[19] 中华人民共和国国家知识产权局

[51] Int. Cl[7]
C12N 15/52
C12N 15/54 C12N 1/21
C12N 9/12 C12P 13/08
//C12N15/52,C12R1:15

[12] 发明专利说明书

[21] ZL 专利号 97120819.0

[45] 授权公告日 2004 年 8 月 11 日

[11] 授权公告号 CN 1161470C

[22] 申请日 1997.12.5 [21] 申请号 97120819.0

[30] 优先权

[32] 1996.12.5 [33] JP [31] 325658/1996

[71] 专利权人 味之素株式会社

地址 日本东京都

[72] 发明人 早川敦 杉本雅一 吉原康彦
中松亘

审查员 丁 海

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 谭明胜

权利要求书 1 页 说明书 52 页 附图 13 页

[54] 发明名称 用于产生 L-赖氨酸的方法

[57] 摘要

一种在棒状杆菌细胞中可自主复制的重组 DNA，该 DNA 包含编码天冬氨酸激酶(其中 L-苏氨酸和 L-赖氨酸的反馈抑制实质上被脱敏)的 DNA 序列和编码二氨基庚二酸脱羧酶的 DNA 序列。一种具有天冬氨酸激酶的棒状杆菌(其中 L-苏氨酸和 L-赖氨酸的反馈抑制实质上被脱敏)，其包含编码二氨基庚二酸脱羧酶的增强 DNA 序列；一种用于产生 L-赖氨酸的方法，该方法包括以下步骤：在适当的培养基中培养棒状杆菌使得在所说的细菌的培养物中产生并积累 L-赖氨酸，并从培养物收集 L-赖氨酸。

ISSN1008-4274

1. 一种具有其中 L-赖氨酸和 L-苏氨酸的反馈抑制通过突变被脱敏的天冬氨酸激酶的棒状杆菌，其中二氨基庚二酸脱羧酶的胞内活性通过增加二氨基庚二酸脱羧酶基因的拷贝数、使用强启动子或者组合这些方法得到提高。

2. 按照权利要求 1 的棒状杆菌，其中磷酸烯醇丙酮酸羧酶的胞内活性通过增加磷酸烯醇丙酮酸羧酶基因的拷贝数、使用强启动子或者组合这些方法得到提高。

3. 一种用于产生 L-赖氨酸的方法，该方法包括以下步骤：在适当的培养基中培养如权利要求 1 所限定的棒状杆菌，使得在所说的细菌的培养物中产生并积累 L-赖氨酸，并从培养物收集 L-赖氨酸。

4. 一种用于产生 L-赖氨酸的方法，该方法包括以下步骤：在适当的培养基中培养如权利要求 2 所限定的棒状杆菌，使得在所说的细菌的培养物中产生并积累 L-赖氨酸，并从培养物收集 L-赖氨酸。

用于产生 L-赖氨酸的方法

本发明涉及通过培养微生物产生 L-赖氨酸的方法，所述的微生物是通过借助基于基因工程的技术修饰用于发酵生产氨基酸的棒状杆菌等获得的。

作为饲料添加剂的赖氨酸通常通过利用属于棒状杆菌的 L-赖氨酸-产生突变菌株用发酵法产生。现在已知各种 L-赖氨酸-产生菌是从属于棒状杆菌的野生型菌株开始经人工突变产生的那些。

就棒状杆菌而言，公开了载体质粒，这种质粒是在细菌细胞中可自主复制的，并具有药物抗性标记基因(参见美国专利 4,514,502)，并公开了用于把基因引入细菌细胞的方法(例如日本专利申请公开 2207791)。也公开了通过利用以上所述的技术培养 L-苏氨酸-或 L-异亮氨酸-产生菌的可能性(参见美国专利 4,452,890)。培养 L-赖氨酸-产生菌的技术是已知的，在其中参与 L-赖氨酸生物合成的基因掺入到载体质粒中，以便在细菌细胞中扩增该基因(例如，日本专利申请公开 56-160997)。

L-赖氨酸生物合成的已知基因包括，例如，二氢二吡啶甲酸还原酶基因(日本专利申请公开 7-75578)和二氨基庚二酸脱氢酶基因(Ishino, S.等，核酸研究, 15, 3917))(其中参与 L-赖氨酸生物合成的基因被克隆)，以及磷酸烯醇丙酮酸羧酶基因(日本专利申请公开 60-87788)，二氢二吡啶甲酸合酶基因(日本专利出版物 6-55149)和二氨基庚二酸脱羧酶基因(日本专利申请公开 60-62994)(其中扩增基因影响 L-赖氨酸产量)。

就参与 L-赖氨酸生物合成的酶而言，在以野生型使用时，经历反馈抑制的酶的情况是已知的。在这种情况下， L-赖氨酸产量通过这样一种方法得到提高，即引入具有脱敏反馈抑制之突变的基因。具体地说，已知这样的基因包括天冬氨酸激酶基因(国际出版物 WO 94/25605)。

如上所述，借助 L-赖氨酸生物合成系统的基因扩增或突变基因的引入获得了成功的结果。例如，具有突变天冬氨酸激酶基因(脱敏了由赖氨酸和苏氨酸引起的固有抑制)的棒状杆菌产生大量的 L-赖氨酸(大约 25g/L)。然而，这种细菌与不具有突变天冬氨酸激酶基因的细菌比较生长速度降低。也报道了通过

除了引入突变天冬氨酸激酶基因之外进一步引入二氢二吡啶甲酸合酶基因可提高 L-赖氨酸产量(应用的和环境的微生物学， 57(6)， 1746-1752(1991))。然而，这种细菌生长速度进一步降低。

也没有报道通过增强 L-赖氨酸生物合成基因改善生长的情况。目前，对棒状杆菌，没有出现任何人已成功地通过组合 L-赖氨酸生物合成的多种基因明显改善 L-赖氨酸产量而不抑制生长的情况。

本发明的目的是通过在棒状杆菌中同时增强 L －赖氨酸生物合成的多种基因来提高 L －赖氨酸的产率，而不障碍棒状杆菌的生长。

当目标产物经采用微生物发酵产生时，与引入的物质相关的目标产物的产生速度和产率是非常重要的因素。通过增加每单位发酵设备的产生速度，可以廉价地生产目标产物。因此，发酵产率和产生速度相互匹配是工业上非常重要的。为了通过采用棒状杆菌经发酵产生 L －赖氨酸，本发明提出了以上所述问题的一种解决方案。

本发明的原理基于这样一种事实: 通过增强编码天冬氨酸激酶(其中实质上脱敏了 L-苏氨酸 L-赖氨酸的反馈抑制)的 DNA 序列和编码二氨基庚二酸脱羧酶的 DNA 序列两者，可以改善棒状杆菌的生长，并且因此可以改善 L －赖氨酸产生速度(与单独增强这两种 DNA 序列之一比较)。

在本发明的第一方面，本发明提供了在棒状杆菌细胞中可自主复制的重组 DNA，该 DNA 包含编码天冬氨酸激酶(其中实质上脱敏了 L-苏氨酸 L-赖氨酸的反馈抑制)的 DNA 序列和编码二氨基庚二酸脱羧酶的 DNA 序列。本发明也提供了还包含编码磷酸烯醇丙酮酸羧酶的重组 DNA 。

在第二方面，本发明提供了一种具有天冬氨酸激酶的棒状杆菌(其中 L-苏氨酸和 L-赖氨酸的反馈抑制实质上被脱敏)，该棒状杆菌包含编码二氨基庚二酸脱羧酶的增强 DNA 序列。本发明也提供了还包含编码磷酸烯醇丙酮酸羧酶的增强 DNA 序列的棒状杆菌。

在第三方面，本发明提供了一种用于产生 L-赖氨酸的方法，该方法包括以下步骤: 在适当的培养基中培养以上所述的任何棒状杆菌使得在所说的细菌的培养物中产生并积累 L-赖氨酸，并从培养物收集 L-赖氨酸。

此后，需要时天冬氨酸激酶称为"AK "，编码 AK 的基因称为"*lysC*")，由 L －赖氨酸和 L －苏氨酸产生的反馈抑制被脱敏的 AK 称为“突变 AK”。编码突变

AK 的基因称为"突变 *lysC*"。同样，需要时二氨基庚二酸脱羧酶称为"DDC "，编码 DDC 的基因称为"*lysA*")，磷酸烯醇丙酮酸羧酶称为"PEPC"，编码 PEPC 的基因称为"*ppc*"。

在本发明中所称的棒状杆菌是在 Bergey's 细菌学手册(第 8 版 p.599(1974))中限定的一组微生物，其是无孢子-形成能力的需氧革兰氏-阳性-非抗酸棒杆状菌。棒状杆菌包括属于棒杆菌属的细菌，属于短杆菌短的细菌(迄今为止分类为短杆菌属但现在与棒杆菌属细菌合并在一起)和与属于棒杆菌的细菌密切相关的短杆菌属的细菌。

按照本发明，棒状杆菌的 L-赖氨酸产生量和产生速度可以得到提高。

本发明具体涉及以下技术方案：

1. 一种具有其中 L-赖氨酸和 L-苏氨酸的反馈抑制通过突变被脱敏的天冬氨酸激酶的棒状杆菌，其中二氨基庚二酸脱羧酶的胞内活性通过增加二氨基庚二酸脱羧酶基因的拷贝数、使用强启动子或者组合这些方法得到提高。

2. 按照 1 的棒状杆菌，其中磷酸烯醇丙酮酸羧酶的胞内活性通过增加磷酸烯醇丙酮酸羧酶基因的拷贝数、使用强启动子或者组合这些方法得到提高。

3. 一种用于产生 L-赖氨酸的方法，该方法包括以下步骤：在适当的培养基中培养如 1 所限定的棒状杆菌，使得在所说的细菌的培养物中产生并积累 L-赖氨酸，并从培养物收集 L-赖氨酸。

4. 一种用于产生 L-赖氨酸的方法，该方法包括以下步骤：在适当的培养基中培养如 2 所限定的棒状杆菌，使得在所说的细菌的培养物中产生并积累 L-赖氨酸，并从培养物收集 L-赖氨酸。

附图简要描述

图 1 说明包含突变 *lysC* 的质粒 p399AK9B 和 p399AKYB 的构建方法

图 2 说明包含 *lysA* 的质粒 p299LYSA 的构建方法。

图 3 说明包含 *lysA* 和 Brevi.-ori 的质粒 pLYSAB 的构建方法。

图 4 说明包含 PEPC 结构基因的质粒 pAKPFds 的构建方法。

图 5 说明棒状杆菌的新克隆载体 pVKG 和 pVK7 的构建方法。

图 6 说明包含野生型高表达 *ppc* 的质粒 pPwm 的构建方法。

图 7 说明包含突变 *lysC*, *lysA* 和 Brevi.-ori 的质粒 pCL 的构建方法。

图 8 说明包含 *dapA* 和 Brevi.-ori 的质粒 pDPSB 的构建方法。

图 9 说明包含 *dapB* 和 Brevi.-ori 的质粒 pDPRB 的构建方法。

图 10 说明包含 *ddh* 和 Brevi.-ori 的质粒 pPK4D 的构建方法。

图 11 说明包含 *lysC*, *dapA* 和 Brevi.-ori 的质粒 pCRCAB 的构建方法。

图 12 说明包含突变 *lysC*, *dapB* 和 Brevi.- ori 的质粒 pCB 的构建方法

图 13 说明包含法突变 *lysC* 和 *ddh* 的质粒 pCD 的构建方法。

发明的详尽描述

1 制备用于本发明的 L-赖氨酸生物合成的基因

用于本发明的 L-赖氨酸生物合成的基因分别通过以下方法获得：从 DNA 供体细菌染色体 DNA 用质粒载体或类似物构建染色体 DNA 文库，选择具有所需基因的菌株，和从所选择的菌株回收已插入了所述基因的重组 DNA。对用于本发

明的 L-赖氨酸生物合成的基因的 DNA 供体没有特别的限制，只要 L-赖氨酸生物合成的所需基因表达酶蛋白质(其在棒状杆菌细胞中起作用)。然而， DNA 供体优选地是棒状杆菌。

来源于棒状杆菌的 *lysC*， *dapA* 和 *ppc* 的所有基因具有已知的序列。因此，可以通过按照聚合酶链反应方法(PCR; see White, T. J.等, Trends Genet., 5, 185(1989))进行扩增来获得它们。

用于本发明的 L-赖氨酸生物合成的各种基因可以按照以下例举的方法获得:

(1)突变 *lysC*的制备

包含突变 *lysC* 的 DNA 片段可以从突变菌株制备，在所述菌株中由 L-苏氨酸和 L-赖氨酸产生的 AK 活性的协同反馈抑制实质上被脱敏(国际出版物 WO 94/25605)。这样一种突变菌株可以通过例如用诱变剂如紫外光和 N-甲基-N-硝基-N-亚硝基胍(NTG)对来源于棒状杆菌野生型菌株的一组细胞进行常规突变处理获得。 AK 活性可以通过用由 Miyajima, R.等在生物化学杂志(1968), 63(2), 139-148 中描述的方法测量。最优选的这样的突变菌株由 L-赖氨酸-产生菌 AJ 3445(FERM 1944)代表，其来源于突变处理乳发酵短杆菌 ATCC 13869 野生型菌株(现在其改变的名称是 *Corynebacterium alutamicum*)。

另外，突变 *lysC*也可由包含野生型 *lysC*的质粒 DNA 体外突变处理获得。另一方面，突变脱敏 L-赖氨酸和 L-苏氨酸产生的对 AK 活性的协同反馈抑制的信息是十分清楚的(国际出版物 WO 94/25605)。因此，在该信息的基础上按照定点突变方法也可以制备突变 *lysC*。

可以按照例如 Saito 和 Miura(H. Saito 和 K. Miura, 生物化学生物物理学报, 72, 619(1963))的方法通过制备染色体 DNA 或按照聚合酶链反应方法(PCR; 参见 White, T. J.等, Trends Genet., 5, 185(1989))扩增 *lysC*从棒状杆菌分离包含 *lysC*的片段。

以单链 DNA 例证性地说明 DNA 引物，所说单链 DNA 是具有 SEQ ID NO:1 和 2 所示的核苷酸序列的 23 链节和 21 链节的，以便扩增例如基于谷氨酸棒杆菌已知序列(参见分子微生物学(1991), 5(5), 1197-1204; Mel. Gen. Genet.(1990), 224, 317-324)的编码 *lysC*的约 1,643 bp 的区域。可以通过采用 Applied Biosystems 生产的 DNA 合成仪 380B 型和采用 phosphoamidite

法(参见 Tetrahedron Letters(1981), 22, 1859)按照常规方法合成 DNA。可以通过用由 Takara Shuzo 生产的 DNA 热循环器 PJ2000 型按照供给者指定的方法和尾端 DNA 聚合酶进行 PCR。

优选的是将经 PCR 扩增的 *lysC* 与在大肠杆菌和/和/或棒状杆菌细胞中可自主复制的载体 DNA 连接制备重组 DNA，把重组 DNA 引入预先的大肠杆菌细胞中。这一过程使随后的操作容易进行。在大肠杆菌细胞中可自主复制的载体优选地是质粒载体，其是在宿主的细胞中优选的可自主复制的，包括例如，pUC19、pUC18、pBR322、pHSG299、pHSG399、pHSG398 和 RSF1010。

当具有使得质粒可以在棒状杆菌中自主复制的能力的 DNA 片段插入到这些载体中时，它们可以用作在大肠杆菌和棒状杆菌中可自主复制的所谓的穿梭载体。

这样穿梭载体包括以下载体。含有各载体的微生物和其在国际保藏单位中的保藏号(在括号)显示如下:

pHC4:	大肠杆菌 AJ12617(FERM3532)
pAJ655:	大肠杆菌 AJ11882(FERM136)
	谷氨酸棒杆菌 SR8201(ATCC39135)
pAJ1844:	大肠杆菌 AJ11883(FERM137)
	谷氨酸棒杆菌 SR8202(ATCC39136)
pAJG11:	大肠杆菌 A511884(FERM138)
pAJ3148:	谷氨酸棒杆菌 SR8203(ATCC39137)
pAJ440:	枯草芽孢杆菌 AJ11901(FERM140)

这些载体可从其后的保藏的微生物获得。采用溶菌酶和 SDS 裂解在对数的生长期所收集的细胞，接着在 30,000 × g 下从裂解物分离以获得上清液。将聚乙二醇添加到上清液中，借助氯化铯-溴化乙锭均衡密度梯度离心进行分级分离和纯化。

可以通过按照以下方法将质粒引入来转化大肠杆菌，所述方法例如 D. M. Morrison (酶学方法, 68, 326 (1979))的方法或其中用氯化钙处理受体细胞以增加 DNA 渗透性的方法(Mandel, M. 和 Higa, A., 分子生物学杂志, 53, 159 (1970))。

当按照以上所述的方法从AK野型菌株分离 *lysC* 时，获得野生型 *lysC*，而从AK 突变菌株分离 *lysC* 时，获得突变 *lysC*。

包含野生型 *lysC* 的DNA片段的核苷酸序列的例子在序列表中SEQ ID NO: 3中显示。野生型 AK 蛋白质的α 亚单位的氨基酸序列是从核苷酸序列推定的，其与DNA序列一起在序列表SEQ ID NO: 4中显示。单独的氨基酸序列在SEQ ID NO: 5中显示。野生型AK蛋白质的β 亚单位的氨基酸序列是从DNA的核苷酸序列推定的，其与DNA序列一起在序列表SEQ ID NO: 6中显示。单独的氨基酸序列在SEQ ID NO: 7中显示。在各亚单位中，GTG 用作起始密码子，相应的氨基酸由甲硫氨酸代表。然而，这种代表涉及到甲硫氨酸、缬氨酸或甲酰甲硫氨酸。

对用于本发明的突变 *lysC* 没有特别的限制，只要其编码AK(其中由L-苏氨酸和L-赖氨酸产生的协同反馈抑制被脱敏)。然而，突变 *lysC* 的一个例证性例子包括这样的突变，其中与279位(从N-末端开始计算)丙氨酸残基相应的氨基酸残基改变成为非丙氨酸和非非野生型 AK 氨基酸序列中的β 亚单位中的酸性氨基酸的氨基酸残基，与30位(从N-末端开始计数)丙氨酸残基相应的氨基酸残基改变成为非丙氨酸和非野生型 AK 氨基酸序列中的β 亚单位中的酸性氨基酸的氨基酸残基。野生型 AK 的氨基酸序列具体来说包括作为α 亚单位的在序列表中SEQ ID NO: 5所示的氨基酸序列，作为β 亚单位的在序列表中SEQ ID NO: 7所示的氨基酸序列。

称为非丙氨酸和非酸性氨基酸的那些氨基酸残基优选的包括苏氨酸、精氨酸、半胱氨酸、苯丙氨酸、脯氨酸、丝氨酸、酪氨酸和缬氨酸残基。

对相应于所要替代的氨基酸残基的密码子没有特别的限制，只要它编码所说的氨基酸残基。可以预计，依据细菌物种和细菌菌株的不同，野生型 AK 的氨基酸序列可以稍微不同。具有基于在一个或多个位置上取代，缺失或插入一个或多个氨基酸残基而不影响以上所述活性的突变的 AK 也可以用于本发明中。 具有自发突变的编码AK的DNA可以通过分离这样一种DNA获得，该DNA在严格条件与例如具有SEQ ID NO: 3所示的一部分核苷酸序列的DNA杂交。本文所说的"严格条件"意指形成特异性杂交而不形成非特异性杂交的条件。用数值难于清楚表达所述条件。然而，例举性的条件是这样一种条件，在该条件下，具有高同源性的核酸(例如具有不低于90 %同源性的DNA)相互杂交或者是

这样一种条件，其温度是从熔点温度 Tm 到(Tm-30)℃，优选地是从 Tm 到(Tm-20)℃，盐浓度相当于 1x SSC，优选地相当于 0.1x SSC。

基于例如，取代、缺失或插入一个或多个氨基酸残基进行过人工突变的其它 AK 也可以使用，只要对 AK 活性实质上没有影响和对 L-赖氨酸和 L-苏氨酸的协同反馈抑制的脱敏没有影响。可以通过例如定点突变修饰核苷酸序列，产生特定位点的缺失或插入，由此来获得编码具有人工突变的 AK 的 DNA。也可以通过已知诱变处理方法获得具有突变的 *lysC*。诱变处理包括用羟胺或类似物体外处理包含 *lysC* 的 DNA，用诱变如紫外光或诱变剂(其是用于常规人工诱变的，如 N-甲基-N-硝基-N-亚硝基胍(NTG)或硝酸)处理具有包含 *lysC* 之 DNA 的微生物的。诱变处理后，可以通过选择编码或产生 AK(其具有 AK 活性，其氨基酸序列从经历诱变处理的 DNA 或者经历诱变处理的微生物突变)的 DNA 或微生物测定引入了突变或者发生了突变的位点。对所引入突变的位点没有特定的限制，只要对 AK 活性实质上没有影响和对反馈抑制的脱敏没有影响。所引入的突变的数目随在蛋白质的立体结构中的氨基酸的位置和种类变化，并且没有特定的限制，只要对 AK 活性实质上没有影响和对反馈抑制的脱敏没有影响。该数目通常是 1 至 20，优选地是 1 至 10。

通过将突变 *lysC* 质粒 p399AK9B 引入乳发酵短杆菌野生型菌株 AJ12036 菌株(FERM BP-734)获得的 AJ12691 菌株已于 1992 年 4 月 10 日以保藏号 FERM P-12918 保藏在国际贸易和工业部工业科学和技术机构的国立生命科学和人体技术研究所(1-3, Higashi L-chome, Tsukuba-shi, Ibaraki-ken, 305, 日本)，基于布达佩斯条约于 1995 年 2 月 10 日转变成国际保藏，以保藏号 FERM BP-4999 保藏。

(2) *lysA* 的制备

可以借助 PCR 从棒状杆菌的染色体制备包含 *lysA* 的 DNA 片段。对 DNA 供体没有特别的限制，然而，由乳发酵短杆菌 ATCC 13869 菌株所例证。

在棒状杆菌中，*lysA* 与 *argS* 一道形成操纵子(精氨酰-tRNA 合酶基因)，*lysA* 存在于 *argS* 下游。*lysA* 的表达由存在于 *argS* 上游的启动子调节(参见生物技术杂志，Nov.，7356-7362(1993))。谷氨酸棒杆菌的这些基因的 DNA 序列是已知的(参见分子微生物学，4(11)，1819-1830(1990)；分子和普通遗传学，212，112-119(1988))，基于该序列可以制备 PCR 的 DNA 引物。这样的

DNA 引物由分别具有序列表 SEQ ID NO: 8(相应于在分子微生物学 4(11), 1819-1830(1990)中描述的核苷酸序列中的 11 至 33 位核苷酸)和 SEQ ID NO: 9(相应于在分子和普通遗传学, 212, 112-119(1988)中描述的核苷酸序列中的 1370 至 1392 位核苷酸)中所示的核苷酸序列的 23-链节 DNA 所具体例证。可以以与以上所述的有关 *lysC* 的那些相同的方式进行 DNA 的合成、PCR 和包含获得的 *lysA* 的质粒的制备。

在以下描述的实施例中，包含启动子、*argS* 和 *lysA* 的 DNA 片段用来增强 *lysA*。然而，*argS* 对本发明不是必要的。其允许使用在其中 *lysA* 就连接在启动子下游的 DNA 片段。

包含 *argS* 和 *lysA* 以及由核苷酸序列编码的推断的氨基酸序列的 DNA 片段的核苷酸序列在 SEQ ID NO: 18 中例证。由 *argS* 编码的氨基酸序列的实例在 SEQ ID NO: 19 中显示，由 *lysA* 编码的氨基酸序列的实例在 SEQ ID NO: 12 中显示。除编码这些氨基酸序列的 DNA 片段之外，如果对 DDC 活性实质上没有影响，本发明可以等同地使用编码实质上与 SEQ ID NO: 12 所示的氨基酸序列相同的氨基酸序列(即具有基于例如取代、缺失或插入一个或多个氨基酸之突变的氨基酸序列)的 DNA 片段。可以用与编码具有突变(只要对 AK 活性实质上没有影响和对 L-赖氨酸和 L-苏氨酸的协同反馈抑制的脱敏没有影响)的 AK 的 DNA 的那些相同的方式获得具有自发和人工突变的 *lysA*。

(2) *ppc* 的制备

可以借助 PCR 从棒状杆菌的染色体制备包含 *ppc* 的 DNA 片段。对 DNA 供体没有特别的限制，然而，由乳发酵短杆菌 ATCC 13869 菌株所例证。

