe Ile Gln Ala Leu Gly Met
115 120 125

tca tat gga ttc aac aat ctg gca aac ggt atg ctg atc gaa aaa tcc 432
Ser Tyr Gly Phe Asn Asn Leu Ala Asn Gly Met Leu Ile Glu Lys Ser
130 135 140

ggt gta tcg aca tat ctt atc att gct tta gtg ctc act ggc gga act 480
Gly Val Ser Thr Tyr Leu Ile Ile Ala Leu Val Leu Thr Gly Gly Thr
145 150 155 160

gcc ttt tta atg tgg ctt ggg gaa caa att act tct cat gga gta ggc 528
Ala Phe Leu Met Trp Leu Gly Glu Gln Ile Thr Ser His Gly Val Gly
165 170 175

aac gga ata tcg atc att atc ttc gcg ggg att gtg tct agt att cca 576
Asn Gly Ile Ser Ile Ile Ile Phe Ala Gly Ile Val Ser Ser Ile Pro
180 185 190

aaa aca att ggg caa ata tat gag act caa ttt gtc ggc agc aac gat 624
Lys Thr Ile Gly Gln Ile Tyr Glu Thr Gln Phe Val Gly Ser Asn Asp
195 200 205

cag ttg ttt att cat att gtg aaa gtc gca ctt ctt gtg att gcg att 672
Gln Leu Phe Ile His Ile Val Lys Val Ala Leu Leu Val Ile Ala Ile
210 215 220

tta gca gtt att gtt gga gtt att ttc att cag caa gcc gta cgg aaa 720
Leu Ala Val Ile Val Gly Val Ile Phe Ile Gln Gln Ala Val Arg Lys
225 230 235 240

att gcg att caa tat gct aaa ggc aca ggt cgt tca cct gct ggc gga 768
Ile Ala Ile Gln Tyr Ala Lys Gly Thr Gly Arg Ser Pro Ala Gly Gly
245 250 255

ggt cag tct aca cac ctt cca ttg aaa gtg aat cct gca ggg gtt att 816
Gly Gln Ser Thr His Leu Pro Leu Lys Val Asn Pro Ala Gly Val Ile
260 265 270

ccg gta atc ttt gcg gtt gcg ttt ttg ata acg ccg cgg acg atc gcg 864
Pro Val Ile Phe Ala Val Ala Phe Leu Ile Thr Pro Arg Thr Ile Ala

275 280 285

tca ttc ttt gga aca aac gat gtg aca aag tgg att caa aac aac ttt 912
Ser Phe Phe Gly Thr Asn Asp Val Thr Lys Trp Ile Gln Asn Asn Phe
290 295 300

gat aat acg cat ccg gtg ggt atg gcg ata tat gtt gcg ttg att att 960
Asp Asn Thr His Pro Val Gly Met Ala Ile Tyr Val Ala Leu Ile Ile
305 310 315 320

gcc ttt acg tac ttt tat gct ttt gta cag gta aac cct gaa caa atg 1008
Ala Phe Thr Tyr Phe Tyr Ala Phe Val Gln Val Asn Pro Glu Gln Met
325 330 335

gct gat aac ctt aaa aaa cag ggt ggc tat atc ccg ggg gtt cgt cca 1056
Ala Asp Asn Leu Lys Lys Gln Gly Gly Tyr Ile Pro Gly Val Arg Pro
340 345 350

ggg aaa atg act caa gat aga att acg agc att ttg tat cga ctt acg 1104
Gly Lys Met Thr Gln Asp Arg Ile Thr Ser Ile Leu Tyr Arg Leu Thr
355 360 365

ttt gtg ggt tct ata ttc tta gcc gtg att tcc att ctt cct atc ttt 1152
Phe Val Gly Ser Ile Phe Leu Ala Val Ile Ser Ile Leu Pro Ile Phe
370 375 380

ttc att caa ttc gct gga ttg cct caa agt gca caa att ggc gga aca 1200
Phe Ile Gln Phe Ala Gly Leu Pro Gln Ser Ala Gln Ile Gly Gly Thr
385 390 395 400

tct ttg tta att gtt gtc ggg gta gcc ttg gag aca atg aaa caa cta 1248
Ser Leu Leu Ile Val Val Gly Val Ala Leu Glu Thr Met Lys Gln Leu
405 410 415

gaa agc cag ttg gtg aaa cga aac tac cgt gga ttt atg aaa aac tag 1296
Glu Ser Gln Leu Val Lys Arg Asn Tyr Arg Gly Phe Met Lys Asn
420 425 430

<210>2
<211>431
<212>PRT
<213> 枯草杆菌

<400>2
Met Phe Lys Thr Ile Ser Asn Phe Met Arg Val Ser Asp Ile Arg Asn
1 5 10 15
Lys Ile Ile Phe Thr Leu Leu Met Leu Ile Val Phe Arg Ile Gly Ala
20 25 30

Phe Ile Pro Val Pro Tyr Val Asn Ala Glu Ala Leu Gln Ala Gln Ser
35 40 45
Gln Met Gly Val Phe Asp Leu Leu Asn Thr Phe Gly Gly Gly Ala Leu
50 55 60
Tyr Gln Phe Ser Ile Phe Ala Met Gly Ile Thr Pro Tyr Ile Thr Ala
65 70 75 80
Ser Ile Ile Ile Gln Leu Leu Gln Met Asp Val Val Pro Lys Phe Thr
85 90 95
Glu Trp Ser Lys Gln Gly Glu Val Gly Arg Arg Lys Leu Ala Gln Phe
100 105 110
Thr Arg Tyr Phe Thr Ile Val Leu Gly Phe Ile Gln Ala Leu Gly Met
115 120 125
Ser Tyr Gly Phe Asn Asn Leu Ala Asn Gly Met Leu Ile Glu Lys Ser
130 135 140
Gly Val Ser Thr Tyr Leu Ile Ile Ala Leu Val Leu Thr Gly Gly Thr
145 150 155 160
Ala Phe Leu Met Trp Leu Gly Glu Gln Ile Thr Ser His Gly Val Gly
165 170 175
Asn Gly Ile Ser Ile Ile Ile Phe Ala Gly Ile Val Ser Ser Ile Pro
180 185 190
Lys Thr Ile Gly Gln Ile Tyr Glu Thr Gln Phe Val Gly Ser Asn Asp
195 200 205
Gln Leu Phe Ile His Ile Val Lys Val Ala Leu Leu Val Ile Ala Ile
210 215 220
Leu Ala Val Ile Val Gly Val Ile Phe Ile Gln Gln Ala Val Arg Lys
225 230 235 240
Ile Ala Ile Gln Tyr Ala Lys Gly Thr Gly Arg Ser Pro Ala Gly Gly
245 250 255
Gly Gln Ser Thr His Leu Pro Leu Lys Val Asn Pro Ala Gly Val Ile
260 265 270
Pro Val Ile Phe Ala Val Ala Phe Leu Ile Thr Pro Arg Thr Ile Ala
275 280 285
Ser Phe Phe Gly Thr Asn Asp Val Thr Lys Trp Ile Gln Asn Asn Phe
290 295 300
Asp Asn Thr His Pro Val Gly Met Ala Ile Tyr Val Ala Leu Ile Ile
305 310 315 320
Ala Phe Thr Tyr Phe Tyr Ala Phe Val Gln Val Asn Pro Glu Gln Met
325 330 335
Ala Asp Asn Leu Lys Lys Gln Gly Gly Tyr Ile Pro Gly Val Arg Pro

