



(12) 发明专利

(10) 授权公告号 CN 102177246 B

(45) 授权公告日 2015.02.11

(21) 申请号 200980134943.0

C12N 9/04 (2006.01)

(22) 申请日 2009.09.04

C12N 15/09 (2006.01)

(30) 优先权数据

2008-229736 2008.09.08 JP

2009-032839 2009.02.16 JP

(56) 对比文件

WO 2006133898 A2, 2006.12.21,

BUJNICKI ET

(85) PCT国际申请进入国家阶段日

2011.03.08

AL.Fed-batch-Prozessentwicklung zur
Produktion und Gewinnung von Cyclitolen mit
rekombinanten Escherichia coli. 《CHEMIE
INGENIEUR TECHNIK》. 2004, 第 76 卷 1239.

(86) PCT国际申请的申请数据

PCT/JP2009/065475 2009.09.04

PANKE ET AL. Advances in biocatalytic
synthesis of pharmaceutical intermediates.
《CURRENT OPINION IN CHEMICAL BIOLOGY》. 2005,
第 9 卷 188-194.

(87) PCT国际申请的公布数据

W02010/027045 JA 2010.03.11

FRANKE ET AL. (S,S)-2,3-dihydro-2,3-
dihydrobenzoic acid: Microbial access
with engineered cells of Escherichia coli
and application as starting material in
natural-product synthesis. 《CHEMISTRY - A
EUROPEAN JOURNAL》. 2003, 第 9 卷 4188-4196.

(83) 生物保藏信息

VKPM B-9246 2005.09.21

FERM BP-7207 1999.08.18

FERM BP-11091 2009.02.04

审查员 徐莉

(73) 专利权人 味之素株式会社

地址 日本东京都

(72) 发明人 泷川里绘 原吉彦 野中源

宅见和浩

(74) 专利代理机构 北京市柳沈律师事务所

11105

代理人 罗天乐

(51) Int. Cl.

C12P 13/04 (2006.01)

权利要求书1页 说明书35页

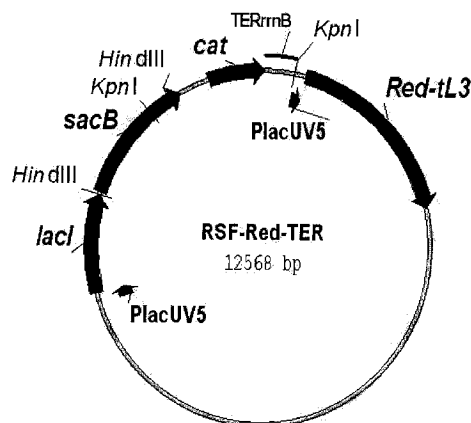
序列表24页 附图3页

(54) 发明名称

生产 L-氨基酸的微生物和 L-氨基酸的生产
方法

(57) 摘要

用培养基培养属于肠杆菌科、且具有 L-氨基
酸生产能力的细菌,在培养物中生产并蓄积 L-氨基
酸,并从该培养物收集 L-氨基酸;其中,所述细菌
是固有地具备以吡咯并喹啉醌为辅酶的葡萄糖
脱氢酶的活性、但经过修饰而降低了该酶的活性
的细菌。



1. 一种生产 L-氨基酸的方法,其特征在于在培养基中培养属于肠杆菌科且具有 L-氨基酸生产能力的细菌,从而在培养物中生产并蓄积 L-氨基酸,并从该培养物收集 L-氨基酸;其中,所述细菌是固有地具备以吡咯并喹啉醌为辅酶的膜结合型葡萄糖脱氢酶的活性,但经过修饰而降低了该葡萄糖脱氢酶的活性的细菌,其中所述细菌是属于泛菌属的细菌 *Pantoea ananatis*。

2. 根据权利要求 1 所述的方法,其中,通过使编码所述葡萄糖脱氢酶 (GCD) 的 *gcd* 基因失活而降低了该葡萄糖脱氢酶的活性。

3. 根据权利要求 2 所述的方法,其中,所述 *gcd* 基因是编码 SEQ ID NO:2 的氨基酸序列的 DNA。

4. 根据权利要求 1 所述的方法,其中,所述 L-氨基酸是选自 L-谷氨酸、L-赖氨酸、L-苏氨酸、L-精氨酸、L-组氨酸、L-异亮氨酸、L-缬氨酸、L-亮氨酸、L-苯丙氨酸、L-酪氨酸、L-色氨酸和 L-半胱氨酸中的一种或二种以上的 L-氨基酸。

5. 根据权利要求 4 所述的方法,其中,所述 L-氨基酸是 L-谷氨酸或 L-半胱氨酸。

6. 根据权利要求 5 所述的方法,其中,所述 L-氨基酸是 L-谷氨酸,且所述细菌中选自柠檬酸合酶、甲基柠檬酸合酶、磷酸烯醇丙酮酸羧化酶和谷氨酸脱氢酶中的 1 种或 2 种以上的酶的活性得到了增强。

7. 根据权利要求 5 所述的方法,其中,所述 L-氨基酸是 L-半胱氨酸,且所述细菌中选自 3-磷酸甘油酸脱氢酶、丝氨酸乙酰转移酶和硫酸盐 / 硫代硫酸盐输送系统中的至少 1 种或 2 种以上的活性和 / 或 *yeaS* 基因的表达得到了增强。

生产 L- 氨基酸的微生物和 L- 氨基酸的生产方法

技术领域

[0001] 本发明涉及使用微生物的 L- 氨基酸生产方法,特别是 L- 谷氨酸、L- 赖氨酸、L- 苏氨酸、L- 色氨酸等 L- 氨基酸的生产方法。L- 谷氨酸作为调味品、L- 赖氨酸、L- 苏氨酸和 L- 色氨酸作为动物饲料用添加剂、健康食品成分或氨基酸输液等,在产业上是有用的 L- 氨基酸。

背景技术

[0002] L- 氨基酸是通过使用各种微生物的发酵方法来进行工业生产的。例如, L- 谷氨酸主要是通过使用属于短杆菌属 (*Brevibacterium*)、棒杆菌属 (*Corynebacterium*)、分枝杆菌属 (*Mycobacterium*) 的称为棒状杆菌型细菌中的 L- 谷氨酸生产菌或其突变株的发酵法来进行生产(参照例如非专利文献 1)。作为通过使用其它微生物的发酵法进行的 L- 谷氨酸生产方法,已知有使用芽孢杆菌属 (*Bacillus*)、链霉菌属 (*Streptomyces*)、青霉菌属 (*Penicillium*) 等的微生物(参照例如专利文献 1),假单胞菌属 (*Pseudomonas*)、节杆菌属 (*Arthrobacter*)、沙雷氏菌属 (*Serratia*)、假丝酵母属 (*Candida*) 等的微生物(参照例如专利文献 2),芽孢杆菌属、假单胞菌属、沙雷氏菌属、产气气杆菌 (*Aerobacter aerogenes*) (现称产气肠杆菌 (*Enterobacter aerogenes*)) 等的微生物(参照例如专利文献 3),大肠杆菌突变株(参照例如专利文献 1) 等的方法。此外,还公开了使用属于克雷伯氏菌属 (*Klebsiella*)、欧文氏菌属 (*Erwinia*) 或泛菌属 (*Pantoea*)、肠杆菌属的微生物的 L- 谷氨酸生产方法(参照例如专利文献 2 ~ 4)。

[0003] 为了通过诸如上述的使用微生物的发酵法生产 L- 氨基酸等目的物质,可以采用使用野生型微生物(野生株)的方法、使用从野生株衍生的营养缺陷型株的方法、使用从野生株作为各种药物抗性突变株衍生的代谢调节突变株的方法、使用具有营养缺陷型株和代谢调节突变株这两者的性质的菌株的方法等。

[0004] 而且,近年来在目的物质的发酵生产中采用了重组 DNA 技术。例如,通过增强 L- 氨基酸生物合成体系酶编码基因的表达(专利文献 5、专利文献 6)、或增强 L- 氨基酸生物合成体系的碳源流入(专利文献 7) 来提高微生物的 L- 氨基酸生产力。

[0005] 葡萄糖脱氢酶大体上可分为 NAD(P) 依赖型和 PQQ(吡咯并喹啉醌) 依赖型。进一步,已知 PQQ 依赖型 (EC1. 1. 5. 2) 又包括可溶型和膜结合型,且后者在肠杆菌科细菌中存在于周质空间(外膜、内膜间的空隙),且在肠杆菌科细菌中广泛存在。以下,将这种存在于周质空间的以 PQQ 为辅酶的葡萄糖脱氢酶也记作“GCD”。

[0006] 已知大肠杆菌等一部分细菌不具有 GCD 辅酶 PQQ 的合成能力,只有通过添加 PQQ 才会表现 GCD 活性(非专利文献 2)。另一方面,泛菌属细菌等细菌具有 PQQ 合成能力,其具有 GCD 全酶。

[0007] 作为涉及 GCD 的技术,已知有利用 GCD 编码基因被缺损的木葡糖杆菌 (*Gluconobacter xylinus*) 从葡萄糖生产纤维素的方法(J. Biosci. Bioeng. 99(4), 415-422, 2005)。此外,还已知利用葡萄糖脱氢酶活性降低的埃希氏菌属细菌、或固有缺失葡萄糖脱

氢酶活性的埃希氏菌属细菌生产 [5S,6S]-5,6-二羟基-1,3-环己二烯-1-羧酸的方法(专利文献 8)。

[0008] 然而,尚不清楚细菌 GCD 活性的降低对 L-氨基酸生产能力的影响。

[0009] 现有技术文献

[0010] 专利文献

[0011] 专利文献 1 :特开平 5-244970 号公报

[0012] 专利文献 2 :美国专利第 3,563,857 号说明书

[0013] 专利文献 3 :特公昭 32-9393 号公报

[0014] 专利文献 4 :特开 2000-189175 号公报

[0015] 专利文献 5 :美国专利第 5168056 号说明书

[0016] 专利文献 6 :美国专利第 5776736 号说明书

[0017] 专利文献 7 :美国专利第 5906925 号说明书

[0018] 专利文献 8 :国际公开第 W02006/133898 号小册子

[0019] 非专利文献

[0020] 非专利文献 1 :明石邦彦等著,アミノ酸発酵,学会出版中心,195 ~ 215 页,1986 年

[0021] 非专利文献 2 :FEMS Microbiol. Lett., 24, 329-333, 1984

发明内容

[0022] 发明所要解决的问题

[0023] 本发明的课题是 :提供提供能够高效生产 L-氨基酸的属于肠杆菌科的微生物,以及使用该微生物高效生产 L-氨基酸的方法。

[0024] 解决问题的方法

[0025] 本发明人等为了解决上述问题进行了深入研究,结果发现 :通过对固有地具备 GCD 活性的肠杆菌科细菌进行修饰以降低 GCD 活性,能够提高 L-氨基酸生产能力,从而完成了本发明。

[0026] 即,本发明如下。

[0027] (1) 一种生产 L-氨基酸的方法,其中,在培养基中培养属于肠杆菌科且具有 L-氨基酸生产能力的细菌,从而在培养物中生产并蓄积 L-氨基酸,并从该培养物收集 L-氨基酸 ;其中,所述细菌是固有地具备以吡咯并喹啉醌为辅酶的葡萄糖脱氢酶的活性、但经过修饰而降低了该酶的活性的细菌。

[0028] (2) 前述方法,其中,通过灭活编码所述酶的 gcd 基因而降低了 GCD 活性。

[0029] (3) 前述方法,其中,所述 gcd 基因是编码 SEQ ID NO :2 的氨基酸序列的 DNA 或其变体。

[0030] (4) 前述方法,其中,所述 L-氨基酸选自 L-谷氨酸、L-赖氨酸、L-苏氨酸、L-精氨酸、L-组氨酸、L-异亮氨酸、L-缬氨酸、L-亮氨酸、L-苯丙氨酸、L-酪氨酸、L-色氨酸以及 L-半胱氨酸。

[0031] (5) 前述方法,其中,所述 L-氨基酸是 L-谷氨酸或 L-半胱氨酸。

[0032] (6) 前述方法,其中,所述 L-氨基酸是 L-谷氨酸,且所述细菌中选自柠檬酸合酶、

甲基柠檬酸合酶、磷酸烯醇丙酮酸羧化酶和谷氨酸脱氢酶中的 1 种或 2 种以上的酶的活性得到了增强。

[0033] (7) 所述方法, 其中, 所述 L- 氨基酸是 L- 半胱氨酸, 且所述细菌中选自 3- 磷酸甘油酸脱氢酶、丝氨酸乙酰转移酶、硫酸盐 / 硫代硫酸盐输送系统中的至少 1 种或 2 种以上的活性, 和 / 或 yeaS 基因的表达得到了增强。

[0034] (8) 前述方法, 其中, 所述细菌是属于选自泛菌属、肠杆菌属、欧文氏菌属、克雷伯氏菌属、普罗威登斯菌属、沙门氏菌属、沙雷氏菌属、摩根氏菌属和耶尔森菌属中的属的细菌。

[0035] 发明的效果

[0036] 通过使用本发明的微生物, 能够高效地发酵生产 L- 谷氨酸、L- 赖氨酸、L- 苏氨酸、L- 精氨酸、L- 组氨酸、L- 异亮氨酸、L- 缬氨酸、L- 亮氨酸、L- 苏氨酸、L- 苯丙氨酸、L- 酪氨酸、L- 色氨酸、或 L- 半胱氨酸等 L- 氨基酸。

附图说明

[0037] [图 1] 显示辅助质粒 RSF-Red-TER 的结构图。

[0038] [图 2] 显示辅助质粒 RSF-Red-TER 的构建图。

[0039] [图 3] 显示启动子 Pnlp 的序列图。

[0040] 发明的具体实施方式

[0041] 以下, 对本发明进行具体说明。

[0042] <1> 本发明中使用的属于肠杆菌科的细菌

[0043] 本发明中使用的细菌是固有地具备 GCD 活性、且具有 L- 氨基酸生产能力的属于肠杆菌科的细菌, 且该细菌经过了修饰而降低了 GCD 活性。本发明的细菌可以通过对固有地具备 GCD 活性、且具有 L- 氨基酸生产能力的属于肠杆菌科的细菌修饰而降低 GCD 活性来获得。此外, 本发明的细菌还可以通过对固有地具备 GCD 活性、但经过了修饰而降低了 GCD 活性的属于肠杆菌科的细菌赋予 L- 氨基酸生产能力、或者增强所述细菌的 L- 氨基酸生产能力来获得。