对乳发酵短杆菌的 *ppc* 序列是已知的 (参见 O'Regan, M. 等, 基因, 77, 237-251 (1989))，基于该序列可以制备 PCR 的 DNA 引物。这样的 DNA 引物由分别具有序列表 SEQ ID NO: 13 和 14 中所示的核苷酸序列的 23-链节 DNA 所具体例证。可以以与以上所述的有关 *lysC* 的那些相同的方式进行 DNA 的合成、PCR 和包含获得的 *ppc* 的质粒的制备。

包含 *ppc* 之 DNA 片段的核苷酸序列和由所述核苷酸序列编码的推断的氨基酸序列在 SEQ ID NO: 15 中显示。单独的氨基酸序列在 SEQ ID NO: 16 中显示。

除编码这种氨基酸序列的 DNA 片段之外，如果对 PEPC 活性实质上没有影响，本发明可以等同地使用编码实质上与 SEQ ID NO: 16 所示的氨基酸序列相

同的氨基酸序列(即具有基于例如取代、缺失或插入一个或多个氨基酸之突变的氨基酸序列)的 DNA 片段。可以用与编码具有突变(只要对 AK 活性实质上没有影响和对 L-赖氨酸和 L-苏氨酸的协同反馈抑制的脱敏没有影响)的 AK 的 DNA 的那些相同的方式获得具有自发和人工突变的 *ppc* 。

棒状杆菌的 *ppc* 与 gap(甘油醛－ 3-磷酸脱氢酶基因)、 pgk(磷酸甘油酸盐激酶基因)和 tpi(丙糖磷酸异构酶基因)一道形成操纵子，并且 *ppc* 存在于 tpi 的下游。 *ppc* 的表达由存在于 pgk 上游的启动子调节(参见 Schwinde, J.W. 等, 细菌学杂志, 175(12), 3905-3908(1993))。因此，类似以上提到的 *lysA* ， *ppc* 可以与 pgk 和 tpi 一道经 PCR 扩增，以使用包含 pgk 、 tpi 和 *ppc* 的 DNA 片段，如以下描述的实施例所示，允许使用在其中合适的启动子就连接在 PEPC 编码区上游的 DNA 片段。所述的启动子包括 *lysC* 启动子、来源于大肠杆菌的 tac 启动子和 trc 启动子。

2 本发明的重组 DNA 和棒状杆菌

所述重组 DNA 包含编码天冬氨酸激酶(其中实质上脱敏了 L-苏氨酸 L-赖氨酸的反馈抑制)的 DNA 序列和编码二氨基庚二酸脱羧酶的 DNA 序列，并且在棒状杆菌细胞中可自主复制。在一个优选的实施方案中，所述重组 DNA 除了包含以上所述的 DNA 序列外还包含编码磷酸烯醇丙酮酸羧酶的 DNA 序列。

本发明的棒状杆菌含有天冬氨酸激酶(突变 AK)，在其中由 L-赖氨酸和 L-苏氨酸产生的反馈抑制实质上被脱敏，其中所说的编码二氨基庚二酸脱羧酶的 DNA 序列被增强。

术语"增强"本文指这一事实: 通过例如，增加基因的拷贝数，用强启动子，使用编码具有高比活性的酶的基因或者组合这些方法，由所说 DNA 编码的酶的胞内活性得到提高。

含有突变 AK 的棒状杆菌可以是作为突变的结果产生突变天冬氨酸激酶的那些或通过引入突变 *lysC* 转化的那些。

用以引入以上所述的 DNA 棒状杆菌的例子包括，例如，下列赖氨酸-产生野生型菌株:

Corynebacterium acetoacidophilum ATCC 13870;

Corynebacterium acetocrlutamicum ATCC 15806;

美棒杆菌 ATCC 15991;

谷氨酸棒杆菌 ATCC 13032;

(*Brevibacterium divaricatum*) ATCC 14020;

(乳发酵短杆菌) ATCC 13869;

(*Corynebacterium lilium*) ATCC 15990;

(*Brevibacterium flavum*) ATCC 14067;

Corynebacterium melassecola ATCC 17965;

Brevibacterium saccharolvticum ATCC 14066;

Brevibacterium immariophilum ATCC 14068;

Brevibacterium roseum ATCC 13825;

Brevibacterium thioaenitalis ATCC 19240;

Microbacterium ammoniaphilum ATCC 15354;

Corynebacterium thermoaminoaenes AJ12340 (FERM BP-1539)

除了以上所述的细菌菌株外，其他可用的宿主包括，例如，来源于以上所述菌株的具有 L-赖氨酸-产生能力的突变菌株。这样的的人工突变菌株包括如下菌株: (2-氨乙基)-半胱氨酸(在下文中简称为"AEC")抗性突变菌株，如乳发酵短杆菌 AJ11082 (NRRL B-1147)，日本专利出版物 56-1914 、 56-1915 、 57-14157 、 57-14158 、 57-30474 、 58-10075 、 59-4993 、 61-35840 、 62-24074 、 62-36673 、 5-11958 、 7-112437 和 7-112438)；突变菌株，其生长需要氨基酸(如 L-高丝氨酸)的突变菌株(日本专利出版物 48-28078 和 56-6499)；显示出对 AEC 的抗性并需要氨基酸如 L-亮氨酸、 L-高丝氨酸、 L-脯氨酸、 L-丝氨酸、 L-精氨酸、 L-丙氨酸和 L-缬氨酸的突变菌株(美国专利 3,708,395 和 3,825,472)；显示出对 DL-α-氨基-ε-己内酰胺、α-氨基-月桂内酰胺、天冬氨酸-类似物、磺胺类药、醌型和 N-月桂酰亮氨酸抗性的 L-赖氨酸-产生突变菌株；显示出对 oxyaloacetate 脱羧酶抑制剂或呼吸系统酶抗性的 L-赖氨酸-产生突变菌株(日本专利申请公开 50-53588 、 50-31093 、 52-102498 、 53-9394 、 53-86089 、 55-9783 、 55-9759 、 56-32995 和 56-39778，以及日本专利出版物 53-43591 和 53-1833)；需要肌醇或醋酸的 L-赖氨酸-产生突变菌株(日本专利申请 55-9784 和 56-8692)；对氟丙酮酸或不低于 34 ℃的温度表现出敏感性的 L-赖氨酸-产生突变菌株(日本专利申请公开 55-9783 和

53-86090)；属于短杆菌属或棒杆菌属的表现出对乙二醇抗性并产生 L-赖氨酸的产生突变菌株(美国专利 4,411,997)。

在一个特定的实施方案中，为了如上所述在宿主中增强 L-赖氨酸生物合成基因，通过采用质粒载体、转座子或噬菌体载体或类似物将所述基因引入到宿主中。一旦引入，就预期会有某种程度的增强，即使使用低拷贝型载体。然而，优选地是采用多拷贝型载体。这样载体包括，例如以上所述的质粒载体，pAJ655 、 pAJ1844 、 pAJG11 、 pAJ3148 和 pAJ440 。此外，在以下文献中描述了来源于棒状杆菌的转座子: 国际出版物 WO02/02627 和 WO93/18151，欧洲专利出版物 445385，日本专利申请公开 6-46867； Vertes, A. A.等，分子微生物学，11，739-746 (1994)，； Bonamy, C.，等，分子微生物学，14，571-581 (1994)； Vertes, A. A.等，Mel. Gen. Genet.，245，397-405 (1994)； Jagar, W.等，FEMS 微生物学通讯，126，1-6 (1995)，日本专利申请公开 7-107976；日本专利申请公开 7-327680 等。

在本发明中，不可缺少的是突变 *lysC* 必需被增强。允许使用在染色体 DNA 上的 *lysC* 的突变的那些或者其中突变 *lysC* 掺入到染色体 DNA 中从那些。另外，突变 *lysC* 可以通过用质粒载体引入。另一方面，为了有效地产生 L-赖氨酸，优选地是增强 *lysA* 和 *ppc* 。

通过分别采用不同的载体，可以成功地将 *lysC* 、 *dapA* 和 *ppc* 各基因引入宿主中。另外，可以用单一载体将二或三种基因一起引入。当使用不同的载体时，基因可以以任何秩序引入，然而，优选地使用这样一些载体，它们具有在宿主细胞中的稳定参与和隐藏机制，并且能够相互共存。

通过例如向宿主棒状杆菌引入在棒状杆菌细胞中可自主复制的包含突变 *lysC* 、 *lysA* 和 *ppc* 的重组 DNA ，获得含有突变 AK 并包含增强的 *lysA* 的棒状杆菌。

除了含有突变 *lysC* 和 *lysA* 之外还含有增强的 *ppc* 的棒状杆菌可以通过将包含突变 *lysC* 、 *lysA* 和 *ppc* 的在棒状杆菌细胞中可自主复制的重组 DNA 引入到宿主棒状杆菌中获得。含有增强的突变 *lysC* ， *lysA* 和 *ppc* 的棒状杆菌也可以通过将包含 *ppc* 的在棒状杆菌细胞中可自主复制的重组 DNA 引入到含有增强的突变 *lysC* 和 *lysA* 的宿主棒状杆菌中获得。

可以通过例如把以上所述的参与 L-赖氨酸生物合成的各基因插入到载体

(如质粒载体，转座子或噬菌体载体)中获得上述重组 DNA 。

在其中质粒用作载体情况下，按照电脉冲方法(Sugimoto 等，日本专利申请公开 2207791) 可以将重组 DNA 引入宿主。用转座子的基因扩增可以通过引入质粒进行，该质粒把转座子携带至宿主细胞，并诱导转座子转座。

3 用于产生 L-赖氨酸的方法

通过在合适的培养基中培养棒状杆菌(包含如上所述的 L-赖氨酸生物合成的增强的基因)，使得 L-赖氨酸在细菌培养物中产生和积累，并从培养物收集 L-赖氨酸，由此可以有效地产生 L-赖氨酸。

所用的培养基的例举性例子是包含碳源、氮源、无机离子和可有可无的其它有机组分的普通培养基。

作为碳源，可以用食糖，如葡萄糖、果糖、蔗糖、糖蜜和淀粉水解物；和有机酸，如富马酸、柠檬酸和琥珀酸。

作为氮源，可以用无机铵盐，如硫酸铵、氯化铵和磷酸铵；有机氮，如大豆水解物；氨气和氨水。

作为有机微量营养源，以合适的量包含所需物质(如维生素 B1 和 L-高丝氨酸或酵母提取物等)是合乎需要的。此外，如果需要，少量添加磷酸钾、硫酸镁、铁离子、锰离子等。

培养优选地在需氧条件下进行大约 30 至 90 小时。在培养期间培养温度优选地控制在 25 ℃到 37 ℃， pH 优选地控制在 5 至 8 。无机或有机，酸性或碱性物质或氨气等可以用于调节 pH 。可以用普通的离子交换树脂法、沉淀法和其它已知方法结合起来从培养物收集赖氨酸。

实施例

以下参照实施例更详细地解释本发明。

实施例 1： 从乳发酵短杆菌制备野生型 lysC 基因和突变 lysC 基因

1 制备野生型和突变 lysC 和制备含有它们的质粒

将乳发酵短杆菌 ATCC 13869 菌株和通过突变处理从 ATCC 13869 菌株获得的 L-赖氨酸-产生突变菌株 AJ3445(FERM P-1944)用作染色体 DNA 供体。AJ3445 菌株已经历突变，使 *lysC* 改变得实质上脱敏了由赖氨酸和苏氨酸产生的协同抑制(生物化学杂志， 68， 701-710(1970))。

用 PCR 法(聚合酶链反应; 参见 White, T. J. 等, Trends Genet., 5, 185 (1989))从染色体 DNA 扩增包含 *lysC* 的 DNA 片段,就用于扩增的 DNA 引物而言,为了扩增编码 *lysC* 的 1,643 bp 的区,基于谷氨酸棒杆菌已知的序列(参见分子微生物学(1991), 5(5), 1197-1204; 和 Mel. Gen. Genet. (1990), 224, 317-324)合成 23 链节和 21 链节(具有 SEQ ID NO: 1 和 2 所示的核苷酸序列)的单链 DNA 序列。经 Applied Biosystems 生产的 DNA 合成仪 380B 型用普通方法和用 phosphoamidite 法(参见 Tetrahedron Letters (1981), 22, 1859)合成 DNA 。

通过用 Takara Shuzo 生产的 DNA 热循环仪 PJ2000 型按照供给者指定的方法和用尾端 DNA 聚合酶经 PCR 扩增所说的基因。由琼脂糖凝胶电泳确认 1,643 bp 的扩增的基因片段。之后,用普通的方法纯化从凝胶上切下的片段,用限制酶 NruI(由 Takara Shuzo 生产)和 EcoRI(由 Takara Shuzo 生产)消化该片段。

将 pHSG399(参见, S.等, 基因 (1987), 61 63-74)用作基因片段的克隆载体。用限制酶 SmaI(由 Takara Shuzo 生产)和 EcoRI 消化 pHSG399 ,将其与扩增的 *lysC* 片段连接。按照指定的方法用 DNA 连接试剂盒(由 Takara Shuzo 生产)连接 DNA 。由此制备质粒,其中从乳发酵短杆菌染色体扩增的 *lysC* 片段分别与 pHSG399 连接。包含 ATCC 13869(野生型菌株) *lysC* 的质粒命名为 p399AKY ,包含 AJ3463(L-赖氨酸-产生菌) *lysC* 的质粒命名为 p399AK9 。

将具有使细菌在属于棒杆菌属的细菌中可自主复制能力的DNA片段(在下文中称为"Brevi.-ori")分别引入到 p399AKY 和 p399AK9 中,以制备携带有 *lysC* 的在属于棒杆菌属的细菌中可自主复制的质粒。从包含 Brevi.-ori 并且在大肠杆菌和属于棒杆菌属的细菌两种细胞中可自主复制的质粒载体 pHK4 制备 Brevi.-ori 。 pHK4 通过以下方法构建 pHK4: 用 KpnI(由 Takara Shuzo 生产)和 BamHI (由 Takara Shuzo 生产)消化 pHC4 ,提取 Brevi.-ori 片段,与 pHSG298(也已由 KpnI 和 BamHI 消化)连接(参见日本专利申请公开 5-7491)。pHK4 使宿主具有卡那霉素抗性。含有 pHK4 的大肠杆菌称命名为大肠杆菌 AJ13136 ,该菌株已于 1995 年 8 月 1 日以保藏号 FERM BP-5186 保藏在国际贸易和工业部工业科学和技术机构的国立生命科学和人体技术研究所(1-3, Higashi L-chome, Tsukuba-shi, Ibaraki-ken, 305, 日本)。

用限制酶 KpnI 和 BamHI 消化 pHK4 ,使切开的末端平端化。按照指定的方

法用平端化试剂盒(由 Takara Shuzo 生产)进行平端的形成。形成平端后，连接上磷酸化 BamHI 接头(由 Takara Shuzo 生产)进行修饰，使相应于 Brevi.-ori 部分的 DNA 片段可以仅经 BamHI 消化从 pHK4 切下。用 BamHI 消化这一质粒，将所产生的 Brevi.-ori DNA 片段与 p399AKY 和 p399AK9(也已分别用 BamHI 消化)连接，以制备各自含有在属于棒杆菌属的细菌中可自主复制的 *lysC* 基因的质粒。

包含源于 p399AKY 的野生型 *lysC* 基因的质粒命名为 p399AKYB，包含源于 p399AK9 的突变 *lysC* 基因的质粒命名为 p399AK9B。 p399AK9B 和 p399AKYB 的构建方法在图 1 中显示。通过把突变 *lysC* 质粒 p399AK9B 引入乳发酵短杆菌野生型菌株所获得的菌株 AJ12691(AJ12036 菌株， FERM 734) 已于 1992 年 4 月 10 日以保藏号 FERM P-12918 保藏在国际贸易和工业部工业科学和技术机构的国立生命科学和人体技术研究所(1-3, Higashi L-chome, Tsukuba-shi, Ibaraki-ken, 305, 日本)，基于布达佩斯条约于 1995 年 2 月 10 日转变成国际保藏，以保藏号 FERM BP-4999 保藏。

2 测定乳发酵短杆菌野生型 lysC 和突变 lysC 的核苷酸序列

从各自的转化体制备包含野生型 *lysC* 的质粒 p399AKY 和包含突变 *lysC* 的质粒 p399AK9，以测定野生型与突变 *lysC* 的核苷酸序列。核苷酸序列测定按照 Sanger 等 (例如 F. Sanger 等, 美国科学院学报, 74, 5463 (1977))的方法进行。

由 p399AKY 编码的野生型 *lysC* 的核苷酸序列在序列表 SEQ ID NO: 3 中显示。另一方面，由 p399AK9 编码的突变 *lysC* 的核苷酸序列仅具有一个核苷酸突变，这样与野生型 *lysC* 比较，在 SEQ ID NO: 3 中的 1051 位 G 改变成 A。已知谷氨酸棒杆菌的 *lysC* 有两个在同一 DNA 链上的同一读框中编码的亚单位(α，β)(参见， Kalinowski, J. 等,分子微生物学 (1991) 5(5), 1197-1204)。从同源性判断，假定此处测序的基因也具有两个在同一 DNA 链上的同一读框中编码的亚单位(α，β)。

从 DNA 的核苷酸序列推断的野生型 AK 蛋白质α-亚单位的氨基酸序列与 DNA 序列一起在 SEQ ID NO: 4 中显示。单独的氨基酸序列在 SEQ ID NO: 5 中显示。从 DNA 的核苷酸序列推断的野生型 AK 蛋白质β-亚单位的氨基酸序列与 DNA 序列一起在 SEQ ID NO: 6 中显示。单独的氨基酸序列在 SEQ ID NO: 7 中显示。

在每一亚单位中，GTG 用作起始密码子，相应的氨基酸由甲硫氨酸代表。然而，这种代表指甲硫氨酸、缬氨酸或甲酰甲硫氨酸。

另一方面，在突变 *lysC* 序列上的突变指出现氨基酸残基取代， 以便在野生型 AK 蛋白质氨基酸序列中α-亚单位的 279 位丙氨酸残基改变成为苏氨酸残基，β-亚单位的 30 位丙氨酸残基改变成为苏氨酸残基(SEQ ID NO: 5、7)。

实施例 2：从短杆菌属制备 lysA

1 制备 lysA 和构建含有 lysA 的质粒

乳发酵短杆菌野生型菌株 ATCC 13869 用作染色体 DNA 供体。按照普通方法从 ATCC 13869 菌株制备染色体 DNA。按照 PCR 从染色体 DNA 扩增包含 *argS*、*lysA* 和操纵子的启动子的 DNA 片段。就用于扩增的 DNA 引物而言，为了扩增编码精氨酰-tRNA 合酶和 DDC 的 3.6 kb 的区，基于谷氨酸棒杆菌已知的序列(参见分子微生物学，4(11)，1819-1830 (1990)；分子和普通遗传学，212，112-119 (1988))，使用 23 链节(分别具有序列表中 SEQ ID NO: 8 和 9 所示的核苷酸序列)的合成 DNA，DNA 合成和 PCR 以与实施例 1 所描述的相同的方式进行。将 pHSG399 用作扩增 3,579 bp 的基因片段的克隆载体，用限制酶 SmaI(由 Takara Shuzo 生产) 消化 pHSG399，该质粒与包含扩增的 *lysA* 的 DNA 片段连接。具有源于 ATCC13869 的 *lysA* 的如上所述获得的质粒命名为 p399*LYSA*。

通过以 KpnI(由 Takara Shuzo 生产)和 BamHI(由 Takara Shuzo 生产)消化 p399*LYSA* 提取包含 *lysA* 的 DNA 片段。将这一片段与 pHSG299(已由 KpnI 和 BamHI 消化)。构建 p399*LYSA* 的方法在图 2 中显示。

将 Brevi.-ori 引入到制备的 p399*LYSA* 中以构建在棒状杆菌中可自主复制的携带 *lysA* 的质粒。用限制酶 KpnI 和 BamHI 消化 pHK4， 使切开的末端平端化。按照指定的方法用平端化试剂盒(由 Takara Shuzo 生产)进行平端的形成。形成平端后，连接上磷酸化 KpnI 接头(由 Takara Shuzo 生产)进行修饰，使相应于 Brevi.-ori 部分的 DNA 片段可以仅经 KpnI 消化从 pHK4 切下。用 KpnI 消化这一质粒，将所产生的 Brevi.-ori DNA 片段与 p299LYSA(也已用 KpnI 消化)连接，以制备含有在棒状杆菌中可自主复制的 *lysA* 基因的质粒。所制备的质粒命名为 pLYSAB。pLYSAB 的构建方法在图 3 中显示。

2 测定乳发酵短杆菌 lysA 的核苷酸序列

制备质粒 p299LYSA 的 DNA，以与实施例 1 中描述的相同的方式测定其核苷酸序列。所测定的核苷酸序列和从该核苷酸序列推断的氨基酸序列在 SEQ ID NO: 10 中显示，有关核苷酸序列，由 *argS* 编码的氨基酸序列和由 *lysA* 编码的氨基酸序列分别在 SEQ ID NO: 11 和 12 中显示。

实施例 3：从短杆菌属制备 ppc

1 制备 ppc

乳发酵短杆菌野生型菌株 ATCC 13869 用作染色体 DNA 供体。按照普通方法从 ATCC 13869 菌株制备染色体 DNA。按照 PCR 从染色体 DNA 扩增包含 *ppc* 的 DNA 片段。就用于扩增的 DNA 引物而言，为了扩增编码 PEPC 的 3.3 kb 的区，基于谷氨酸棒杆菌已知的序列(参见 O'Regan, M.等, 基因, 77, 237-251 (1989))，使用 23 链节的分别具有序列表中 SEQ ID NO: 13 和 14 所示的核苷酸序列的合成 DNA，DNA 合成和 PCR 以与实施例 1 所描述的相同的方式进行。

由琼脂糖凝胶电泳确认 3, 300 bp 的扩增的基因片段。之后，用普通的方法纯化从凝胶上抽提的片段，用限制酶 SalI(由 Takara Shuzo 生产)消化该片段。将 pHSG399 用作 *ppc* 的克隆载体。以限制酶 SalI (由 Takara Shuzo 生产)消化 pHSG399，将其与包含扩增的 DNA 片段的质粒连接，具有来源于 ATCC 13869 的 *ppc* 的如上所述的获得的质粒命名为 *ppc*F。

2 将 ppc 与具有 lysC 启动子的基因连接

用限制酶 DraI (由 Takara Shuzo 生产)消化如上所述的所获得的 *ppc*F。在除去 PEPC 结构基因上游的约 150 bp 的 DNA 片段后，进行自连接，以获得质粒 ppcFds。以限制酶 SalI (由 Takara Shuzo 生产)消化 ppcFds，使切开的末端平端化。按照指定的方法经采用 DNA 平端化试剂盒(由 Takara Shuzo 生产)完成平端形成。

用限制酶 ApaLI 和 PstI(两者都由 Takara Shuzo 生产)消化实施例 1 中所获得的含有野生型 *lysC* 的 p399AKYB。以与以上类似的方式使切开的末端平端化。在所获得的两个 DNA 片段中的较小的一个含有 Brevi.-ori 和 *lysC* 启动子。将这一片段与以上提到的片段(通过用 SalI 消化 ppcFds 并经采用 DNA 连接试剂盒(由 Takara Shuzo 生产)平端化获得)连接。

用电脉冲法(Sugimoto 等，日本专利申请公开 2-207791)将在连接溶液中的

DNA 引入到乳发酵短杆菌 ATCC 13969 中。在含有 5 μg/ml 氯霉素的完全培养基上选择转化体。从转化体收集质粒 DNA，用 EcoRI 消化以获得其中 *lysC* 启动子以正常的方向与 *ppc* 的结构基因连接的质粒。所获得的质粒命名为 pAKPFds 。 pAKPFds 的构建方法在图 4 中显示。与 *lysC* 启动子连接的 *ppc* 此后称为"野生型高表达 *ppc*"。

3 将野生型高表达 *ppc* 插入到载体中

以上获得的野生型高表达 *ppc* 经 PCR 扩增，以便将其插入到具有在棒状杆菌中可自主复制的复制起点而不是 Brevi.-ori 的载体中。就 DNA 引物而言，其是相应于在谷氨酸棒杆菌已知的 *lysC* 序列(参见分子微生物学, 5(5)， 1197-1204(1991)； Mel. Gen. Genet.， 224， 317-324(1990))基础上合成的 *lysC* 启动子部分的寡核苷酸(SEQ ID NO: 7)，和相应于在谷氨酸棒杆菌已知的 *ppc* 序列(参见 O'Regan, M.等, 基因, 77, 237-251 (1989))基础上合成的 *ppc* 部分的寡核苷酸(SEQ ID NO: 8)。设计这些引物，以便可以扩增含有野生型高表达 *ppc* 的约 3,150 bp 的片段，并且可以用限制酶 KpnI 消化最终的扩增 DNA 片段。 DNA 合成和 PCR 可以以与实施例 1 所描述的相同的方式进行。