340 345 350

Gly Lys Met Thr Gln Asp Arg Ile Thr Ser Ile Leu Tyr Arg Leu Thr
355 360 365

Phe Val Gly Ser Ile Phe Leu Ala Val Ile Ser Ile Leu Pro Ile Phe
370 375 380

Phe Ile Gln Phe Ala Gly Leu Pro Gln Ser Ala Gln Ile Gly Gly Thr
385 390 395 400

Ser Leu Leu Ile Val Val Gly Val Ala Leu Glu Thr Met Lys Gln Leu
405 410 415

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

<210>3

<211>1140

<212>DNA

<213> 芽孢杆菌属 KSM-K16

<220>

<221>CDS

<222>(1)..(1140)

<400>3

atg aag aaa ccg ttg ggg aaa att gtc gca agc acc gca cta ctc att 48
Met Lys Lys Pro Leu Gly Lys Ile Val Ala Ser Thr Ala Leu Leu Ile
1 5 10 15

tct gtt gct ttt agt tca tcg atc gca tcg gct gct gag gaa gca aaa 96
Ser Val Ala Phe Ser Ser Ser Ile Ala Ser Ala Ala Glu Glu Ala Lys
20 25 30

gaa aaa tat tta att ggc ttt aat gag cag gaa gca gtt agt gag ttt 144
Glu Lys Tyr Leu Ile Gly Phe Asn Glu Gln Glu Ala Val Ser Glu Phe
35 40 45

gta gag caa ata gag gca aat gac gat gtc gcg att ctc tct gag gaa 192
Val Glu Gln Ile Glu Ala Asn Asp Asp Val Ala Ile Leu Ser Glu Glu
50 55 60

gag gaa gtc gaa att gaa ttg ctt cat gag ttt gaa acg att cct gtt 240
Glu Glu Val Glu Ile Glu Leu Leu His Glu Phe Glu Thr Ile Pro Val
65 70 75 80

tta tct gtt gag tta agt cca gaa gat gtg gac gcg ctt gag ctc gat 288
Leu Ser Val Glu Leu Ser Pro Glu Asp Val Asp Ala Leu Glu Leu Asp

85 90 95

cca acg att tcg tat att gaa gag gat gca gaa gta acg aca atg gcg 336
Pro Thr Ile Ser Tyr Ile Glu Glu Asp Ala Glu Val Thr Thr Met Ala
100 105 110

caa tca gtg cca tgg gga att agc cgt gta caa gcc cca gct gcc cat 384
Gln Ser Val Pro Trp Gly Ile Ser Arg Val Gln Ala Pro Ala Ala His
115 120 125

aac cgt gga ttg aca ggt tct ggt gta aaa gtt gct gtc ctc gat acg 432
Asn Arg Gly Leu Thr Gly Ser Gly Val Lys Val Ala Val Leu Asp Thr
130 135 140

ggt att tcc acc cat cca gac tta aat att cgc ggt ggt gct agc ttt 480
Gly Ile Ser Thr His Pro Asp Leu Asn Ile Arg Gly Gly Ala Ser Phe
145 150 155 160

gtg cca gga gaa cca tcc act caa gat gga aat gga cat ggc acg cat 528
Val Pro Gly Glu Pro Ser Thr Gln Asp Gly Asn Gly His Gly Thr His
165 170 175

gtg gca ggg acg att gct gct tta aac aat tcg att ggc gtt ctg ggc 576
Val Ala Gly Thr Ile Ala Ala Leu Asn Asn Ser Ile Gly Val Leu Gly
180 185 190

gta gca ccg agc gcg gaa cta tac gct gta aaa gta tta ggc gcg agc 624
Val Ala Pro Ser Ala Glu Leu Tyr Ala Val Lys Val Leu Gly Ala Ser
195 200 205

ggt tca ggt tcg gtc agc tcg att gcc caa gga ttg gaa tgg gca ggg 672
Gly Ser Gly Ser Val Ser Ser Ile Ala Gln Gly Leu Glu Trp Ala Gly
210 215 220

aac aat ggc atg cac gtt gcg aat ttg agt tta gga agc ccg tcg ccg 720
Asn Asn Gly Met His Val Ala Asn Leu Ser Leu Gly Ser Pro Ser Pro
225 230 235 240

agt gca aca ctt gag caa gct gtt aat agc gct act tct aga ggc gtt 768
Ser Ala Thr Leu Glu Gln Ala Val Asn Ser Ala Thr Ser Arg Gly Val
245 250 255

ctt gtc gta gca gca tct ggt aat tca ggt gca ggc tca atc agc tat 816
Leu Val Val Ala Ala Ser Gly Asn Ser Gly Ala Gly Ser Ile Ser Tyr
260 265 270

ccg gcc cgt tat gcg aac gca atg gca gtc gga gcg act gac caa aac 864
Pro Ala Arg Tyr Ala Asn Ala Met Ala Val Gly Ala Thr Asp Gln Asn
275 280 285

aac aac cgc gct agc ttt tca cag tat gga gct ggg ctt gac att gtc 912
Asn Asn Arg Ala Ser Phe Ser Gln Tyr Gly Ala Gly Leu Asp Ile Val

290 295 300

gcg cca ggt gtc aat gtg cag agc aca tac cca ggt tca aca tat gcc 960
Ala Pro Gly Val Asn Val Gln Ser Thr Tyr Pro Gly Ser Thr Tyr Ala
305 310 315 320

agc tta aac ggt aca tcg atg gct act cct cat gtt gca ggt gta gca 1008
Ser Leu Asn Gly Thr Ser Met Ala Thr Pro His Val Ala Gly Val Ala
325 330 335

gcc ctt gtt aaa caa aag aat cca tct tgg tcc aat gta caa atc cgc 1056
Ala Leu Val Lys Gln Lys Asn Pro Ser Trp Ser Asn Val Gln Ile Arg
340 345 350

aat cat cta aag aat acg gca acg ggt tta gga aac acg aac ttg tat 1104
Asn His Leu Lys Asn Thr Ala Thr Gly Leu Gly Asn Thr Asn Leu Tyr
355 360 365

gga agc ggg ctt gtc aat gca gaa gcg gca aca cgc 1140
Gly Ser Gly Leu Val Asn Ala Glu Ala Ala Thr Arg
370 375 380

<210>4
<211>380
<212>PRT
<213> 芽孢杆菌属 KSM-K16

<400>4

Met Lys Lys Pro Leu Gly Lys Ile Val Ala Ser Thr Ala Leu Leu Ile
1 5 10 15
Ser Val Ala Phe Ser Ser Ser Ile Ala Ser Ala Ala Glu Glu Ala Lys
20 25 30
Glu Lys Tyr Leu Ile Gly Phe Asn Glu Gln Glu Ala Val Ser Glu Phe
35 40 45
Val Glu Gln Ile Glu Ala Asn Asp Asp Val Ala Ile Leu Ser Glu Glu
50 55 60
Glu Glu Val Glu Ile Glu Leu Leu His Glu Phe Glu Thr Ile Pro Val
65 70 75 80
Leu Ser Val Glu Leu Ser Pro Glu Asp Val Asp Ala Leu Glu Leu Asp
85 90 95
Pro Thr Ile Ser Tyr Ile Glu Glu Asp Ala Glu Val Thr Thr Met Ala
100 105 110
Gln Ser Val Pro Trp Gly Ile Ser Arg Val Gln Ala Pro Ala Ala His
115 120 125

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

<210>5
<211>3150
<212>DNA
<213> 芽孢杆菌属 KSM-S237
<220>
<221>CDS

<222>(573)..(3044)

<220>
<221>sig_peptide
<222>(573)..(659)

<220>
<221>mat_peptide
<222>(660)..(3044)