[0044] 对于 L- 氨基酸的种类没有特殊限制, 可以列举出 L- 赖氨酸、L- 鸟氨酸、L- 精氨酸、L- 组氨酸、L- 瓜氨酸等碱性氨基酸, L- 异亮氨酸、L- 丙氨酸、L- 缬氨酸、L- 亮氨酸、L- 甘氨酸等脂肪族氨基酸, L- 苏氨酸、L- 丝氨酸等羟基单氨基羧酸形式的氨基酸, L- 脯氨酸等环式氨基酸, L- 苯丙氨酸、L- 酪氨酸、L- 色氨酸等芳香族氨基酸, L- 半胱氨酸、L- 胱氨酸、L- 甲硫氨酸等含硫氨基酸, L- 谷氨酸、L- 天冬氨酸、L- 谷氨酰胺、L- 天冬酰胺等酸性氨基酸, 特别优选 L- 谷氨酸、L- 赖氨酸、L- 苏氨酸、L- 色氨酸。本发明的微生物可以具有 2 种以上氨基酸的生产能力。

[0045] 此外, 在本发明中, L- 氨基酸包括游离体的 L- 氨基酸和 / 或其盐, 例如硫酸盐、盐酸盐、碳酸盐、铵盐、钠盐、钾盐。

[0046] 在本发明中, 具有 L- 氨基酸生产能力的细菌, 是指具有当在培养基中培养时, 产生并向培养基中分泌 L- 氨基酸的能力的细菌。此外, 优选的是能够使培养基中蓄积优选为 0.5g/L 以上、更优选为 1.0g/L 以上的量的目的 L- 氨基酸的细菌。

[0047] 以下, 将用于进行修饰以降低 GCD 活性的、作为本发明的细菌的亲本株使用的细

菌以及赋予或增强 L- 氨基酸生产能力的方法示例如下。

[0048] <2-1> 本发明的细菌

[0049] 本发明的细菌是属于肠杆菌科的细菌。

[0050] 肠杆菌科中包括属于埃希氏菌 (*Escherichia*)、肠杆菌 (*Enterobacter*)、欧文氏菌 (*Erwinia*)、克雷伯氏菌 (*Klebsiella*)、泛菌 (*Pantoea*)、光状杆菌 (*Photobacterium*)、普罗威登斯菌 (*Providencia*)、拉乌尔菌 (*Raoultella*)、沙门氏菌 (*Salmonella*)、沙雷氏菌 (*Serratia*)、志贺氏菌 (*Shigella*)、摩根氏菌 (*Morganella*)、耶尔森菌 (*Yersinia*) 等属的细菌。特别是优选根据 NCBI (National Center for Biotechnology Information) 的数据库 (<http://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=91347>) 中使用的分类法被分类为肠杆菌科的细菌。

[0051] 所谓属于埃希氏菌属的细菌没有特别的限制,是指该细菌根据微生物学家所知的分类被归类为埃希氏菌属。作为属于埃希氏菌属的细菌的例子,包括但不限于大肠杆菌 (*E. coli*)。

[0052] 属于埃希氏菌属的细菌例如包括 Neidhardt 等的著作 (Neidhardt, F. C. Ed. 1996. *Escherichia coli and Salmonella: Cellular and Molecular Biology/Second Edition* pp. 2477-2483. Table 1. American Society for Microbiology Press, Washington, D. C.) 中记述的种系的微生物。具体而言,可以列举源自原型野生株 K12 株的大肠杆菌 W3110 (ATCC 27325)、大肠杆菌 MG1655 (ATCC47076) 等。

[0053] 这些菌株可以从例如美国典型培养物保藏中心 (地址 P. O. Box 1549 Manassas, VA 20108, United States of America) 通过分售获得。即,对各菌株给予了对应的登录号,可以利用该登录号通过分售获得。各菌株的相应登录号见美国典型培养物保藏中心的目录。

[0054] 作为肠杆菌属细菌,可以列举成团肠杆菌 (*Enterobacter agglomerans*)、产气肠杆菌 (*Enterobacter aerogenes*) 等。具体而言,可以使用欧洲专利申请公开 952221 号说明书中示例的菌株。而且,一些成团肠杆菌最近基于 16S rRNA 碱基序列分析等被重新分类为成团泛菌 (*Pantoea agglomerans*) 或 *Pantoea ananatis*、斯氏泛菌 (*Pantoea stewartii*)。在本发明中,只要是分类为肠杆菌科的,可以是属于肠杆菌属或泛菌属中任何一个属的细菌。

[0055] 作为肠杆菌属的代表株,可以列举出成团肠杆菌 ATCC12287 株。

[0056] 作为泛菌属细菌的代表性细菌,可以列举出 *Pantoea ananatis*、斯氏泛菌 (*Pantoea stewartii*)、成团泛菌、柠檬泛菌 (*Pantoea citrea*)。具体而言,可以列举出下述菌株。

[0057] *Pantoea ananatis* AJ13355 株 (FERM BP-6614) (欧洲专利申请公开 0952221 号说明书)

[0058] *Pantoea ananatis* AJ13356 株 (FERM BP-6615) (欧洲专利申请公开 0952221 号说明书)

[0059] *Pantoea ananatis* AJ13601 株 (FERM BP-7207) (欧洲专利申请公开 0952221 号说明书)

[0060] 这些菌株在分离出来时被鉴定为成团肠杆菌,并作为成团肠杆菌保藏。但是,如上所述,现在通过 16S rRNA 的碱基序列解析等被重新分类为 *Pantoea ananatis*。

[0061] 作为欧文氏菌属细菌,可以列举出解淀粉欧文氏菌 (*Erwinia amylovora*)、胡萝卜软腐欧文氏菌 (*Erwinia carotovora*);作为克雷伯氏菌属细菌,可以列举出植生克雷伯氏菌 (*Klebsiella planticola*)。具体地,可以列举出下述菌株。

[0062] 解淀粉欧文氏菌 ATCC15580 株

[0063] 胡萝卜软腐欧文氏菌 ATCC15713 株

[0064] 植生克雷伯氏菌 AJ13399 株 (FERM BP-6600) (欧洲专利申请公开 955368 号说明书)

[0065] 植生克雷伯氏菌 AJ13410 株 (FERM BP-6617) (欧洲专利申请公开 955368 号说明书)

[0066] 本发明的细菌是诸如上述的肠杆菌科细菌,并且,该细菌是固有地具备 GCD 活性、且具有 L-氨基酸生产能力的属于肠杆菌科的细菌。所谓固有地具备 GCD 活性的肠杆菌科细菌,是指野生株或 *gcd* 基因非修饰株具有 GCD 活性的细菌。作为这样的属于肠杆菌科的细菌,可以列举出属于泛菌、肠杆菌属、欧文氏菌属、克雷伯氏菌属、普罗威登斯菌属、沙门氏菌属、沙雷氏菌属、摩根氏菌属、以及耶尔森菌属、柠檬酸杆菌属 (*Citrobacter*)、变形杆菌属 (*Proteus*) 等属的细菌。更具体地,可以列举出 *Int. J. Syst. Bacteriol.*, 39(1), 61-67, 1989 中描述的细菌。

[0067] 大肠杆菌携带 *gcd* 基因,能够产生 GCD 脱辅酶,但其不具备 PQQ 产生能力,未添加 PQQ 的条件下其不具有 GCD 活性。然而已经知道,如果使其表达某外源基因,则会生成代替 PQQ 的物质,从而表现 GCD 活性 (W02006/133898)。这样,像埃希氏菌属细菌那样通常不具有 GCD 活性、但处于表现 GCD 活性的状态的细菌,包含于本发明所称的“固有地具备 GCD 活性的肠杆菌科细菌”中。关于 GCD 活性,参见后述。

[0068] 以下,就赋予如所述的细菌以 L-氨基酸生产能力的方法、或增强如所述的细菌的 L-氨基酸生产能力的方法进行示例说明。

[0069] 为了赋予 L-氨基酸生产能力,可以使用在棒状杆菌型细菌或埃希氏菌属细菌等的氨基酸生产菌的选育中沿用至今的方法,如获取营养缺陷突变株、L-氨基酸的类似物抗性株或代谢调节突变株,创建 L-氨基酸生物合成系统酶表达得到增强的重组菌株等等 (参见《アミノ酸発酵》, (株) 学会出版中心, 1986 年 5 月 30 日首次出版发行, 第 77-100 页)。这里,在 L-氨基酸生产菌的选育中,所赋予的营养缺陷性、类似物抗性 or 代谢调节突变等性质可以是单独一种,也可以是两种或者三种以上。此外,表达被增强的 L-氨基酸生物合成系统酶可以是单独一种,也可以是两种或三种以上。此外,营养缺陷性、类似物抗性 or 代谢调节突变等性质的赋予可以与生物合成系统酶的增强组合来进行。

[0070] 具有 L-氨基酸生产能力的营养缺陷突变体菌株、类似物抗性菌株或代谢调节突变菌株可如下获得:对亲本菌株或野生型菌株施以常规突变处理,即 X-射线或紫外线照射或用诱变剂例如 N-甲基-N'-硝基-N-亚硝基胍等处理;然后从所得的突变菌株中选择显示营养缺陷性、类似物抗性 or 代谢调节突变并且具有 L-氨基酸生产能力的菌株。

[0071] 此外,可以通过进行基因重组来增强酶活性从而赋予或增强 L-氨基酸生产能力。酶活性的增强,可以列举出例如通过修饰细菌从而增强编码参与 L-氨基酸的生物合成的基因的表达式的方法。用于增强基因的表达式的方法可以这样来实现:导入通过将包含这些基因的 DNA 片段导入适当的质粒 (例如至少包含负责质粒在微生物内的复制增殖功能的基因

的质粒载体)而形成的扩增质粒;或者,通过接合、基因转移等使这些基因在染色体上多拷贝化;抑或在这些基因的启动子区域导入突变(参考国际公开小册子 W095/34672 号)。

[0072] 当向上述扩增质粒或者染色体上导入目的基因时,用于表达这些基因的启动子可以是能够在属于肠杆菌科的细菌中发挥功能的任何启动子,可以是所使用的基因自身的启动子,也可以是经修饰的启动子。也可以通过适当地选择在 L-氨基酸生产菌中强力发挥功能的启动子,或者通过使启动子的 -35、-10 区域与共有序列接近,来进行基因表达量的调节。如上所述的增强酶基因表达的方法,在 W000/18935 号小册子、欧洲专利申请公开 1010755 号说明书等中有记载。

[0073] 下面,就赋予细菌以 L-氨基酸生产能力的具体方法、以及被赋予了 L-氨基酸生产能力的细菌进行示例。以下的描述主要涉及埃希氏菌属细菌,但以下方法也适用于本发明中使用的肠杆菌科细菌。

[0074] L-苏氨酸生产菌

[0075] 作为具有 L-苏氨酸生产能力的微生物优选实例可以列举出 L-苏氨酸生物合成系统酶的 1 种或 2 种以上的活性得到了增强的细菌。作为 L-苏氨酸生物合成系统酶,可以列举出天冬氨酸激酶 III (lysC)、天冬氨酸半醛脱氢酶基因 (asd)、thr 操纵子所编码的天冬氨酸激酶 I 基因 (thrA)、高丝氨酸激酶基因 (thrB)、苏氨酸合酶 (thrC)、天冬氨酸氨基转移酶(天冬氨酸转氨酶)(aspC)。括号里为该基因的简写符号(下同)。这些酶中特别优选天冬氨酸半醛脱氢酶、天冬氨酸激酶 I、高丝氨酸激酶、天冬氨酸氨基转移酶以及苏氨酸合酶。也可以将 L-苏氨酸生物合成系统基因导入苏氨酸分解受到抑制的细菌中。作为苏氨酸分解受到抑制的埃希氏菌属细菌,例如可以列举出苏氨酸脱氢酶活性缺损的 TDH6 菌株(特开 2001-346578 号)。

[0076] L-苏氨酸生物合成系统酶的酶活性受最终产物 L-苏氨酸的抑制。因此,为了构建 L-苏氨酸生产菌,优选对 L-苏氨酸生产菌合成系统基因进行修饰以使所述酶不受 L-苏氨酸的反馈抑制。此外,上述 thrA、thrB 和 thrC 基因组成苏氨酸操纵子,苏氨酸操纵子形成弱化子(attenuator)结构,苏氨酸操纵子的表达受到培养液中异亮氨酸和苏氨酸的抑制,并且表达受到弱化作用的抑制。这种修饰可通过去除弱化区中的前导序列或弱化子来实现(参照 Lynn, S. P., Burton, W. S., Donohue, T. J., Gould, R. M., Gumport, R. I., and Gardner, J. F. J. Mol. Biol. 194 :59-69(1987);国际公开第 02/26993 号小册子;国际公开第 2005/049808 号小册子)。

[0077] 苏氨酸操纵子的上游存在固有启动子,但可以用非天然启动子来取代它(参考 W098/04715 号小册子),也可以构建苏氨酸操纵子使得参与苏氨酸生物合成的基因的表达受到 λ 噬菌体的阻遏物(repressor)和启动子的调控(参考欧洲专利第 0593792 号说明书)。此外,为了对细菌进行修饰使其不受 L-苏氨酸的反馈抑制,可以通过选择对 α -氨基- β -羟基戊酸(AHV)有抗性的菌株来实现。

[0078] 对于如上所述经修饰而不受 L-苏氨酸的反馈抑制的苏氨酸操纵子而言,优选的是其在宿主内的拷贝数有所提高、或者是其连接于强启动子而表达量有所增加。除了通过使用质粒的扩增来提高拷贝数之外,还可以通过使用转座子、Mu 噬菌体等向基因组中转移苏氨酸操纵子来提高拷贝数。

[0079] 理想的是,除了 L-苏氨酸生物合成系统酶之外,还增强与糖酵解系统、TCA 循环或

呼吸链相关的基因或调控基因表达的基因,或者糖摄取基因。这些在 L- 苏氨酸生产中有效的基因的实例可以列举出,转氢酶基因 (pntAB) (欧洲专利 733712 号说明书)、磷酸烯醇丙酮酸羧化酶基因 (pepC) (国际公开 95/06114 号小册子)、磷酸烯醇丙酮酸合酶基因 (pps) (欧洲专利 877090 号说明书)、棒状杆菌型细菌或者芽孢杆菌属细菌的丙酮酸羧化酶基因 (国际公开 99/18228 号小册子、欧洲申请公开 1092776 号说明书)。

[0080] 此外,强化赋予 L- 苏氨酸抗性的基因、赋予 L- 高丝氨酸抗性的基因的表达,或对宿主赋予 L- 苏氨酸抗性、L- 高丝氨酸抗性,也是合适的。作为给予抗性的基因的实例,可以列举出 rhtA 基因 (Res. Microbiol. 154 :123-135 (2003))、rhtB 基因 (欧洲专利申请公开第 0994190 号说明书)、rhtC 基因 (欧洲专利申请公开第 1013765 号说明书)、yfiK、yeaS 基因 (欧洲专利申请公开第 1016710 号说明书)。此外,关于对宿主赋予 L- 苏氨酸抗性的方法,可以参考欧洲专利申请公开第 0994190 号说明书、国际公开第 90/04636 号小册子中所记载的方法。