作为用于把野生型高表达 *ppc* 引入棒状杆菌的载体，使用新构建的棒状杆菌的克隆载体 pVK7 。如以下描述通过将大肠杆菌的载体 pHSG299(Km^r；Takeshita， S.等，基因， 61， 63-74(1987))与乳发酵短杆菌的隐蔽性质粒 pAM330 构建 pVK7 。以形成一个切割位点的限制酶 AvaII(由 Takara Shuzo 生产)消化 pHSG299，用 T4 DNA 聚合酶平端化，并与已用 HindIII (由 Takara Shuzo 生产)消化和用 T4 DNA 聚合酶平端化的 pAM330 连接。根据在 pHSG299 中所插入的 pAM330 的方向，两个获得的质粒命名为 pVK6 和 pVK7， pVK7 用于下列实验中。 pVK7 是在大肠杆菌和乳发酵短杆菌两者中可自主复制的，并且具有来源于 pHSG299 和 lacZ'的多克隆位点。 pVKG 和 pVK7 的构建方法在图 5 中显示。

由琼脂糖凝胶电泳确认 3,150 bp 的扩增的基因片段。之后，用普通的方法纯化从凝胶上切下的片段，用限制酶 KpnI (由 Takara Shuzo 生产)消化该片段。将该 DNA 片段与已由限制酶 KPnI 消化的 pVK7 连接。所制备的质粒命名为 pPwm 。 pPwm 的构建方法在图 6 中显示。

实施例 4： 制备包含组合突变的 lysC 和 lysA 的质粒

从含有突变 *lysC* 和 Brevi.-ori 的质粒 p399AK9B 和含有 *lysA* 的质粒 p299LYSA 制备含有 *lysC*、*lysA* 和棒状杆菌的复制起点的质粒。以限制酶 BamHI 和 KpnI (两者都由 Takara Shuzo 生产)消化 p299LYSA，并平端化。按照指定的方法用 DNA 平端化试剂盒(由 Takara Shuzo 生产) 进行平端形成。将所获得的 DNA 片段与 p399AK9B(已由 SalI 消化并平端化)连接。由此制备在棒状杆菌中可自主复制的含有突变 *lysC* 和 *lysA* 的质粒。该质粒命名为 pCL。pCL 的构建方法在图 7 中显示。

比较实施例 1：从乳发酵短杆菌制备 dapA、dapB 和 ddh

按如下方法获得 L-赖氨酸生物合成相关的非 *lysC*，*lysA* 和 *ppc* 的基因：dapA(二氢二吡啶甲酸合酶基因)、*dapB*(二氢二吡啶甲酸还原酶基因)和 *ddh*(二氨基庚二酸脱氢酶基因)。

1 制备 dapA 和构建含有 dapA 的质粒

乳发酵短杆菌野生型菌株 ATCC 13869 用作染色体 DNA 供体。按照普通方法从 ATCC 13869 菌株制备染色体 DNA。按照 PCR 从染色体 DNA 扩增包含 *dapA* 的 DNA 片段。就用于扩增的 DNA 引物而言，为了扩增编码 DDPS 的 1.5 kb 的区，基于谷氨酸棒杆菌已知的序列(核酸研究，18(21)，6421(1990)；EMBL 保藏号 X53993)合成 23 链节的分别具有序列表中 SEQ ID NO: 21 和 22 所示的核苷酸序列的 DNA，DNA 合成和 PCR 以与实施例 1 所描述的相同的方式进行。将 pCR-1000(由 Invitrogen 生产，参见生物技术，9，657-663(1991)) 用作扩增 1,411 bp 的基因片段的克隆载体，将其与扩增的 *dapA* 片段连接。DNA 的连接按照指定的方法用 DNA 连接试剂盒进行。由此构建质粒，其中从乳发酵短杆菌染色体扩增的 1,411 bp *dapA* 片段与 pCR-1000 连接。具有源于 ATCC13869 的 *dapA* 的如上所述获得的质粒命名为 pCRDAPA。

基于布达佩斯条约，通过把 pCRDAPA 引入大肠杆菌 JM109 菌株获得的转化菌株 AJ13106 已从 1995 年 5 月 26 日以保藏号 FERM BP-5113 国际保藏在国际贸易和工业部工业科学和技术机构的国立生命科学和人体技术研究所(1-3, Higashi L-chome, Tsukuba-shi, Ibaraki-ken, 305, 日本)。

将 Brevi.-ori 引入到制备的 pCRDAPA 中以构建在棒状杆菌中可自主复制的携带 *dapA* 的质粒。用限制酶 KpnI 和 BamHI(由 Takara Shuzo 生产)消化 pHK4，

使切开的末端平端化。按照指定的方法用平端化试剂盒(由 Takara Shuzo 生产)进行平端的形成。形成平端后，连接上磷酸化 SmaI 接头(由 Takara Shuzo 生产)进行修饰，使相应于 Brevi.-ori 部分的 DNA 片段可以仅经 SmaI 消化从 pHK4 切下。用 SmaI 消化这一质粒，将所产生的 Brevi.-ori DNA 片段与 pCRDAPA(也已用 SmaI 消化)连接，以制备含有在棒状杆菌中可自主复制的 *dapA* 基因的质粒。所制备的质粒命名为 pDPSB。pDPSB(Km^r)的构建方法在图 8 中显示。

2 制备 dapB 和构建含有 dapB 的质粒

乳发酵短杆菌野生型菌株 ATCC 13869 用作染色体 DNA 供体。按照普通方法从 ATCC 13869 菌株制备染色体 DNA。按照 PCR 从染色体 DNA 扩增包含 *dapB* 的 DNA 片段。就用于扩增的 DNA 引物而言，为了扩增编码 DDPR 的 2.0 kb 的区，基于乳发酵短杆菌已知的序列(参见细菌学杂志，175(9)，2743-2749 (1993))分别合成 23 链节(分别具有序列表中 SEQ ID NO:*19* 和 *20* 所示的核苷酸序列)的 DNA，DNA 合成和 PCR 以与实施例 1 所描述的相同的方式进行。将 pCR-Script(由 Invitrogen 生产)用作扩增 2,001 bp 的基因片段的克隆载体，将其与扩增的 *dapB* 片段连接。由此构建质粒，其中从乳发酵短杆菌染色体扩增的 2,001 bp *dapB* 片段与 pCR-Script 连接。具有源于 ATCC13869 的 *dapB* 的如上所述获得的质粒命名为 pCRDAPB。基于布达佩斯条约，通过把 pCRDAPB 引入大肠杆菌 JM109 菌株获得的转化菌株 AJ13107 已从 1995 年 5 月 26 日以保藏号 FERM BP-5114 国际保藏在国际贸易和工业部工业科学和技术机构的国立生命科学和人体技术研究所(1-3, Higashi L-chome, Tsukuba-shi, Ibaraki-ken, 305, 日本)。

通过以 EcoRV 和 SphI 消化 pCRDAPB 提取包含 DDPR 结构基因的 1,101 bp 片段。将这一片段与 pHSG399(已由 HincII 和 SghI 消化)以制备质粒。该制备的质粒命名为 p399DPR。

将 Brevi.-ori 引入到制备的 p399DPR 中以构建在棒状杆菌中可自主复制的携带 *dapB* 的质粒。用限制酶 KpnI(由 Takara Shuzo 生产)消化 pHK4，使切开的末端平端化。按照指定的方法用平端化试剂盒(由 Takara Shuzo 生产)进行平端的形成。形成平端后，连接上磷酸化 BamHI 接头(由 Takara Shuzo 生产)进行修饰，使相应于 Brevi.-ori 部分的 DNA 片段可以仅经 BamHI 消化从 pHK4 切下。用 BamHI 消化这一质粒，将所产生的 Brevi.-ori DNA 片段与

p399DPR(也已用 BamHI 消化)连接，以制备含有在棒状杆菌中可自主复制的 *dapB* 基因的质粒。所制备的质粒命名为 pDPRB 。 pDPRB 的构建方法在图 9 中显示。

3 制备 *ddh* 和构建包含 *ddh* 的质粒

按照 PCR 方法采用两种寡核苷酸引物(SEQ ID NO: 23 , 24)从乳发酵短杆菌 ATCC 13869 染色体 DNA *ddh* 基因扩增来获得 *ddh* 基因，所述引物是基于谷氨酸棒杆菌 *ddh* 基因的已知核苷酸序列(Ishino , S.等，核酸研究， 15, 3917(1987))制备的。将所获得的扩增 DNA 片段用 EcoT22I 和 AvaI 消化，使切开末端平端化。之后，将该片段插入到 pMW119 的 SmaI 位点中以便获得质粒 p*DDH* 。

接着，用 SalI 和 EcoRI 消化 p*DDH*，并形成平端。之后，将所获得的片段与已由 SmaI 消化的 pUC18 连接。这样获得的质粒命名为 pUC18DDH 。

将 Brevi.-ori 引入 pUC18DDH，以构建在棒状杆菌中可自主复制的携带 *ddh* 的质粒。以限制酶 KpnI 和 BamHI 消化 pHK4 。使切开的末端平端化。按照指定的方法用平端化试剂盒(由 Takara Shuzo 生产)进行平端的形成。形成平端后，连接上磷酸化 PstI 接头(由 Takara Shuzo 生产)，以便将其插入到 pHSG299 的 PstI 位点中。如以上所述构建的质粒命名为 pP4 。接着，以 XbaI 和 KpnI 消化 pUC18DDH , 将所产生的片段与已由 KpnI 和 XbaI 消化的 pPK4 连接。这样，构建了在棒状杆菌中可自主复制的包含 *ddh* 的质粒。这一质粒命名为 pPK4D 。pPK4D 的构建方法在图 10 中显示。

比较实施例 2： 构建含有组合的突变 lysC 和 dapA 、 dapB 或 ddh 的质粒

1 构建含有突变 lysC 和 dapA 的质粒

从包含 *dapA* 的质粒 pCRDAPA 和包含突变 *lysC* 和 Brevi.-ori 的质粒 p399AK9B 构建包含突变 *lysC* 、 *dapA* 和棒状杆菌复制源的质粒。以 SalI 完全消化 p399AK9B , 然后平端化，并与 EcoRI 接头连接，以构建其中 SalI 位点修饰成 EcoRI 位点的质粒。所获得的质粒命名为 p399AK9BSE 。通过以 EcoRI 部分消化 p399AK9BSE 作为一个片段切下突变 *lysC* 和 Brevi.-ori 。将这一片段与已用 EcoRI 消化的 pCRDAPA 连接。所获得的质粒命名为 pCRCAB 。这一质粒在大肠杆菌和棒状杆菌中可自主复制，其使宿主具有对卡那霉素的抗性，该质

粒包含组合的突变 *lysC*和 *dapA*。pCRCAB 的构建方法在图 11 中显示。

2 构建包含组合的突变 lysC 和 dapB 的质粒

从具有突变 *lysC*的质粒 p399AK9 和具有 *dapB* 的质粒 p399DPR 构建包含突变 *lysC*和 *dapB* 的质粒。通过以 EcoRV 和 SphI 消化 p399DPR 提取包含 DDPR 结构基因的 1,101 bp 的片段。将这一片段与已用 SalI 消化，然后平端化，并进一步由 SphI 消化的 p399AK9 连接，以构建包含组合的突变 *lysC*和 *dapB*的质粒。这一质粒命名为 p399AKDDPR。

其次，将 Brevi.-ori 引入所获得的 p399AKDDPR。以限制酶 KpnI (由 Takara Shuzo 生产)消化包含 Brevi.-ori 的质粒 pHK4，使切下的边缘平端化。按照指定的方法用平端化试剂盒(由 Takara Shuzo 生产)进行平端的形成。形成平端后，连接上磷酸化 BamHI 接头(由 Takara Shuzo 生产)进行修饰，使相应于 Brevi.-ori 部分的 DNA 片段可以仅经 BamHI 消化从 pHK4 切下。用 BamHI 消化这一质粒，将所产生的 Brevi.-ori DNA 片段与 p399AKDDPR(也已用 BamHI 消化)连接，以构建在棒状杆菌中可自主复制的含有突变 *lysC*和 *dapB* 的质粒。构建的质粒命名为 pCB。pCB 的构建方法在图 12 中显示。

3 构建含有组合突变 lysC 和 ddh 的质粒

自包含 *ddh* 的质粒 pUC18DDH 和包含突变 *lysC* 和 Brevi.-ori 的质粒 p399AK9B 制备包含突变 *lysC*、*ddh* 的和棒状杆菌的复制起点的质粒。 以限制酶 EcoRI(由 Takara Shuzo 生产)消化 pUC18DDH，并平端化，将其在末端与 SalI 多接头连接，以便将 EcoRIJ 位点改变成 SalI 位点。用 SalI 消化所获得的质粒，以获得含有 *ddh* 的 DNA 片段。

然后，用限制酶 SalI 消化 p399AK9B，将其与包含 *ddh* 的 DNA 片段连接。这样，由此制备在棒状杆菌中可自主复制的包含突变 *lysC*、*ddh* 和 Brevi.-ori 的质粒，该质粒命名为 pCD。pCD 的构建方法在图 13 中显示。

实施例 5：将包含 L-赖氨酸生物合成基因的质粒引入到乳发酵短杆菌 L-赖氨酸-产生菌中

将如上所述构建的质粒分别引入乳发酵短杆菌 L-赖氨酸-产生菌 AJ11082(NRRL B-11470)中，所述质粒即 p399AK9B(Cm^r)、pLYSAB(Cm^r)、pPwm(Km^r)、pCRCAB(Km^r)、pCB(Cm^r)、pCD(Cm^r)和 pCL(Cm^r)。AJ11082 菌株具有 AEC 抗性。按照电脉冲方法引入质粒(Sugimoto 等，日本专利申请公开 2-

207791)，基于质粒具有的药物抗性标记选择转化体，当包含氯霉素抗性基因的质粒被引入时，在包含 5 μg/ml 氯霉素完全培养基上选择转化体，当包含卡那霉素抗性基因的质粒被引入时，在包含 25 μg/ml 卡那霉素完全培养基上选择转化体。

向所获得的转化体中的突变 *lysC* 和 *lysA* 被增强的菌株中引入 pPwm(Km^r)，以获得其中三种突变 *lysC*、*lysA* 和 *ppc* 被增强的菌株(AJ11082/pCL/pPwm)。在含有 5μg/ml 氯霉素和 25μg/ml 卡那霉素的完全培养基上选择转化体。

实施例 6：L-赖氨酸的产生

在 L-赖氨酸-产生培养基中培养实施例 5 中所获得的各种转化体，以估价 L-赖氨酸产量。L-赖氨酸-产生培养基具有下列组成。

[L-赖氨酸-产生培养基]

除了碳酸钙外，溶解下列组分(在 1 升中)，用 KOH 调节 pH 为 8.0。于 115℃灭菌 15 分钟，将已经以干燥状态在热气流中分别灭菌过的碳酸钙(50 g)加入其中。

葡萄糖	130 g
$(NH_4)_2SO_4$	55 g
KH_2PO_4	1 g
$MgSO_4 \cdot 7H_2O$	1 g
生物素	500 μg
硫胺	2000 μg
$FeSO_4 \cdot 7H20$	0.01 g
$MnSO_4 \cdot 7H20$	0.01 g
烟酰胺	5 mg
蛋白质水解物(Mamenou)	30 ml
碳酸钙	50 g

将各种类型的转化体和亲本菌株接种到具有以上组成的培养基中，以进行在 31.5℃下的往复振荡培养。在培养 40 或 72 小时和生长 72 小时(OD_{562})后产生的 L-赖氨酸的量在表 1 中显示。在表中，*lysC** 代表突变 *lysC*。

表 1

在40或72小时培养后积累的L-赖氨酸

细菌菌株/质粒	引入的基因	产生的L-赖氨酸量(g/l)	
		40 h后	72 h 后
AJ 11082		22.0	29.8
AJ 11082/p399AK9B	*lysC**	16.8	34.5
AJ 11082/pLYSAB	*lysA*	19.8	32.5
AJ11082/pPwm	*ppc*	20.7	28.9
AJ 11082/pCRCAB	*lysC**, *dapA*	19.7	36.5
AJ 11082/pCB	*lysC**, *dapB*	23.3	35.0
AJ 11082/pCD	*lysC**, *ddh*	15.0	27.0
AJ 11082/pCL	*lysC**, *lysA*	24.0	44.0
AJ 11082/pCL/pPwm	*lysC**, *lysA*, *ppc*	25.0	45.2

如上所示，当突变 *lysC*、*lysA* 或 *ppc* 单独增强时或者当突变 *lysC* 和 *dapA* 或 *ddh* 一起增强时，培养72 小时后产生的L-赖氨酸的量大于由亲本菌株产生的量，然而，培养 40 小时后产生的 L-赖氨酸的量小于由亲本菌株产生的量。即，在短的培养时间内，L-赖氨酸-产生速度降低。同样地，当突变 *lysC* 和 *ddh* 一起增强时，培养40和72小时后产生的L-赖氨酸的量大于由亲本菌株产生的量。相反，在其中 *dapB* 与突变 *lysC* 一道增强的菌株的情况下，在短的培养时间内 L-赖氨酸产生速度成功地恢复，在长的培养时间内所积累的 L-赖氨酸的量也得到提高。

在其中三种突变 *lysC*、*lysA* 和 *ppc* 同时增强的情况下，L-赖氨酸的产量得到进一步的提高。

序 列 表

(1) 一般信息:

(i) 申请人: AJINOMOTO 有限公司

(ii) 发明名称: 用于产生 L-赖氨酸的方法

(iii) 序列数: 24 个

(iv) 通讯地址:

(A) 收信人:

(B) 街道:

(C) 城市:

(E) 国家:

(F) ZIP :

(v) 计算机可读形式:

(A) 介质类型: 软盘

(B) 计算机: IBM PC 兼容机

(C) 操作系统: PC-DOS/MS-DOS

(D) 软件: Patentln Release#, 版本# 1.30

(vi) 当前申请的数据:

(A) 申请号:

(B) 申请日:

(C) 分类号:

(vii) 在先申请的数据:

(A) 申请号: JP 8-325659

(B) 申请日: 1996 年 12 月 5 日

(viii) 律师/代理人信息:

(A) 姓名:

(B) 登记号:

(ix) 电信信息:

(A) 电话:

(B) 传真:

(2) SEQ ID NO: 1 的信息:

(i) 序列特征:

(A) 长度: 23 个碱基

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其它核酸

(A) 描述: /desc = "合成的 DNA"

(iv) 反义: 否

(xi) 序列描述: SEQ ID NO: 1:

TCGAAGTA GCACCTGTCA CTT 23

(2) SEQ ID NO: 2 的信息:

(i) 序列特征:

(A) 长度: 21 个碱基

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其它核酸

(A) 描述: desc = "合成的 DNA"

(iv) 反义: 是

(xi) 序列描述: SEQ ID NO: 2:

ACGGAATTCA ATCTTACGGC C 21

(2) SEQ ID NO: 3 的信息:

(i) 序列特征:

(A) 长度: 1643 个碱基

(B) 类型: 核酸

(C) 链型: 双链

(D) 拓扑结构: 线性

(ii)分子类型: 基因组 DNA

(vi)原始来源:

(A)有机体: 乳发酵短杆菌

(B)菌株: ATCC 13869

(xi)序列描述: SEQ ID NO: 3:

```
TCGCGAAGTA GCACCTGTCA CTTTTGTCTC AAATATTAAA TCGAATATCA ATATACGGTC     60
TGTTTATTGG AACGCATCCC AGTGGCTGAG ACGCATCCGC TAAAGCCCCA GGAACCCTGT    120
GCAGAAAGAA AACACTCCTC TGGCTAGGTA GACACAGTTT ATAAAGGTAG AGTTGAGCGG    180
GTAACTGTCA GCACGTAGAT CGAAAGGTGC ACAAAGGTGG CCCTGGTCGT ACAGAAATAT    240
GGCGGTTCCT CGCTTGAGAG TGCGGAACGC ATTAGAAACG TCGCTGAACG GATCGTTGCC    300
ACCAAGAAGG CTGGAAATGA TGTCGTGGTT GTCTGCTCCG CAATGGGAGA CACCACGGAT    360
GAACTTCTAG AACTTGCAGC GGCAGTGAAT CCCGTTCCGC CAGCTCGTGA AATGGATATG    420
CTCCTGACTG CTGGTGAGCG TATTTCTAAC GCTCTCGTCG CCATGGCTAT TGAGTCCCTT    480
GGCGCAGAAG CTCAATCTTT CACTGGCTCT CAGGCTGGTG TGCTCACCAC CGAGCGCCAC    540
GGAAACGCAC GCATTGTTGA CGTCACACCG GGTCGTGTGC GTGAAGCACT CGATGAGGGC    600
AAGATCTGCA TTGTTGCTGG TTTTCAGGGT GTTAATAAAG AAACCCGCGA TGTCACCACG    660
TTGGGTCGTG GTGGTTCTGA CACCACTGCA GTTGCGTTGG CAGCTGCTTT GAACGCTGAT    720
GTGTGTGAGA TTTACTCGGA CGTTGACGGT GTGTATACCG CTGACCCGCG CATCGTTCCT    780
AATGCACAGA AGCTGGAAAA GCTCAGCTTC GAAGAAATGC TGGAACTTGC TGCTGTTGGC    840
TCCAAGATTT TGGTGCTGCG CAGTGTTGAA TACGCTCGTG CATTCAATGT GCCACTTCGC    900
GTACGCTCGT CTTATAGTAA TGATCCCGGC ACTTTGATTG CCGGCTCTAT GGAGGATATT    960
CCTGTGGAAG AAGCAGTCCT TACCGGTGTC GCAACCGACA AGTCCGAAGC CAAAGTAACC   1020
GTTCTGGGTA TTTCCGATAA GCCAGGCGAG GCTGCCAAGG TTTTCCGTGC GTTGGCTGAT   1080
GCAGAAATCA ACATTGACAT GGTTCTGCAG AACGTCTCCT CTGTGGAAGA CGGCACCACC   1140
GACATCACGT TCACCTGCCC TCGCGCTGAC GGACGCCGTG CGATGGAGAT CTTGAAGAAG   1200
CTTCAGGTTC AGGGCAACTG GACCAATGTG CTTTACGACG ACCAGGTCGG CAAAGTCTCC   1260
CTCGTGGGTG CTGGCATGAA GTCTCACCCA GGTGTTACCG CAGAGTTCAT GGAAGCTCTG   1320
CGCGATGTCA ACGTGAACAT CGAATTGATT TCCACCTCTG AGATCCGCAT TTCCGTGCTG   1380
ATCCGTGAAG ATGATCTGGA TGCTGCTGCA CGTGCATTGC ATGAGCAGTT CCAGCTGGGC   1440
GGCGAAGACG AAGCCGTCGT TTATGCAGGC ACCGGACGCT AAAGTTTTAA AGGAGTAGTT   1500
TTACAATGAC CACCATCGCA GTTGTTGGTG CAACCGGCCA GGTCGGCCAG GTTATGCGCA   1560
CCCTTTTGGA AGAGCGCAAT TTCCCAGCTG ACACTGTTCG TTTCTTTGCT TCCCCGCGTT   1620
CCGCAGGCCG TAAGATTGAA TTC                                           1643
```

(2) SEQ ID NO: 4 的信息:

(i)序列特征:

(A)长度: 1643 个碱基

(B)类型: 核酸

(C)链型: 双链

(D)拓扑结构: 线性

(ii)分子类型: 基因组 DNA

(vi)原始来源:

(A)有机体: 乳发酵短杆菌

(B)菌株: ATCC 13869

(ix)特征:

(A)名称/关键词: CDS

(B)位置: 217..1482

(xi)序列描述: SEQ ID NO: 4:

5

```
TCGCGAAGTA GCACCTGTCA CTTTTGTCTC AAATATTAAA TCGAATATCA ATATACGGTC      60
TGTTTATTGG AACGCATCCC AGTGGCTGAG ACGCATCCGC TAAAGCCCCA GGAACCCTGT     120
GCAGAAAGAA AACACTCCTC TGGCTAGGTA GACACAGTTT ATAAAGGTAG AGTTGAGCGG     180
GTAACTGTCA GCACGTAGAT CGAAAGGTGC ACAAAG GTG GCC CTG GTC GTA CAG       234
                                        Met Ala Leu Val Val Gln
                                          1               5
```

```
AAA TAT GGC GGT TCC TCG CTT GAG AGT GCG GAA CGC ATT AGA AAC GTC      282
Lys Tyr Gly Gly Ser Ser Leu Glu Ser Ala Glu Arg Ile Arg Asn Val
                10                  15                  20
GCT GAA CGG ATC GTT GCC ACC AAG AAG GCT GGA AAT GAT GTC GTG GTT      330
Ala Glu Arg Ile Val Ala Thr Lys Lys Ala Gly Asn Asp Val Val Val
            25                  30                  35
GTC TGC TCC GCA ATG GGA GAC ACC ACG GAT GAA CTT CTA GAA CTT GCA      378
Val Cys Ser Ala Met Gly Asp Thr Thr Asp Glu Leu Leu Glu Leu Ala
        40                  45                  50
GCG GCA GTG AAT CCC GTT CCG CCA GCT CGT GAA ATG GAT ATG CTC CTG      426
Ala Ala Val Asn Pro Val Pro Pro Ala Arg Glu Met Asp Met Leu Leu
    55                  60                  65                  70
ACT GCT GGT GAG CGT ATT TCT AAC GCT CTC GTC GCC ATG GCT ATT GAG      474
Thr Ala Gly Glu Arg Ile Ser Asn Ala Leu Val Ala Met Ala Ile Glu
                75                  80                  85
TCC CTT GGC GCA GAA GCT CAA TCT TTC ACT GGC TCT CAG GCT GGT GTG      522
Ser Leu Gly Ala Glu Ala Gln Ser Phe Thr Gly Ser Gln Ala Gly Val
            90                  95                  100
CTC ACC ACC GAG CGC CAC GGA AAC GCA CGC ATT GTT GAC GTC ACA CCG      570
Leu Thr Thr Glu Arg His Gly Asn Ala Arg Ile Val Asp Val Thr Pro
        105                 110                 115
GGT CGT GTG CGT GAA GCA CTC GAT GAG GGC AAG ATC TGC ATT GTT GCT      618
Gly Arg Val Arg Glu Ala Leu Asp Glu Gly Lys Ile Cys Ile Val Ala
    120                 125                 130
GGT TTT CAG GGT GTT AAT AAA GAA ACC CGC GAT GTC ACC ACG TTG GGT      666
Gly Phe Gln Gly Val Asn Lys Glu Thr Arg Asp Val Thr Thr Leu Gly
135                 140                 145                 150
CGT GGT GGT TCT GAC ACC ACT GCA GTT GCG TTG GCA GCT GCT TTG AAC      714
Arg Gly Gly Ser Asp Thr Thr Ala Val Ala Leu Ala Ala Ala Leu Asn
                155                 160                 165
GCT GAT GTG TGT GAG ATT TAC TCG GAC GTT GAC GGT GTG TAT ACC GCT      762
Ala Asp Val Cys Glu Ile Tyr Ser Asp Val Asp Gly Val Tyr Thr Ala
            170                 175                 180
GAC CCG CGC ATC GTT CCT AAT GCA CAG AAG CTG GAA AAG CTC AGC TTC      810
Asp Pro Arg Ile Val Pro Asn Ala Gln Lys Leu Glu Lys Leu Ser Phe
        185                 190                 195
GAA GAA ATG CTG GAA CTT GCT GCT GTT GGC TCC AAG ATT TTG GTG CTG      858
Glu Glu Met Leu Glu Leu Ala Ala Val Gly Ser Lys Ile Leu Val Leu
    200                 205                 210
CGC AGT GTT GAA TAC GCT CGT GCA TTC AAT GTG CCA CTT CGC GTA CGC      906
Arg Ser Val Glu Tyr Ala Arg Ala Phe Asn Val Pro Leu Arg Val Arg
215                 220                 225                 230

TCG TCT TAT AGT AAT GAT CCC GGC ACT TTG ATT GCC GGC TCT ATG GAG      954
Ser Ser Tyr Ser Asn Asp Pro Gly Thr Leu Ile Ala Gly Ser Met Glu
                235                 240                 245
GAT ATT CCT GTG GAA GAA GCA GTC CTT ACC GGT GTC GCA ACC GAC AAG     1002
Asp Ile Pro Val Glu Glu Ala Val Leu Thr Gly Val Ala Thr Asp Lys
            250                 255                 260
TCC GAA GCC AAA GTA ACC GTT CTG GGT ATT TCC GAT AAG CCA GGC GAG     1050
Ser Glu Ala Lys Val Thr Val Leu Gly Ile Ser Asp Lys Pro Gly Glu
        265                 270                 275
GCT GCC AAG GTT TTC CGT GCG TTG GCT GAT GCA GAA ATC AAC ATT GAC     1098
Ala Ala Lys Val Phe Arg Ala Leu Ala Asp Ala Glu Ile Asn Ile Asp
    280                 285                 290
ATG GTT CTG CAG AAC GTC TCC TCT GTG GAA GAC GGC ACC ACC GAC ATC     1146
Met Val Leu Gln Asn Val Ser Ser Val Glu Asp Gly Thr Thr Asp Ile
295                 300                 305                 310
```

```
ACG TTC ACC TGC CCT CGC GCT GAC GGA CGC CGT GCG ATG GAG ATC TTG      1194
Thr Phe Thr Cys Pro Arg Ala Asp Gly Arg Arg Ala Met Glu Ile Leu
                315                 320                 325
AAG AAG CTT CAG GTT CAG GGC AAC TGG ACC AAT GTG CTT TAC GAC GAC      1242
Lys Lys Leu Gln Val Gln Gly Asn Trp Thr Asn Val Leu Tyr Asp Asp
            330                 335                 340
CAG GTC GGC AAA GTC TCC CTC GTG GGT GCT GGC ATG AAG TCT CAC CCA      1290
Gln Val Gly Lys Val Ser Leu Val Gly Ala Gly Met Lys Ser His Pro
        345                 350                 355
GGT GTT ACC GCA GAG TTC ATG GAA GCT CTG CGC GAT GTC AAC GTG AAC      1338
Gly Val Thr Ala Glu Phe Met Glu Ala Leu Arg Asp Val Asn Val Asn
    360                 365                 370
ATC GAA TTG ATT TCC ACC TCT GAG ATC CGC ATT TCC GTG CTG ATC CGT      1386
Ile Glu Leu Ile Ser Thr Ser Glu Ile Arg Ile Ser Val Leu Ile Arg
375                 380                 385                 390
GAA GAT GAT CTG GAT GCT GCT GCA CGT GCA TTG CAT GAG CAG TTC CAG      1434
Glu Asp Asp Leu Asp Ala Ala Ala Arg Ala Leu His Glu Gln Phe Gln
                395                 400                 405
CTG GGC GGC GAA GAC GAA GCC GTC GTT TAT GCA GGC ACC GGA CGC TAA      1482
Leu Gly Gly Glu Asp Glu Ala Val Val Tyr Ala Gly Thr Gly Arg
            410                 415                 420
AGTTTTAAAG GAGTAGTTTT ACAATGACCA CCATCGCAGT TGTTGGTGCA ACCGGCCAGG   1542
TCGGCCAGGT TATGCGCACC CTTTTGGAAG AGCGCAATTT CCCAGCTGAC ACTGTTCGTT   1602
TCTTTGCTTC CCCGCGTTCC GCAGGCCGTA AGATTGAATT C                       1643
```

(2) SEQ ID NO: 5 的信息:

(i) 序列特征:

(A) 长度: 421 个氨基酸

(B) 类型: 氨基酸

(D) 拓扑结构: 线性

(ii) 分子类型: 蛋白质

(xi) 序列描述: SEQ ID NO: 3:

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

```
Leu Ala Ala Ala Leu Asn Ala Asp Val Cys Glu Ile Tyr Ser Asp Val
                165                 170                 175
Asp Gly Val Tyr Thr Ala Asp Pro Arg Ile Val Pro Asn Ala Gln Lys
            180                 185                 190
Leu Glu Lys Leu Ser Phe Glu Glu Met Leu Glu Leu Ala Ala Val Gly
        195                 200                 205

Ser Lys Ile Leu Val Leu Arg Ser Val Glu Tyr Ala Arg Ala Phe Asn
    210                 215                 220
Val Pro Leu Arg Val Arg Ser Ser Tyr Ser Asn Asp Pro Gly Thr Leu
225                 230                 235                 240
Ile Ala Gly Ser Met Glu Asp Ile Pro Val Glu Glu Ala Val Leu Thr
                245                 250                 255
Gly Val Ala Thr Asp Lys Ser Glu Ala Lys Val Thr Val Leu Gly Ile
            260                 265                 270
Ser Asp Lys Pro Gly Glu Ala Ala Lys Val Phe Arg Ala Leu Ala Asp
        275                 280                 285
Ala Glu Ile Asn Ile Asp Met Val Leu Gln Asn Val Ser Ser Val Glu
    290                 295                 300
Asp Gly Thr Thr Asp Ile Thr Phe Thr Cys Pro Arg Ala Asp Gly Arg
305                 310                 315                 320
Arg Ala Met Glu Ile Leu Lys Lys Leu Gln Val Gln Gly Asn Trp Thr
                325                 330                 335
Asn Val Leu Tyr Asp Asp Gln Val Gly Lys Val Ser Leu Val Gly Ala
            340                 345                 350
Gly Met Lys Ser His Pro Gly Val Thr Ala Glu Phe Met Glu Ala Leu
        355                 360                 365
Arg Asp Val Asn Val Asn Ile Glu Leu Ile Ser Thr Ser Glu Ile Arg
    370                 375                 380
Ile Ser Val Leu Ile Arg Glu Asp Asp Leu Asp Ala Ala Ala Arg Ala
385                 390                 395                 400
Leu His Glu Gln Phe Gln Leu Gly Gly Glu Asp Glu Ala Val Val Tyr
                405                 410                 415
Ala Gly Thr Gly Arg
            420
```

(2) SEQ ID NO: 6 的信息:

(i) 序列特征:

(A) 长度: 1643 个碱基

(B) 类型: 核酸

(C) 链型: 双链

(D) 拓扑结构: 线性

(ii) 分子类型: 基因组 DNA

(vi) 原始来源:

(A) 有机体: 乳发酵短杆菌

(B) 菌株: ATCC 13869

(ix) 特征:

(A)名称/关键词: CDS

(B)位置: 964..1482

(xi)序列描述: SEQ ID NO: 6:

```
TCGCGAAGTA GCACCTGTCA CTTTTGTCTC AAATATTAAA TCGAATATCA ATATACGGTC      60
TGTTTATTGG AACGCATCCC AGTGGCTGAG ACGCATCCGC TAAAGCCCCA GGAACCCTGT     120
GCAGAAAGAA AACACTCCTC TGGCTAGGTA GACACAGTTT ATAAAGGTAG AGTTGAGCGG     180
GTAACTGTCA GCACGTAGAT CGAAAGGTGC ACAAAGGTGG CCCTGGTCGT ACAGAAATAT     240
GGCGGTTCCT CGCTTGAGAG TGCGGAACGC ATTAGAAACG TCGCTGAACG GATCGTTGCC     300
ACCAAGAAGG CTGGAAATGA TGTCGTGGTT GTCTGCTCCG CAATGGGAGA CACCACGGAT     360
GAACTTCTAG AACTTGCAGC GGCAGTGAAT CCCGTTCCGC CAGCTCGTGA AATGGATATG     420
CTCCTGACTG CTGGTGAGCG TATTTCTAAC GCTCTCGTCG CCATGGCTAT TGAGTCCCTT     480
GGCGCAGAAG CTCAATCTTT CACTGGCTCT CAGGCTGGTG TGCTCACCAC CGAGCGCCAC     540
GGAAACGCAC GCATTGTTGA CGTCACACCG GGTCGTGTGC GTGAAGCACT CGATGAGGGC     600
AAGATCTGCA TTGTTGCTGG TTTTCAGGGT GTTAATAAAG AAACCCGCGA TGTCACCACG     660
TTGGGTCGTG GTGGTTCTGA CACCACTGCA GTTGCGTTGG CAGCTGCTTT GAACGCTGAT     720
GTGTGTGAGA TTTACTCGGA CGTTGACGGT GTGTATACCG CTGACCCGCG CATCGTTCCT     780
AATGCACAGA AGCTGGAAAA GCTCAGCTTC GAAGAAATGC TGGAACTTGC TGCTGTTGGC     840
TCCAAGATTT TGGTGCTGCG CAGTGTTGAA TACGCTCGTG CATTCAATGT GCCACTTCGC     900

GTACGCTCGT CTTATAGTAA TGATCCCGGC ACTTTGATTG CCGGCTCTAT GGAGGATATT     960
CCT GTG GAA GAA GCA GTC CTT ACC GGT GTC GCA ACC GAC AAG TCC GAA       1008
    Met Glu Glu Ala Val Leu Thr Gly Val Ala Thr Asp Lys Ser Glu
      1               5                  10                  15
GCC AAA GTA ACC GTT CTG GGT ATT TCC GAT AAG CCA GGC GAG GCT GCC      1056
Ala Lys Val Thr Val Leu Gly Ile Ser Asp Lys Pro Gly Glu Ala Ala
                20                  25                  30
AAG GTT TTC CGT GCG TTG GCT GAT GCA GAA ATC AAC ATT GAC ATG GTT      1104
Lys Val Phe Arg Ala Leu Ala Asp Ala Glu Ile Asn Ile Asp Met Val
            35                  40                  45
CTG CAG AAC GTC TCC TCT GTG GAA GAC GGC ACC ACC GAC ATC ACG TTC      1152
Leu Gln Asn Val Ser Ser Val Glu Asp Gly Thr Thr Asp Ile Thr Phe
        50                  55                  60
ACC TGC CCT CGC GCT GAC GGA CGC CGT GCG ATG GAG ATC TTG AAG AAG      1200
Thr Cys Pro Arg Ala Asp Gly Arg Arg Ala Met Glu Ile Leu Lys Lys
    65                  70                  75
CTT CAG GTT CAG GGC AAC TGG ACC AAT GTG CTT TAC GAC GAC CAG GTC      1248
Leu Gln Val Gln Gly Asn Trp Thr Asn Val Leu Tyr Asp Asp Gln Val
80                  85                  90                  95
GGC AAA GTC TCC CTC GTG GGT GCT GGC ATG AAG TCT CAC CCA GGT GTT      1296
Gly Lys Val Ser Leu Val Gly Ala Gly Met Lys Ser His Pro Gly Val
                100                 105                 110
ACC GCA GAG TTC ATG GAA GCT CTG CGC GAT GTC AAC GTG AAC ATC GAA      1344
Thr Ala Glu Phe Met Glu Ala Leu Arg Asp Val Asn Val Asn Ile Glu
            115                 120                 125
TTG ATT TCC ACC TCT GAG ATC CGC ATT TCC GTG CTG ATC CGT GAA GAT      1392
Leu Ile Ser Thr Ser Glu Ile Arg Ile Ser Val Leu Ile Arg Glu Asp
        130                 135                 140
GAT CTG GAT GCT GCT GCA CGT GCA TTG CAT GAG CAG TTC CAG CTG GGC      1440
Asp Leu Asp Ala Ala Ala Arg Ala Leu His Glu Gln Phe Gln Leu Gly
    145                 150                 155
GGC GAA GAC GAA GCC GTC GTT TAT GCA GGC ACC GGA CGC TAAAGTTTTAA      1490
Gly Glu Asp Glu Ala Val Val Tyr Ala Gly Thr Gly Arg
160                 165                 170
AGGAGTAGTT TTACAATGAC CACCATCGCA GTTGTTGGTG CAACCGGCCA GGTCGGCCAG    1550
GTTATGCGCA CCCTTTTGGA AGAGCGCAAT TTCCCAGCTG ACACTGTTCG TTTCTTTGCT    1610
TCCCCGCGTT CCGCAGGCCG TAAGATTGAA TTC                                 1643
```

(2) SEQ ID NO: 7 的信息:

(i) 序列特征:

(A) 长度: 172 个氨基酸

(B) 类型: 氨基酸

(D) 拓扑结构: 线性

(ii) 分子类型: 蛋白质

(xi) 序列描述: SEQ ID NO: 7:

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

(2) SEQ ID NO: 8 的信息:

(i) 序列特征:

(A) 长度: 23 个碱基

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其它核酸

(A) 描述: desc = "合成的 DNA"

(iv) 反义: 否

(xi) 序列描述: SEQ ID NO: 8:

GTGGAGCCGA CCATTCCGCG AGG 23

(2) SEQ ID NO: 的信息: 9:

(i)序列特征:

(A)长度: 23 个碱基

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: 其它核酸

(A)描述: desc = "合成的 DNA"

(iv)反义: 是

(xi)序列描述: SEQ ID NO: 9:

CCAAACCGC CCTCCACGGC GAA

(2) SEQ ID NO: 10 的信息:

(i)序列特征:

(A)长度: 3579 个碱基

(B)类型: 核酸

(C)链型: 双链

(D)拓扑结构: 线性

(ii)分子类型: 基因组 DNA

(vi)原始来源:

(A)有机体: 乳发酵短杆菌

(B)菌株: ATCC 13869

(ix)特征:

(A)名称/关键词: CDS

(B)位置: 533..2182

(ix)特征:

(A)名称/关键词: CDS

(B)位置: 2188..3522

(xi)序列描述: SEQ ID NO: 10:

```
GTGGAGCCGA CCATTCCGCG AGGCTGCACT GCAACGAGGT CGTAGTTTTG GTACATGGCT     60
TCTGGCCAGT TCATGGATTG GCTGCCGAAG AAGCTATAGG CATCGCACCA GGGCCACCGA    120
GTTACCGAAG ATGGTGCCGT GCTTTTCGCC TTGGGCAGGG ACCTTGACAA AGCCCACGCT    180
GATATCGCCA AGTGAGGGAT CAGAATAGTG CATGGGCACG TCGATGCTGC CACATTGAGC    240
GGAGGCAATA TCTACCTGAG GTGGGCATTC TTCCCAGCGG ATGTTTTCTT GCGCTGCTGC    300
```

```
AGTGGGCATT GATACCAAAA AGGGGCTAAG CGCAGTCGAG GCGGCAAGAA CTGCTACTAC    360
CCTTTTTATT GTCGAACGGG GCATTACGGC TCCAAGGACG TTTGTTTTCT GGGTCAGTTA    420
CCCCAAAAAG CATATACAGA GACCAATGAT TTTTCATTAA AAAGGCAGGG ATTTGTTATA    480
AGTATGGGTC GTATTCTGTG CGACGGGTGT ACCTCGGCTA GAATTTCTCC CC ATG        535
                                                              Met
                                                               1
ACA CCA GCT GAT CTC GCA ACA TTG ATT AAA GAG ACC GCG GTA GAG GTT      583
Thr Pro Ala Asp Leu Ala Thr Leu Ile Lys Glu Thr Ala Val Glu Val
                5                  10                  15
TTG ACC TCC CGC GAG CTC GAT ACT TCT GTT CTT CCG GAG CAG GTA GTT      631
Leu Thr Ser Arg Glu Leu Asp Thr Ser Val Leu Pro Glu Gln Val Val
        20                  25                  30
GTG GAG CGT CCG CGT AAC CCA GAG CAC GGC GAT TAC GCC ACC AAC ATT      679
Val Glu Arg Pro Arg Asn Pro Glu His Gly Asp Tyr Ala Thr Asn Ile
    35                  40                  45
GCA TTG CAG GTG GCT AAA AAG GTC GGT CAG AAC CCT CGG GAT TTG GCT      727
Ala Leu Gln Val Ala Lys Lys Val Gly Gln Asn Pro Arg Asp Leu Ala
50                  55                  60                  65
ACC TGG CTG GCA GAG GCA TTG GCT GCA GAT GAC GCC ATT GAT TCT GCT      775
Thr Trp Leu Ala Glu Ala Leu Ala Ala Asp Asp Ala Ile Asp Ser Ala
                70                  75                  80
GAA ATT GCT GGC CCA GGC TTT TTG AAC ATT CGC CTT GCT GCA GCA GCA      823
Glu Ile Ala Gly Pro Gly Phe Leu Asn Ile Arg Leu Ala Ala Ala Ala
            85                  90                  95
CAG GGT GAA ATT GTG GCC AAG ATT CTG GCA CAG GGC GAG ACT TTC GGA      871
Gln Gly Glu Ile Val Ala Lys Ile Leu Ala Gln Gly Glu Thr Phe Gly
        100                 105                 110
AAC TCC GAT CAC CTT TCC CAC TTG GAC GTG AAC CTC GAG TTC GTT TCT      919
Asn Ser Asp His Leu Ser His Leu Asp Val Asn Leu Glu Phe Val Ser
    115                 120                 125
GCA AAC CCA ACC GGA CCT ATT CAC CTT GGC GGA ACC CGC TGG GCT GCC      967
Ala Asn Pro Thr Gly Pro Ile His Leu Gly Gly Thr Arg Trp Ala Ala
130                 135                 140                 145
GTG GGT GAC TCT TTG GGT CGT GTG CTG GAG GCT TCC GGC GCG AAA GTG     1015
Val Gly Asp Ser Leu Gly Arg Val Leu Glu Ala Ser Gly Ala Lys Val
                150                 155                 160
ACC CGC GAA TAC TAC TTC AAC GAT CAC GGT CGC CAG ATC GAT CGT TTC     1063
Thr Arg Glu Tyr Tyr Phe Asn Asp His Gly Arg Gln Ile Asp Arg Phe
            165                 170                 175
GCT TTG TCC CTT CTT GCA GCG GCG AAG GGC GAG CCA ACG CCA GAA GAC     1111
Ala Leu Ser Leu Leu Ala Ala Ala Lys Gly Glu Pro Thr Pro Glu Asp
        180                 185                 190
GGT TAT GGC GGC GAA TAC ATT AAG GAA ATT GCG GAG GCA ATC GTC GAA     1159
Gly Tyr Gly Gly Glu Tyr Ile Lys Glu Ile Ala Glu Ala Ile Val Glu
    195                 200                 205
AAG CAT CCT GAA GCG TTG GCT TTG GAG CCT GCC GCA ACC CAG GAG CTT     1207
Lys His Pro Glu Ala Leu Ala Leu Glu Pro Ala Ala Thr Gln Glu Leu
210                 215                 220                 225
TTC CGC GCT GAA GGC GTG GAG ATG ATG TTC GAG CAC ATC AAA TCT TCC     1255
Phe Arg Ala Glu Gly Val Glu Met Met Phe Glu His Ile Lys Ser Ser
                230                 235                 240
CTG CAT GAG TTC GGC ACC GAT TTC GAT GTC TAC TAC CAC GAG AAC TCC     1303
Leu His Glu Phe Gly Thr Asp Phe Asp Val Tyr Tyr His Glu Asn Ser
            245                 250                 255
CTG TTC GAG TCC GGT GCG GTG GAC AAG GCC GTG CAG GTG CTG AAG GAC     1351
Leu Phe Glu Ser Gly Ala Val Asp Lys Ala Val Gln Val Leu Lys Asp
        260                 265                 270
```

```
AAC GGC AAC CTG TAC GAA AAC GAG GGC GCT TGG TGG CTG CGT TCC ACC      1399
Asn Gly Asn Leu Tyr Glu Asn Glu Gly Ala Trp Trp Leu Arg Ser Thr
    275                 280                 285
GAA TTC GGC GAT GAC AAA GAC CGC GTG GTG ATC AAG TCT GAC GGC GAC      1447
Glu Phe Gly Asp Asp Lys Asp Arg Val Val Ile Lys Ser Asp Gly Asp
290                 295                 300                 305
GCA GCC TAC ATC GCT GGC GAT ATC GCG TAC GTG GCT GAT AAG TTC TCC      1495
Ala Ala Tyr Ile Ala Gly Asp Ile Ala Tyr Val Ala Asp Lys Phe Ser
                310                 315                 320
CGC GGA CAC AAC CTA AAC ATC TAC ATG TTG GGT GCT GAC CAC CAT GGT      1543
Arg Gly His Asn Leu Asn Ile Tyr Met Leu Gly Ala Asp His His Gly
    325                 330                 335
TAC ATC GCG CGC CTG AAG GCA GCG GCG GCG GCA CTT GGC TAC AAG CCA      1591
Tyr Ile Ala Arg Leu Lys Ala Ala Ala Ala Ala Leu Gly Tyr Lys Pro
340                 345                 350
GAA GGC GTT GAA GTC CTG ATT GGC CAG ATG GTG AAC CTG CTT CGC GAC      1639
Glu Gly Val Glu Val Leu Ile Gly Gln Met Val Asn Leu Leu Arg Asp
355                 360                 365
GGC AAG GCA GTG CGT ATG TCC AAG CGT GCA GGC ACC GTG GTC ACC CTA      1687
Gly Lys Ala Val Arg Met Ser Lys Arg Ala Gly Thr Val Val Thr Leu
370                 375                 380                 385
GAT GAC CTC GTT GAA GCA ATC GGC ATC GAT GCG GCG CGT TAC TCC CTG      1735
Asp Asp Leu Val Glu Ala Ile Gly Ile Asp Ala Ala Arg Tyr Ser Leu
                390                 395                 400
ATC CGT TCC TCC GTG GAT TCT TCC CTG GAT ATC GAT CTC GGC CTG TGG      1783
Ile Arg Ser Ser Val Asp Ser Ser Leu Asp Ile Asp Leu Gly Leu Trp
            405                 410                 415
GAA TCC CAG TCC TCC GAC AAC CCT GTG TAC TAC GTG CAG TAC GGA CAC      1831
Glu Ser Gln Ser Ser Asp Asn Pro Val Tyr Tyr Val Gln Tyr Gly His
        420                 425                 430
GCT CGT CTG TGC TCC ATC GCG CGC AAG GCA GAG ACC TTG GGT GTC ACC      1879
Ala Arg Leu Cys Ser Ile Ala Arg Lys Ala Glu Thr Leu Gly Val Thr
    435                 440                 445
GAG GAA GGC GCA GAC CTA TCT CTA CTG ACC CAC GAC CGC GAA GGC GAT      1927
Glu Glu Gly Ala Asp Leu Ser Leu Leu Thr His Asp Arg Glu Gly Asp
450                 455                 460                 465
CTC ATC CGC ACA CTC GGA GAG TTC CCA GCA GTG GTG AAG GCT GCC GCT      1975
Leu Ile Arg Thr Leu Gly Glu Phe Pro Ala Val Val Lys Ala Ala Ala
                470                 475                 480
GAC CTA CGT GAA CCA CAC CGC ATT GCC CGC TAT GCT GAG GAA TTA GCT      2023
Asp Leu Arg Glu Pro His Arg Ile Ala Arg Tyr Ala Glu Glu Leu Ala
            485                 490                 495
GGA ACT TTC CAC CGC TTC TAC GAT TCC TGC CAC ATC CTT CCA AAG GTT      2071
Gly Thr Phe His Arg Phe Tyr Asp Ser Cys His Ile Leu Pro Lys Val
        500                 505                 510
GAT GAG GAT ACG GCA CCA ATC CAC ACA GCA CGT CTG GCA CTT GCA GCA      2119
Asp Glu Asp Thr Ala Pro Ile His Thr Ala Arg Leu Ala Leu Ala Ala
    515                 520                 525
GCA ACC CGC CAG ACC CTC GCT AAC GCC CTG CAC CTG GTT GGC GTT TCC      2167
Ala Thr Arg Gln Thr Leu Ala Asn Ala Leu His Leu Val Gly Val Ser
530                 535                 540                 545
GCA CCG GAG AAG ATG TAACA ATG GCT ACA GTT GAA AAT TTC AAT GAA        2214
Ala Pro Glu Lys Met       Met Ala Thr Val Glu Asn Phe Asn Glu
                550         1                 5
CTT CCC GCA CAC GTA TGG CCA CGC AAT GCC GTG CGC CAA GAA GAC GGC      2262
Leu Pro Ala His Val Trp Pro Arg Asn Ala Val Arg Gln Glu Asp Gly
 10                  15                  20                  25
```

```
GTT GTC ACC GTC GCT GGT GTG CCT CTG CCT GAC CTC GCT GAA GAA TAC      2310
Val Val Thr Val Ala Gly Val Pro Leu Pro Asp Leu Ala Glu Glu Tyr
                30                  35                  40
GGA ACC CCA CTG TTC GTA GTC GAC GAG GAC GAT TTC CGT TCC CGC TGT      2358
Gly Thr Pro Leu Phe Val Val Asp Glu Asp Asp Phe Arg Ser Arg Cys
            45                  50                  55
CGC GAC ATG GCT ACC GCA TTC GGT GGA CCA GGC AAT GTG CAC TAC GCA      2406
Arg Asp Met Ala Thr Ala Phe Gly Gly Pro Gly Asn Val His Tyr Ala
        60                  65                  70
TCT AAA GCG TTC CTG ACC AAG ACC ATT GCA CGT TGG GTT GAT GAA GAG      2454
Ser Lys Ala Phe Leu Thr Lys Thr Ile Ala Arg Trp Val Asp Glu Glu
    75                  80                  85
GGG CTG GCA CTG GAC ATT GCA TCC ATC AAC GAA CTG GGC ATT GCC CTG      2502
Gly Leu Ala Leu Asp Ile Ala Ser Ile Asn Glu Leu Gly Ile Ala Leu
90                  95                  100                 105
GCC GCT GGT TTC CCC GCC AGC CGT ATC ACC GCG CAC GGC AAC AAC AAA      2550
Ala Ala Gly Phe Pro Ala Ser Arg Ile Thr Ala His Gly Asn Asn Lys
                110                 115                 120
GGC GTA GAG TTC CTG CGC GCG TTG GTT CAA AAC GGT GTG GGA CAC GTG      2598
Gly Val Glu Phe Leu Arg Ala Leu Val Gln Asn Gly Val Gly His Val
            125                 130                 135
GTG CTG GAC TCC GCA CAG GAA CTA GAA CTG TTG GAT TAC GTT GCC GCT      2646
Val Leu Asp Ser Ala Gln Glu Leu Glu Leu Leu Asp Tyr Val Ala Ala
        140                 145                 150
GGT GAA GGC AAG ATT CAG GAC GTG TTG ATC CGC GTA AAG CCA GGC ATC      2694
Gly Glu Gly Lys Ile Gln Asp Val Leu Ile Arg Val Lys Pro Gly Ile
    155                 160                 165
GAA GCA CAC ACC CAC GAG TTC ATC GCC ACT AGC CAC GAA GAC CAG AAG      2742
Glu Ala His Thr His Glu Phe Ile Ala Thr Ser His Glu Asp Gln Lys
170                 175                 180                 185
TTC GGA TTC TCC CTG GCA TCC GGT TCC GCA TTC GAA GCA GCA AAA GCC      2790
Phe Gly Phe Ser Leu Ala Ser Gly Ser Ala Phe Glu Ala Ala Lys Ala
                190                 195                 200
GCC AAC AAC GCA GAA AAC CTG AAC CTG GTT GGC CTG CAC TGC CAC GTT      2838
Ala Asn Asn Ala Glu Asn Leu Asn Leu Val Gly Leu His Cys His Val
            205                 210                 215
GGT TCC CAG GTG TTC GAC GCC GAA GGC TTC AAG CTG GCA GCA GAA CGC      2886
Gly Ser Gln Val Phe Asp Ala Glu Gly Phe Lys Leu Ala Ala Glu Arg
        220                 225                 230
GTG TTG GGC CTG TAC TCA CAG ATC CAC AGC GAA CTG GGC GTT GCC CTT      2934
Val Leu Gly Leu Tyr Ser Gln Ile His Ser Glu Leu Gly Val Ala Leu
    235                 240                 245
CCT GAA CTG GAT CTC GGT GGC GGA TAC GGC ATT GCC TAT ACC GCA GCT      2982
Pro Glu Leu Asp Leu Gly Gly Gly Tyr Gly Ile Ala Tyr Thr Ala Ala
250                 255                 260                 265
GAA GAA CCA CTC AAC GTC GCA GAA GTT GCC TCC GAC CTG CTC ACC GCA      3030
Glu Glu Pro Leu Asn Val Ala Glu Val Ala Ser Asp Leu Leu Thr Ala
                270                 275                 280
GTC GGA AAA ATG GCA GCG GAA CTA GGC ATC GAC GCA CCA ACC GTG CTT      3078
Val Gly Lys Met Ala Ala Glu Leu Gly Ile Asp Ala Pro Thr Val Leu
            285                 290                 295
GTT GAG CCC GGC CGC GCT ATC GCA GGC CCC TCC ACC GTG ACC ATC TAC      3126
Val Glu Pro Gly Arg Ala Ile Ala Gly Pro Ser Thr Val Thr Ile Tyr
        300                 305                 310
GAA GTC GGC ACC ACC AAA GAC GTC CAC GTA GAC GAC GAC AAA ACC CGC      3174
Glu Val Gly Thr Thr Lys Asp Val His Val Asp Asp Asp Lys Thr Arg
    315                 320                 325
```

```
CGT TAC ATC GCC GTG GAC GGA GGC ATG TCC GAC AAC ATC CGC CCA GCA      3222
Arg Tyr Ile Ala Val Asp Gly Gly Met Ser Asp Asn Ile Arg Pro Ala
330                 335                 340                 345
CTC TAC GGC TCC GAA TAC GAC GCC CGC GTA GTA TCC CGC TTC GCC GAA      3270
Leu Tyr Gly Ser Glu Tyr Asp Ala Arg Val Val Ser Arg Phe Ala Glu
                350                 355                 360
GGA GAC CCA GTA AGC ACC CGC ATC GTG GGC TCC CAC TGC GAA TCC GGC      3318
Gly Asp Pro Val Ser Thr Arg Ile Val Gly Ser His Cys Glu Ser Gly
            365                 370                 375
GAT ATC CTG ATC AAC GAT GAA ATC TAC CCA TCT GAC ATC ACC AGC GGC      3366
Asp Ile Leu Ile Asn Asp Glu Ile Tyr Pro Ser Asp Ile Thr Ser Gly
        380                 385                 390
GAC TTC CTT GCA CTC GCA GCC ACC GGC GCA TAC TGC TAC GCC ATG AGC      3414
Asp Phe Leu Ala Leu Ala Ala Thr Gly Ala Tyr Cys Tyr Ala Met Ser
    395                 400                 405
TCC CGC TAC AAC GCC TTC ACA CGG CCC GCC GTC GTG TCC GTC CGC GCT      3462
Ser Arg Tyr Asn Ala Phe Thr Arg Pro Ala Val Val Ser Val Arg Ala
410                 415                 420                 425
GGC AGC TCC CGC CTC ATG CTG CGC CGC GAA ACG CTC GAC GAC ATC CTC      3510
Gly Ser Ser Arg Leu Met Leu Arg Arg Glu Thr Leu Asp Asp Ile Leu
                430                 435                 440
TCA CTA GAG GCA TAACGCTTTT CGACGCCTGA CCCCGCCCTT CACCTTCGCC          3562
Ser Leu Glu Ala
            445
GTGGAGGGCG GTTTTGG                                                  3579
```

(2) SEQ ID NO: 11 的信息:

(i) 序列特征:

(A) 长度: 550 个氨基酸

(B) 类型: 氨基酸

(D) 拓扑结构: 线性

(ii) 分子类型: 蛋白质

(xi) 序列描述: SEQ ID NO: 11 :

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

```
Val Thr Arg Glu Tyr Tyr Phe Asn Asp His Gly Arg Gln Ile Asp Arg
                165                 170                 175
Phe Ala Leu Ser Leu Leu Ala Ala Ala Lys Gly Glu Pro Thr Pro Glu
            180                 185                 190
Asp Gly Tyr Gly Gly Glu Tyr Ile Lys Glu Ile Ala Glu Ala Ile Val
        195                 200                 205
Glu Lys His Pro Glu Ala Leu Ala Leu Glu Pro Ala Ala Thr Gln Glu
    210                 215                 220
Leu Phe Arg Ala Glu Gly Val Glu Met Met Phe Glu His Ile Lys Ser
225                 230                 235                 240
Ser Leu His Glu Phe Gly Thr Asp Phe Asp Val Tyr Tyr His Glu Asn
                245                 250                 255
Ser Leu Phe Glu Ser Gly Ala Val Asp Lys Ala Val Gln Val Leu Lys
            260                 265                 270
Asp Asn Gly Asn Leu Tyr Glu Asn Glu Gly Ala Trp Trp Leu Arg Ser
        275                 280                 285
Thr Glu Phe Gly Asp Asp Lys Asp Arg Val Val Ile Lys Ser Asp Gly
    290                 295                 300
Asp Ala Ala Tyr Ile Ala Gly Asp Ile Ala Tyr Val Ala Asp Lys Phe
305                 310                 315                 320
Ser Arg Gly His Asn Leu Asn Ile Tyr Met Leu Gly Ala Asp His His
                325                 330                 335
Gly Tyr Ile Ala Arg Leu Lys Ala Ala Ala Ala Ala Leu Gly Tyr Lys
            340                 345                 350
Pro Glu Gly Val Glu Val Leu Ile Gly Gln Met Val Asn Leu Leu Arg
        355                 360                 365
Asp Gly Lys Ala Val Arg Met Ser Lys Arg Ala Gly Thr Val Val Thr
    370                 375                 380
Leu Asp Asp Leu Val Glu Ala Ile Gly Ile Asp Ala Ala Arg Tyr Ser
385                 390                 395                 400
Leu Ile Arg Ser Ser Val Asp Ser Ser Leu Asp Ile Asp Leu Gly Leu
                405                 410                 415
Trp Glu Ser Gln Ser Ser Asp Asn Pro Val Tyr Tyr Val Gln Tyr Gly
            420                 425                 430
His Ala Arg Leu Cys Ser Ile Ala Arg Lys Ala Glu Thr Leu Gly Val
        435                 440                 445
Thr Glu Glu Gly Ala Asp Leu Ser Leu Leu Thr His Asp Arg Glu Gly
    450                 455                 460
Asp Leu Ile Arg Thr Leu Gly Glu Phe Pro Ala Val Val Lys Ala Ala
465                 470                 475                 480
Ala Asp Leu Arg Glu Pro His Arg Ile Ala Arg Tyr Ala Glu Glu Leu
                485                 490                 495
Ala Gly Thr Phe His Arg Phe Tyr Asp Ser Cys His Ile Leu Pro Lys
            500                 505                 510
Val Asp Glu Asp Thr Ala Pro Ile His Thr Ala Arg Leu Ala Leu Ala
        515                 520                 525
Ala Ala Thr Arg Gln Thr Leu Ala Asn Ala Leu His Leu Val Gly Val
    530                 535                 540
Ser Ala Pro Glu Lys Met
545                 550
```

(2) SEQ ID NO: 12 的信息:

(i) 序列特征:

(A) 长度: 445 个氨基酸

(B) 类型: 氨基酸

(D) 拓扑结构: 线性

(ii) 分子类型: 蛋白质

(xi) 序列描述: SEQ ID NO: 12:

```
Met Ala Thr Val Glu Asn Phe Asn Glu Leu Pro Ala His Val Trp Pro
 1               5                  10                  15
Arg Asn Ala Val Arg Gln Glu Asp Gly Val Val Thr Val Ala Gly Val
            20                  25                  30
Pro Leu Pro Asp Leu Ala Glu Glu Tyr Gly Thr Pro Leu Phe Val Val
        35                  40                  45
Asp Glu Asp Asp Phe Arg Ser Arg Cys Arg Asp Met Ala Thr Ala Phe
    50                  55                  60
Gly Gly Pro Gly Asn Val His Tyr Ala Ser Lys Ala Phe Leu Thr Lys
65                  70                  75                  80
Thr Ile Ala Arg Trp Val Asp Glu Glu Gly Leu Ala Leu Asp Ile Ala
                85                  90                  95
Ser Ile Asn Glu Leu Gly Ile Ala Leu Ala Ala Gly Phe Pro Ala Ser
            100                 105                 110
Arg Ile Thr Ala His Gly Asn Asn Lys Gly Val Glu Phe Leu Arg Ala
        115                 120                 125
Leu Val Gln Asn Gly Val Gly His Val Val Leu Asp Ser Ala Gln Glu
    130                 135                 140
Leu Glu Leu Leu Asp Tyr Val Ala Ala Gly Glu Gly Lys Ile Gln Asp
145                 150                 155                 160
Val Leu Ile Arg Val Lys Pro Gly Ile Glu Ala His Thr His Glu Phe
                165                 170                 175
Ile Ala Thr Ser His Glu Asp Gln Lys Phe Gly Phe Ser Leu Ala Ser
            180                 185                 190
Gly Ser Ala Phe Glu Ala Ala Lys Ala Ala Asn Asn Ala Glu Asn Leu
        195                 200                 205
Asn Leu Val Gly Leu His Cys His Val Gly Ser Gln Val Phe Asp Ala
    210                 215                 220
Glu Gly Phe Lys Leu Ala Ala Glu Arg Val Leu Gly Leu Tyr Ser Gln
225                 230                 235                 240
Ile His Ser Glu Leu Gly Val Ala Leu Pro Glu Leu Asp Leu Gly Gly
                245                 250                 255
Gly Tyr Gly Ile Ala Tyr Thr Ala Ala Glu Glu Pro Leu Asn Val Ala
            260                 265                 270
Glu Val Ala Ser Asp Leu Leu Thr Ala Val Gly Lys Met Ala Ala Glu
        275                 280                 285
Leu Gly Ile Asp Ala Pro Thr Val Leu Val Glu Pro Gly Arg Ala Ile
    290                 295                 300
Ala Gly Pro Ser Thr Val Thr Ile Tyr Glu Val Gly Thr Thr Lys Asp
305                 310                 315                 320
Val His Val Asp Asp Asp Lys Thr Arg Arg Tyr Ile Ala Val Asp Gly
                325                 330                 335
Gly Met Ser Asp Asn Ile Arg Pro Ala Leu Tyr Gly Ser Glu Tyr Asp
            340                 345                 350
Ala Arg Val Val Ser Arg Phe Ala Glu Gly Asp Pro Val Ser Thr Arg
        355                 360                 365
Ile Val Gly Ser His Cys Glu Ser Gly Asp Ile Leu Ile Asn Asp Glu
    370                 375                 380
Ile Tyr Pro Ser Asp Ile Thr Ser Gly Asp Phe Leu Ala Leu Ala Ala
385                 390                 395                 400
Thr Gly Ala Tyr Cys Tyr Ala Met Ser Ser Arg Tyr Asn Ala Phe Thr
                405                 410                 415
Arg Pro Ala Val Val Ser Val Arg Ala Gly Ser Ser Arg Leu Met Leu
            420                 425                 430
Arg Arg Glu Thr Leu Asp Asp Ile Leu Ser Leu Glu Ala
        435                 440                 445
```

(2) SEQ ID NO: 13 的信息:

(i) 序列特征:

(A) 长度: 23 个碱基

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其它核酸

(A) 描述: desc = "合成的 DNA"

(iv) 反义: 否

(xi) 序列描述: SEQ ID NO: 13:

TCGTCGGTCA GCCTGACGTC GAC 20

(2) SEQ ID NO: 14 的信息:

(i) 序列特征:

(A) 长度: 23 个碱基

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其它核酸

(A) 描述: desc = "合成的 DNA"

(iv) 反义: 是

(xi) 序列描述: SEQ ID NO: 14:

TCTTGGTGTCGAAAGTGCACACC 20

(2) SEQ ID NO: 的信息: 15:

(i) 序列特征:

(A) 长度: 3533 个碱基

(B) 类型: 核酸

(C) 链型: 双链

(D) 拓扑结构: 线性

(ii)分子类型: 基因组 DNA

(vi)原始来源:

(A)有机体: 乳发酵短杆菌

(B)菌株: ATCC 13869

(ix)特征:

(A)名称/关键词: CDS

(B)位置: 321..3077

(xi)序列描述: SEQ ID NO: 15:

```
GGTGGTTCTG TTAAGGCAGA AACCGTCGCT GAGATCGTCG GTCAGCCTGA CGTCGACGGC      60
GGACTTGTCG GTGGCGCTTC CCTCGACGGT GAAGCATTCG CCAAGCTGGC TGCCAACGCT     120
GCGAGCGTTG CTTAAAGTAC AGAGCTTTAA AGCACAGCCT TAAAGCACAG CCTTAAAGCA     180
CAAGCACTGT AGAAGTGCGG TTTTGATGAG CCCATGAAAG CCATCGAAAT CAATCGCCCA     240
GCTAAACACC TGTTTTGCTG GGTGATTTTT TATCTCATGC ACGCCAACAC CCTCAATGTG     300
AAAGAGTGTT TAAAGTAGTT ATG ACT GAT TTT TTA CGC GAT GAC ATC AGG          350
                      Met Thr Asp Phe Leu Arg Asp Asp Ile Arg
                        1               5                  10
TTC CTC GGT CAA ATC CTC GGT GAG GTA ATT GCG GAA CAA GAA GGC CAG       398
Phe Leu Gly Gln Ile Leu Gly Glu Val Ile Ala Glu Gln Glu Gly Gln
                   15                  20                  25
GAG GTT TAT GAA CTG GTC GAA CAA GCG CGC CTG ACT TCT TTT GAT ATC       446
Glu Val Tyr Glu Leu Val Glu Gln Ala Arg Leu Thr Ser Phe Asp Ile
               30                  35                  40
GCC AAG GGC AAC GCC GAA ATG GAT AGC CTG GTT CAG GTT TTC GAC GGC       494
Ala Lys Gly Asn Ala Glu Met Asp Ser Leu Val Gln Val Phe Asp Gly
           45                  50                  55
ATT ACT CCA GCC AAG GCA ACA CCG ATT GCT CGC GCA TTT TCC CAC TTC       542
Ile Thr Pro Ala Lys Ala Thr Pro Ile Ala Arg Ala Phe Ser His Phe
       60                  65                  70
GCT CTG CTG GCT AAC CTG GCG GAA GAC CTC TAC GAT GAA GAG CTT CGT       590
Ala Leu Leu Ala Asn Leu Ala Glu Asp Leu Tyr Asp Glu Glu Leu Arg
   75                  80                  85                  90
```

```
GAA CAG GCT CTC GAT GCA GGC GAC ACC CCT CCG GAC AGC ACT CTT GAT      638
Glu Gln Ala Leu Asp Ala Gly Asp Thr Pro Pro Asp Ser Thr Leu Asp
                95                  100                 105
GCC ACC TGG CTG AAA CTC AAT GAG GGC AAT GTT GGC GCA GAA GCT GTG      686
Ala Thr Trp Leu Lys Leu Asn Glu Gly Asn Val Gly Ala Glu Ala Val
            110                 115                 120
GCC GAT GTG CTG CGC AAT GCT GAG GTG GCG CCG GTT CTG ACT GCG CAC      734
Ala Asp Val Leu Arg Asn Ala Glu Val Ala Pro Val Leu Thr Ala His
        125                 130                 135
CCA ACT GAG ACT CGC CGC CGC ACT GTT TTT GAT GCG CAA AAG TGG ATC      782
Pro Thr Glu Thr Arg Arg Arg Thr Val Phe Asp Ala Gln Lys Trp Ile
    140                 145                 150
ACC ACC CAC ATG CGT GAA CGC CAC GCT TTG CAG TCT GCG GAG CCT ACC      830
Thr Thr His Met Arg Glu Arg His Ala Leu Gln Ser Ala Glu Pro Thr
155                 160                 165                 170
GCT CGT ACG CAA AGC AAG TTG GAT GAG ATC GAG AAG AAC ATC CGC CGT      878
Ala Arg Thr Gln Ser Lys Leu Asp Glu Ile Glu Lys Asn Ile Arg Arg
                175                 180                 185
CGC ATC ACC ATT TTG TGG CAG ACC GCG TTG ATT CGT GTG GCC CGC CCA      926
Arg Ile Thr Ile Leu Trp Gln Thr Ala Leu Ile Arg Val Ala Arg Pro
            190                 195                 200
CGT ATC GAG GAC GAG ATC GAA GTA GGG CTG CGC TAC TAC AAG CTG AGC      974
Arg Ile Glu Asp Glu Ile Glu Val Gly Leu Arg Tyr Tyr Lys Leu Ser
        205                 210                 215
CTT TTG GAA GAG ATT CCA CGT ATC AAC CGT GAT GTG GCT GTT GAG CTT     1022
Leu Leu Glu Glu Ile Pro Arg Ile Asn Arg Asp Val Ala Val Glu Leu
    220                 225                 230
CGT GAG CGT TTC GGC GAG GAT GTT CCT TTG AAG CCC GTG GTC AAG CCA     1070
Arg Glu Arg Phe Gly Glu Asp Val Pro Leu Lys Pro Val Val Lys Pro
235                 240                 245                 250
GGT TCC TGG ATT GGT GGA GAC CAC GAC GGT AAC CCT TAT GTC ACC GCG     1118
Gly Ser Trp Ile Gly Gly Asp His Asp Gly Asn Pro Tyr Val Thr Ala
                255                 260                 265
GAA ACA GTT GAG TAT TCC ACT CAC CGC GCT GCG GAA ACC GTG CTC AAG     1166
Glu Thr Val Glu Tyr Ser Thr His Arg Ala Ala Glu Thr Val Leu Lys
            270                 275                 280
TAC TAT GCA CGC CAG CTG CAT TCC CTC GAG CAT GAG CTC AGC CTG TCG     1214
Tyr Tyr Ala Arg Gln Leu His Ser Leu Glu His Glu Leu Ser Leu Ser
        285                 290                 295
GAC CGC ATG AAT AAG GTC ACC CCG CAG CTG CTT GCG CTG GCA GAT GCC     1262
Asp Arg Met Asn Lys Val Thr Pro Gln Leu Leu Ala Leu Ala Asp Ala
    300                 305                 310
GGG CAC AAC GAC GTG CCA AGC CGC GTG GAT GAG CCT TAT CGA CGC GCC     1310
Gly His Asn Asp Val Pro Ser Arg Val Asp Glu Pro Tyr Arg Arg Ala
315                 320                 325                 330
GTC CAT GGC GTT CGC GGA CGT ATC CTC GCG ACG ACG GCC GAG CTG ATC     1358
Val His Gly Val Arg Gly Arg Ile Leu Ala Thr Thr Ala Glu Leu Ile
                335                 340                 345
GGC GAG GAC GCC GTT GAG GGC GTG TGG TTC AAG GTC TTT ACT CCA TAC     1406
Gly Glu Asp Ala Val Glu Gly Val Trp Phe Lys Val Phe Thr Pro Tyr
            350                 355                 360
GCA TCT CCG GAA GAA TTC TTA AAC GAT GCG TTG ACC ATT GAT CAT TCT     1454
Ala Ser Pro Glu Glu Phe Leu Asn Asp Ala Leu Thr Ile Asp His Ser
        365                 370                 375
CTG CGT GAA TCC AAT GAC GTT CTC ATT GCC GAT GAT CGT TTG TCT GTG     1502
Leu Arg Glu Ser Asn Asp Val Leu Ile Ala Asp Asp Arg Leu Ser Val
    380                 385                 390
```

```
CTG ATT TCT GCC ATC GAG AGC TTT GGA TTC AAC CTT TAC GCA CTG GAT     1550
Leu Ile Ser Ala Ile Glu Ser Phe Gly Phe Asn Leu Tyr Ala Leu Asp
395                 400                 405                 410
CTG CGC CAA AAC TCC GAA AGC TAC GAG GAC GTC CTC ACC GAG CTT TTC     1598
Leu Arg Gln Asn Ser Glu Ser Tyr Glu Asp Val Leu Thr Glu Leu Phe
                415                 420                 425
GAA CGC GCC CAA GTC ACC GCA AAC TAC CGC GAG CTG TCT GAA GCA GAG     1646
Glu Arg Ala Gln Val Thr Ala Asn Tyr Arg Glu Leu Ser Glu Ala Glu
            430                 435                 440
AAG CTT GAG GTG CTG CTG AAG GAA CTG CGC AGC CCT CGT CCG CTG ATC     1694
Lys Leu Glu Val Leu Leu Lys Glu Leu Arg Ser Pro Arg Pro Leu Ile
        445                 450                 455
CCG CAC GGT TCA GAT GAA TAC AGC GAG GTC ACC GAC CGC GAG CTC GGC     1742
Pro His Gly Ser Asp Glu Tyr Ser Glu Val Thr Asp Arg Glu Leu Gly
    460                 465                 470
ATC TTC CGC ACC GCG TCG GAG GCT GTT AAG AAA TTC GGG CCA CGG ATG     1790
Ile Phe Arg Thr Ala Ser Glu Ala Val Lys Lys Phe Gly Pro Arg Met
475                 480                 485                 490
GTG CCT CAC TGC ATC ATC TCC ATG GCA TCA TCG GTC ACC GAT GTG CTC     1838
Val Pro His Cys Ile Ile Ser Met Ala Ser Ser Val Thr Asp Val Leu
                495                 500                 505
GAG CCG ATG GTA TTG CTC AAG GAA TTC GGC CTC ATT GCA GCC AAC GGC     1886
Glu Pro Met Val Leu Leu Lys Glu Phe Gly Leu Ile Ala Ala Asn Gly
            510                 515                 520
GAC AAC CCA CGC GGC ACC GTC GAT GTC ATC CCA CTG TTC GAA ACC ATC     1934
Asp Asn Pro Arg Gly Thr Val Asp Val Ile Pro Leu Phe Glu Thr Ile
        525                 530                 535
GAA GAT CTC CAG GCC GGC GCC GGA ATC CTC GAC GAA CTG TGG AAA ATT     1982
Glu Asp Leu Gln Ala Gly Ala Gly Ile Leu Asp Glu Leu Trp Lys Ile
    540                 545                 550
GAT CTT TAC CGC AAC TAC CTC CTG CAG CGC GAC AAC GTC CAG GAA GTC     2030
Asp Leu Tyr Arg Asn Tyr Leu Leu Gln Arg Asp Asn Val Gln Glu Val
555                 560                 565                 570
ATG CTC GGT TAC TCC GAT TCC AAC AAG GAT GGC GGA TAT TTC TCC GCA     2078
Met Leu Gly Tyr Ser Asp Ser Asn Lys Asp Gly Gly Tyr Phe Ser Ala
                575                 580                 585
AAC TGG GCG CTT TAC GAC GCG GAA CTG CAG CTC GTC GAA CTA TGC CGA     2126
Asn Trp Ala Leu Tyr Asp Ala Glu Leu Gln Leu Val Glu Leu Cys Arg
            590                 595                 600
TCA GCC GGG GTC AAG CTT CGC CTG TTC CAC GGC CGT GGT GGC ACC GTC     2174
Ser Ala Gly Val Lys Leu Arg Leu Phe His Gly Arg Gly Gly Thr Val
        605                 610                 615
GGC CGC GGT GGC GGA CCT TCC TAC GAC GCG ATT CTT GCC CAG CCC AGG     2222
Gly Arg Gly Gly Gly Pro Ser Tyr Asp Ala Ile Leu Ala Gln Pro Arg
    620                 625                 630
GGG GCT GTC CAA GGT TCC GTG CGC ATC ACC GAG CAG GGC GAG ATC ATC     2270
Gly Ala Val Gln Gly Ser Val Arg Ile Thr Glu Gln Gly Glu Ile Ile
635                 640                 645                 650
TCC GCT AAG TAC GGC AAC CCC GAA ACC GCG CGC CGA AAC CTC GAA GCT     2318
Ser Ala Lys Tyr Gly Asn Pro Glu Thr Ala Arg Arg Asn Leu Glu Ala
                655                 660                 665
CTG GTC TCA GCA ACG CTT GAG GCA TCG CTT CTC GAC GTC TCC GAA CTC     2366
Leu Val Ser Ala Thr Leu Glu Ala Ser Leu Leu Asp Val Ser Glu Leu
            670                 675                 680
ACC GAT CAC CAA CGC GCG TAC GAC ATC ATG AGT GAG ATC TCT GAG CTC     2414
Thr Asp His Gln Arg Ala Tyr Asp Ile Met Ser Glu Ile Ser Glu Leu
        685                 690                 695
```

```
AGC TTG AAG AAG TAC GCC TCC TTG GTG CAC GAG GAT CAA GGC TTC ATC      2462
Ser Leu Lys Lys Tyr Ala Ser Leu Val His Glu Asp Gln Gly Phe Ile
    700                 705                 710
GAT TAC TTC ACC CAG TCC ACG CCG CTG CAG GAG ATT GGA TCC CTC AAC      2510
Asp Tyr Phe Thr Gln Ser Thr Pro Leu Gln Glu Ile Gly Ser Leu Asn
715                 720                 725                 730
ATC GGA TCC AGG CCT TCC TCA CGC AAG CAG ACC TCC TCG GTG GAA GAT      2558
Ile Gly Ser Arg Pro Ser Ser Arg Lys Gln Thr Ser Ser Val Glu Asp
                735                 740                 745
TTG CGA GCA ATC CCG TGG GTG CTC AGT TGG TCC CAG TCT CGT GTC ATG      2606
Leu Arg Ala Ile Pro Trp Val Leu Ser Trp Ser Gln Ser Arg Val Met
            750                 755                 760
CTG CCG GGC TGG TTT GGT GTC GGC ACC GCA CTT GAG CAA TGG ATT GGC      2654
Leu Pro Gly Trp Phe Gly Val Gly Thr Ala Leu Glu Gln Trp Ile Gly
        765                 770                 775
GAA GGG GAG CAG GCC ACC CAG CGC ATT GCC GAG CTA CAA ACA CTC AAC      2702
Glu Gly Glu Gln Ala Thr Gln Arg Ile Ala Glu Leu Gln Thr Leu Asn
    780                 785                 790
GAG TCC TGG CCA TTT TTC ACC TCA GTG TTG GAT AAC ATG GCT CAG GTG      2750
Glu Ser Trp Pro Phe Phe Thr Ser Val Leu Asp Asn Met Ala Gln Val
795                 800                 805                 810
ATG TCC AAG GCA GAG CTG CGT TTG GCA AAG CTC TAC GCA GAC CTG ATC      2798
Met Ser Lys Ala Glu Leu Arg Leu Ala Lys Leu Tyr Ala Asp Leu Ile
                815                 820                 825
CCA GAT AGG GAA GTA GCT GAG CGC GTT TAT GCC GTC ATC CGC GAG GAA      2846
Pro Asp Arg Glu Val Ala Glu Arg Val Tyr Ala Val Ile Arg Glu Glu
            830                 835                 840
TAC TTC CTG ACC AAG AAG ATG TTC TGC GTA ATC ACC GGT TCT GAT GAT      2894
Tyr Phe Leu Thr Lys Lys Met Phe Cys Val Ile Thr Gly Ser Asp Asp
        845                 850                 855
CTG CTT GAT GAC AAC CCG CTT CTC GCA CGA TCC GTC CAG CGC CGA TAC      2942
Leu Leu Asp Asp Asn Pro Leu Leu Ala Arg Ser Val Gln Arg Arg Tyr
    860                 865                 870
CCC TAC CTG CTT CCA CTC AAC GTG ATC CAG GTA GAG ATG ATG CGA CGC      2990
Pro Tyr Leu Leu Pro Leu Asn Val Ile Gln Val Glu Met Met Arg Arg
875                 880                 885                 890
TAC CGA AAA GGC GAC CAA AGC GAG CAA GTA TCC CGC AAC ATC CAG CTG      3038
Tyr Arg Lys Gly Asp Gln Ser Glu Gln Val Ser Arg Asn Ile Gln Leu
                895                 900                 905
ACC ATG AAC GGT CTT TCC ACT GCA CTG CGC AAC TCT GGC TAGTCCTGCT       3087
Thr Met Asn Gly Leu Ser Thr Ala Leu Arg Asn Ser Gly
            910                 915
GGGTAGGTAG TACTCGTGTA TACTGTCTAA AGTTATTCGA AATCAGGTGG GAATAAGGTT    3147
CACCTGGGTT CTCAAACGGC AAAGGAACAT TTTCCACATG GCATTGACGC TTCAAATCAT    3207
CCTCGTCGTC GCCAGCCTGC TCATGACGGT TTTCGTCTTG CTGCACAAGG GCAAAGGCGG    3267
CGGACTCTCC AGCCTCTTCG GTGGCGGTGT GCAGTCCAAT CTTTCGGGCT CCACTGTTGT    3327
TGAAAAGAAC CTGGATCGCG TCACCATTTT GGTTGCCGTT ATCTGGATTG TGTGCATTGT    3387
CGCACTCAAC CTCATCCAGA CTTATTCATA AGACACGAGC TTAAAAAGAG CGGTTCCCTT    3447
TTCATAGGGG AGCCGCTTTT TTGGGTTTTG TCGACCTGTT GTCTCCCCAC TGTTCCTCGG    3507
TGTGCACTTT CGACACCAAG ATTTCG                                         3533
```

(2) SEQ ID NO: 16 的信息:

(i) 序列特征:

(A) 长度: 919 个氨基酸

(B) 类型: 氨基酸

(D)拓扑结构: 线性

(ii)分子类型: 蛋白质

(xi)序列描述: SEQ ID NO: 16:

```
Met Thr Asp Phe Leu Arg Asp Asp Ile Arg Phe Leu Gly Gln Ile Leu
 1               5                  10                  15
Gly Glu Val Ile Ala Glu Gln Glu Gly Gln Glu Val Tyr Glu Leu Val
            20                  25                  30
Glu Gln Ala Arg Leu Thr Ser Phe Asp Ile Ala Lys Gly Asn Ala Glu
        35                  40                  45
Met Asp Ser Leu Val Gln Val Phe Asp Gly Ile Thr Pro Ala Lys Ala
    50                  55                  60
Thr Pro Ile Ala Arg Ala Phe Ser His Phe Ala Leu Leu Ala Asn Leu
65                  70                  75                  80
Ala Glu Asp Leu Tyr Asp Glu Glu Leu Arg Glu Gln Ala Leu Asp Ala
                85                  90                  95
Gly Asp Thr Pro Pro Asp Ser Thr Leu Asp Ala Thr Trp Leu Lys Leu
            100                 105                 110
Asn Glu Gly Asn Val Gly Ala Glu Ala Val Ala Asp Val Leu Arg Asn
        115                 120                 125
Ala Glu Val Ala Pro Val Leu Thr Ala His Pro Thr Glu Thr Arg Arg
    130                 135                 140
Arg Thr Val Phe Asp Ala Gln Lys Trp Ile Thr Thr His Met Arg Glu
145                 150                 155                 160
Arg His Ala Leu Gln Ser Ala Glu Pro Thr Ala Arg Thr Gln Ser Lys
                165                 170                 175
Leu Asp Glu Ile Glu Lys Asn Ile Arg Arg Arg Ile Thr Ile Leu Trp
            180                 185                 190
Gln Thr Ala Leu Ile Arg Val Ala Arg Pro Arg Ile Glu Asp Glu Ile
        195                 200                 205
Glu Val Gly Leu Arg Tyr Tyr Lys Leu Ser Leu Leu Glu Glu Ile Pro
    210                 215                 220
Arg Ile Asn Arg Asp Val Ala Val Glu Leu Arg Glu Arg Phe Gly Glu
225                 230                 235                 240
Asp Val Pro Leu Lys Pro Val Val Lys Pro Gly Ser Trp Ile Gly Gly
                245                 250                 255
Asp His Asp Gly Asn Pro Tyr Val Thr Ala Glu Thr Val Glu Tyr Ser
            260                 265                 270
Thr His Arg Ala Ala Glu Thr Val Leu Lys Tyr Tyr Ala Arg Gln Leu
        275                 280                 285
His Ser Leu Glu His Glu Leu Ser Leu Ser Asp Arg Met Asn Lys Val
    290                 295                 300
Thr Pro Gln Leu Leu Ala Leu Ala Asp Ala Gly His Asn Asp Val Pro
305                 310                 315                 320
Ser Arg Val Asp Glu Pro Tyr Arg Arg Ala Val His Gly Val Arg Gly
                325                 330                 335
Arg Ile Leu Ala Thr Thr Ala Glu Leu Ile Gly Glu Asp Ala Val Glu
            340                 345                 350
Gly Val Trp Phe Lys Val Phe Thr Pro Tyr Ala Ser Pro Glu Glu Phe
        355                 360                 365
Leu Asn Asp Ala Leu Thr Ile Asp His Ser Leu Arg Glu Ser Asn Asp
    370                 375                 380
Val Leu Ile Ala Asp Asp Arg Leu Ser Val Leu Ile Ser Ala Ile Glu
385                 390                 395                 400
Ser Phe Gly Phe Asn Leu Tyr Ala Leu Asp Leu Arg Gln Asn Ser Glu
                405                 410                 415
Ser Tyr Glu Asp Val Leu Thr Glu Leu Phe Glu Arg Ala Gln Val Thr
            420                 425                 430
Ala Asn Tyr Arg Glu Leu Ser Glu Ala Glu Lys Leu Glu Val Leu Leu
        435                 440                 445
Lys Glu Leu Arg Ser Pro Arg Pro Leu Ile Pro His Gly Ser Asp Glu
    450                 455                 460
```

```
Tyr Ser Glu Val Thr Asp Arg Glu Leu Gly Ile Phe Arg Thr Ala Ser
465                 470                 475                 480
Glu Ala Val Lys Lys Phe Gly Pro Arg Met Val Pro His Cys Ile Ile
                485                 490                 495
Ser Met Ala Ser Ser Val Thr Asp Val Leu Glu Pro Met Val Leu Leu
            500                 505                 510
Lys Glu Phe Gly Leu Ile Ala Ala Asn Gly Asp Asn Pro Arg Gly Thr
        515                 520                 525
Val Asp Val Ile Pro Leu Phe Glu Thr Ile Glu Asp Leu Gln Ala Gly
    530                 535                 540
Ala Gly Ile Leu Asp Glu Leu Trp Lys Ile Asp Leu Tyr Arg Asn Tyr
545                 550                 555                 560
Leu Leu Gln Arg Asp Asn Val Gln Glu Val Met Leu Gly Tyr Ser Asp
                565                 570                 575
Ser Asn Lys Asp Gly Gly Tyr Phe Ser Ala Asn Trp Ala Leu Tyr Asp
            580                 585                 590
Ala Glu Leu Gln Leu Val Glu Leu Cys Arg Ser Ala Gly Val Lys Leu
        595                 600                 605
Arg Leu Phe His Gly Arg Gly Gly Thr Val Gly Arg Gly Gly Gly Pro
    610                 615                 620
Ser Tyr Asp Ala Ile Leu Ala Gln Pro Arg Gly Ala Val Gln Gly Ser
625                 630                 635                 640
Val Arg Ile Thr Glu Gln Gly Glu Ile Ile Ser Ala Lys Tyr Gly Asn
                645                 650                 655
Pro Glu Thr Ala Arg Arg Asn Leu Glu Ala Leu Val Ser Ala Thr Leu
            660                 665                 670
Glu Ala Ser Leu Leu Asp Val Ser Glu Leu Thr Asp His Gln Arg Ala
        675                 680                 685
Tyr Asp Ile Met Ser Glu Ile Ser Glu Leu Ser Leu Lys Lys Tyr Ala
    690                 695                 700
Ser Leu Val His Glu Asp Gln Gly Phe Ile Asp Tyr Phe Thr Gln Ser
705                 710                 715                 720
Thr Pro Leu Gln Glu Ile Gly Ser Leu Asn Ile Gly Ser Arg Pro Ser
                725                 730                 735
Ser Arg Lys Gln Thr Ser Ser Val Glu Asp Leu Arg Ala Ile Pro Trp
            740                 745                 750
Val Leu Ser Trp Ser Gln Ser Arg Val Met Leu Pro Gly Trp Phe Gly
        755                 760                 765
Val Gly Thr Ala Leu Glu Gln Trp Ile Gly Glu Gly Glu Gln Ala Thr
    770                 775                 780
Gln Arg Ile Ala Glu Leu Gln Thr Leu Asn Glu Ser Trp Pro Phe Phe
785                 790                 795                 800
Thr Ser Val Leu Asp Asn Met Ala Gln Val Met Ser Lys Ala Glu Leu
                805                 810                 815
Arg Leu Ala Lys Leu Tyr Ala Asp Leu Ile Pro Asp Arg Glu Val Ala
            820                 825                 830
Glu Arg Val Tyr Ala Val Ile Arg Glu Glu Tyr Phe Leu Thr Lys Lys
        835                 840                 845
Met Phe Cys Val Ile Thr Gly Ser Asp Asp Leu Leu Asp Asp Asn Pro
    850                 855                 860
Leu Leu Ala Arg Ser Val Gln Arg Arg Tyr Pro Tyr Leu Leu Pro Leu
865                 870                 875                 880
Asn Val Ile Gln Val Glu Met Met Arg Arg Tyr Arg Lys Gly Asp Gln
                885                 890                 895
Ser Glu Gln Val Ser Arg Asn Ile Gln Leu Thr Met Asn Gly Leu Ser
            900                 905                 910
Thr Ala Leu Arg Asn Ser Gly
        915
```

SEQ ID NO:17 的信息:

(i)序列特征:

(A)长度: 20 个碱基

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: 其它核酸

(A)描述: desc = "合成的 DNA"

(iv)反义: 否

(xi)序列描述: SEQ ID NO: 17:

CGCGAGGTAC CACCTGTCAC 20

(2)SEQ ID NO: 18 的信息:

(i)序列特征:

(A)长度: 20 个碱基

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: 其它核酸

(A)描述: desc = "合成的 DNA"

(iv)反义: 是

(xi)序列描述: SEQ ID NO: 18:

CAATCCAGGT ACCGGCAACC

(2)SEQ ID NO: 19 的信息:

(i)序列特征:

(A)长度: 23 个碱基

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: 其它核酸

(A)描述: desc = "合成的 DNA"

(iv)反义: 否

(xi)序列描述: SEQ ID NO: 19:

GGATCCCCAA TCGATACCTG GAA 23

(2)SEQ ID NO: 20 的信息:

(i)序列特征:

(A)长度: 23 个碱基

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: 其它核酸

(A)描述: desc = "合成的 DNA"

(iv)反义: 是

(xi)序列描述: SEQ ID NO: 20:

CGGTTCATCG CCAAGTTTTT CTT

(2)SEQ ID NO: 21 的信息:

(i)序列特征:

(A)长度: 23 个碱基

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: 其它核酸

(A)描述: desc = "合成的 DNA"

(iv)反义: 否

(xi)序列描述: SEQ ID NO: 21:

GTCGACGGAT CGCAAATGGC AAC

(2) SEQ ID NO: 22 的信息:

(i) 序列特征:

(A) 长度: 23 个碱基

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其它核酸

(A) 描述: dese = "合成的 DNA"

(iv) 反义: 是

(xi) 序列描述: SEQ ID NO: 22

GGATCCTTGA GCACCTTGCG CAG 23

(2) SEQ ID NO: 23 的信息:

(i) 序列特征:

(A) 长度: 20 个碱基

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 其它核酸

(A) 描述: dese = "合成的 DNA"

(iv) 反义: 否

(xi) 序列描述: SEQ ID NO: 23:

CATCTAAGTA TGCATCTCGG 20

(2) SEQ ID NO: 24 的信息:

(i) 序列特征:

(A) 长度: 20 个碱基

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii)分子类型: 其它核酸

(A)描述: desc =, "合成的 DNA"

(iv)反义: 是

(xi)序列描述: SEQID.NO: 24:

TGCCCCTCGA GCTAAATTAG 20

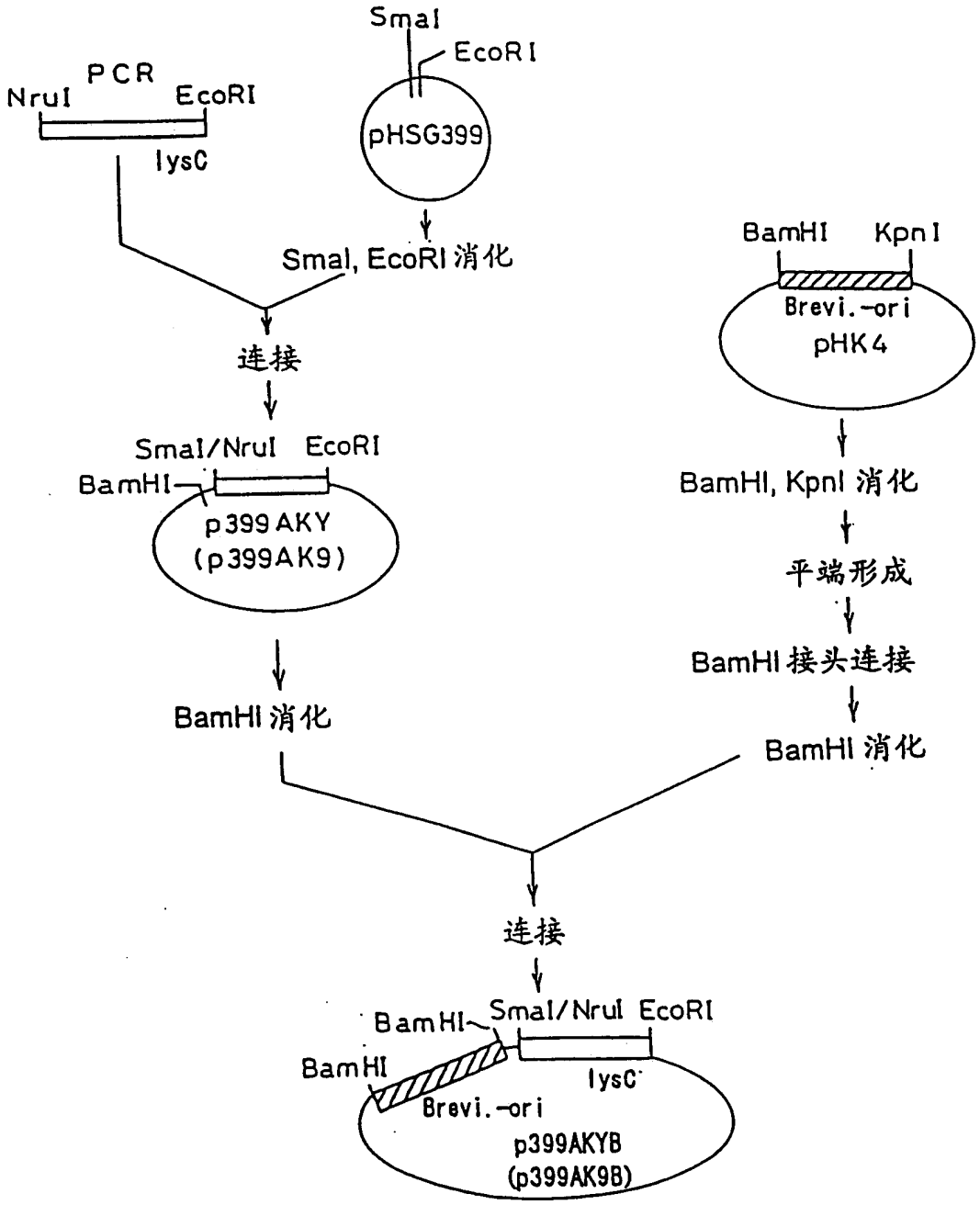

SmaI
EcoRI
PCR
NruI
EcoRI
lysC
pHSG399
SmaI, EcoRI 消化
连接
SmaI/NruI EcoRI
BamHI
p399AKY
(p399AK9)
BamHI 消化
BamHI KpnI
Brevi.-ori
pHK4
BamHI, KpnI 消化
平端形成
BamHI 接头连接
BamHI 消化
连接
BamHI
SmaI/NruI EcoRI
BamHI
lysC
Brevi.-ori
p399AKYB
(p399AK9B)