<400>5

```
gatttgccga tgcaacaggc ttatatttag aggaaatttc tttttaaatt gaatacggaa     60
taaaatcagg taaacaggtc ctgattttat ttttttgagt tttttagaga actgaagatt    120
gaaataaaag tagaagacaa aggacataag aaaattgcat tagttttaat tatagaaaac    180
gccttttttat aattatttat acctagaacg aaaatactgt ttcgaaagcg gtttactata    240
aaaccttata ttccggctct tttttaaaac agggggtaaa aattcactct agtattctaa    300
tttcaacatg ctataataaa tttgtaagac gcaatatgca tctcttttttt tacgatatat    360
gtaagcggtt aaccttgtgc tatatgccga tttaggaagg ggggtagatt gagtcaagta    420
gtaataatat agataactta taagttgttg agaagcagga gagcatctgg gttactcaca    480
agtttttttta aaactttaac gaaagcactt tcggtaatgc ttatgaattt agctatttga    540
ttcaattact ttaaaaatat 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 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 1121
Asp Phe Phe Arg Glu Ile Ala Ala Leu Tyr Pro Asn Asn Pro His Ile
140 145 150

att tat gag tta gcg aat gag ccg agt agt aat aat aat ggt gga gca 1169
Ile Tyr Glu Leu Ala Asn Glu Pro Ser Ser Asn Asn Asn Gly Gly Ala
155 160 165 170

ggg att ccg aat aac gaa gaa ggt tgg aaa gcg gta aaa gaa tat gct 1217
Gly Ile Pro Asn Asn Glu Glu Gly Trp Lys Ala Val Lys Glu Tyr Ala
175 180 185

gat cca att gta gaa atg tta cgt aaa agc ggt aat gca gat gac aac 1265
Asp Pro Ile Val Glu Met Leu Arg Lys Ser Gly Asn Ala Asp Asp Asn
190 195 200

att atc att gtt ggt agt cca aac tgg agt cag cgt ccg gac tta gca 1313
Ile Ile Ile Val Gly Ser Pro Asn Trp Ser Gln Arg Pro Asp Leu Ala
205 210 215

gct gat aat cca att gat gat cac cat aca atg tat act gtt cac ttc 1361
Ala Asp Asn Pro Ile Asp Asp His His Thr Met Tyr Thr Val His Phe
220 225 230

tac act ggt tca cat gct gct tca act gaa agc tat ccg tct gaa act 1409
Tyr Thr Gly Ser His Ala Ala Ser Thr Glu Ser Tyr Pro Ser Glu Thr
235 240 245 250

cct aac tct gaa aga gga aac gta atg agt aac act cgt tat gcg tta 1457
Pro Asn Ser Glu Arg Gly Asn Val Met Ser Asn Thr Arg Tyr Ala Leu
255 260 265

gaa aac gga gta gcg gta ttt gca aca gag tgg gga acg agt caa gct 1505

Glu Asn Gly Val Ala Val Phe Ala Thr Glu Trp Gly Thr Ser Gln Ala
270 275 280

agt gga gac ggt ggt cct tac ttt gat gaa gca gat gta tgg att gaa 1553
Ser Gly Asp Gly Gly Pro Tyr Phe Asp Glu Ala Asp Val Trp Ile Glu
285 290 295

ttt tta aat gaa aac aac att agc tgg gct aac tgg tct tta acg aat 1601
Phe Leu Asn Glu Asn Asn Ile Ser Trp Ala Asn Trp Ser Leu Thr Asn
300 305 310

aaa aat gaa gta tct ggt gca ttt aca cca ttc gag tta ggt aag tct 1649
Lys Asn Glu Val Ser Gly Ala Phe Thr Pro Phe Glu Leu Gly Lys Ser
315 320 325 330

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
635 640 645 650

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 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 aactagttat agggttatct aaaggtctga tgtagatctt 3094
Lys
795

ttagataacc tttttcttgc ataactggac acagagttgt tattaaagaa agtaag 3150

<210>6

<211>824

<212>PRT

<213> 芽孢杆菌属 KSM-S237

<400>6

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>7
<211>3332
<212>DNA
<213> 芽孢杆菌属 KSM-64

<220>
<221>CDS
<222>(610)..(3075)

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

<220>
<221>mat_peptide
<222>(697)..(3075)

<400>7
agtacttacc attttagagt caaaagatag aagccaagca ggatttgccg atgcaaccgg 60
cttatattta gagggaattt ctttttaaat tgaatacgga ataaaatcag gtaaacaggt 120
cctgatttta ttttttgaa ttttttgag aactaaagat tgaaatagaa gtagaagaca 180

```
acggacataa gaaaattgta ttagttttaa ttatagaaaa cgcttttcta taattattta   240
tacctagaac gaaaatactg tttcgaaagc ggtttactat aaaaccttat attccggctc   300
ttttttaaa caggggtga aaattcactc tagtattcta atttcaacat gctataataa   360
atttgtaaga cgcaatatac atctttttt tatgatattt gtaagcggtt aaccttgtgc   420
tatatgccga tttaggaagg gggtagattg agtcaagtag tcataattta gataacttat   480
aagttgttga gaagcaggag agaatctggg ttactcacaa gtttttaaa 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 1419
His His Thr Met Tyr Thr Val His Phe Tyr Thr Gly Ser His Ala Ala
230 235 240

tca act gaa agc tat ccg cct gaa act cct aac tct gaa aga gga aac 1467
Ser Thr Glu Ser Tyr Pro Pro Glu Thr Pro Asn Ser Glu Arg Gly Asn
245 250 255

gta atg agt aac act cgt tat gcg tta gaa aac gga gta gca gta ttt 1515
Val Met Ser Asn Thr Arg Tyr Ala Leu Glu Asn Gly Val Ala Val Phe
260 265 270

gca aca gag tgg gga act agc caa gca aat gga gat ggt ggt cct tac 1563
Ala Thr Glu Trp Gly Thr Ser Gln Ala Asn Gly Asp Gly Gly Pro Tyr
275 280 285

ttt gat gaa gca gat gta tgg att gag ttt tta aat gaa aac aac att 1611
Phe Asp Glu Ala Asp Val Trp Ile Glu Phe Leu Asn Glu Asn Asn Ile
290 295 300 305

agc tgg gct aac tgg tct tta acg aat aaa aat gaa gta tct ggt gca 1659
Ser Trp Ala Asn Trp Ser Leu Thr Asn Lys Asn Glu Val Ser Gly Ala
310 315 320

ttt aca cca ttc gag tta ggt aag tct aac gca aca agt ctt gac cca 1707
Phe Thr Pro Phe Glu Leu Gly Lys Ser Asn Ala Thr Ser Leu Asp Pro
325 330 335

ggg cca gac caa gta tgg gta cca gaa gag tta agt ctt tct gga gaa 1755
Gly Pro Asp Gln Val Trp Val Pro Glu Glu Leu Ser Leu Ser Gly Glu