[0081] 作为 L- 苏氨酸生产菌或用于衍生 L- 苏氨酸生产菌的亲本株的例子,可以列举:大肠杆菌 TDH-6/pVIC40 (VKPM B-3996) (美国专利第 5,175,107 号、美国专利第 5,705,371 号)、大肠杆菌 472T23/pYN7 (ATCC 98081) (美国专利第 5,631,157 号)、大肠杆菌 NRRL-21593 (美国专利第 5,939,307 号)、大肠杆菌 FERM BP-3756 (美国专利第 5,474,918 号)、大肠杆菌 FERM BP-3519 和 FERM BP-3520 (美国专利第 5,376,538 号)、大肠杆菌 MG442 (Gusyatiner et al., Genetika (俄语), 14, 947-956 (1978))、大肠杆菌 VL643 和 VL2055 (EP1149911A) 等属于埃希氏菌属的菌株,但不限于此。

[0082] TDH-6 菌株在 thrC 基因有缺陷,是蔗糖同化型,并且 ilvA 基因有泄漏突变 (leaky mutation)。该菌株的 rhtA 基因也有赋予对高浓度苏氨酸或高丝氨酸的抗性的突变。B-3996 菌株携带 pVIC40, pVIC40 是通过将含有突变的 thrA 基因的 thrA*BC 操纵子插入到由 RSF1010 来源的载体中而获得的。这种突变的 thrA 基因编码基本上已对苏氨酸反馈抑制脱敏的天冬氨酸激酶高丝氨酸脱氢酶 I。B-3996 菌株在 1987 年 11 月 19 日被保藏于 All-Union Scientific Center of Antibiotics (全联盟抗生素科学中心) (Nagatinskaya Street 3-A, 117105 Moscow, Russia), 保藏号为 RIA 1867。该菌株也在 1987 年 4 月 7 日于俄罗斯国立工业微生物保藏中心 (Russian National Collection of Industrial Microorganisms) (VKPM) (1 Dorozhny proezd., 1 Moscow 117545, Russia) 进行了国际保藏,保藏号为 B-3996。

[0083] 还可以使用大肠杆菌 VKPM B-5318 (EP 0593792B) 作为 L- 苏氨酸生产菌或用于衍生 L- 苏氨酸生产菌的亲本株。B-5318 菌株是异亮氨酸原养型的,并且 pVIC40 质粒中的苏氨酸操纵子的调控区域被替换为温度敏感的 λ 噬菌体 C1 阻遏物和 PR 启动子。VKPM B-5318 菌株在 1990 年 5 月 3 日于俄罗斯国立工业微生物保藏中心 (VKPM, 1 Dorozhny proezd., 1 Moscow 117545, Russia) 进行了国际保藏,保藏号为 VKPM B-5318。

[0084] 编码天冬氨酸激酶高丝氨酸脱氢酶 I 的大肠杆菌 thrA 基因已经阐明 (核苷酸编号 337 至 2799, GenBank 登录号 NC_000913.2, gi :49175990)。thrA 基因位于大肠杆菌 K-12 染色体上 thrL 基因和 thrB 基因之间。大肠杆菌的编码高丝氨酸激酶的 thrB 基因已经阐明 (核苷酸编号 2801 至 3733, GenBank 登录号 NC_000913.2, gi :49175990)。thrB 基因位于大肠杆菌 K-12 染色体上的 thrA 基因和 thrC 基因之间。大肠杆菌的编码苏氨酸合酶的 thrC

基因已经阐明(核苷酸编号 3734 至 5020, GenBank 登录号 NC_000913.2, gi :49175990)。thrC 基因位于大肠杆菌 K-12 染色体上 thrB 基因和 yaaX 开放阅读框之间。所述 3 个基因全部作为一个单独的苏氨酸操纵子发挥功能。为增强苏氨酸操纵子的表达,优选从操纵子中将影响转录的弱化子区域去除(WO2005/049808, WO2003/097839)。

[0085] 编码对苏氨酸的反馈抑制有抗性的天冬氨酸激酶高丝氨酸脱氢酶 I 的突变 thrA 基因,以及 thrB 和 thrC 基因可以以一个操纵子的形式从存在于苏氨酸生产株大肠杆菌 VKPM B-3996 菌株中的公知的 pVIC40 质粒获得。pVIC40 质粒在美国专利 5,705,371 中有详述。

[0086] rhtA 基因存在于大肠杆菌染色体上的 18min 处,靠近编码谷氨酰胺转运系统的成分的 glnHPQ 操纵子。rhtA 基因与 ORF1(ybiF 基因,核苷酸编号 764 至 1651, GenBank 登录号 AAA218541, gi :440181) 相同,位于 pexB 与 ompX 基因之间。表达由 ORF1 编码的蛋白的单元称为 rhtA 基因(rht:对高丝氨酸和苏氨酸抗性)。此外,已经证明 rhtA23 突变是在相对于 ATG 起始密码子的-1 位置用 G→A 取代(ABSTRACTS of the 17th International Congress of Biochemistry and Molecular Biology in conjugation with Annual Meeting of the American Society for Biochemistry and Molecular Biology(第 17 界国际生物化学与分子生物学大会暨美国生物化学与分子生物学协会年会摘要), San Francisco, California, 1997 年 8 月 24-29 日,摘要号 457, EP 1013765 A)。

[0087] 大肠杆菌的 asd 基因已经阐明(核苷酸编号 3572511 至 3571408, GenBank 登录号 NC_000913.1, gi :16131307),并且可以根据该基因的碱基序列制备引物通过 PCR 来获得(参见 White, T. J. 等, Trends Genet., 5, 185(1989))。其它微生物的 asd 基因可以通过类似方式获得。

[0088] 此外,大肠杆菌的 aspC 基因已经阐明(核苷酸编号 983742 至 984932, GenBank 登录号 NC_000913.1, gi :16128895),并且可以通过 PCR 获得。其它微生物的 aspC 基因可以通过类似方式获得。

[0089] L-赖氨酸生产菌

[0090] 属于埃希氏菌属的 L-赖氨酸生产菌的实例可以列举出对 L-赖氨酸类似物具有抗性的突变株。L-赖氨酸类似物可抑制属于埃希氏菌属的细菌的生长,但是这种抑制在当 L-赖氨酸共存于培养基中时会被全部或部分解除。L-赖氨酸类似物的实例可以列举出、但不限于,氧代赖氨酸,赖氨酸氧肟酸(lysine hydroxamate), S-(2-氨基乙基)-L-半胱氨酸(AEC), γ -甲基赖氨酸, α -氯代己内酰胺等。这些对赖氨酸类似物具有抗性的突变株可以通过对属于埃希氏菌属的细菌进行常规的人工诱变处理来得到。可用于 L-赖氨酸生产的菌株的具体实例可以列举出大肠杆菌 AJ11442(FERM BP-1543, NRRLB-12185; 参见美国专利第 4,346,170 号)和大肠杆菌 VL611。在这些微生物中,天冬氨酸激酶已经对 L-赖氨酸的反馈抑制脱敏。

[0091] 作为 L-赖氨酸生产菌或用于衍生 L-赖氨酸生产菌的亲本株,可以列举出 L-赖氨酸生物合成系统酶的 1 种或者 2 种以上的活性得到了增强的菌株。相关的酶包括,但不限于:二氢吡啶二羧酸合酶(dapA)、天冬氨酸激酶(lysC)、二氢吡啶二羧酸还原酶(dapB)、二氨基庚二酸脱羧酶(lysA)、二氨基庚二酸脱氢酶(ddh)(美国专利第 6,040,160 号)、磷酸烯醇丙酮酸羧化酶(ppc)、天冬氨酸半醛脱氢酶基因、二氨基庚二酸差向异构酶(dapF)、

四氢吡啶二羧酸琥珀酰酶 (dapD)、琥珀酰二氨基庚二酸脱酰基酶 (dapE) 以及天冬氨酸酶 (aspA) (EP 1253195 A)。在这些酶中,特别优选为二氢吡啶二羧酸还原酶、二氨基庚二酸脱羧酶、二氨基庚二酸脱氢酶、磷酸烯醇丙酮酸羧化酶、天冬氨酸氨基转移酶、二氨基庚二酸差向异构酶、天冬氨酸半醛脱氢酶、四氢吡啶二羧酸琥珀酰酶以及琥珀酰二氨基庚二酸脱酰基酶。此外,亲本株中下述基因的表达水平可以得到增强:与能量效率相关的基因 (cyo) (EP 1170376 A)、编码烟酰胺核苷酸转氢酶的基因 (pntAB) (美国专利第 5,830,716 号)、ybjE 基因 (WO2005/073390)、或所述基因的组合。

[0092] 作为 L-赖氨酸生产菌或用于衍生 L-氨基酸生产菌的亲本株的例子,还可以列举催化通过从 L-氨基酸生物合成途径改道而生成 L-赖氨酸以外的化合物的反应的酶的活性减少或者缺损的菌株。作为催化通过从 L-赖氨酸生物合成途径分支而生成 L-赖氨酸以外的化合物的反应的酶的例子,可以列举出高丝氨酸脱氢酶、赖氨酸脱羧酶 (美国专利第 5,827,698 号) 和苹果酸酶 (WO2005/010175)。

[0093] 作为优选的 L-赖氨酸生产菌,可以列举大肠杆菌 WC196 Δ cadA Δ ldcC/pCABD2 (WO2006/078039)。该菌株是通过在编码赖氨酸脱羧酶的 cadA 和 ldcC 基因被破坏了的 WC196 株中导入美国专利第 6040160 记载的质粒 pCABD2 而得到的菌株。WC196 株是由大肠杆菌 K-12 来源的 W3110 株出发,用突变型 lysC 基因 (美国专利第 5,661,012 号) 替换 W3110 株染色体上的野生型 lysC 基因后赋予 AEC 抗性而选育得到的菌株 (美国专利第 5,827,698 号),所述突变型 lysC 基因编码第 352 位的苏氨酸被替换成异亮氨酸从而对 L-赖氨酸反馈抑制脱敏的天冬氨酸激酶 III。WC196 株被命名为大肠杆菌 AJ13069,已于 1994 年 12 月 6 日保藏在工业技术院生命工学工业技术研究所 (现为独立行政法人产业技术综合研究所专利生物保藏中心、〒305-8566 日本国茨城县筑波市东 1 丁目 1 番地 1 中央第 6),保藏号 FERMP-14690;并于 1995 年 9 月 29 日转为基于布达佩斯条约的国际保藏,保藏号 FERM BP-5252 (美国专利第 5,827,698 号)。WC196 Δ cadA Δ ldc 本身也是优选的 L-赖氨酸生产菌。pCABD2 包含:编码具有解除了 L-赖氨酸反馈抑制的突变的大肠杆菌来源二氢吡啶二羧酸合成酶 (DDPS) 的突变型 dapA 基因、编码具有解除了 L-赖氨酸反馈抑制的突变的大肠杆菌来源天冬氨酸激酶 III 的突变型 lysC 基因、编码大肠杆菌来源二氢吡啶二羧酸还原酶的 dapB 基因、以及编码乳发酵短杆菌 (*Brevibacterium lactofermentum*) 来源二氨基庚二酸脱氢酶的 ddh 基因。

[0094] L-半胱氨酸生产菌

[0095] 细菌的 L-半胱氨酸生产能力可以通过增强 L-半胱氨酸生物合成途径的酶的活性,或增强参与该途径的底物化合物如 L-丝氨酸等的生成的酶,例如 3-磷酸甘油酸脱氢酶或丝氨酸乙酰转移酶等的活性来提高。3-磷酸甘油酸脱氢酶受丝氨酸的反馈抑制,通过使细菌携带编码该反馈抑制被减弱或解除的突变型 3-磷酸甘油酸脱氢酶的突变型 serA 基因,可以增强该酶活性。

[0096] 此外,丝氨酸乙酰转移酶受 L-半胱氨酸的反馈抑制。因此,通过使细菌携带编码该反馈抑制被减弱或解除的丝氨酸乙酰转移酶的突变型 cysE 基因,可以增强该酶活性。作为编码大肠杆菌的 SAT 的基因, cysE 已经从野生株以及 L-半胱氨酸分泌突变株中被克隆出来,其碱基序列已经阐明 (Denk, D. and Boeck, A., J. General Microbiol., 133, 515-525 (1987))。该碱基序列以及该碱基序列编码的氨基酸序列如 SEQ ID NO:37 和 38 所

示。

[0097] 此外,通过增强硫酸盐/硫代硫酸盐输送系统的活性,也可以提高 L-半胱氨酸生产能力。硫酸盐/硫代硫酸盐输送系统蛋白质群由 cysPTWA 基因簇编码(特开 2005-137369 号公报,EP1528108 号说明书)。

[0098] 此外,细菌的 L-半胱氨酸生产能力还可以通过提高 yeaS 基因(欧洲专利申请公开第 1016710 号说明书)的表达来提高。yeaS 基因的碱基序列以及该基因编码的氨基酸序列如 SEQ ID NO:39 和 40 所示。已知在细菌中,除了 ATG 以外,还会使用 GTG 等各种密码子作为起始密码子(<http://depts.washington.edu/agro/genomes/students/stanstart.htm>)。在 SEQ IDNO:39 和 40 中标出的是与最初的密码子 gtg 对应的氨基酸 Val,但实际上很可能是 Met。

[0099] 具有 L-半胱氨酸生产能力的埃希氏菌属细菌或用于衍生它的亲本株的例子包括但不限于下述属于埃希氏菌属的菌株,例如,用编码反馈抑制抗性丝氨酸乙酰转移酶的不同 cysE 等位基因转化的大肠杆菌 JM15(美国专利 6,218,168,俄罗斯专利申请 2003121601),具有过表达编码适于向细胞分泌有毒物质的蛋白的基因的大肠杆菌 W3110(美国专利 5,972,663),半胱氨酸脱巯基酶(cysteine desulfohydase)活性减少的大肠杆菌菌株(JP11155571A2);由 cysB 基因编码的半胱氨酸调节子的正向转录调控物活性增强的大肠杆菌 W3110(W00127307A1),等等。

[0100] 细菌的 L-半胱氨酸生产能力可以通过进行修饰降低 yhaM 基因编码的蛋白质(以下也记作“YhaM”)的活性来加以提高。yhaM 基因与 ECK3099、b4470、yhaN 基因同义,以前也称作 b3109 或 b3108。