图 1

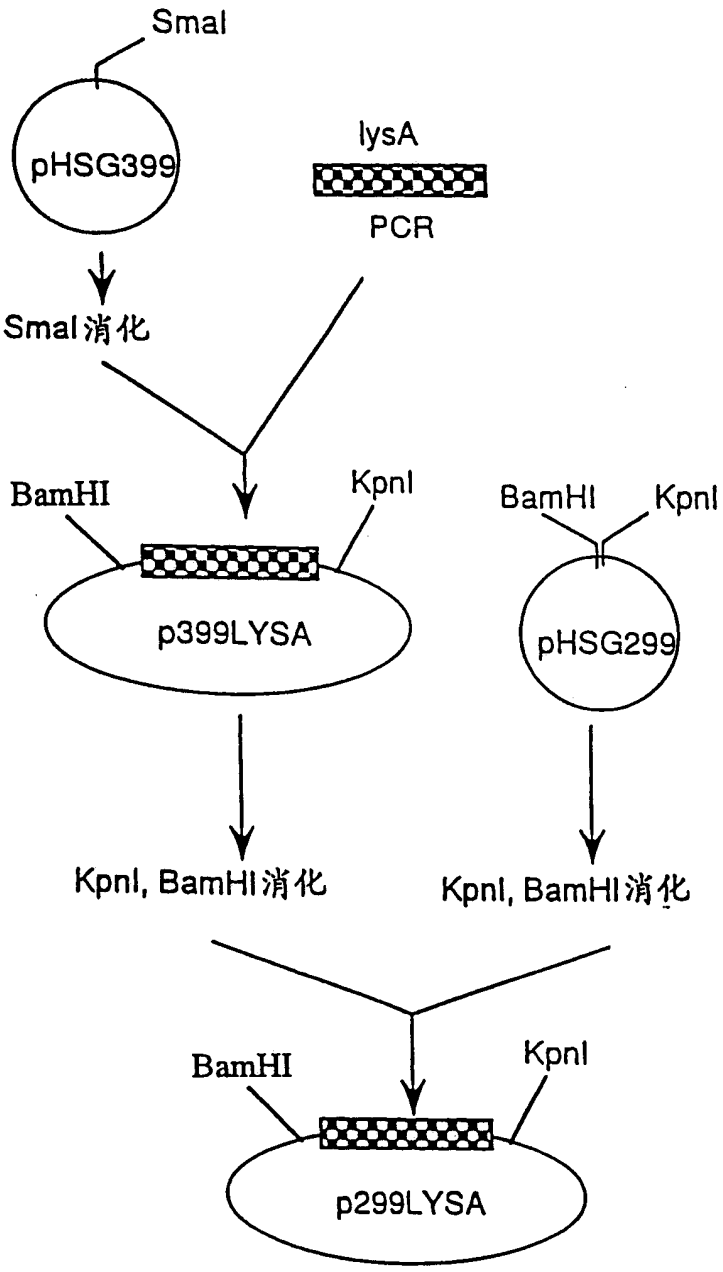

SmaI
pHSG399
lysA
PCR
SmaI消化
BamHI
KpnI
p399LYSA
BamHI
KpnI
pHSG299
KpnI, BamHI消化
KpnI, BamHI消化
BamHI
KpnI
p299LYSA

图 2

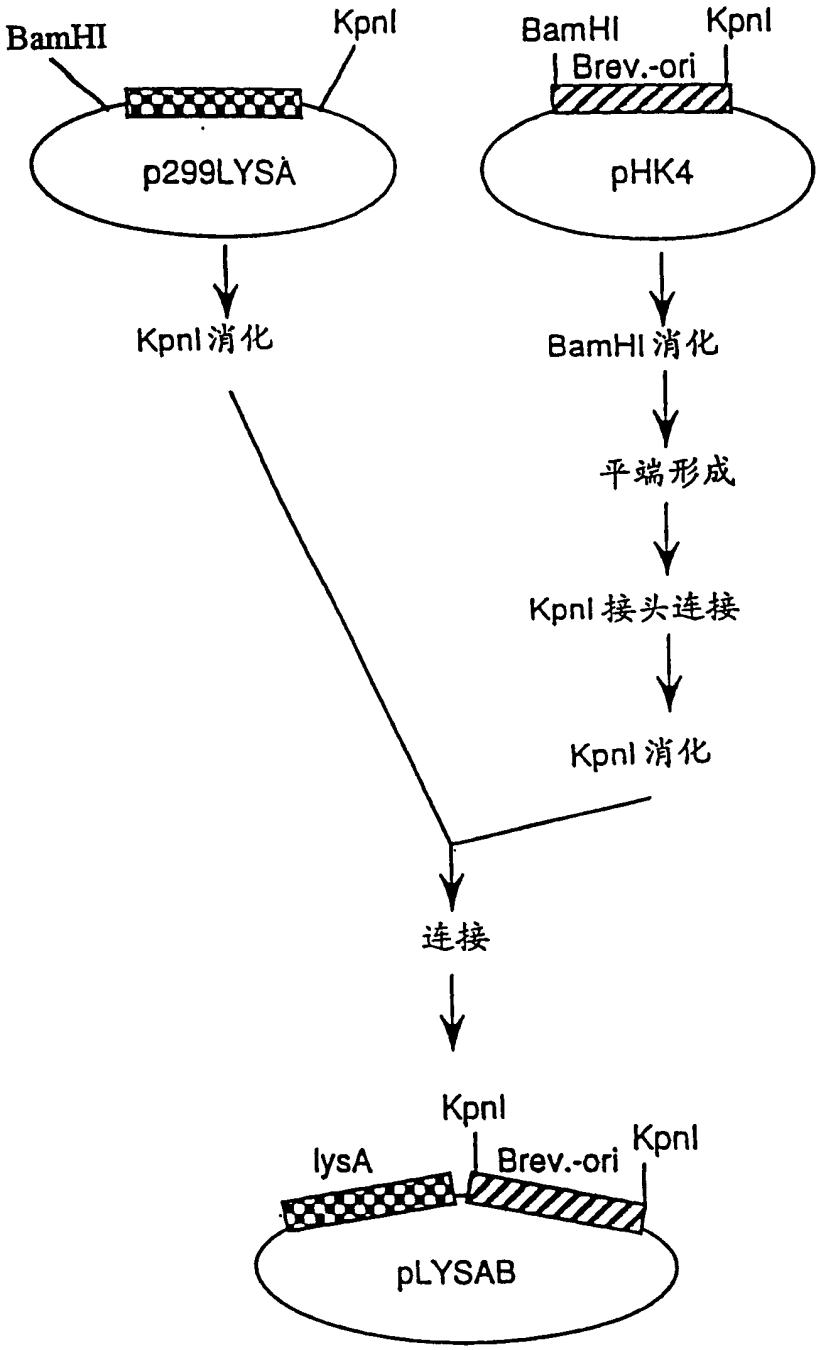

BamHI
KpnI
p299LYSA
KpnI消化
BamHI
KpnI
Brev.-ori
pHK4
BamHI消化
平端形成
KpnI接头连接
KpnI消化
连接
KpnI
lysA
Brev.-ori
KpnI
pLYSAB

图 3

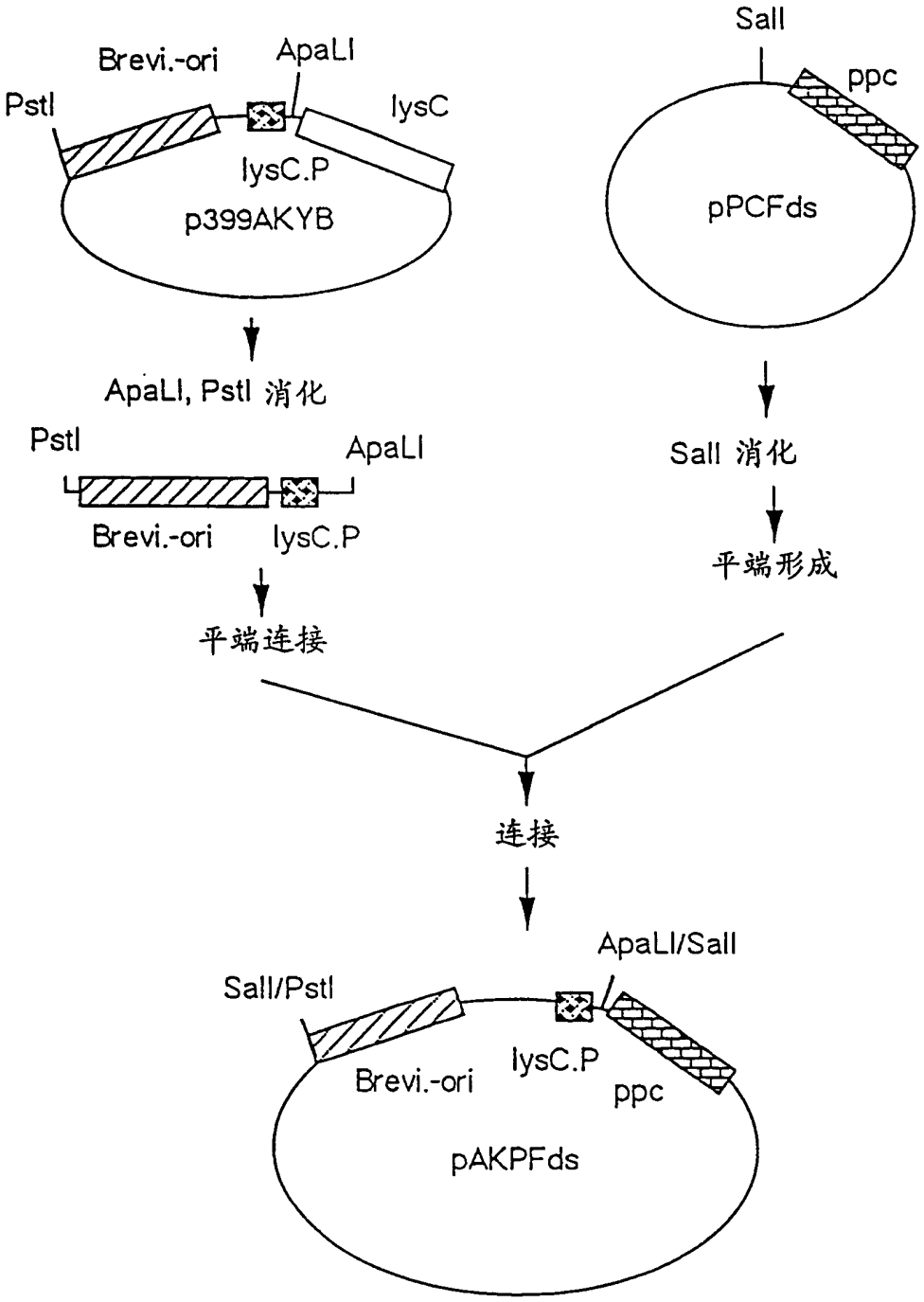

Brevi.-ori
ApaLI
PstI
lysC
lysC.P
p399AKYB
ApaLI, PstI 消化
PstI
ApaLI
Brevi.-ori
lysC.P
平端连接
SalI
ppc
pPCFds
SalI 消化
平端形成
连接
ApaLI/SalI
SalI/PstI
Brevi.-ori
lysC.P
ppc
pAKPFds

图 4

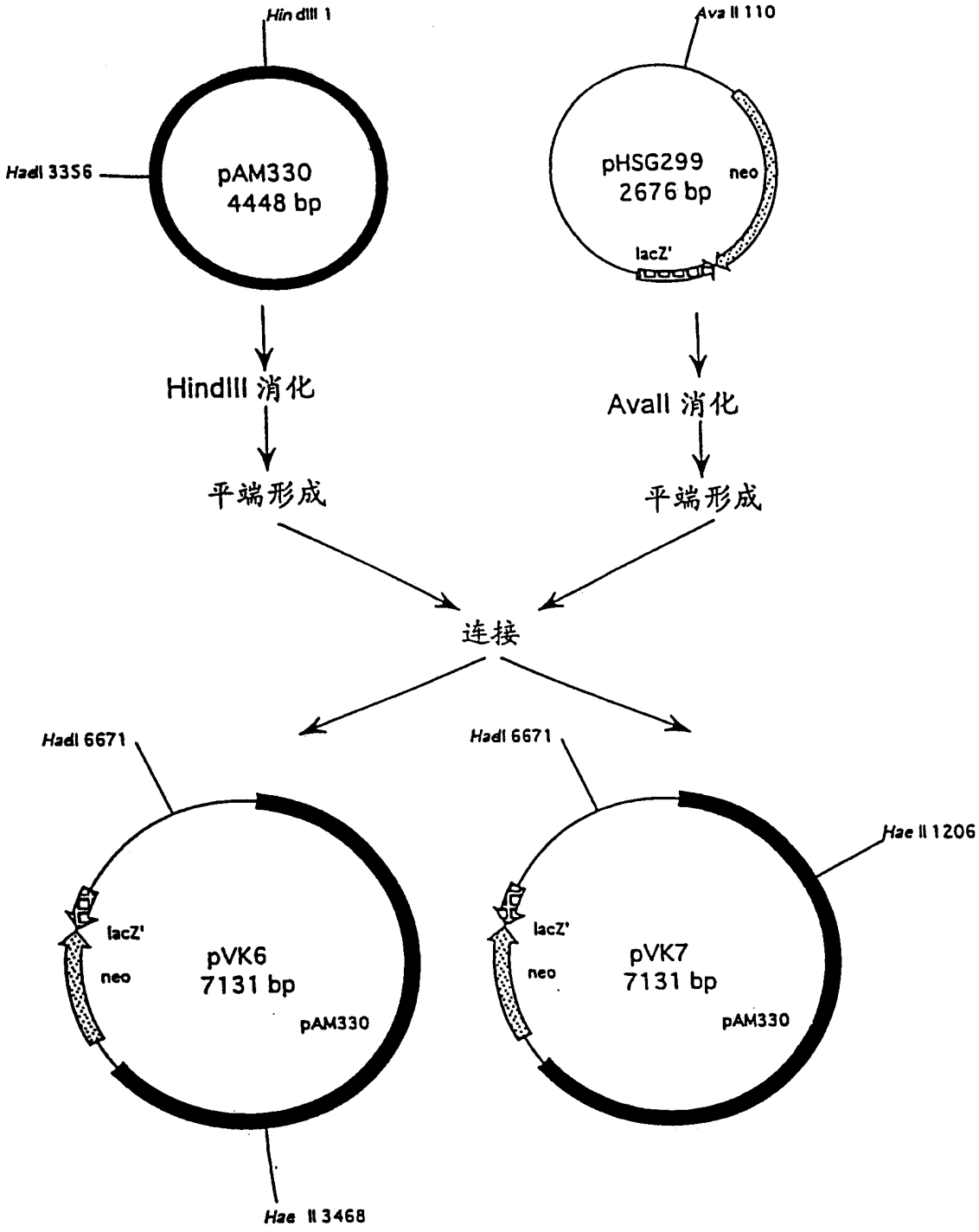

Hin dIII 1
HaeI 3356
pAM330
4448 bp
Ava II 110
pHSG299
2676 bp
neo
lacZ'
HindIII 消化
AvaII 消化
平端形成
平端形成
连接
HaeI 6671
HaeI 6671
Hae II 1206
lacZ'
neo
pVK6
7131 bp
pAM330
lacZ'
neo
pVK7
7131 bp
pAM330
Hae II 3468

图 5

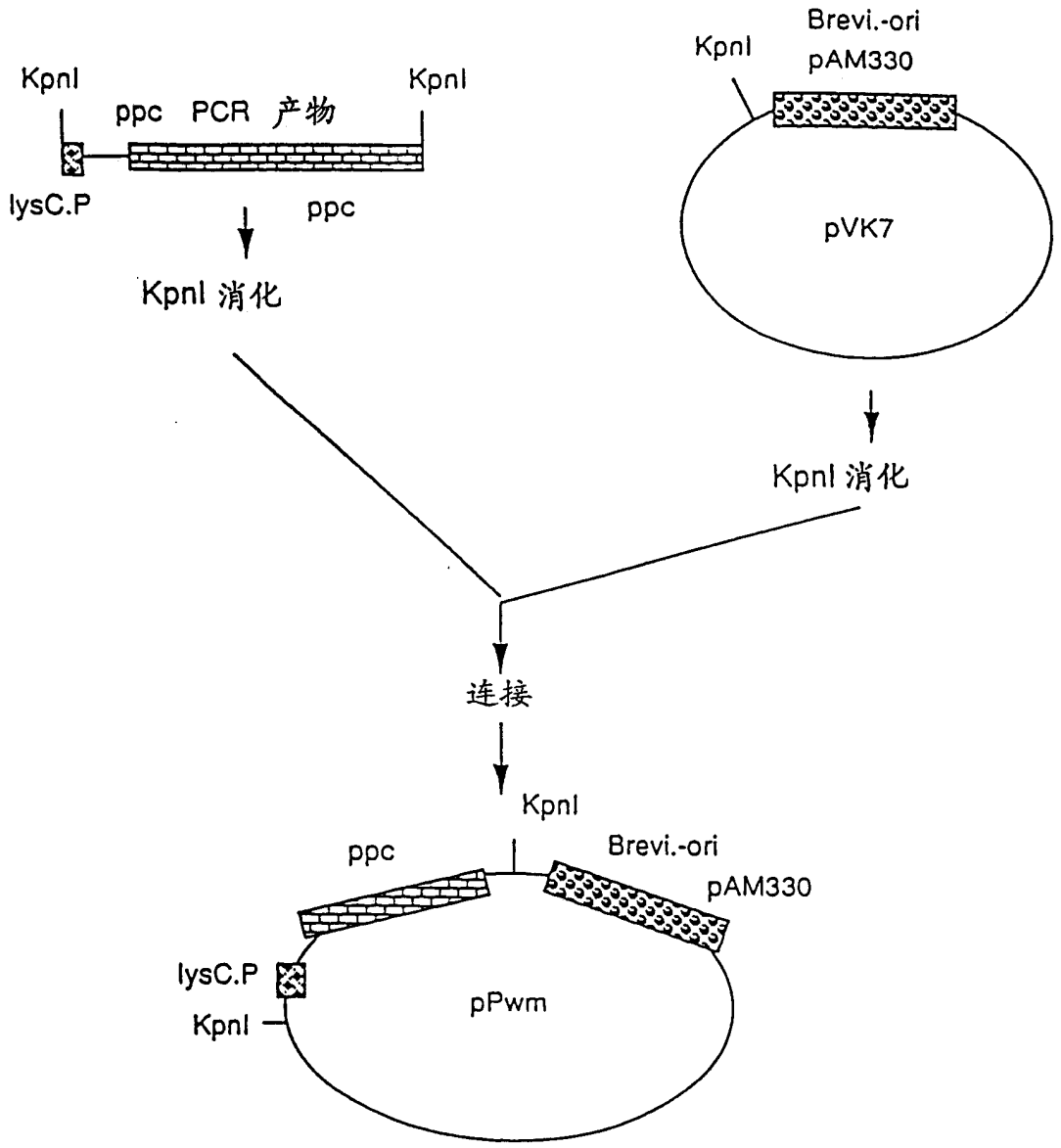

KpnI
KpnI
ppc PCR 产物
lysC.P
ppc
KpnI 消化
KpnI
Brevi.-ori
pAM330
pVK7
KpnI 消化
连接
KpnI
ppc
Brevi.-ori
pAM330
lysC.P
KpnI
pPwm

图 6

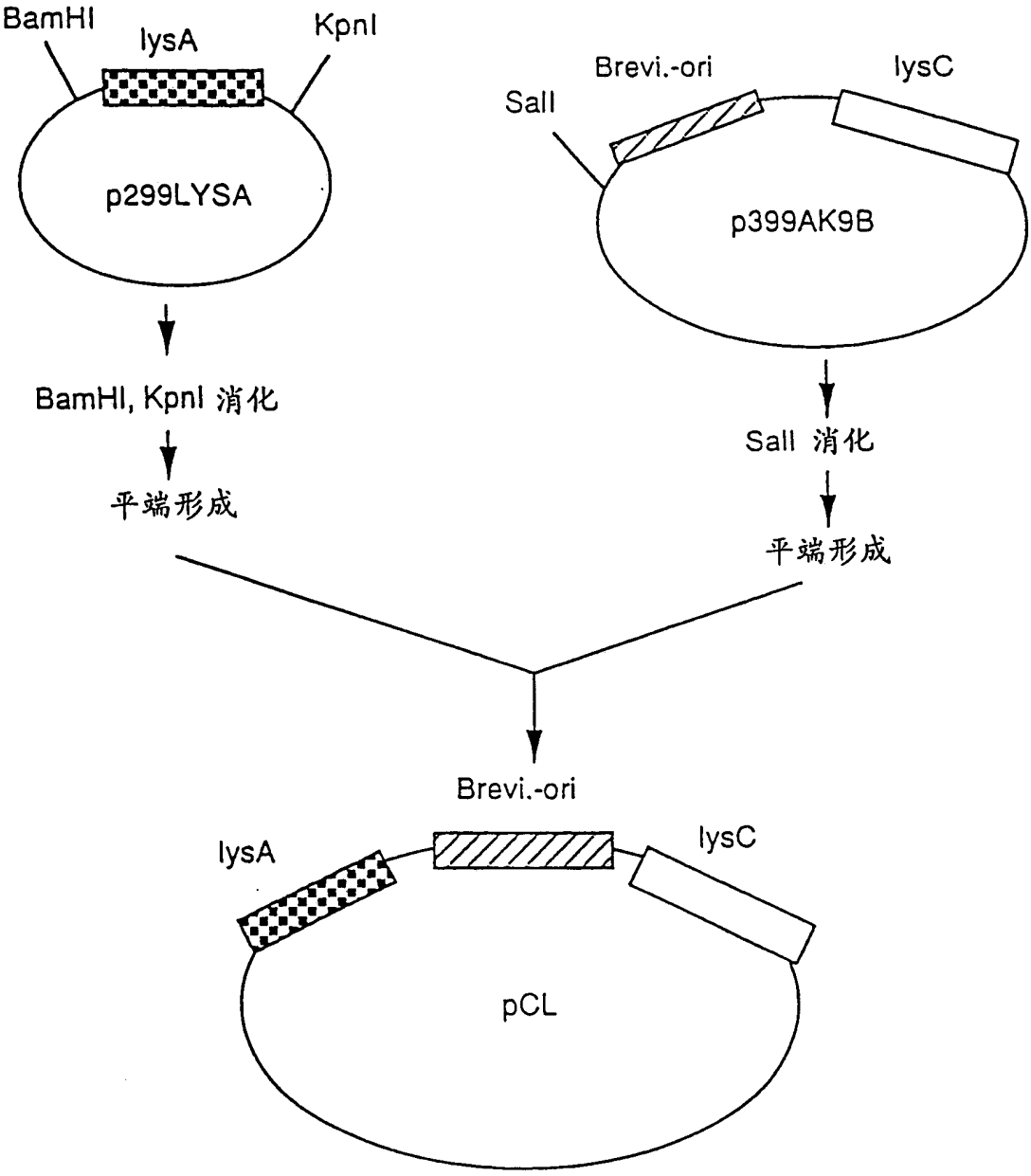

BamHI
lysA
KpnI
p299LYSA
BamHI, KpnI 消化
平端形成
Brevi.-ori
lysC
SalI
p399AK9B
SalI 消化
平端形成
Brevi.-ori
lysA
lysC
pCL