340 345 350

tat gta cgt gct cgt att aaa ggt gtg aac tat gag cca atc gac cgt 1803
Tyr Val Arg Ala Arg Ile Lys Gly Val Asn Tyr Glu Pro Ile Asp Arg
355 360 365

aca aaa tac acg aaa gta ctt tgg gac ttt aat gat gga acg aag caa 1851
Thr Lys Tyr Thr Lys Val Leu Trp Asp Phe Asn Asp Gly Thr Lys Gln
370 375 380 385

gga ttt gga gtg aat gga gat tct cca gtt gaa gat gta gtt att gag 1899
Gly Phe Gly Val Asn Gly Asp Ser Pro Val Glu Asp Val Val Ile Glu
390 395 400

aat gaa gcg ggc gct tta aaa ctt tca gga tta gat gca agt aat gat 1947
Asn Glu Ala Gly Ala Leu Lys Leu Ser Gly Leu Asp Ala Ser Asn Asp
405 410 415

gtt tct gaa ggt aat tac tgg gct aat gct cgt ctt tct gcc gac ggt 1995
Val Ser Glu Gly Asn Tyr Trp Ala Asn Ala Arg Leu Ser Ala Asp Gly
420 425 430

tgg gga aaa agt gtt gat att tta ggt gct gaa aaa ctt act atg gat 2043
Trp Gly Lys Ser Val Asp Ile Leu Gly Ala Glu Lys Leu Thr Met Asp
435 440 445

gtg att gtt gat gag ccg acc acg gta tca att gct gca att cca caa 2091
Val Ile Val Asp Glu Pro Thr Thr Val Ser Ile Ala Ala Ile Pro Gln
450 455 460 465

ggg cca tca gcc aat tgg gtt aat cca aat cgt gca att aag gtt gag 2139
Gly Pro Ser Ala Asn Trp Val Asn Pro Asn Arg Ala Ile Lys Val Glu
470 475 480

cca act aat ttc gta ccg tta gga gat aag ttt aaa gcg gaa tta act 2187
Pro Thr Asn Phe Val Pro Leu Gly Asp Lys Phe Lys Ala Glu Leu Thr
485 490 495

ata act tca gct gac tct cca tcg tta gaa gct att gcg atg cat gct 2235
Ile Thr Ser Ala Asp Ser Pro Ser Leu Glu Ala Ile Ala Met His Ala
500 505 510

gaa aat aac aac atc aac aac atc att ctt ttt gta gga act gaa ggt 2283
Glu Asn Asn Asn Ile Asn Asn Ile Ile Leu Phe Val Gly Thr Glu Gly
515 520 525

gct gat gtt atc tat tta gat aac att aaa gta att gga aca gaa gtt 2331
Ala Asp Val Ile Tyr Leu Asp Asn Ile Lys Val Ile Gly Thr Glu Val
530 535 540 545

gaa att cca gtt gtt cat gat cca aaa gga gaa gct gtt ctt cct tct 2379
Glu Ile Pro Val Val His Asp Pro Lys Gly Glu Ala Val Leu Pro Ser

550 555 560

gtt ttt gaa gac ggt aca cgt caa ggt tgg gac tgg gct gga gag tct 2427
Val Phe Glu Asp Gly Thr Arg Gln Gly Trp Asp Trp Ala Gly Glu Ser
565 570 575

ggt gtg aaa aca gct tta aca att gaa gaa gca aac ggt tct aac gcg 2475
Gly Val Lys Thr Ala Leu Thr Ile Glu Glu Ala Asn Gly Ser Asn Ala
580 585 590

tta tca tgg gaa ttt gga tac cca gaa gta aaa cct agt gat aac tgg 2523
Leu Ser Trp Glu Phe Gly Tyr Pro Glu Val Lys Pro Ser Asp Asn Trp
595 600 605

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 ttttttttgca taactggaca tagaatggtt 3165
attaaagaaa gcaaggtgtt tatacgatat taaaaaggta gcgattttaa attgaaacct 3225
ttaataatgt cttgtgatag aatgatgaag taatttaaga gggggaaacg aagtgaaaac 3285
ggaaatttct agtagaagaa aaacagacca agaaatactg caagctt 3332

<210>8
<211>822
<212>PRT
<213> 芽孢杆菌属 KSM-64

<400>8

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>9

<211>230

<212>DNA

<213> 枯草杆菌

<400>9

gttagtcgag atcgaagtta ttgcactggt gaaataataa gaaaagtgat tctgggagag 60

ccgggatcac ttttttattt accttatgcc cgaaatgaaa gctttatgac ctaattgtgt 120

aactatatcc tatttttca aaaaatattt taaaaacgag caggatttca gaaaaaatcg 180

tggaattgat acactaatgc ttttatatag ggaaaaggtg gtgaactact 230

<210>10

<211>20

<212>DNA

<213> 人工序列

<220>

<223> 作为由枯草杆菌基因 spoVG 的 5'- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>10

gttagtcgag atcgaagtta 20

<210>11

<211>20

<212>DNA

<213> 人工序列

<220>

<223> 作为由枯草杆菌基因 spoVG 的 5'- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>11

agtagttcac cacctttcc 20

<210>12
<211>43
<212>DNA
<213> 人工序列

<220>
<223> 作为 PCR 引物的寡核苷酸，3’- 部分由枯草杆菌基因 secY 的核苷酸序列设计，且 5’- 部分由枯
草杆菌基因 spoVG 的 5’- 侧翼区的核苷酸序列设计
<400>12
ggaaaaggtg gtgaactact atgttgttta aaacaatctc caa 43

<210>13
<211>40
<212>DNA
<213> 人工序列

<220>
<223> 作为 PCR 引物的寡核苷酸，3’- 部分由枯草杆菌基因 secY 的核苷酸序列设计，且 5’- 部分由氯
霉素抗性基因的核苷酸序列设计

<400>13
atgggtgctt tagttgaaga ctagtttttc ataaatccac 40

<210>14
<211>19
<212>DNA
<213> 人工序列

<220>
<223> 作为扩增氯霉素抗性基因用的正向 PCR 引物的寡核苷酸

<400>14
caactaaagc acccattag 19

<210>15
<211>18

<212>DNA
<213> 人工序列

<220>
<223> 作为扩增氯霉素抗性基因用的正向 PCR 引物的寡核苷酸

<400>15
cttcaactaa cggggcag 18

<210>16
<211>20
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 spoVG 的 5'- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸
<400>16
taagaaaagt gattctggga 20

<210>17
<211>21
<212>DNA
<213> 人工序列

<220>
<223> 作为扩增氯霉素抗性基因用的反向 PCR 引物的寡核苷酸
<400>17
ctcatattat aaaagccagt c 21

<210>18
<211>20
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 amyE 的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>18

ggagtgtcaa gaatgtttgc 20

<210>19
<211>40
<212>DNA
<213> 人工序列

<220>
<223> 作为 PCR 引物的寡核苷酸，3’- 部分由枯草杆菌基因 amyE 的核苷酸序列设计，且 5’- 部分由枯
草杆菌基因 spoVG 的 5’- 侧翼区的核苷酸序列设计

<400>19
tcccagaatc actttctta atcatcgctc atccatgtcg 40

<210>20
<211>41
<212>DNA
<213> 人工序列

<220>
<223> 作为 PCR 引物的寡核苷酸，3’- 部分由枯草杆菌基因 amyE 的核苷酸序列设计，且 5’- 部分由氯
霉素抗性基因的核苷酸序列设计

<400>20
gactggcttt tataatatga ggtttaggct gggcggtgat a 41
<210>21
<211>18
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 amyE 的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>21
tcaatgggga agagaacc 18

<210>22

<211>20
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 amyE 的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>22
tcaaaacctc tttactgccg 20

<210>23
<211>20
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 amyE 的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>23
cacgtaatca aagccaggct 20

<210>24
<211>17
<212>DNA
<213> 人工序列

<220>
<223> 作为用于从 pDG1727 扩增壮观霉素抗性基因的正向 PCR 引物的寡核苷酸

<400>24
atcgattttc gttcgtg 17

<210>25
<211>19
<212>DNA
<213> 人工序列

<220>
<223> 作为用于从 pDG1727 扩增壮观霉素抗性基因的反向 PCR 引物的寡核苷酸

<400>25
catatgcaag ggtttattg 19

<210>26
<211>20
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 sigF 的 5’- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>26
gaagaaagcc gggtttatca 20