[0101] 作为具有 L-半胱氨酸生产能力的泛菌属细菌,可以列举出:经过修饰而降低了半胱氨酸脱巯基酶活性的 *Pantoea ananatis* 菌株,以及再进一步一具有编码 L-半胱氨酸反馈抑制减弱的突变型丝氨酸乙酰转移酶的基因的 *Pantoea ananatis* 菌株。作为用于这样的 L-半胱氨酸生产菌的选育的亲本株,可以列举出 *Pantoea ananatis* AJ13355 株、SC17 株以及 SC17(0) 株。AJ13355 株是作为低 pH 下能够在含 L-谷氨酸和碳源的培养基中增殖的菌株从静岡県磐田市的土壤中分离出来的菌株,而 SC17 株是从 AJ13355 株出发作为粘液质低生产突变株选择出来的菌株(美国专利第 6,596,517 号)。AJ13355 株于平成 10 年 2 月 19 日保藏在通产省工业技术院生命工学工业技术研究所(现名产业技术综合研究所专利生物保藏中心,地址:邮政编码 305-8566,茨城县筑波市东 1 丁目 1 番地 1 中央第 6),保藏号 FERM P-16644,并于平成 11 年 1 月 11 日转为基于布达佩斯条约的国际保藏,保藏号 FERM BP-6614。SC17(0) 株是为了对 *Pantoea ananatis* 进行基因破坏,作为对 λ Red 基因产物具有抗性的菌株而构建的菌株(参照参考例 1)。

[0102] SC17 株的内部编号是 AJ416,其于平成 21 年 2 月 4 日国际保藏在产业技术综合研究所专利生物保藏中心(地址 邮政编码 305-8566 茨城县筑波市东 1 丁目 1 番地 1 中央第 6),保藏号 FERM BP-11091。此外,SC17(0) 株于 2005 年 9 月 21 日国际保藏在俄罗斯国立工业微生物保藏中心(Russian National Collection of Industrial Microorganisms(VKPM),GNII Genetika)(地址:Russia,117545 Moscow,1 Dorozhny proezd. 1),保藏号 VKPM B-9246。

[0103] 细菌生产的 L-半胱氨酸在培养基中通过二硫键部分转化为 L-胱氨酸。此外,

如后述,有时 L- 半胱氨酸通过和培养基中含有的硫代硫酸反应生成 S- 磺酸半胱氨酸 (Szczepkowski T.W., Nature, vol. 182 (1958))。进一步,细菌的细胞内生成的 L- 半胱氨酸有时还和细胞中存在的酮或醛、例如丙酮酸缩合,以硫代半酮缩醇 (hemithioketal) 为中间体生成噻唑烷 (thiazolidine) 衍生物 (参考日本专利第 2992010)。这些噻唑烷衍生物以及硫代半酮缩醇有时作为平衡混合物存在。因此,在本发明中, L- 半胱氨酸生产能力,不仅限于在培养基中或者菌体内蓄积 L- 半胱氨酸的能力,还包括在培养基中蓄积除 L- 半胱氨酸之外, L- 胱氨酸或其衍生物,例如 S- 磺酸半胱氨酸、噻唑烷衍生物或硫代半酮缩醇或者它们的混合物的能力。

[0104] L- 亮氨酸生产菌

[0105] L- 亮氨酸生产菌或用于衍生它的亲本株的例子包括,但不限于下述属于埃希氏菌属的菌株:对亮氨酸有抗性的大肠杆菌菌株 (例如,菌株 57 (VKPMB-7386, 美国专利第 6, 124, 121 号)) 或对亮氨酸类似物如 β -2- 噻吩丙氨酸、3- 羟基亮氨酸、4- 氮亮氨酸、5, 5, 5- 三氟亮氨酸等有抗性的大肠杆菌菌株 (特公昭 62-34397 号和特开平 8-70879 号); 通过 W096/06926 中描述的基因工程方法获得的大肠杆菌菌株; 大肠杆菌 H-9068 (特开平 8-70879 号), 等等。

[0106] 用于本发明的细菌可以是已通过增加参与 L- 亮氨酸生物合成的一种以上的基因的表达得到改良的。这样的基因的实例可优选例举以编码解除了 L- 亮氨酸反馈抑制的异丙基苹果酸合酶的突变 leuA 基因 (美国专利第 6, 403, 342 号) 为代表的 leuABCD 操纵子中的基因。此外,本发明的细菌可以是已通过增加一种以上编码从细菌细胞分泌 L- 氨基酸的蛋白的基因的表达而得以改良的。这样的基因的实例包括 b2682 和 b2683 基因 (ygaZH 基因) (EP1239041A2)。

[0107] L- 组氨酸生产菌

[0108] L- 组氨酸生产菌或用于衍生它的亲本株的例子包括但不限于下述属于埃希氏菌属的菌株:大肠杆菌 24 株 (VKPM B-5945、RU2003677)、大肠杆菌 80 株 (VKPM B-7270、RU2119536)、大肠杆菌 NRRL B-12116-B12121 (美国专利第 4, 388, 405 号)、大肠杆菌 H-9342 (FERM BP-6675) 以及 H-9343 (FERM BP-6676) (美国专利第 6, 344, 347 号)、大肠杆菌 H-9341 (FERMBP-6674) (EP1085087)、大肠杆菌 AI80/pFM201 (美国专利第 6, 258, 554 号), 等等。

[0109] L- 组氨酸生产菌或用于衍生它的亲本株的例子还包括编码 L- 组氨酸生物合成系统酶的 1 种以上的基因的表达已得到增强的菌株。相关基因的实例包括 ATP 磷酸核糖基转移酶基因 (hisG)、磷酸核糖基 AMP 环化水解酶基因 (hisI)、磷酸核糖基 -ATP 焦磷酸水解酶基因 (hisI)、磷酸核糖基亚氨基甲基 -5- 氨基咪唑氨甲酰核糖核苷酸异构酶基因 (phosphoribosylformimino-5-aminoimidazole carboxamide ribotide isomerase) (hisA)、酰胺转移酶基因 (hisH)、组氨醇磷酸氨基转移酶基因 (hisC)、组氨醇磷酸酶基因 (hisB)、组氨醇脱氢酶基因 (hisD) 等。

[0110] 已知 hisG 和 hisBHAFI 编码的 L- 组氨酸生物合成系统酶被 L- 组氨酸所抑制,因此还可以通过向 ATP 磷酸核糖基转移酶基因 (hisG) 中引入诱导可赋予对反馈抑制的抗性的突变,来有效地增加产生 L- 组氨酸的能力 (俄罗斯专利 2003677 和 2119536)。

[0111] 具有 L- 组氨酸生产能力的菌株具体实例包括导入了携带编码 L- 组氨酸生物合成

系统酶的 DNA 的载体的大肠杆菌 FERM-P 5038 和 5048 (特开昭 56-005099 号), 导入了氨基酸输送的基因的大肠杆菌菌株 (EP1016710A), 赋予了磺胺胍、DL-1, 2, 4- 三唑-3- 丙氨酸和链霉素抗性的大肠杆菌 80 菌株 (VKPM B-7270, 俄罗斯专利 2119536 号), 等等。

[0112] L- 谷氨酸生产菌

[0113] L- 谷氨酸生产菌或用于衍生它的亲本株的例子包括但不限于大肠杆菌 (*E. coli*) VL334thrC⁺ (EP 1172433) 等属于埃希氏菌属的菌株。大肠杆菌 VL334 (VKPM B-1641) 是 L- 异亮氨酸和 L- 苏氨酸营养缺陷型菌株, 其在 thrC 和 ilvA 基因中有突变 (美国专利 4, 278, 765)。利用可在野生型大肠杆菌菌株 K-12 (VKPM B-7) 细胞中增殖的噬菌体 P1, 通过常规转化方法来转入 thrC 基因的野生型等位基因。结果, 获得了 L- 异亮氨酸营养缺陷型的 L- 谷氨酸生产菌 VL334thrC⁺ (VKPM B-8961)。

[0114] L- 谷氨酸生产菌或用于衍生它的亲本株的例子包括但不限于 L- 谷氨酸生物合成系统酶中的 1 种或 2 种以上的活性得到增强的菌株。作为所述基因的例子, 可以列举: 谷氨酸脱氢酶 (gdhA)、谷氨酰胺合成酶 (glnA)、谷氨酸合成酶 (gltAB)、异柠檬酸脱氢酶 (icdA)、顺乌头酸水合酶 (acnA, acnB)、柠檬酸合酶 (gltA)、甲基柠檬酸合酶 (prpC)、磷酸烯醇丙酮酸羧化酶 (ppc)、丙酮酸脱氢酶 (aceEF, lpdA)、丙酮酸激酶 (pykA, pykF)、磷酸烯醇丙酮酸合酶 (ppsA)、烯醇化酶 (eno)、磷酸甘油酸变位酶 (pgmA, pgmI)、磷酸甘油酸激酶 (pgk)、甘油醛-3- 磷酸脱氢酶 (gapA)、丙糖磷酸异构酶 (tpiA)、果糖二磷酸醛缩酶 (fbp)、磷酸果糖激酶 (pfkA, pfkB)、葡萄糖磷酸异构酶 (pgi) 等。所述酶中, 优选谷氨酸脱氢酶、柠檬酸合酶、磷酸烯醇丙酮酸羧化酶以及甲基柠檬酸合酶。

[0115] 作为经过了修饰使得柠檬酸合酶基因、磷酸烯醇丙酮酸羧化酶基因、和 / 或谷氨酸脱氢酶基因的表达增加的菌株的例子, 可以列举 EP1078989A、EP955368A 以及 EP952221A 中公开的菌株。

[0116] 细菌的 L- 谷氨酸合成能力可以通过提高参与呼吸链的酶、例如氰抗性呼吸终端氧化酶 (cioA, cioB) 的活性来提高。

[0117] 作为 L- 谷氨酸生产菌或用于衍生它的亲本株的例子, 还可以列举下述酶的活性降低或缺损的菌株, 所述酶催化从 L- 谷氨酸生物合成途径改道而合成 L- 谷氨酸以外的化合物, 或者催化分解或消耗 L- 谷氨酸的反应。作为这样的酶的例子, 可以列举: 异柠檬酸裂合酶 (aceA)、 α - 酮戊二酸脱氢酶 (sucA)、磷酸乙酰转移酶 (pta)、乙酸激酶 (ack)、乙酰羟酸合酶 (ilvG)、乙酰乳酸合酶 (ilvI)、甲酸乙酰转移酶 (pfI)、乳酸脱氢酶 (ldh)、谷氨酸脱羧酶 (gadAB)、 γ - 谷氨酰基转移酶 (ggt)、 γ - 谷氨酸半胱氨酸合成酶 (gshA)、 γ - 谷氨酸腐胺合成酶 (ycjK) 等。 α - 酮戊二酸脱氢酶活性缺损、或 α - 酮戊二酸脱氢酶活性降低的属于埃希氏菌属的细菌及其获取方法在美国专利第 5, 378, 616 号和第 5, 573, 945 号中有记载。

[0118] 作为具体例子, 可以列举出如下。

[0119] 大肠杆菌 W3110sucA::Kmr

[0120] 大肠杆菌 AJ12624 (FERM BP-3853)

[0121] 大肠杆菌 AJ12628 (FERM BP-3854)

[0122] 大肠杆菌 AJ12949 (FERM BP-4881)

[0123] 大肠杆菌 W3110sucA::Km^r 是通过破坏大肠杆菌 W3110 的 α - 酮戊二酸脱氢酶基

因（以下也称“sucA 基因”）而得到的菌株。该菌株的 α -酮戊二酸脱氢酶完全缺损。

[0124] 作为 L-谷氨酸生产菌的其它例子，可以列举属于埃希氏菌属、且对天冬氨酸代谢拮抗物质有抗性的菌株。所述菌株可以是缺损 α -酮戊二酸脱氢酶的，例如大肠杆菌 AJ13199 (FERM BP-5807) (美国专利第 5,908,768 号)，以及还降低了 L-谷氨酸分解能力的 FFRM P-12379 (美国专利第 5,393,671 号)；AJ13138 (FERM BP-5565) (美国专利第 6,110,714 号) 等。

[0125] 作为 *Pantoea ananatis* 的 L-谷氨酸生产菌的例子，可以列举出所述 *Pantoea ananatis* AJ13355 株。

[0126] 此外，作为 *Pantoea ananatis* 的 L-谷氨酸生产菌，可以列举 α -酮戊二酸脱氢酶 (α KGDH) 活性缺损、或者 α KGDH 活性降低的属于泛菌属的细菌。作为这样的菌株有：缺损 AJ13355 株的 α KGDH-E1 亚基基因 (sucA) 而得到的 AJ13356 (美国专利第 6,331,419 号)，以及从 AJ13355 株出发作为粘液质低生产突变菌株选择出来的 SC17 株衍生的 sucA 基因缺损株 SC17sucA (美国专利第 6,596,517 号)。AJ13356 于 1998 年 2 月 19 日保藏于工业技术院生命工学工业技术研究所（现为独立行政法人产业技术综合研究所专利生物保藏中心、〒305-8566 日本国茨城县筑波市东 1 丁目 1 番地 1 中央第 6），保藏号 FERM P-16645，并于 1999 年 1 月 11 日转为基于布达佩斯条约的国际保藏，保藏号 FERM BP-6616。AJ13355 以及 AJ13356 在上述保藏机构是作为成团肠杆菌保藏的，但是在本说明书中，将其描述为 *Pantoea ananatis*。此外，SC17sucA 株的内部编号为 AJ417，已于 2004 年 2 月 26 日保藏于产业技术综合研究所专利生物保藏中心，保藏号 FERM BP-08646。

[0127] 而且，作为 *Pantoea ananatis* 的 L-谷氨酸生产菌，可以列举 SC17sucA/RSFCPG+pSTVCB 株、AJ13601 株、NP106 株 以及 NA1 株。SC17sucA/RSFCPG+pSTVCB 株是在 SC17sucA 株中导入大肠杆菌来源的包含柠檬酸合酶基因 (gltA)、磷酸烯醇丙酮酸羧化酶基因 (ppsA) 和谷氨酸脱氢酶基因 (gdhA) 的质粒 RSFCPG 以及乳发酵短杆菌 (*Brevibacterium lactofermentum*) 来源的包含柠檬酸合酶基因 (gltA) 的质粒 pSTVCB 而得到的菌株。AJ13601 株是从该 SC17sucA/RSFCPG+pSTVCB 株出发作为低 pH 下显示出对高浓度 L-谷氨酸的抗性的菌株而选择出的菌株。此外，如实施例所述，NP106 株是 AJ13601 株脱去质粒 RSFCPG+pSTVCB 而得到的菌株。AJ13601 株于 1999 年 8 月 18 日保藏在独立行政法人产业技术综合研究所专利生物保藏中心（〒305-8566 日本国茨城县筑波市东 1 丁目 1 番地 1 中央第 6），保藏号 FERM P-17516，并于 2000 年 7 月 6 日转为基于布达佩斯条约的国际保藏，保藏号 FERM BP-7207。此外，携带将所述 RSFCPG 的 gltA 基因替换为 prpC 而得到的 RSFPFG (W02008/020654，参照后述实施例) 的 NP106 株也是优选的 L-谷氨酸生产菌。