图 7

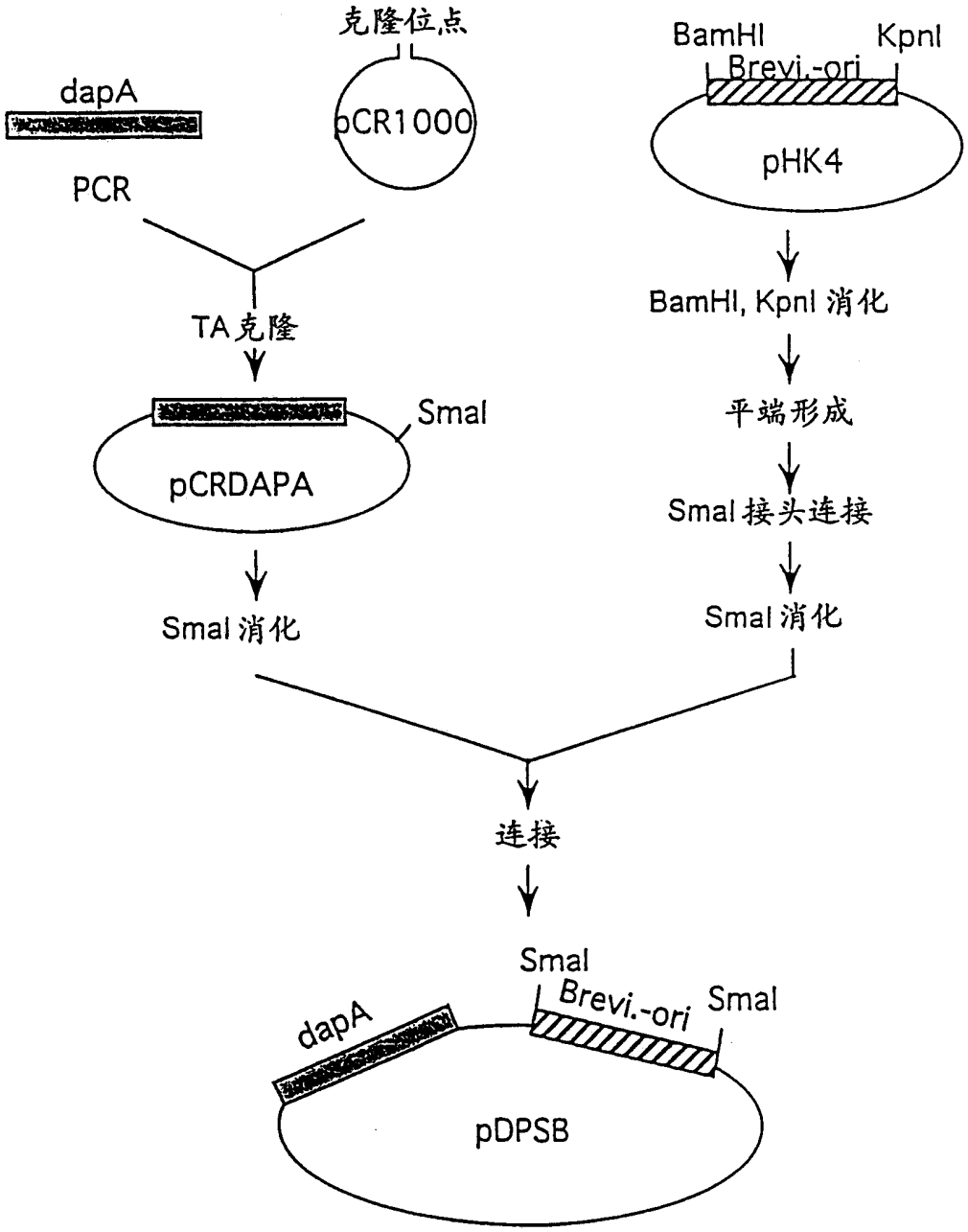

克隆位点
dapA
pCR1000
PCR
TA克隆
SmaI
pCRDAPA
SmaI消化
BamHI
KpnI
Brevi.-ori
pHK4
BamHI, KpnI消化
平端形成
SmaI接头连接
SmaI消化
连接
SmaI
Brevi.-ori
SmaI
dapA
pDPSB

图 8

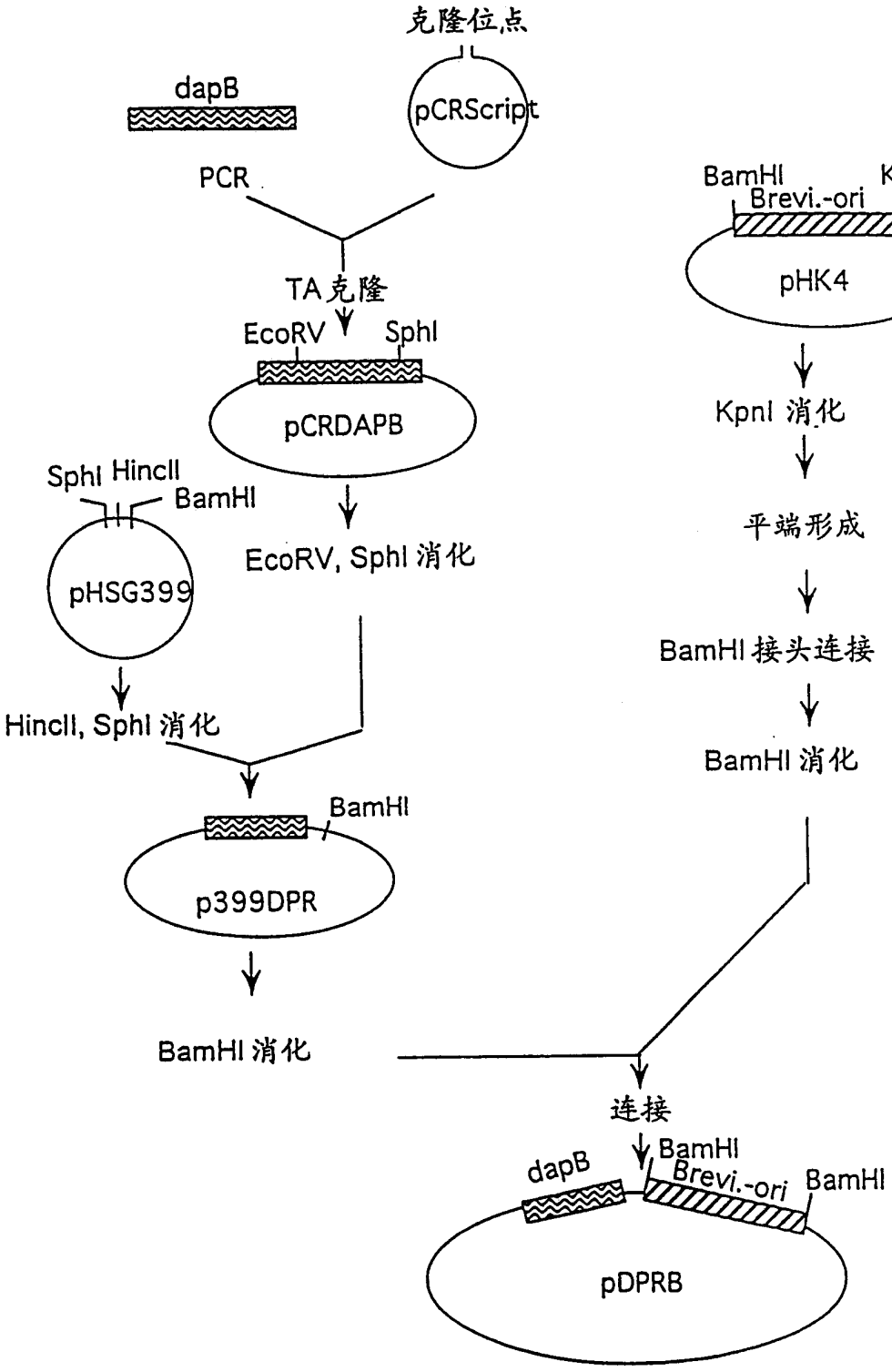

克隆位点
dapB
pCRScript
PCR
TA克隆
EcoRV
SphI
pCRDAPB
SphI HincII
BamHI
pHSG399
EcoRV, SphI 消化
HincII, SphI 消化
BamHI
p399DPR
BamHI 消化
BamHI
KpnI
Brevi.-ori
pHK4
KpnI 消化
平端形成
BamHI 接头连接
BamHI 消化
连接
BamHI
dapB
Brevi.-ori
BamHI
pDPRB

图 9

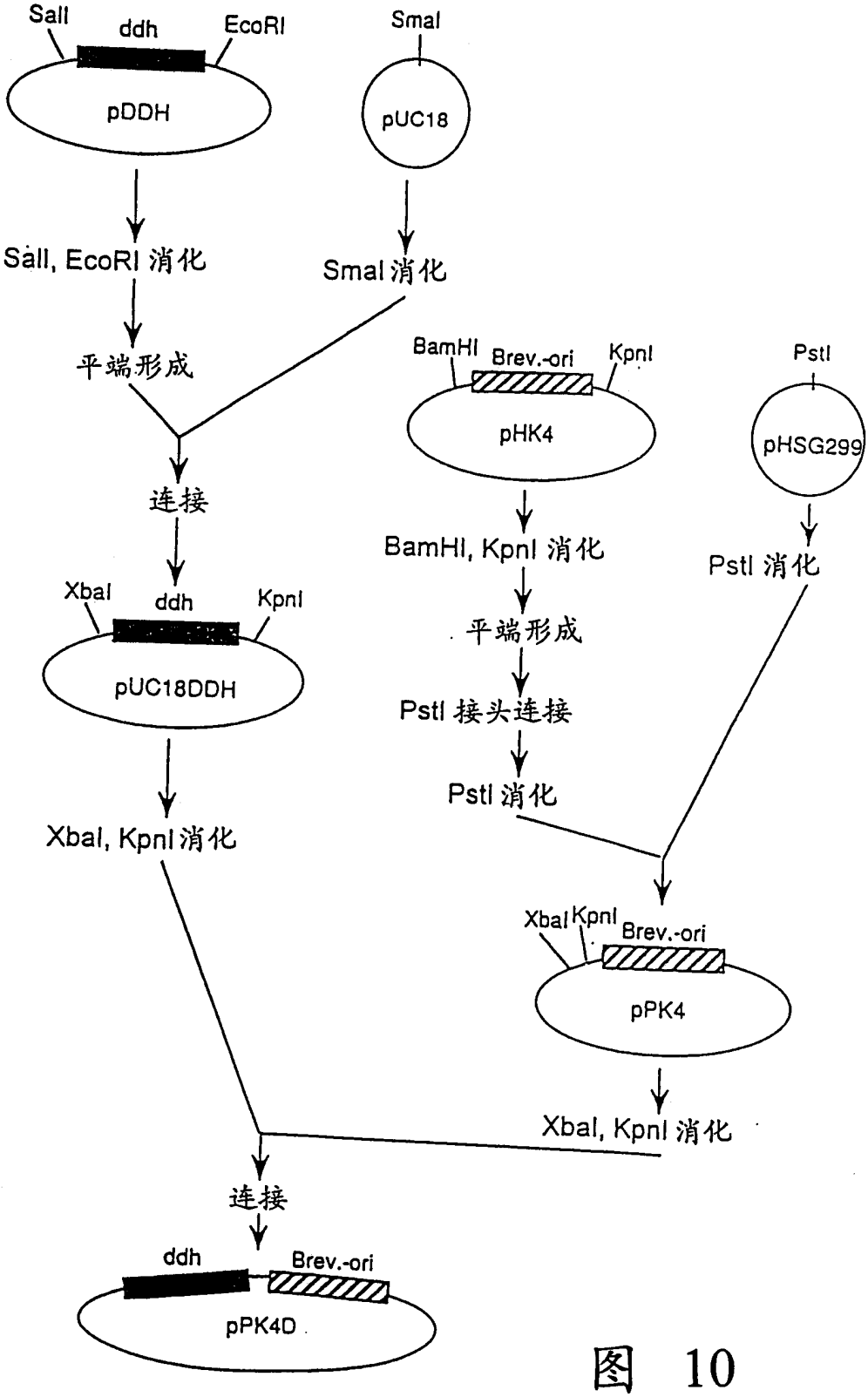

SalI
ddh
EcoRI
pDDH
SmaI
pUC18
SalI, EcoRI 消化
SmaI 消化
平端形成
连接
BamHI
Brev.-ori
KpnI
pHK4
PstI
pHSG299
BamHI, KpnI 消化
PstI 消化
平端形成
PstI 接头连接
PstI 消化
XbaI
ddh
KpnI
pUC18DDH
XbaI, KpnI 消化
XbaI
KpnI
Brev.-ori
pPK4
XbaI, KpnI 消化
连接
ddh
Brev.-ori
pPK4D

图 10

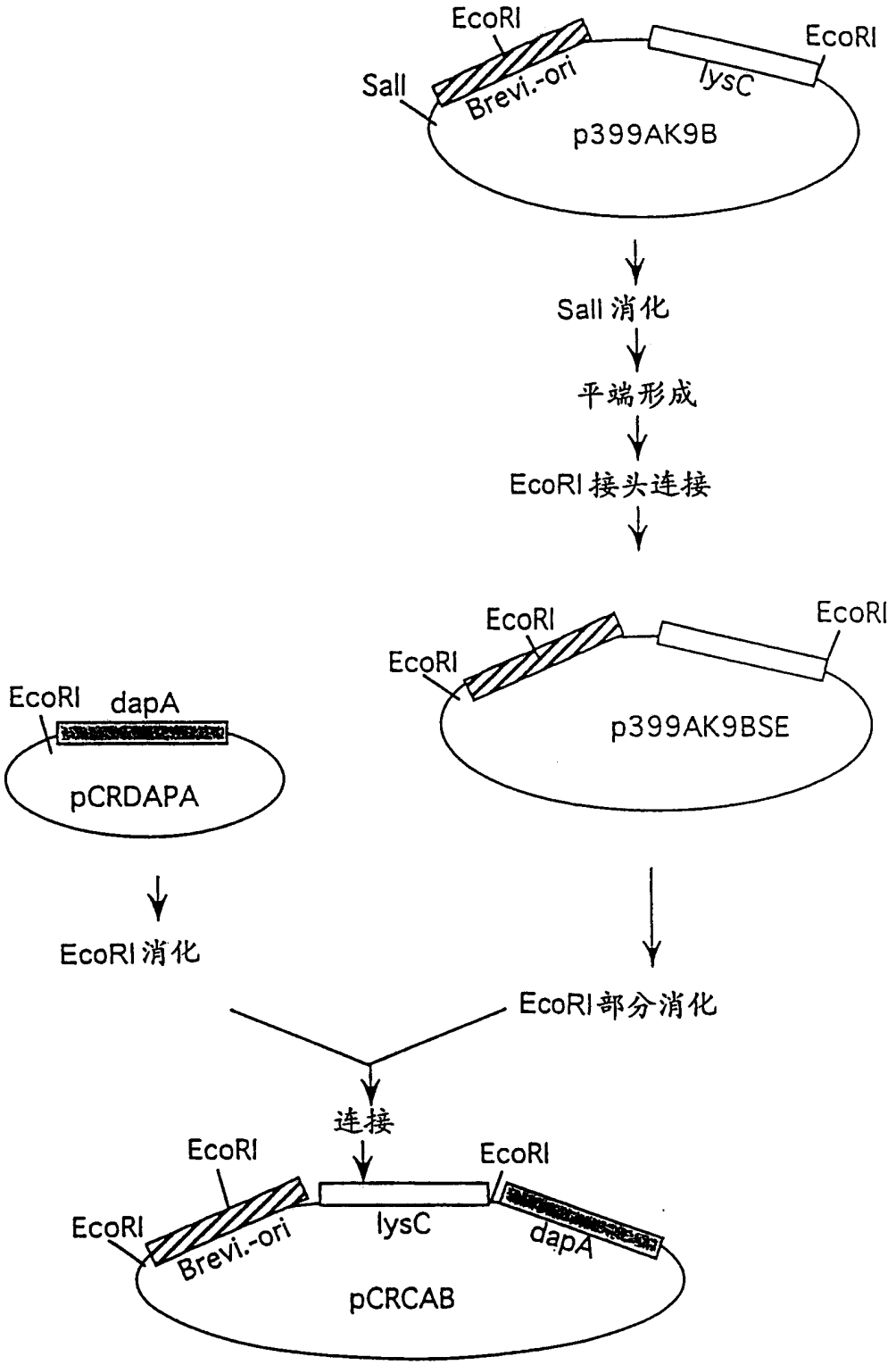

EcoRI
EcoRI
SalI
Brevi.-ori
lysC
p399AK9B
SalI 消化
平端形成
EcoRI 接头连接
EcoRI
EcoRI
EcoRI
EcoRI
p399AK9BSE
EcoRI
dapA
pCRDAPA
EcoRI 消化
EcoRI 部分消化
连接
EcoRI
EcoRI
EcoRI
Brevi.-ori
lysC
dapA
pCRCAB

图 11

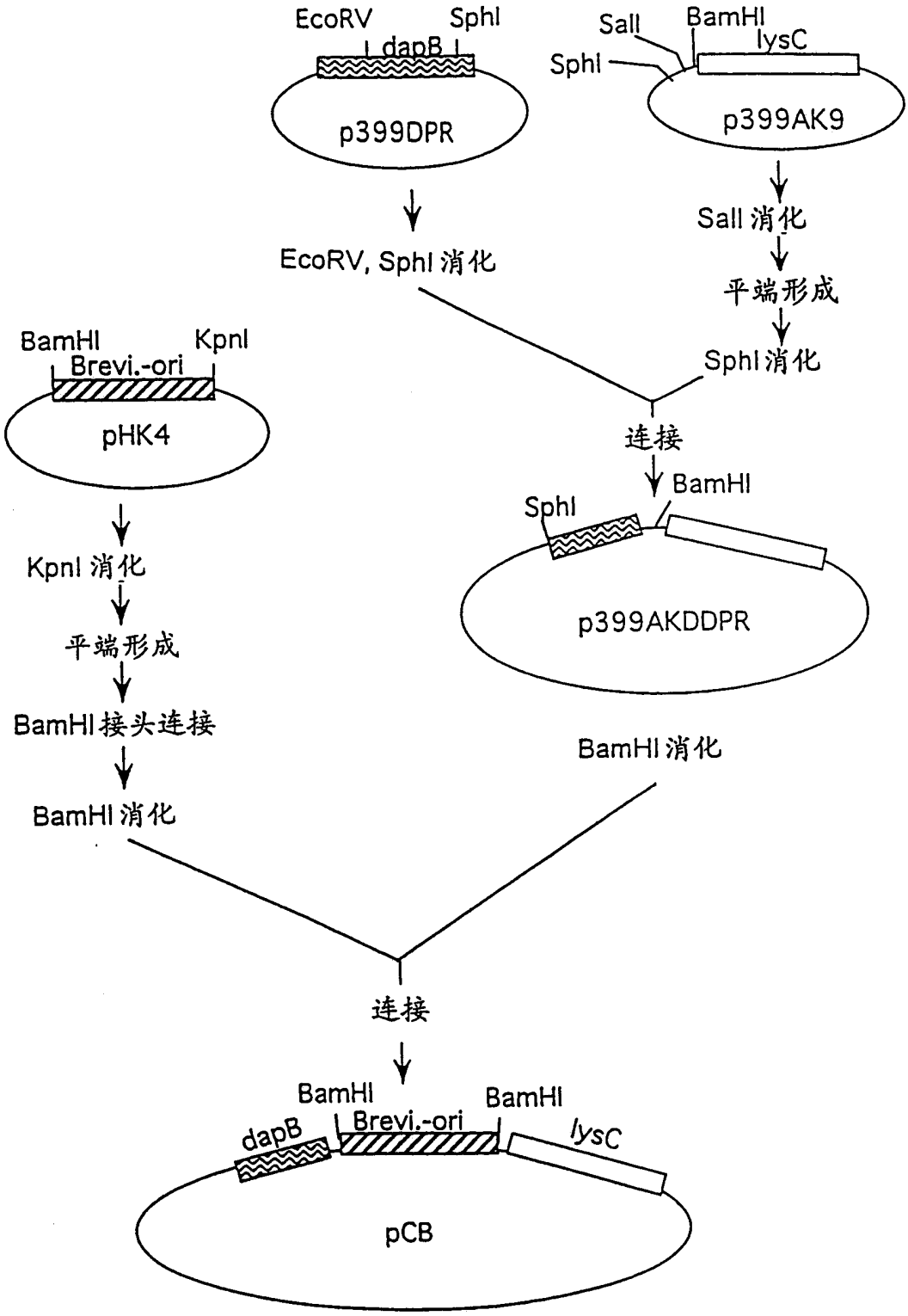

EcoRV
SphI
dapB
p399DPR
SalI
BamHI
lysC
SphI
p399AK9
SalI 消化
平端形成
SphI 消化
EcoRV, SphI 消化
连接
BamHI
SphI
p399AKDDPR
BamHI
KpnI
Brevi.-ori
pHK4
KpnI 消化
平端形成
BamHI 接头连接
BamHI 消化
BamHI 消化
连接
BamHI
BamHI
dapB
Brevi.-ori
lysC
pCB

图 12

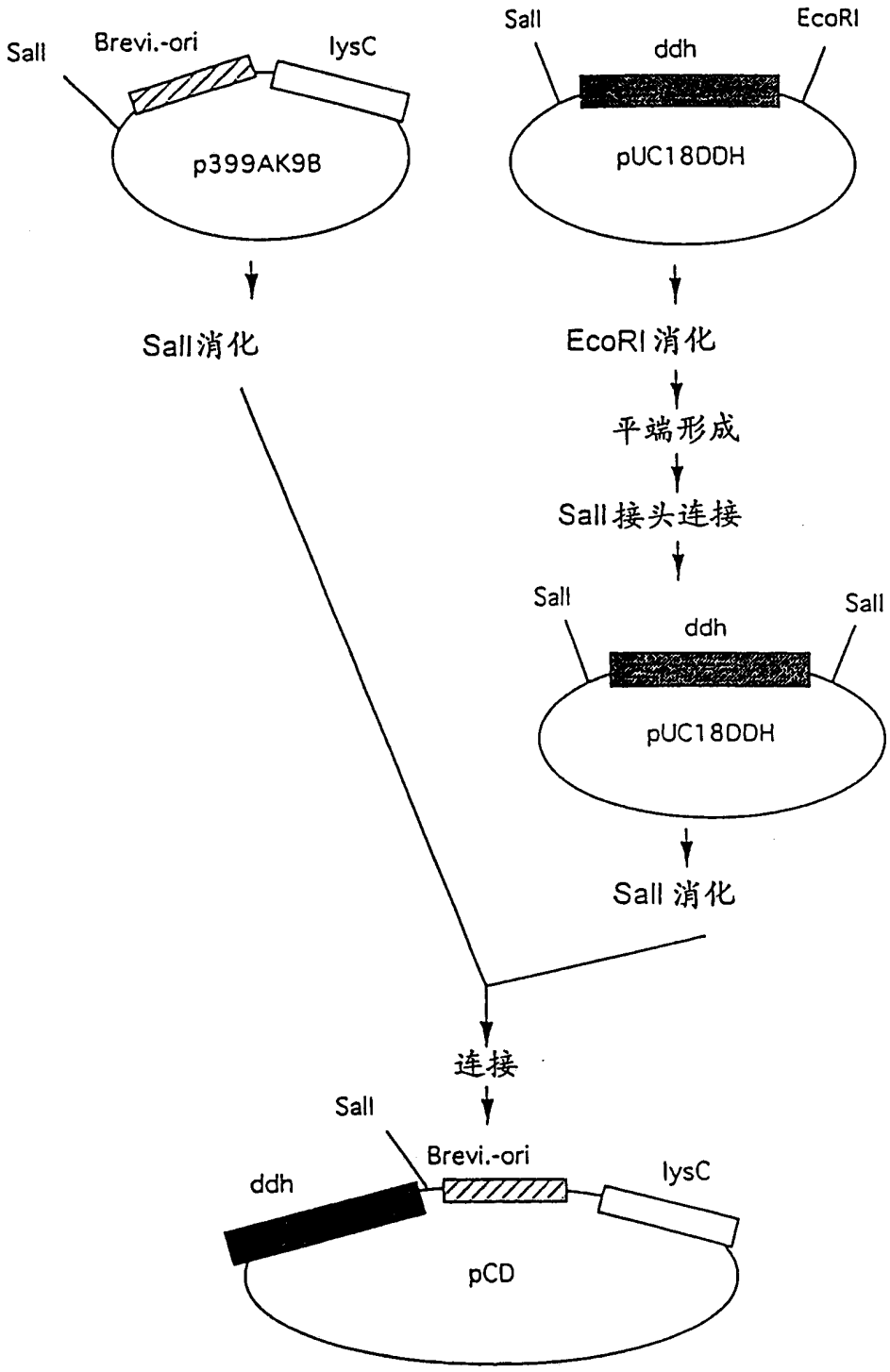

SalI
Brevi.-ori
lysC
p399AK9B
SalI
ddh
EcoRI
pUC18DDH
SalI消化
EcoRI消化
平端形成
SalI接头连接
SalI
ddh
SalI
pUC18DDH
SalI 消化
连接
SalI
Brevi.-ori
ddh
lysC
pCD

图 13