<210>27
<211>37
<212>DNA
<213> 人工序列

<220>
<223> 作为 PCR 引物的寡核苷酸,3’- 部分由枯草杆菌基因 sigF 的 5’- 侧翼区的核苷酸序列设计,且 5’-
部分由质粒 pDG1727 的核苷酸序列设计

<400>27
cacgaacgaa aatcgatctg agcgtttttg ccgtttt 37

<210>28
<211>39
<212>DNA
<213> 人工序列

<220>
<223> 作为 PCR 引物的寡核苷酸,3’- 部分由枯草杆菌基因 sigF 的 3’- 侧翼区的核苷酸序列设计,且
5’- 部分由质粒 pDG1727 的核苷酸序列设计

<400>28
caataaaccc ttgcatatgt ctgcagtgca ggctagctt 39

<210>29
<211>20
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 sigF 的 3’- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>29
cccgacgaac aaacctgcca 20

<210>30
<211>20
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 sigF 的 5’- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>30
cgaatgacca ctagttttgt 20

<210>31
<211>20
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 sigF 的 3’- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>31
tgaagcgtct cccatccccc 20

<210>32
<211>21
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 sigE 的 5’- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>32
agtcagatgt gaaaatctat t 21

<210>33
<211>37
<212>DNA
<213> 人工序列
<220>
<223> 作为 PCR 引物的寡核苷酸,3’- 部分由枯草杆菌基因 sigE 的 5’- 侧翼区的核苷酸序列设计,且
5’- 部分由质粒 pDG1727 的核苷酸序列设计

<400>33
cacgaacgaa aatcgatctt cctctccctt ctaaatg 37

<210>34
<211>40
<212>DNA
<213> 人工序列

<220>
<223> 作为 PCR 引物的寡核苷酸,3’- 部分由枯草杆菌基因 sigE 的 3’- 侧翼区的核苷酸序列设计,且
5’- 部分由质粒 pDG1727 的核苷酸序列设计

<400>34
caataaaccc ttgcatatga aaattttatg gttagaaccc 40

<210>35
<211>20

<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 sigE 的 3’- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>35
ccttactttt tccaaaacgt 20

<210>36
<211>21
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 sigE 的 5’- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>36
ctcacggcat ttattttaaa a 21

<210>37
<211>22
<212>DNA
<213> 人工序列
<220>
<223> 作为由枯草杆菌基因 sigE 的 3’- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>37
gcttttcatt attgatgaat at 22

<210>38
<211>20
<212>DNA
<213> 人工序列

<220>

<223> 作为由枯草杆菌基因 phrA 的 5’- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>38
agaagaccaa gatttgctgc 20

<210>39
<211>37
<212>DNA
<213> 人工序列

<220>
<223> 作为 PCR 引物的寡核苷酸,3’- 部分由枯草杆菌基因 phrA 的 5’- 侧翼区的核苷酸序列设计,且 5’-
部分由质粒 pDG1727 的核苷酸序列设计

<400>39
cacgaacgaa aatcgatatg aaatgttttc ccttctg 37

<210>40
<211>37
<212>DNA
<213> 人工序列

<220>
<223> 作为 PCR 引物的寡核苷酸,3’- 部分由枯草杆菌基因 phrA 的 3’- 侧翼区的核苷酸序列设计,且
5’- 部分由质粒 pDG1727 的核苷酸序列设计

<400>40
caataaaccc ttgcatatgg gttcatgcag gtgaaac 37

<210>41
<211>20
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 phrA 的 3’- 侧翼区的核苷酸序列设计的 PCR 引物的寡核

苷酸

<400>41
actggccccg tgtgatgcgg 20

<210>42
<211>21
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 phrA 的 5’- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>42
gagttttcag aattgttaga a 21

<210>43
<211>19
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 phrA 的 3’- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>43
gaagagactg cagcttttt 19

<210>44
<211>46
<212>DNA
<213> 人工序列

<220>
<223> 作为由克劳氏芽孢杆菌 (Bacillus clausii)KSM-K16 碱性蛋白酶基因的核苷酸序列和芽孢
杆菌属 KSM-S237 碱性纤维素酶基因的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>44

actttaaaaa tatttaggag gtaatatgaa gaaaccgttg gggaaa 46

<210>45

<211>32

<212>DNA

<213> 人工序列

<220>

<223> 作为由克劳氏芽孢杆菌 KSM-K16 碱性蛋白酶基因下游区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>45

gggagatctt cagcgatcta tttctctttt tc 32

<210>46

<211>25

<212>DNA

<213> 人工序列

<220>

<223> 作为由芽孢杆菌属 KSM-S237 碱性纤维素酶基因上游区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>46

cccggatcca acaggcttat attta 25

<210>47

<211>46

<212>DNA

<213> 人工序列

<220>

<223> 作为由芽孢杆菌属 KSM-S237 碱性纤维素酶基因的核苷酸序列和克劳氏芽孢杆菌 KSM-K16 碱性蛋白酶基因的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>47
tttccccaac ggtttcttca tattacctcc taaatatttt taaagt 46

<210>48
<211>37
<212>DNA
<213> 人工序列

<220>
<223> 作为由芽孢杆菌属 KSM-S237 中碱性纤维素酶基因的上游区的核苷酸序列设计的 PCR 引物的寡
核苷酸,其中在 5’- 端插入有 BamHI 限制性内切酶位点

<400>48
ttgcggatcc aacaggctta tatttagagg aaatttc 37

<210>49
<211>36
<212>DNA
<213> 人工序列

<220>
<223> 作为由芽孢杆菌属 KSM-S237 中碱性纤维素酶基因的下游区的核苷酸序列设计的 PCR 引物的寡
核苷酸,其中在 5’- 端插入有 BamHI 限制性内切酶位点

<400>49
ttgcggatcc aacaactctg tgtccagtta tgcaag 36

<210>50
<211>25
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 rsiX 的 5’- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>50

attccagtta ctcgtaatat agttg 25

<210>51
<211>39
<212>DNA
<213> 人工序列

<220>
<223> 作为 PCR 引物的寡核苷酸,3'- 部分由枯草杆菌基因 rsiX 的 5'- 侧翼区的核苷酸序列设计,且
5'- 部分由质粒 pC194 的核苷酸序列设计

<400>51
ctaatgggtg ctttagttga cttcatcatc cattagctc 39

<210>52
<211>38
<212>DNA
<213> 人工序列

<220>
<223> 作为 PCR 引物的寡核苷酸,3'- 部分由枯草杆菌基因 rsiX 的 3'- 侧翼区的核苷酸序列设计,且
5'- 部分由质粒 pC194 的核苷酸序列设计

<400>52
ctgccccgtt agttgaagct gctccaaatc cgatttcc 38
<210>53
<211>23
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 rsiX 的 3'- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>53
gtcctgcatt tttcgaagtc tgg 23

<210>54
<211>20
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 rsiX 的 5’- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>54
actccgggtc tggcataccg 20

<210>55
<211>21
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 rsiX 的 3’- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>55
acatctggaa gataaaattg t 21

<210>56
<211>20
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 yacP 的 5’- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>56
caggctgaga tcctatttttt 20
<210>57
<211>39
<212>DNA
<213> 人工序列