[0128] L-苯丙氨酸生产菌

[0129] L-苯丙氨酸生产菌或者用于衍生 L-苯丙氨酸生产菌的亲本株的实例，可以列举，但不限于下述属于埃希氏菌属的菌株：分支酸变位酶-预苯酸脱氢酶以及酪氨酸阻遏物缺损的大肠杆菌 AJ12739 (tyrA::Tn10, tyrR) (VKPMB-8197) (W003/044191)，携带编码解除了反馈抑制的分支酸变位酶-预苯酸脱氢酶的突变型 pheA34 基因的大肠杆菌 HW1089 (ATCC 55371) (美国专利第 5,354,672 号)，大肠杆菌 MWEC101-b (KR8903681)、大肠杆菌 NRRLB-12141、NRRL B-12145、NRRL B-12146 以及 NRRL B-12147 (美国专利第 4,407,952 号) 等。此外，也可以使用携带编码解除了反馈抑制的分支酸变位酶-预苯酸脱氢酶基因的

大肠杆菌 K-12[W3110(tyrA)/pPHAB] (FERMBP-3566)、大肠杆菌 K-12[W3110(tyrA)/pPHAD] (FERM BP-12659)、大肠杆菌 K-12[W3110(tyrA)/pPHATerm] (FERM BP-12662) 以及命名为 AJ 12604 的大肠杆菌 K-12[W3110(tyrA)/pBR-aroG4,pACMAB] (FERM BP-3579) 作为亲本株 (EP 488424 B1)。此外,还可使用增加了由 yedA 基因或 yddG 基因编码的蛋白质的活性的属于埃希氏菌属的 L- 苯丙氨酸生产菌 (美国专利申请公开 2003/0148473 号以及 2003/0157667、WO 03/044192)。

[0130] L- 色氨酸生产菌

[0131] L- 色氨酸生产菌或用于诱导 L- 色氨酸生产菌的亲本株的例子包括、但不限于下述属于埃希氏菌属的菌株:大肠杆菌 JP4735/pMU3028 (DSM10122) 和 JP6015/pMU91 (DSM10123),其色氨酸-tRNA 合成酶缺陷,该酶由突变的 trpS 基因编码 (美国专利 5,756,345);大肠杆菌 SV164 (pGH5),其具有编码不受丝氨酸反馈抑制性的磷酸甘油酸脱氢酶的 serA 等位基因和编码不受色氨酸反馈抑制的邻氨基苯甲酸合酶 (anthranilate synthase) 的 trpE 等位基因 (美国专利 6,180,373);色氨酸酶缺陷的大肠杆菌 AGX17 (pGX44) (NRRLB-12263) 和 AGX6 (pGX50) aroP (NRRL B-12264) (美国专利 4,371,614);磷酸烯醇丙酮酸的生产能力增加的大肠杆菌 AGX17/pGX50, pACKG4-pps (W09708333, 美国专利 6,319,696) 等等。此外,还可以使用增强了由 yedA 基因或 yddG 基因编码的蛋白的活性的属于埃希氏菌属的 L- 色氨酸生产菌 (美国专利申请公开 2003/0148473 A1 和 2003/0157667 A1)。

[0132] 作为 L- 色氨酸生产菌或用于衍生它的亲本株的例子,还可以列举选自邻氨基苯甲酸合酶 (trpE)、磷酸甘油酸脱氢酶 (serA)、3- 脱氧-D- 阿拉伯庚酮糖酸-7- 磷酸合酶 (aroG)、3- 脱氢奎宁酸合酶 (aroB)、莽草酸脱氢酶 (aroE)、莽草酸激酶 (aroL)、5- 烯醇酸丙酮酰莽草酸 3- 磷酸合酶 (aroA,5-enolpyruvylshikimate-3-phosphate synthase)、分支酸合酶 (aroC)、预苯酸脱水酶、分支酸变位酶以及色氨酸合酶 (trpAB) 中的 1 种或 2 种以上酶的活性增强的菌株。预苯酸脱水酶和分支酸变位酶以双功能酶 (CM-PD) 的形式被 pheA 基因编码。所述酶之中,特别优选磷酸甘油酸脱氢酶、3- 脱氧-D- 阿拉伯庚酮糖酸-7- 磷酸合酶、3- 脱氢奎宁酸合酶、莽草酸脱水酶、莽草酸激酶、5- 烯醇酸丙酮酰莽草酸 3- 磷酸合酶、分支酸合酶、预苯酸脱水酶、分支酸变位酶-预苯酸脱氢酶。邻氨基苯甲酸合酶和磷酸甘油酸脱氢酶两者都受 L- 色氨酸和 L- 丝氨酸的反馈抑制,因此可以向这些酶中引入使反馈抑制解除的突变。具有此类突变的菌株的具体实例包括具有脱敏型邻氨基苯甲酸合酶的大肠杆菌 SV164、和通过将质粒 pGH5 (WO 94/08031) 导入大肠杆菌 SV164 而获得的转化菌株,所述质粒 pGH5 包含编码解除了反馈抑制的磷酸甘油酸脱氢酶的突变 serA 基因。

[0133] L- 色氨酸生产菌或用于诱导 L- 色氨酸生产菌的亲本株的例子还包括导入了包含编码抑制解除型邻氨基苯甲酸合酶的基因的色氨酸操纵子的菌株 (特开昭 57-71397 号,特开昭 62-244382 号,美国专利第 4,371,614)。此外,可以通过增加色氨酸操纵子 (trpBA) 中编码色氨酸合酶的基因的表达式来赋予 L- 色氨酸生产能力。色氨酸合酶由 α 和 β 亚单位组成,这两种亚单位分别由 trpA 和 trpB 基因编码。另外,可以通过增强异柠檬酸裂合酶-苹果酸合酶操纵子的表达式来改进 L- 色氨酸生产能力 (W02005/103275)。

[0134] L- 脯氨酸生产菌

[0135] 作为 L- 脯氨酸生产菌或用于衍生它的亲本株的例子,可以列举 ilvA 基因缺损、且

能够生产 L- 脯氨酸的大肠杆菌 702ilvA(VKPM B-8012) (EP1172433) 等属于埃希氏菌属的菌株,但不限于此。

[0136] 本发明中使用的细菌可以通过增加一种以上参与 L- 脯氨酸生物合成的基因的表达来进行改良。作为 L- 脯氨酸生产菌所优选的基因的例子,可以列举编码解除了 L- 脯氨酸反馈抑制的谷氨酸激酶的 proB 基因(德国专利第 3127361 号)。而且,本发明中使用的细菌,可以通过增加一种以上编码从细菌细胞中分泌 L- 氨基酸的蛋白质的基因的表达来进行改良。作为这样的基因,可以列举 b2682 基因和 b2683 基因(ygaZH 基因)(EP1239041 A2)。

[0137] 作为具有 L- 脯氨酸生产能力的属于埃希氏菌属的细菌的例子,可以列举:NRRL B-12403 以及 NRRL B-12404(英国专利第 2075056 号)、VKPMB-8012(俄罗斯专利申请 2000124295)、德国专利第 3127361 号所述的质粒突变体、Bloom F.R 等(The 15th Miami winter symposium,1983,p.34) 所述的质粒突变体等大肠杆菌株。

[0138] L- 精氨酸生产菌

[0139] L- 精氨酸生产菌或用于衍生它的亲本株的例子包括、但不限于下述属于埃希氏菌属的菌株,例如大肠杆菌 237 菌株(VKPM B-7925)(美国专利申请公开 2002/058315A1) 及其带有突变的 N- 乙酰谷氨酸合酶(N-acetylglutamatesynthase) 的衍生菌株(俄罗斯专利申请 2001112869),大肠杆菌 382 菌株(VKPM B-7926)(EP1170358A1),引入了编码 N- 乙酰谷氨酸合成酶(N-acetylglutamate synthetase) 的 argA 基因的精氨酸生产株(EP1170361A1),等等。

[0140] L- 精氨酸生产菌或用于衍生它的亲本株的例子还包括编码 L- 精氨酸生物合成系统酶的 1 种以上基因的表达增加的菌株。相关基因的实例包括 N- 乙酰谷氨酰磷酸还原酶基因(argC)、鸟氨酸乙酰转移酶基因(argJ)、N- 乙酰谷氨酸激酶基因(argB)、乙酰鸟氨酸转氨酶基因(argD)、鸟氨酸氨甲酰基转移酶基因(argF)、精氨酸琥珀酸合成酶基因(argG)、精氨酸琥珀酸裂合酶基因(argH) 和氨甲酰基磷酸合成酶基因(carAB)。

[0141] L- 缬氨酸生产菌

[0142] L- 缬氨酸生产菌或用于衍生它的亲本株的例子包括但不限于经修饰而过表达 ilvGMEDA 操纵子的菌株(美国专利 5,998,178)。优选将衰减作用所必需的 ilvGMEDA 操纵子的区域移除以使操纵子的表达不会被产生的 L- 缬氨酸所削弱。此外,优选将操纵子中的 ilvA 基因破坏从而降低苏氨酸脱氨酶的活性。

[0143] L- 缬氨酸生产菌或用于衍生它的亲本株的例子还包括具有氨酰 t-RNA 合成酶突变的突变株(美国专利 5,658,766)。例如,可以使用大肠杆菌 VL1970,该菌株在编码异亮氨酸 tRNA 合成酶的 ileS 基因中具有突变。已将大肠杆菌 VL1970 在 1988 年 6 月 24 日保藏在俄罗斯国立工业微生物保藏中心(VKPM)(1 Dorozhny Proezd.,1 Moscow 117545, Russia),保藏号为 VKPM B-4411。

[0144] 此外,还可以使用生长需要硫辛酸(lipoic acid) 的突变株和/或缺乏 H⁺-ATP 酶的突变株(W096/06926) 作为亲本株。

[0145] L- 异亮氨酸生产菌

[0146] L- 异亮氨酸生产菌或用于衍生 L- 异亮氨酸生产菌的亲本株的例子包括,但不限于对 6- 二甲基氨基嘌呤有抗性的突变株(特开平 5-304969 号),对异亮氨酸类似物如硫

代异亮氨酸 (thiaisoleucine) 和异亮氨酸氧肟酸等具有抗性的突变株, 以及对 DL- 乙硫氨酸 (DL-ethionine) 和 / 或精氨酸氧肟酸具有抗性的突变株 (特开平 5-130882 号)。此外, 还可以使用经过编码涉及 L- 异亮氨酸生物合成的蛋白 (如苏氨酸脱氨酶、乙酰羟酸合酶 (acetohydroxate synthase) 等) 的基因转化的重组菌株作为亲本株 (特开平 2-458 号, FR0356739, 和美国专利第 5, 998, 178 号)。

[0147] L- 酪氨酸生产菌

[0148] 作为酪氨酸生产菌, 可以列举出具有不受酪氨酸的抑制的脱敏型预苯酸脱水酶基因 (tyrA) 的埃希氏菌属细菌 (欧洲专利申请公开 1616940 号公报)。

[0149] 在通过基因重组来选育上述 L- 氨基酸生产菌时, 所使用的基因不限于具有上述基因信息的基因或具有公知序列的基因, 还可以使用这些基因的变体, 即以不损害所编码的蛋白质的功能为限, 该基因的同源物、人工修饰体等具有保守突变的基因。具体地说, 还可以是下述基因, 所述基因编码具有在公知的蛋白质的氨基酸序列中包含一个或数个位置上的一个或数个氨基酸的取代、缺失、插入或添加等而成的序列的蛋白质。

[0150] 这里, 所谓“一个或数个”因氨基酸残基的种类或其在蛋白质的立体结构中的位置而异, 但具体地是指优选 1 ~ 20 个、更优选 1 ~ 10 个、更优选 1 ~ 5 个。作为保守突变, 当取代部位为芳族氨基酸时, 是指在 Phe、Trp、Tyr 之间相互取代的突变, 当取代部位为疏水性氨基酸时, 是指在 Leu、Ile、Val 之间相互取代的突变, 当取代部位为极性氨基酸时, 是指在 Gln、Asn 之间相互取代的突变, 当取代部位为碱性氨基酸时, 是指在 Lys、Arg、His 之间相互取代的突变, 当取代部位为酸性氨基酸时, 是指在 Asp、Glu 之间相互取代的突变, 当取代部位为带有羟基的氨基酸时, 是指在 Ser、Thr 之间进行相互取代的突变。保守突变的代表是保守取代, 作为可视作保守取代的取代, 具体地可以列举: 从 Ala 到 Ser 或 Thr 的取代, 从 Arg 到 Gln、His 或 Lys 的取代, 从 Asn 到 Glu、Gln、Lys、His 或 Asp 的取代, 从 Asp 到 Asn、Glu 或 Gln 的取代, 从 Cys 到 Ser 或 Ala 的取代, 从 Gln 到 Asn、Glu、Lys、His、Asp 或 Arg 的取代, 从 Glu 到 Gly、Asn、Gln、Lys 或 Asp 的取代, 从 Gly 到 Pro 的取代, 从 His 到 Asn、Lys、Gln、Arg 或 Tyr 的取代, 从 Ile 到 Leu、Met、Val 或 Phe 的取代, 从 Leu 到 Ile、Met、Val 或 Phe 的取代, 从 Lys 到 Asn、Glu、Gln、His 或 Arg 的取代, 从 Met 到 Ile、Leu、Val 或 Phe 的取代, 从 Phe 到 Trp、Tyr、Met、Ile 或 Leu 的取代, 从 Ser 到 Thr 或 Ala 的取代, 从 Thr 到 Ser 或 Ala 的取代, 从 Trp 到 Phe 或 Tyr 的取代, 从 Tyr 到 His、Phe 或 Trp 的取代, 以及从 Val 到 Met、Ile 或 Leu 的取代。此外, 如上所述的氨基酸的取代、缺失、插入、添加或倒位等还包括因基于基因所来源的微生物的个体差异、种间差异等情形而天然发生的突变 (mutant 或 variant) 而产生的那些取代、缺失、插入、添加或倒位。这样的基因例如可以这样获得: 采用定点突变法对公知基因的碱基序列进行修饰, 使得被编码的蛋白质的特定部位处的氨基酸残基包含取代、缺失、插入或增加。

[0151] 而且, 具有如上所述的保守突变的基因还可以是编码下述蛋白质的基因: 就编码的氨基酸序列全长而言, 其与野生型蛋白质具有 80% 以上、优选 90% 以上、更优选 95% 以上、特别优选 97% 以上的同源性, 且具有与野生型蛋白质等同的功能。而且, 在本说明书中, 有些情况下“同源性” (homology) 是指“同一性” (identity)。