<220>

<223> 作为 PCR 引物的寡核苷酸，3’- 部分由枯草杆菌基因 yacP 的 5’- 侧翼区的核苷酸序列设计，且

5’- 部分由质粒 pC194 的核苷酸序列设计

<400>57

ctaatgggtg ctttagttgg ggtctttatt ctcccacag 39

<210>58

<211>38

<212>DNA

<213> 人工序列

<220>

<223> 作为 PCR 引物的寡核苷酸，3’- 部分由枯草杆菌基因 yacP 的 3’- 侧翼区的核苷酸序列设计，且

5’- 部分由质粒 pC194 的核苷酸序列设计

<400>58

ctgccccgtt agttgaaggt tgacgctttt ttgcccaa 38

<210>59

<211>20

<212>DNA

<213> 人工序列

<220>

<223> 作为由枯草杆菌基因 yacP 的 3’- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>59

acgcatgtaa aagacctcca 20

<210>60

<211>20

<212>DNA

<213> 人工序列

<220>

<223> 作为由枯草杆菌基因 yacP 的 5’-侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>60

gaggcagaaa tgccaagtca 20

<210>61

<211>20

<212>DNA

<213> 人工序列

<220>

<223> 作为由枯草杆菌基因 yacP 的 3’-侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>61

ttgcaagtac tgcagtattt 20

<210>62

<211>20

<212>DNA

<213> 人工序列

<220>

<223> 作为由枯草杆菌基因 yvdE 的 5’-侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>62

cttcctccat taaaaagccg 20

<210>63

<211>39

<212>DNA

<213> 人工序列

<220>

<223> 作为 PCR 引物的寡核苷酸，3’-部分由枯草杆菌基因 yvdE 的 5’-侧翼区的核苷酸序列设计，且

5’-部分由质粒 pC194 的核苷酸序列设计

<400>63
ctaatgggtg ctttagttgt ttcatcccct ccttatctg 39

<210>64
<211>38
<212>DNA
<213> 人工序列
<220>
<223> 作为 PCR 引物的寡核苷酸,3'- 部分由枯草杆菌基因 yvdE 的 3'- 侧翼区的核苷酸序列设计,且
5'- 部分由质粒 pC194 的核苷酸序列设计

<400>64
ctgccccgtt agttgaaggc gccttattct gttatcgg 38

<210>65
<211>20
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 yvdE 的 3'- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>65
cggcatatca gctgtaaaag 20

<210>66
<211>21
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 yvdE 的 5'- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>66
tttcatccat ttttctgcat c 21

<210>67
<211>20
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 yvdE 的 3’- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>67
cagtccttat agcgggattg 20

<210>68
<211>20
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 yurK 的 5’- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>68
cttcagccgc tttgcttttt 20

<210>69
<211>39
<212>DNA
<213> 人工序列

<220>
<223> 作为 PCR 引物的寡核苷酸，3’- 部分由枯草杆菌基因 yurK 的 5’- 侧翼区的核苷酸序列设计，且
5’- 部分由质粒 pC194 的核苷酸序列设计

<400>69
ctaatgggtg ctttagttga gggtagcctc cttttaacc 39

<210>70

<211>38
<212>DNA
<213> 人工序列

<220>
<223> 作为 PCR 引物的寡核苷酸,3'- 部分由枯草杆菌基因 yurK 的 3'- 侧翼区的核苷酸序列设计,且
5'- 部分由质粒 pC194 的核苷酸序列设计

<400>70
ctgccccgtt agttgaagca ggcataaaaa acgagaca 38

<210>71
<211>20
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 yurK 的 3'- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>71
gtcctgctgg cggggttaac 20
<210>72
<211>20
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 yurK 的 5'- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>72
tgctgctgtt ctatgatgcc 20

<210>73
<211>20
<212>DNA
<213> 人工序列

<220>

<223> 作为由枯草杆菌基因 yurK 的 3’- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>73

ttgtccgcgg gattgcaagc 20

<210>74

<211>20

<212>DNA

<213> 人工序列

<220>

<223> 作为由枯草杆菌基因 yhdQ 的 5’- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>74

tcacaaatcc aagcgttcga 20

<210>75

<211>40

<212>DNA

<213> 人工序列

<220>

<223> 作为 PCR 引物的寡核苷酸，3’- 部分由枯草杆菌基因 yhdQ 的 5’- 侧翼区的核苷酸序列设计，且

5’- 部分由质粒 pC194 的核苷酸序列设计

<400>75

ctaatgggtg ctttagttgc acgttatagt tatgagaata 40

<210>76

<211>39

<212>DNA

<213> 人工序列

<220>

<223> 作为 PCR 引物的寡核苷酸，3’- 部分由枯草杆菌基因 yhdQ 的 3’- 侧翼区的核苷

酸序列设计，且

5’－部分由质粒 pC194 的核苷酸序列设计

<400>76
ctgccccgtt agttgaagaa ccattttatc taacaggag 39

<210>77
<211>20
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 yhdQ 的 3’－侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>77
tgtggaccct ctcttttgc 20

<210>78
<211>20
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 yhdQ 的 5’－侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>78
gtccaatccg atatacccga 20

<210>79
<211>20
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 yhdQ 的 3’－侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸
<400>79

agggttgacg aattgagaaa 20

<210>80
<211>20
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 glcT 的 5’- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>80
aagccggtgt ctctgttaca 20

<210>81
<211>39
<212>DNA
<213> 人工序列

<220>
<223> 作为 PCR 引物的寡核苷酸,3’- 部分由枯草杆菌基因 glcT 的 5’- 侧翼区的核苷酸序列设计,且
5’- 部分由质粒 pC194 的核苷酸序列设计

<400>81
ctaatgggtg ctttagttgt caatacctca tatcgtaca 39

<210>82
<211>41
<212>DNA
<213> 人工序列

<220>
<223> 作为 PCR 引物的寡核苷酸,3’- 部分由枯草杆菌基因 glcT 的 3’- 侧翼区的核苷酸序列设计,且 5’-
部分由质粒 pC194 的核苷酸序列设计

<400>82
ctgccccgtt agttgaagaa tttcataaat tcagtttatc c 41

<210>83
<211>21
<212>DNA
<213> 人工序列
<220>
<223> 作为由枯草杆菌基因 glcT 的 3’- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>83
cttatagctg aagaattcat a 21

<210>84
<211>20
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 glcT 的 5’- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>84
aaaaagagtg tttgaggcaa 20

<210>85
<211>20
<212>DNA
<213> 人工序列

<220>
<223> 作为由枯草杆菌基因 glcT 的 3’- 侧翼区的核苷酸序列设计的 PCR 引物的寡核苷酸

<400>85
gttcaatcac cccgaagata 20

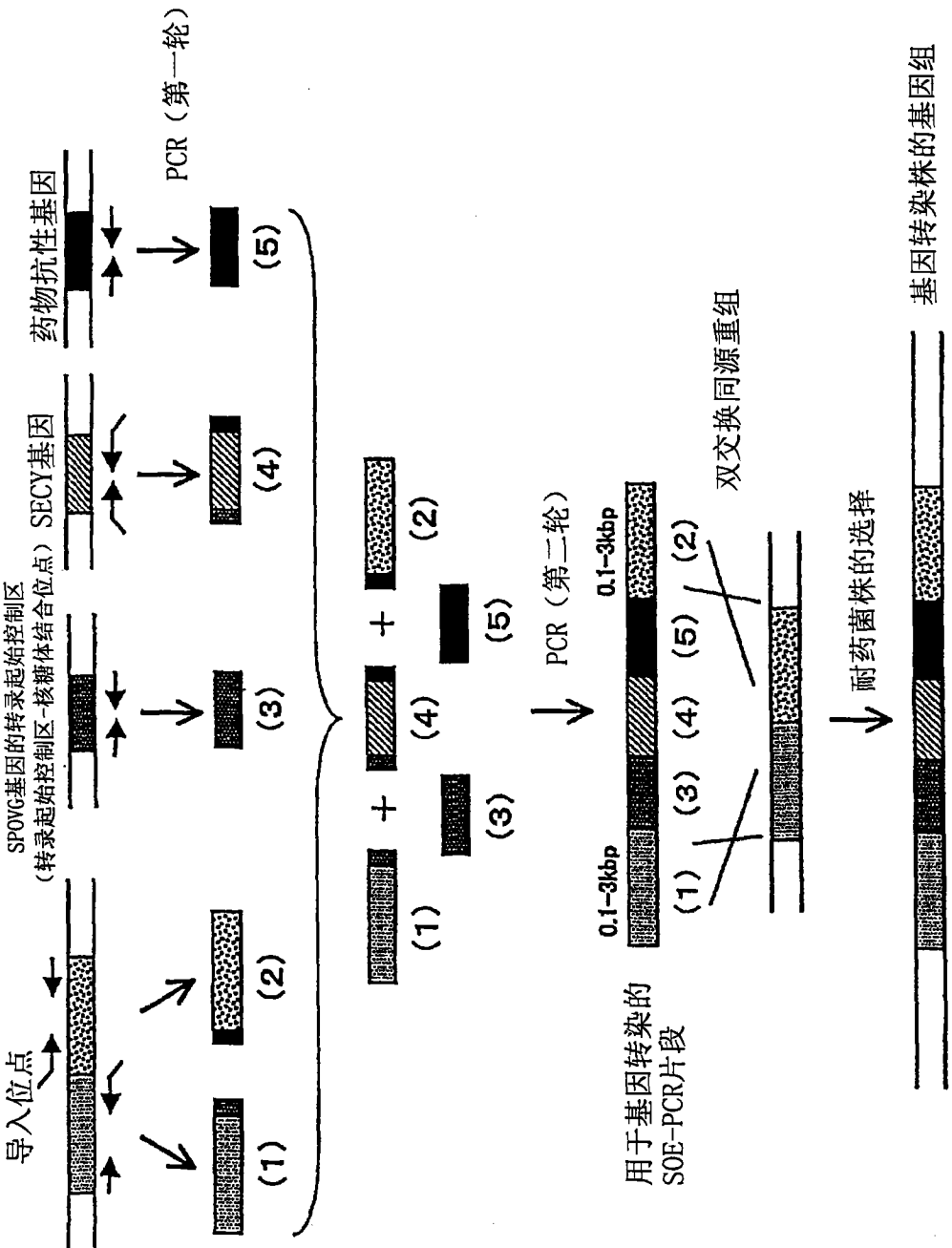

导入位点
SPOVG基因的转录起始控制区
（转录起始控制区-核糖体结合位点）
SECY基因
药物抗性基因
PCR（第一轮）
(1)
(2)
(3)
(4)
(5)
PCR（第二轮）
0.1-3kbp
用于基因转染的
SOE-PCR片段
双交换同源重组
耐药菌株的选择
基因转染株的基因组

图 1

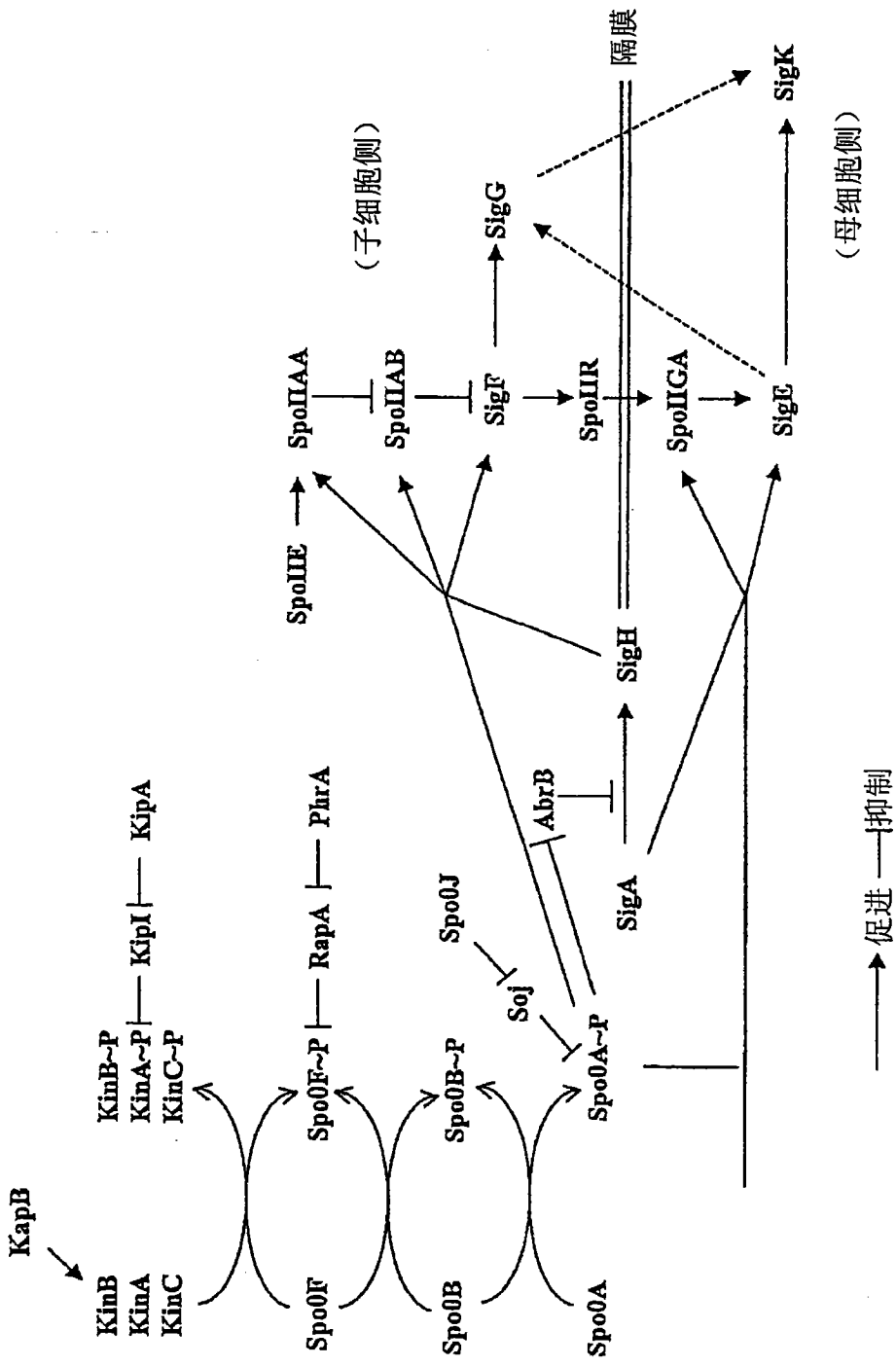

KapB
KinB
KinA
KinC
KinB~P
KinA~P
KinC~P
KipI
KipA
Spo0F
Spo0F~P
RapA
PhrA
Spo0B
Spo0B~P
Spo0J
Soj
Spo0A
Spo0A~P
AbrB
SigA
SigH
SpoIIE
SpoIIAA
SpoIIAB
SigF
SigG
（子细胞侧）
SpoIIR
SpoIIGA
SigE
SigK
隔膜
（母细胞侧）
促进
抑制