[0152] 此外, 可以将基因序列中各自的密码子替换为易于在基因所导入的宿主中使用的密码子。

[0153] 具有保守突变的基因可以通过突变剂处理等常规突变处理中使用的方法来获得。

[0154] 此外,基因还可以是下述 DNA,所述 DNA 与公知基因序列的互补序列或可由所述互补序列制备的探针在严格条件下杂交,且编码具有与公知基因产物等同的功能的蛋白质。这里,“严格条件”,是指所谓特异性杂交体形成但非特异性杂交体不形成的条件。举例说明:在具有高同源性的 DNA 之间、例如具有 80%以上、优选 90%以上、更优选 95%以上、特别优选 97%以上的同源性的 DNA 之间发生杂交,而同源性低于上述的 DNA 之间则不发生杂交的条件;或者在与常规 Southern 杂交的洗涤条件即 60℃、1×SSC、0.1% SDS,优选 0.1×SSC、0.1% SDS,更优选 68℃、0.1×SSC、0.1% SDS 相当的盐浓度、温度下,洗涤 1 次更优选 2~3 次的条件。

[0155] 作为探针,也可以使用基因的互补序列的一部分。这种探针可以通过 PCR 制备,其中 PCR 以基于公知的基因序列制备的寡核苷酸为引物,以包含这些碱基序列的 DNA 片段为模板来进行。例如,当使用 300bp 左右长度的 DNA 片段作为探针时,杂交的洗涤条件可以是例如 50℃、2×SSC、0.1% SDS。

[0156] 上述的关于基因的变体的描述,同样适用于下述的 gcd 基因。

[0157] <2-2>GCD 活性的降低

[0158] 接下来,针对使属于肠杆菌科的细菌的 GCD 活性降低的修饰进行说明。

[0159] GCD 活性是指催化下述反应的活性。

[0160] β -D-葡萄糖 + 氧化型 PQQ $\rightarrow$ D- δ -葡糖酸内酯 + 还原型 PQQ

[0161] GCD 活性例如可以通过采用 600nm 的吸光度测定检测基于以下反应生成的还原型 DCPIP,来测定(特开 2007-129965)。

[0162] D-葡萄糖 + 氧化型 PMS $\rightarrow$ D-葡糖酸-1,5-内酯 + 还原型 PMS

[0163] 还原型 PMS + 氧化型 DCPIP $\rightarrow$ 氧化型 PMS + 还原型 DCPIP

[0164] PMS:吩嗪甲基硫酸酯(phenazine methosulfate)

[0165] DCPIP:2,6-二氯酚靛酚(2,6-dichlorophenol-indophenol)

[0166] “经过修饰而降低了 GCD 活性”是指:与非修饰株、例如野生型的属于肠杆菌科的菌株相比,平均每个细菌细胞的 GCD 活性变低。例如,相当于平均每个细胞的 GCD 分子数降低或平均每个分子的 GCD 活性降低等情况。平均每个细胞的 GCD 活性的比较,可以通过例如,比较相同条件下培养的细菌的细胞抽提液中所含的 GCD 活性来进行。而且,活性的“降低”也包括活性完全消失的情况。用于比较的对照的野生型泛菌属细菌包括例如 *Pantoea ananatis* AJ13355 株(FERM BP-6614)等。

[0167] GCD 活性的降低可以通过使编码 GCD 的基因(gcd)失活来实现。使 gcd 基因“失活”,是指通过基因重组对该基因进行修饰或者向该基因导入突变,使得该基因编码的 GCD 的活性降低或消失。

[0168] 作为 gcd 基因,可以列举出具有 SEQ ID NO:1 所示的碱基序列的 *Pantoea ananatis* 的 gcd 基因。该 gcd 基因编码的 GCD 的氨基酸序列如 SEQ ID NO:2 所示。gcd 基因可以通过基于这些序列合成寡核苷酸,并以 *Pantoea ananatis* 的染色体为模板进行 PCR 反应来克隆。此外,在通过同源重组缺损 gcd 基因的情况下,也可以使用与染色体上的 gcd 基因具有一定程度以上的同源性、例如 80%以上、优选 90%以上、更优选 95%以上的同源性的基因。此外,还可以使用与染色体上的 gcd 基因在严格条件下杂交的基因。作为严

格条件,可以列举出例如:在与 60℃、1×SSC、0.1% SDS 优选 0.1×SSC、0.1% SDS 相当的盐浓度下,进行 1 次更优选 2~3 次洗涤的条件。

[0169] 使 gcd 基因失活具体地可以通过例如,使染色体上的 gcd 基因的编码区域的一部分或全部缺失,或者在编码区域中插入其它序列来实现。这些方法也称作基因破坏。

[0170] 此外,通过对 gcd 基因的启动子或 shine-dalgarno (SD) 序列等表达调节序列进行修饰而降低 gcd 基因的表达,也可以使 gcd 基因失活。表达的降低包括转录的降低和翻译的降低。此外,通过对表达调节序列以外的非翻译区域进行修饰,也能够降低基因的表达。

[0171] 而且,还可以将包含染色体上的目标基因前后的序列在内的目标基因整体缺失。此外, gcd 基因的灭活还可以这样实现:向染色体上的 gcd 基因的编码区域导入氨基酸取代(错义突变)、或者导入终止密码子(无义突变)、或者导入添加或缺失一~二个碱基的移码突变(Journal of Biological Chemistry 272:8611-8617(1997)Proceedings of the National Academy of Sciences, USA 955511-5515(1998), Journal of Biological Chemistry 266,20833-20839(1991))。

[0172] 各基因的修饰优选通过基因重组来进行。作为基于基因重组的方法,具体可以列举出:利用同源重组,使染色体上目标基因的表达调节序列、例如启动子区域或编码区域或非编码区域的一部分或全部缺损,或者在所述区域中插入其它序列。

[0173] 表达调节序列的修饰优选是 1 个碱基以上,更优选是 2 个碱基以上,特别优选是 3 个碱基以上。此外,在缺失编码区域时,只要各基因产生的蛋白质的功能降低或缺失即可,缺失的区域可以是 N 末端区域、内部区域、C 末端区域中的任何区域,也可以是整个编码区域。通常,缺失的区域长能够切实地使目标基因失活。此外,缺失区域的上游和下游的读码框优选是不一致的。

[0174] 在编码区域中插入其他序列时,插入的位置可以为目标基因的任何区域,但插入的序列长能够切实地使目标基因失活。插入部位的前后序列优选读码框不一致。对于其他序列没有特殊限制,只要是能够降低或缺损目标基因编码的蛋白质的功能即可,可以列举例如携带抗生素抗性基因或对 L-氨基酸生产有用的基因的转座子等。

[0175] 为了如上所述地修饰染色体上的目标基因,可以通过下述方法实现:例如,制作缺失目标基因的部分序列从而不产生发挥正常功能的蛋白质的经修饰的缺失型基因,用含有该基因的 DNA 转化细菌,使缺失型基因和染色体上的目标基因之间发生同源重组,将染色体上的目标基因替换为缺失型基因。由缺失型目标基因编码的蛋白质即使产生,也具有与野生型蛋白质不同的立体结构,其功能降低或消失。

[0176] 如上所述的基于利用同源重组的基因替换的基因破坏已经确立,包括下述方法:称为“Red 驱动整合(Red-driven integration)”的方法(Datsenko, K. A, and Wanner, B. L. Proc. Natl. Acad. Sci. U S A. 97:6640-6645(2000))、Red 驱动整合与 λ 噬菌体来源的切出系统(Cho, E. H., Gumpert, R. I., Gardner, J. F. J. Bacteriol. 184:5200-5203(2002))相结合的方法(参照 W02005/010175 号)等使用线性 DNA 的方法,或者使用含有温度敏感复制起点的质粒、可接合转移的质粒的方法,利用在宿主内不具有复制起点的自杀载体的方法,等等(美国专利第 6303383 号说明书或特开平 05-007491 号公报)。

[0177] 目标基因转录量的降低可以通过将从目标基因转录的 mRNA 的量与野生株或非修饰株进行比较来确认。评价 mRNA 量的方法包括例如 Northern 杂交、RT-PCR 等(Molecular

cloning(Cold spring Harbor Laboratory Press,Cold spring Harbor(USA),2001))。转录量的降低只要是与野生株或非修饰株相比降低即可,可以是任何形式的降低,理想的是例如,与野生株、非修饰株相比,降低至至少 75%以下、50%以下、25%以下或 10%以下,特别优选完全不表达。

[0178] 目标基因所编码的蛋白质的量降低可以通过使用与该蛋白质结合的抗体进行 western 印迹来确认 (Molecular cloning(Cold spring Harbor Laboratory Press, Cold spring Harbor(USA),2001))。蛋白质量的降低只要是与野生株或非修饰株相比降低即可,可以是任何形式的降低,理想的是例如,与野生株、非修饰株相比,减少至至少 75%以下、50%以下、25%以下或 10%以下,特别优选完全不产生蛋白质(活性完全消失)。

[0179] 除了上述基因操作方法以外,为了降低 GCD 的活性,还可以列举出例如下述方法:对于泛菌属细菌等肠杆菌科细菌,利用紫外线照射,或者利用 N-甲基-N'-硝基-N-亚硝基胍(NTG)或亚硝酸等等常规的突变处理中使用的突变剂进行处理,然后选择 GCD 的活性降低的菌株。

[0180] GCD 的活性的降低还可以通过降低 PQQ 的合成能力来进行。PQQ 的合成能力可以通过例如,缺失 PQQ 生物合成所必需的操纵子 pqqABCDEF 的一部分或全部来降低(J. S. Velterop, P. W. Postma, J. Bacteriology 177(17):5088-5098(1995))。

[0181] 在大肠杆菌、棒状杆菌型细菌等不具有 GCD 活性的微生物中,葡萄糖是利用称作葡萄糖 PTS(糖磷酸转移酶系统)或葡萄糖通透酶的转运蛋白来摄取的。PTS 是与 PEP(磷酸烯醇丙酮酸)转化为 Pyr(丙酮酸)的反应相偶联以葡萄糖 6-磷酸的形式摄入细胞内的。葡萄糖 6-磷酸转变为果糖 6-磷酸,通过所称的糖酵解系统(EMP:Embden-Meyerhof 途径)代谢,而生成丙酮酸。

[0182] 另一方面,在具有 GCD 活性的微生物中,葡萄糖在周质空间中暂时转变为葡糖酸后,在葡糖酸通透酶的作用下被摄入,通过磷酸化反应而生成 6-磷酸葡糖酸。

[0183] 6-磷酸葡糖酸通过戊糖磷酸循环或 Entner-Doudoroff(ED)途径代谢,生成甘油醛 3-磷酸、丙酮酸等。

[0184] 已知醋酸菌等具有 GCD 的微生物具备特有的糖代谢特性,即:将一部分葡萄糖在周质中暂时转变为葡糖酸后加以摄入。细胞内的 EMP 途径、ED 途径、戊糖磷酸循环的容量因微生物而异,因此预想若缺损 GCD 而改变糖代谢,其下游的代谢模式会发生变化。

[0185] 对于 *Pantoea ananatis* 而言,认为在通常的培养温度例如 34℃下,并非全部的葡萄糖都通过 GCD 途径同化,而是存在相当量通过 PTS 途径同化的葡萄糖。另一方面,若在高温例如 38℃下进行培养,则可以预想到因为 GCD 的最适温度高,GCD 活性提高,由此 GCD 途径的糖消费增加。*Pantoea ananatis* 中不存在 ED 途径,因而 6-磷酸葡糖酸由 6-磷酸葡糖酸脱氢酶脱氢化,通过戊糖磷酸循环代谢。在被 6-磷酸葡糖酸脱氢酶脱氢时,一分子的 6-磷酸葡糖酸释放一分子的二氧化碳,因此,可以预想若 GCD 路径的糖消费增加,则氨基酸生成量降低。此外还可以推断,当高温培养中 GCD 途径的糖消费增加时,戊糖磷酸循环的容量出现不足,溢出的代谢物流向副产物方向,结果是 L-氨基酸收率降低。可以认为,在本发明中,通过降低 GCD 活性,特别是在高温培养时,消除了二氧化碳的释放和向戊糖磷酸循环的溢出,从而 L-氨基酸生产力提高。

[0186] 此外,可以推断,是在导入了 ED 途径的 *Pantoea ananatis*(特开 2003-274988)中,

若 GCD 路径的糖消费增加,也会因基于 6-磷酸葡萄糖酸脱氢酶的 6-磷酸葡萄糖酸脱氢反应中二氧化碳的释放、以及 ED 途径、戊糖磷酸循环的容量不足,而导致 L-氨基酸收率降低。因此,可以认为,在导入了 ED 途径的 *Pantoea ananatis* 中,通过降低 GCD 活性,也能够提高 L-氨基酸生产力。

[0187] 本发明中使用的细菌可以是降低了 GCD 活性、且强化了糖的摄入活性的细菌。为了提高糖的摄入活性,例如可以提高葡萄糖 PTS 或葡萄糖通透酶的活性。此外,还已知半乳糖通透酶 (Flores et al. J Mol Microbiol Biotechnol 2007;13:105-116)、木糖通透酶 (EP1807445A1)、阿拉伯糖通透酶等易化扩散载体超家族 (Major Facilitator superfamily (MFS)) (Griffith, J. K. et al, Curr. Opin. Cell Biol. 4(4);684-95(1992)) 成员的转运蛋白也具有摄入葡萄糖等的活性。因此,在 GCD 活性降低的细菌中,通过强化这些转运蛋白的活性,也能够提高葡萄糖等糖的摄入,从而提高 L-氨基酸生产力。

[0188] <2> 本发明的 L-氨基酸的生产方法

[0189] 可以通过用培养基培养本发明的微生物,在该培养基中生成并蓄积 L-氨基酸,并且从该培养基收集 L-氨基酸,来生产 L-氨基酸。

[0190] 用于培养的培养基可以使用含有碳源、氮源、无机盐类以及视需要的氨基酸、维生素等有机微量营养元素的常规培养基。可以使用合成培养基或者天然培养基。培养基中使用的碳源和氮源可以是任意种类的,只要所培养的菌株能够利用即可。

[0191] 碳源可以使用糖类,如葡萄糖、甘油、果糖、蔗糖、麦芽糖、甘露糖、半乳糖、淀粉水解物和糖蜜,此外也可以单独使用或者与其它碳源组合使用有机酸,如乙酸和柠檬酸,以及醇,如乙醇。作为氮源,可以使用氨、铵盐如硫酸铵、碳酸铵、氯化铵、磷酸铵或乙酸铵,或者使用硝酸盐。作为有机微量营养物,可以使用氨基酸、维生素、脂肪酸和核酸,以及包含这些营养物的蛋白胨、酪蛋白氨基酸、酵母提取物、大豆蛋白水解物等,并且在使用生长要求氨基酸等的营养缺陷型突变菌株时,优选补充所需要的营养物。

[0192] 特别地,当使用调整为能够使 L-谷氨酸析出的条件的液体培养基时,在培养基中添加泛酸,能够更有效地晶析 (W02004/111258 号小册子)。无机盐类可以使用磷酸盐、镁盐、钙盐、铁盐、锰盐等。

[0193] 培养优选控制在发酵温度 20~45℃、pH 3-9 的条件下进行通气培养。当培养中 pH 下降时,可以例如添加碳酸钙,或用氨气等碱来中和。在这样的条件下,优选通过进行 10~120 小时左右的培养,在培养液中蓄积目的氨基酸。

[0194] 在本发明中,通过在适于细菌生长的温度下进行培养,能够高效地生产 L-氨基酸,特别是在高温进行培养时,效果显著。例如,对于 *Pantoea ananatis* 等泛菌属细菌而言,通常 34℃ 左右是生长优选的温度, GCD 活性降低的细菌与非修饰株相比在该培养温度下的 L-氨基酸生产能力更高,而在例如 36℃ 或 38℃, L-氨基酸生产能力进一步提高。

[0195] 此外,可以使用调整为能够析出 L-谷氨酸的条件的液体培养基,在使 L-谷氨酸在培养基中析出的同时进行培养。作为 L-谷氨酸析出的条件,可以列举出例如 pH 5.0~4.0、优选 pH 4.5~4.0、更优选 pH 4.3~4.0、特别优选 pH 4.0。

[0196] 此外,当使 L-谷氨酸在培养基中析出时,预先添加 L-谷氨酸、L-赖氨酸晶体作为晶种,能够更有效地进行晶析 (欧洲专利 1233069 号、欧洲专利申请公开 1624069 号)。

[0197] 作为从培养结束后的培养液中收集 L-氨基酸的方法,可以采用公知的回收方法

来进行。例如,可以采用从培养液中除去菌体后进行浓缩晶析的方法,或者通过离子交换色谱等来收集。当在能够使 L- 谷氨酸析出的条件下进行培养时,析出到培养液中的 L- 谷氨酸可以采用离心分离或过滤等来收集。这种情况下,可以在使溶解在培养基中的 L- 谷氨酸晶析后,一并进行分离。

[0198] 而且,在生产碱性氨基酸时,可以采用通过下述方法进行发酵、并回收目的碱性氨基酸的方法来进行生产:将培养中的 pH 调控为 6.5 ~ 9.0、将培养结束时的培养基的 pH 调控为 7.2 ~ 9.0,将发酵中发酵罐内压力调控为正,或者向培养基中供给二氧化碳气或含二氧化碳气的混合气体,使得存在培养基中的碳酸氢根离子和 / 或碳酸根离子至少为 2g/L 以上的培养期,将所述碳酸氢根离子和 / 或碳酸根离子作为以碱性氨基酸为主的阳离子的抗衡离子(参照特开 2002-065287 号、美国专利申请公开第 2002025564 号)。

[0199] 本发明中收集的 L- 氨基酸除了含有作为目标的 L- 氨基酸之外,还可以含有微生物菌体、培养基成分、水分以及微生物的代谢副产物等。收集的 L- 氨基酸的纯度为 50% 以上,优选为 85% 以上,特别优选为 95% 以上(US5,431,933, JP1214636B, US4,956,471, US4,777,051, US4946654, US5,840358, US6,238,714, US2005/0025878)。

[0200] 当采用本发明的方法生产 L- 半胱氨酸时,得到的 L- 半胱氨酸可以用来生产 L- 半胱氨酸的衍生物。L- 半胱氨酸的衍生物包括甲基半胱氨酸、乙基半胱氨酸、羧甲基半胱氨酸(carbocysteine)、磺酸半胱氨酸、乙酰半胱氨酸等。

[0201] 此外,当 L- 半胱氨酸的噻唑烷衍生物在培养基中积累时,从培养基中采集噻唑烷衍生物,使得噻唑烷衍生物和 L- 半胱氨酸之间的反应平衡向 L- 半胱氨酸一侧移动,从而可以生产 L- 半胱氨酸。此外,当 S- 磺酸半胱氨酸在培养基中积累时,例如可以使用二硫苏糖醇等还原剂通过进行还原转化得到 L- 半胱氨酸。

实施例

[0202] 以下,通过实施例对本发明进行更具体的说明。

[0203] (参考例 1) λ Red 基因产物抗性的 *Pantoea ananatis* 菌株的构建

[0204] 为了在 *Pantoea ananatis* 中进行基因缺损,构建了用于高效率地进行称作“Red 驱动整合”或“Red 介导整合”的方法(Proc. Natl. Acad. Sci. USA. 97. 6640-6645(2000))的受体菌。

[0205] 首先,构建了表达 λ 的 gam、bet 和 exo 诸基因(以下称“ λ Red 基因”)的新辅助质粒 RSF-Red-TER(图 1)。详细描述见参考例 2。

[0206] 该质粒可以用于具有不同基因背景的广域宿主。理由是:1) 其具有在许多革兰氏阴性菌和革兰氏阳性菌甚至植物中都能够稳定保持的 RSF1010 广宿主域质粒的复制子(Scholz, 等, 1989; Buchanan-Wollaston 等, 1987), 2) λ Red 基因、gam、bet 和 exo 基因受到能够被许多细菌的 RNA 聚合酶识别的 PlacUV5 启动子的调节(例如 Brunschwig, E. 和 Darzins, A., Gene, 111, 1, 35-41(1992); Dehio, M. 等, Gene, 215, 2, 223-229(1998)), 3) 自调节因子 P_{lacUV5} -lacI、以及大肠杆菌 rrnB 操纵子的 ρ 非依赖性转录终止子(TrrnB)降低 λ Red 基因的本底表达水平(Skorokhodova, A. Yu 等, Biotekhnologiya(俄语), 5, 3-21(2004))。而且,RSF-Red-TER 质粒包含果聚糖蔗糖酶(levansucrase)基因(sacB),利用该基因,可以在含蔗糖的培养基中从细胞回收质粒。

[0207] 在大肠杆菌中, PCR 反应生成的 DNA 片段与由 RSF-Red-TER 质粒提供的短侧翼区域一起发生整合的频率与使用 pKD46 辅助质粒 (Datsenko, K. A., Wanner, B. L., Proc. Natl. Acad. Sci. USA, 97, 6640-6645, (2000)) 时一样高。然而, λ Red 基因的表达对 *Pantoea ananatis* 显示毒性。用 RSF-Red-TER 辅助质粒转化的细胞在含 IPTG (异丙基- β -D-硫代半乳糖苷, 1mM) 和适当抗生素 (氯霉素 25 μ g/ml 或卡那霉素 40 μ g/ml) 的 LB 培养基上显示非常低的生长速度, λ Red 介导的重组 (λ Red-mediated recombination) 的效率即使能够观察到, 也是非常低的 (10^{-8})。

[0208] 选择了对 λ Red 基因的全部 3 个基因的表达均具有抗性的 *Pantoea ananatis* 突变菌株。为此, 通过电穿孔将 RSF-Red-TER 质粒导入 *Pantoea ananatis* SC17 株 (美国专利第 6, 596, 517 号)。培养 18 小时后, 得到了约 10^6 个转化株, 有 10 个克隆的菌落尺寸大, 其余的菌落都非常小。培养 18 小时后, 大菌落为约 2mm, 小菌落为约 0.2mm。将培养延长至 24 小时, 小菌落不再继续生长, 而大菌落继续生长。将对 λ Red 基因的全部 3 个基因 (gam、bet 和 exo) 的表达均具有抗性的大菌落 *Pantoea ananatis* 突变菌株中的一个用于进一步分析。

[0209] 从 1 个大菌落克隆和数个小菌落克隆中分离出 RSF-Red-TER 质粒 DNA, 用其再次转化大肠杆菌 MG1655, 考察了质粒合成 Red 基因的活性产物的能力。由所得转化体中 Red 依赖性整合的对照实验可知, 只有从大菌落克隆分离的质粒可引起 Red 依赖性整合所必需的 λ Red 基因的表达。为了考察所选的大菌落克隆中是否发生 Red 介导的整合, 使用 PCR 反应生成的下述线性 DNA 片段进行了电穿孔, 所述 DNA 片段设计为包含 Km^R 标记和与 hisD 基因同源的 40bp 侧翼区域, 并且整合到 *Pantoea ananatis* 的 hisD 基因的 SmaI 识别位点。将 2 个小菌落的克隆用作对照。 *Pantoea ananatis* 的 hisD 基因的核苷酸序列示于 SEQ ID NO:3。PCR 中使用 SEQ ID NO:4 和 5 的寡核苷酸作为引物, 使用 pMW118-(λ attL- Km^R - λ attR) 质粒作为模板。使用对 λ Red 基因不具有抗性的 2 个小菌落的克隆作为对照。pMW118-(λ attL- Km^R - λ attR) 质粒的构建在参考例 3 中详细描述。

[0210] RSF-Red-TER 质粒能够通过该质粒上的 lacI 基因诱导 Red 基因的表达。对两种诱导条件进行了考察。在第 1 组中, 电穿孔 1 小时前添加 IPTG (1mM); 在第 2 组中, 在用于制备可电穿孔细胞的培养开始时添加 IPTG。携带来自大菌落克隆的 RSF-Red-TER 的细胞的后代的生长速度并未显著地低于不带有该质粒的菌株。添加 IPTG 仅使这些培养物的生长速度有些许降低。另一方面, 小菌落克隆的后代在未添加 IPTG 的条件下生长非常缓慢, 而加以诱导后生长基本上停止。对大菌落克隆的后代的细胞进行电穿孔后, 有许多 Km^R 克隆 (短诱导时间为 18 个克隆, 延长诱导时间后为约 100 个克隆) 生长。考察的全部 100 个克隆均具有 His⁻ 表现型, 通过 PCR 对 20 个克隆进行了确认, 结果确认了这些细胞的染色体结构与预想的相同。另一方面, 对小菌落克隆的后代的细胞进行电穿孔仍未能得到整合的菌株。

[0211] 使所得大菌落克隆在含 7% 蔗糖的平板上生长以使质粒脱落, 用 RSF-Red-TER 再次进行转化。将不携带质粒的菌株命名为 SC17 (0)。

[0212] 上述再次转化后生长出的全部克隆均与亲本菌株克隆 SC17 (0) 一样具有较大的菌落尺寸。在用 RSF-Red-TER 质粒再次转化的 SC17 (0) 株中进行了 Red 介导的整合实验。对于所得的 3 个独立转化株, 使用与前述实验所用的相同的 DNA 片段进行了考察。采用了短诱导时间 (电穿孔 1 小时前)。在各实验中, 生长出超过 10 个的 Km^R 克隆。测试的全部

克隆均具有 His⁻ 表现型。这样,选择出突变菌株,将其命名为对 λ Red 基因的表达具有抗性的 SC17(0)。该菌株可以作为用于对 *Pantoea ananatis* 染色体的 Red 依赖性整合的合适受体菌。

[0213] (参考例 2) 辅助质粒 RSF-Red-TER 的构建

[0214] 辅助质粒 RSF-Red-TER 的构建方案如图 2 所示。

[0215] 作为构建的最初步骤,设计了 RSFsacBP_{lac}MCS 载体。为此,分别使用 SEQ ID NO: 6、7、8、9 的寡核苷酸通过 PCR 扩增了包含 pACYC184 质粒的 cat 基因和枯草芽孢杆菌的 sacB 基因的结构部分的 DNA 片段。这些寡核苷酸的 5' 末端分别包含对进一步克隆而言必需且便利的 BglII、SacI、XbaI、和 BamHI 限制酶位点。将所得 1.5kb 的 sacB 片段克隆到此前获得的 pMW119-P_{lac}lacI 载体的 XbaI-BamHI 位点。该载体是依照与 pMW118-P_{lac}lacI 载体的相关描述 (Skorokhodova, A. Yu 等, Biotekhnologiya (俄语), 5, 3-21 (2004)) 相同的方法构建的。只是该载体包含来自 pMW219 而不是 pMW218 质粒的多接头位点。

[0216] 接下来,将上述的 1.0kb 的 cat 片段用 BglII 和 SacI 处理,并克隆到在先步骤中所得的 RSF-P_{lac}lacIsacB 质粒的 BamHI-SacI 位点。所得质粒 pMW-P_{lac}lacIsacBcat 包含 PlacUV5-lacI-sacB-cat 片段。为了将该片段亚克隆到 RSF1010 载体,将 pMW-P_{lac}lacIsacBcat 用 BglII 消化,并用 DNA 聚合酶 I 克列诺片段处理使末端平滑化,然后用 SacI 切割。从 1% 琼脂糖凝胶中洗提出 pMW-P_{lac}lacIsacBcat 质粒的 3.8kb 的 BglII-SacI 片段,将其与用 PstI 和 SacI 处理的 RSF1010 载体连接。用连接混合液转化大肠杆菌 TG1, 将其在含氯霉素 (50mg/L) 的 L 培养基 (纯水 1L 中包含细菌用胰蛋白胨 10g、酵母提取物 5g、NaCl 5g、琼脂 15g 而成的培养基, pH7.0) 上铺板。对从生长出的克隆中分离的质粒进行限制酶分析,得到了 RSFsacB 质粒。为了构建 RSFsacBP_{lac}MCS 载体,以 SEQ ID NO:10 以及 11 的寡核苷酸为引物,以 pMW119-P_{lac}lacI 质粒为模板,通过 PCR 扩增了包含 P_{lacUV5} 启动子的 DNA 片段。将所得的 146bp 片段用 SacI 和 NotI 消化,将其与 RSFsacB 质粒的 SacI-NotI 大片段连接。然后,以 SEQ ID NO:12 和 13 的寡核苷酸为引物,以 pKD46 质粒 (Datsenko, K. A., Wanner, B. L., Proc. Natl. Acad. Sci. USA, 97, 6640-6645, (2000)) 为模板,通过 PCR 扩增了包含 λ Red α β γ 基因和转录终止子 tL3 的 2.3kb 的 DNA 片段。将所得片段克隆到 RSFsacBP_{lac}MCS 载体的 PvuI-NotI 位点。这样,设计了 RSFRed 质粒。