图 2

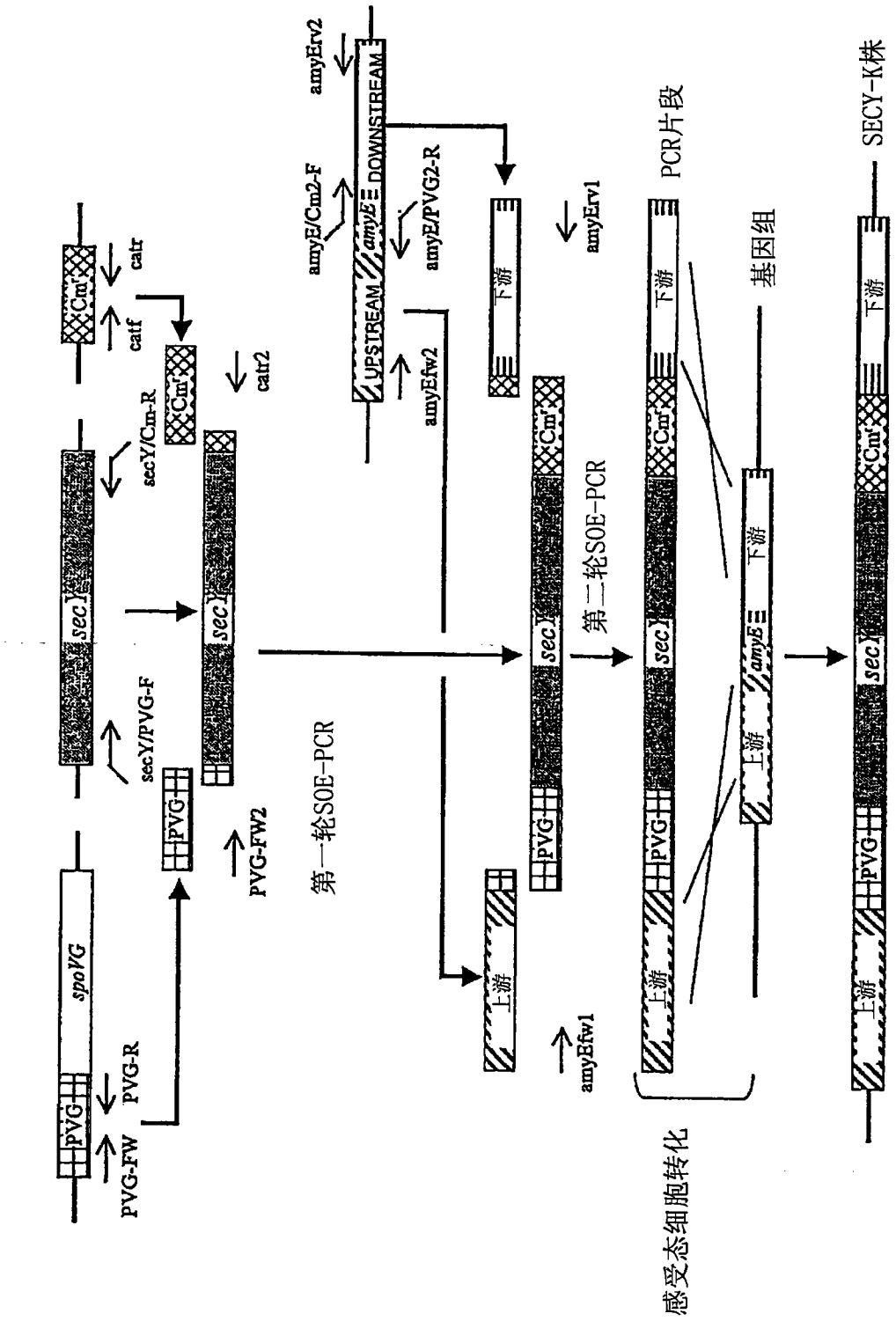

PVG
spoVG
PVG-FW
PVG-R
secY
Cm'
catf
catr
secY/Cm-R
catr2
secY/PVG-F
PVG-FW2
第一轮SOE-PCR
amyE/Cm2-F
amyErv2
amyE
UPSTREAM
DOWNSTREAM
amyEfw2
amyE/PVG2-R
下游
上游
amyErv1
amyEfw1
第二轮SOE-PCR
PCR片段
基因组
感受态细胞转化
SECY-K株

图 3

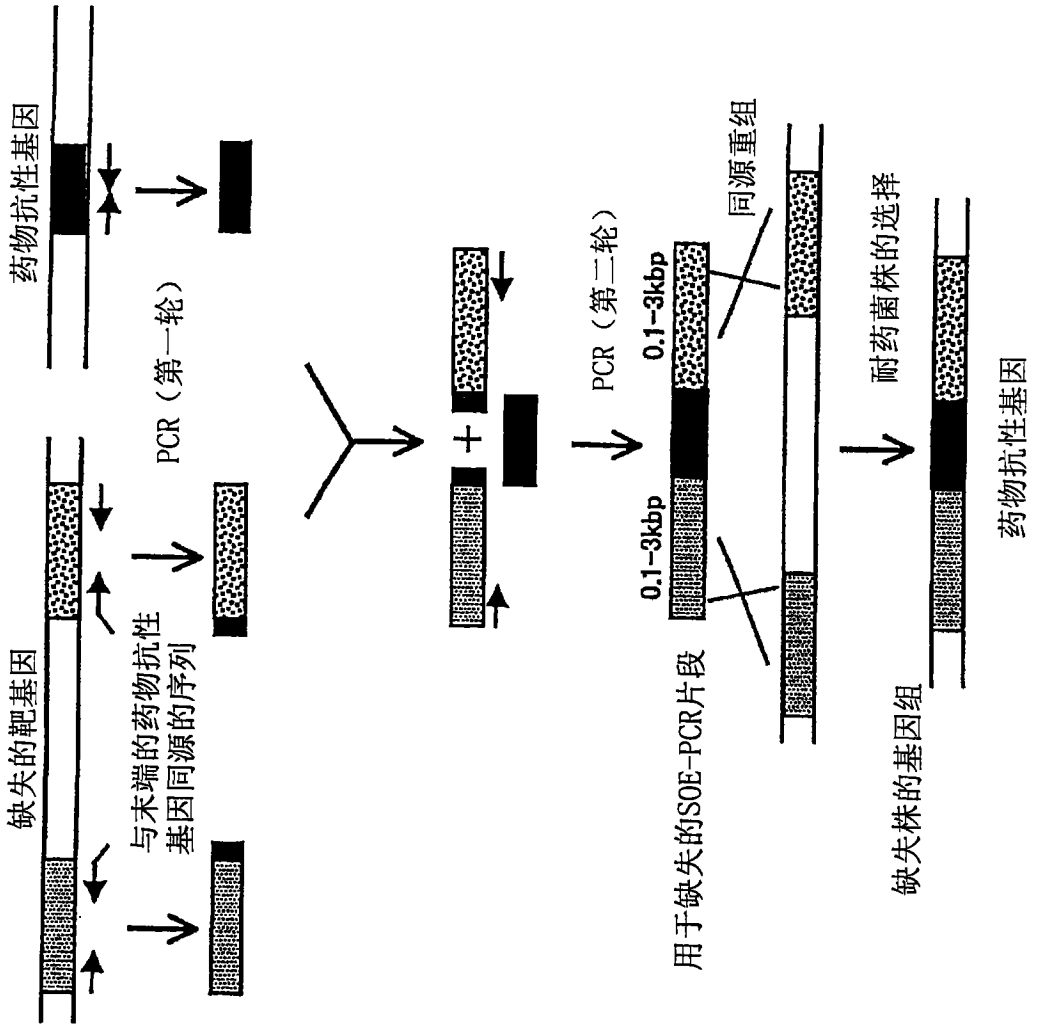

药物抗性基因
缺失的靶基因
与末端的药物抗性基因同源的序列
PCR（第一轮）
PCR（第二轮）
0.1-3kbp
0.1-3kbp
用于缺失的SOE-PCR片段
同源重组
耐药菌株的选择
缺失株的基因组
药物抗性基因

图 4