[0217] 为了排除 Red 基因的通读转录,将大肠杆菌 rrnB 操纵子的 ρ -依赖性转录终止子插入到 cat 基因与 P_{lacUV5} 启动子之间。为此,以 SEQ ID NO:14 和 11 的寡核苷酸为引物,以大肠杆菌 BW3350 的染色体为模板,通过 PCR 扩增了包含 P_{lacUV5} 启动子与 TrnB 终止子的 DNA 片段。将所得的这些片段用 KpnI 处理,并进行连接。然后,以 SEQ ID NO:11 和 15 的寡核苷酸为引物,通过使用寡核苷酸的 PCR 扩增了包含 P_{lacUV5} 和 TrnB 这两者的 0.5kb 片段。将所得 DNA 片段用 EcoRI 消化,并用 DNA 聚合酶 I 克列诺片段处理使末端平滑化,然后用 BamHI 切割,将其与 RSFsacBP_{lac}MCS 载体的 Ec1136II-BamHI 大片段连接。将所得质粒命名为 RSF-Red-TER。

[0218] (参考例 3) pMW118-(λ attL-Km^r- λ attR) 质粒的构建

[0219] pMW118-(λ attL-Km^r- λ attR) 质粒通过用 pUC4K 质粒的卡那霉素抗性基因取代 pMW118-attL-Tc-attR(W02005/010175) 质粒中的四环素抗性标记基因来构建。为此,将 pMW118-attL-Tc-attR 质粒的 EcoRI-HindIII 大片段与 pUC4K 质粒的 HindIII-PstI (676bp)

和 EcoRI-HindIII (585bp) 这 2 个片段连接。作为基础的 pMW118-attL-Tc-attR 是通过连接以下 4 个片段而得到的。

[0220] 1) 使用引物 P1 (SEQ ID NO :16) 和 P2 (SEQ ID NO :17) 从大肠杆菌 W3350 (包含 λ 原噬菌体) 染色体的相当于 attL 的区域通过 PCR 扩增得到的具有 attL (SEQ ID NO :18) 的 BglIII-EcoRI 片段 (114bp)。这些引物包含 BglIII 和 EcoRI 的附加识别位点。

[0221] 2) 从大肠杆菌 W3350 (包含 λ 原噬菌体) 染色体的相当于 attR 的区域使用引物 P3 (SEQ ID NO :19) 和 P2 (SEQ ID NO :20) 通过 PCR 扩增得到的具有 attR (SEQ ID NO :21) 的 PstI-HindIII 片段 (182bp)。这些引物包含 PstI 和 HindIII 的附加识别位点。

[0222] 3) pMW118-ter_rrnB 的 BglIII-HindIII 大片段 (3916bp)。质粒 pMW118-ter_rrnB 是通过连接以下 3 个 DNA 片段而得到的。

[0223] • pMW118 的具有 AatII-EcoRI 片段的大片段 (2359bp)。该片段是通过用 EcoRI 消化 pMW118、并用 DNA 聚合酶 I 克列诺片段进行处理、再用 AatII 进行消化而得到的。

[0224] • 具有氨苄青霉素抗性 (Ap^R) 的基因 bla 的 pUC19 的 AatII-BglIII 小片段 (1194bp)。该片段是使用引物 P5 和 P6 (SEQ ID NO :22 和 23) 对 pUC19 质粒的相应区域进行 PCR 扩增而得到的。这些引物中包含 PstI、AatII 和 BglIII 的附加识别位点。

[0225] • 转录终止子 ter_rrnB 的 BglIII-PstI 小片段 (363bp)。该片段是通过使用引物 P7 和 P8 (SEQ ID NO :24 和 25) 对大肠杆菌 MG1655 染色体的相应区域进行 PCR 扩增而得到的。这些引物包含 PstI、BglIII 和 PstI 的附加识别位点。

[0226] 4) 具有四环素抗性基因和 ter_thrL 转录终止子的 pML-Tc-ter_thrL 的 EcoRI-PstI 小片段 (1388bp) (SEQ ID NO :26)。pML-Tc-ter_thrL 质粒是通过如下的 2 个步骤获得的。

[0227] • 将 pML-MCS 质粒 (Mashko, S. V. 等, Biotekhnologiya (俄语), 2001, no. 5, 3-20) 用 XbaI 和 BamHI 消化, 然后将大片段 (3342bp) 与包含 ter_thrL 终止子的 XbaI-BamHI 片段 (68bp) 连接。该包含 ter_thrL 终止子的片段是使用引物 P9 和 P10 (SEQ ID NO :27 和 28) 对大肠杆菌 MG1655 染色体的相应区域进行 PCR 而得到的。这样, 得到了 pML-ter_thrL 质粒。这些引物包含 PstI、XbaI 和 BamHI 的附加识别位点。

[0228] • 将 pML-ter_thrL 质粒用 KpnI 和 XbaI 消化, 然后用 DNA 聚合酶 I 克列诺片段处理, 再与具有四环素抗性基因的 pBR322 的 EcoRI-Van91I 小片段 (1317bp) 连接, 从而得到了 pML-Tc-ter_thrL 质粒。而且, 将 pBR322 用 EcoRI 和 Van91I 消化, 然后再用 DNA 聚合酶 I 克列诺片段进行了处理。

[0229] (参考例 4) 谷氨酸生产质粒 RSFPPG 的构建

[0230] 构建了 L-谷氨酸生物合成系统基因, prpC 基因 (国际公布 2006/051660 号小册子)、ppc 基因、gdh 基因 (欧洲申请公布 0999282 号说明书) 被扩增的质粒 RSFPPG (WO2008/020654)。

[0231] 设计了引物 1 (SEQ ID NO :29) 和引物 2 (SEQ ID NO :30), 用于扩增 RSFCPG (欧洲申请公开 1233068 号说明书) 的 g1tA 基因的 ORF 以外的部分。使用该引物, 以 RSFCPG 为模板进行 PCR, 获得了约 14.9kb 的片段。另一方面, 对于 prpC, 使用引物 3 (SEQ ID NO :31) 和引物 4 (SEQ ID NO :32), 以大肠杆菌 W3110 株的染色体 DNA 为模板进行 PCR, 获得了约 1.2kb 的片段。将两 PCR 产物分别用 BglIII 和 KpnI 处理并进行连接, 然后转化大肠杆菌 JM109 株。

收集出现的所有菌落,以混合物的形式提取质粒。用该质粒混合物转化柠檬酸合酶(CS) 缺损株大肠杆菌 ME8330 株,将其涂布于含 50mg/L 尿嘧啶、5mg/L 硫胺素-HCl 的 M9 极限培养基(1L 纯水中含葡萄糖 5g、硫酸镁 2mM、磷酸二氢钾 3g、氯化钠 0.5g、氯化铵 1g、磷酸氢二钠 6g 而成的培养基)。收集出现的所有菌落,以混合物的形式提取质粒,用该质粒混合物转化 *P. Ananatis* L- 谷氨酸生产菌 NP106 株。对于出现的克隆,在中性条件下进行试管培养,将显示与 G106S 株等同的 L- 谷氨酸收率的菌株作为 NA1。此外,从该菌株抽提质粒,将其作为 prpC、gdh、ppc 强化用质粒 RSFPPG。在作为 L- 谷氨酸生产菌的 *Pantoea ananatis* NP106 株中导入所述质粒 RSFPPG,构建了 L- 谷氨酸生产菌 NP106/RSFPPG (该菌株称作“NA1 株”)。

[0232] NP106 株是如下述地获得的。将前文例示的 *Pantoea ananatis* AJ13601 株在 LBGM9 培养基中于 34℃ 振荡培养一夜,稀释到每个平板为 100 ~ 200 个菌落,将其涂布在含四环素 12.5mg/L 的 LBGM9 平板上,所述 LBGM9 培养基是在 L 培养基(细菌用胰蛋白胨 10g/L、酵母提取物 5g/L、NaCl 5g/L、pH7.0)中添加极限培养基成分(葡萄糖 5g/L、硫酸镁 2mM、磷酸二氢钾 3g/L、氯化钠 0.5g/L、氯化铵 1g/L、磷酸氢二钠 6g/L)而得到的液体培养基。对于出现的菌落,将其转印到含四环素 12.5mg/L 和氯霉素 25mg/L 的 LBGM9 平板上,选择氯霉素敏感性的菌株,从而获得了 pSTVCB 脱落的菌株,将其命名为 G106S。而且,将 G106S 株在 LBGM9 液体培养基中于 34℃ 振荡培养一夜,稀释到每个平板为 100 ~ 200 个菌落,将其涂布在不含药物的 LBGM9 平板上。对于出现的菌落,将其转印在含四环素 12.5mg/L 的 LBGM9 平板和不含药物的 LBGM9 平板上,选择四环素敏感性菌株,从而获得了 RSFCPG 脱落的菌株,将其命名为 NP106。这样获得的 NP106 株不再具有 AJ13601 株所携带的 RSFCPG 和 pSTVCB 这 2 个质粒。

[0233] G106S 株是像上述那样地在 AJ13601 株中仅脱落 pSTVCB 而得到的菌株。

[0234] (实施例 1) 基于 gcd 基因缺损株的 L- 谷氨酸生产

[0235] (1) gcd 基因缺损株的构建

[0236] 采用常规方法合成了 SEQ ID NO :33 和 34 所示的 2 个合成 DNA 引物。

[0237] SEQ ID NO :33 所示的引物具有 *Pantoea ananatis* 的 gcd 基因上游的同源序列与 λ attL-Km^r- λ attR 的 5' 端的同源序列相接而成的结构。SEQ ID NO :34 的引物具有 *Pantoea ananatis* 的 gcd 基因下游的互补序列与 λ attL-Km^r- λ attR 的 3' 端的互补序列相接而成的结构。通过使用这些引物,以 pMW118-(λ attL-Km^r- λ attR) 为模板进行 PCR,扩增了约 1.5kbp 的片段,其中在 λ attL-Km^r- λ attR 序列的 5' 端连接 gcd 基因上游的同源序列,在 λ attL-Km^r- λ attR 序列的 3' 端连接 gcd 基因下游的同源序列。

[0238] 纯化上述 PCR 片段,将其用于针对 *Pantoea ananatis* 染色体的 λ 依赖性整合。将辅助质粒 RSF-Red-TER 用作 λ 噬菌体 Red 基因的载体。为了得到 *Pantoea ananatis* 的电转化感受态细胞,用 RSF-Red-TER 质粒对 SC17(0) 株进行转化,然后在含 50 μ g/ml 氯霉素的 LB 培养基上于 34℃ 培养一夜。接下来,将培养液用含 50 μ g/ml 氯霉素的新鲜 LB 培养基稀释 100 倍,34℃ 通气下进行生长使得 OD₆₀₀ 为 0.3。然后,添加 1mM 的 IPTG,继续培养至 OD₆₀₀ 为 0.7。将 10mL 培养液中的菌体用等量的冷的 10% 甘油洗涤 3 次,悬浮在 80 μ l 的冷 10% 甘油中。将上述 PCR 产物溶解在 10 μ l 去离子水中,将 PCR 片段 100-200ng 添加到细胞悬浮液中。操作使用细菌电穿孔装置(“BioRad”美国、商品号 165-2089,版本 2-89) 进

行。使用的脉冲的参数为电场强度:18kV/cm,脉冲时间:5 毫秒。

[0239] 电穿孔后,立即在细胞悬浮液中添加补充有葡萄糖(0.5%)的 LB 培养基 1ml。然后,使细胞在 34℃通气下生长 2 小时,用含 40 μg/ml 的卡那霉素的 L 培养基(纯水 1L 中包含细菌用胰蛋白胨 10g、酵母提取物 5g、NaCl 5g、琼脂 15g 的培养基, pH7.0)进行选择,获得了约 20 个菌落作为转化体。使用 SEQ ID NO:35 和 SEQ ID NO:36 所示的 2 个合成 DNA 引物,通过 PCR 确认了 gcd 基因区域中插入了卡那霉素抗性基因的片段,将能够确认片段插入的菌株命名为 SC17(0)::Δgcd。从该菌株抽提基因组 DNA,通过电穿孔转化了 NA1 株。

[0240] 将导入了 SC17(0)::Δgcd 的基因组 DNA 的 NA1 株用添加了 40mg/L 卡那霉素、12.5mg/L 四环素盐酸盐和琼脂 15g/L 的 LBGM9 培养基平板进行选择,获得了约 20 个的菌落作为转化体。这些菌株的 gcd 基因区域中均插入了 λ attL-Km^r-λ attR 的片段,选出其中的 1 个克隆,将其命名为 NA1::Δgcd。

[0241] (2)gcd 基因缺损株的 L-谷氨酸生产能力评价

[0242] 为了研究 gcd 基因的缺损对 L-谷氨酸生产的影响,使用 NA1::Δgcd 株和 NA1 株,进行了 L-谷氨酸生产培养。

[0243] 培养分为形成菌体的种子培养和生成 L-谷氨酸的主培养这两个阶段进行。

[0244] 种子培养的培养基组成如下。

[0245]

〔种子培养培养基组成〕

蔗糖	50g/L
MgSO ₄ ·7H ₂ O	0.4g/L
GD113(消泡剂)	0.1mL/L
(NH ₄) ₂ SO ₄	4.0g/L
KH ₂ PO ₄	2.0g/L
酵母提取物	4.0g/L
FeSO ₄ ·7H ₂ O	0.01g/L
MnSO ₄ ·5H ₂ O	0.01g/L
柠檬酸	0.02g/L
L-赖氨酸盐酸盐	0.4g/L
DL-甲硫氨酸	0.4g/L
ε-二氨基庚二酸	0.4g/L
泛酸钙	18mg/L

[0246]

四环素盐酸盐	12.5mg/L
--------	----------

[0247] 120℃、蒸气灭菌 20 分钟。

[0248] 将 NA1::Δgcd 株和 NA1 株用添加了 12.5mg/L 四环素和琼脂 15g/L 的 LBGM9 培养

基平板进行前培养,将相当于1张平板的量接种到装有上述组成的培养基300mL的1L容量的微型罐中,34℃、pH6.0进行约12小时搅拌培养,并进行控制使得通气为1/1vvm、氧浓度为3%以上。添加氨气进行调整使得培养中的pH为6.0。以培养基中的糖的耗尽为指标,结束种子培养。

[0249] 主培养培养基组成如下。

[0250]

〔培养培养基组成〕 (接菌20%种子培养液后的浓度)

葡萄糖	100g/L
MgSO ₄ ·7H ₂ O	0.4g/L
GD113	0.1mL/L
(NH ₄) ₂ SO ₄	5.0g/L
KH ₂ PO ₄	6.0g/L
酵母提取物	6.0g/L
FeSO ₄ ·7H ₂ O	0.02g/L
MnSO ₄ ·5H ₂ O	0.02g/L
柠檬酸	0.02g/L
甜菜碱 [*]	2.0g/L
L-赖氨酸盐酸盐	0.8g/L
DL-甲硫氨酸	0.6g/L
ε-二氨基庚二酸	0.6g/L
泛酸钙	18mg/L
四环素盐酸盐	25mg/L

[0251] ^{*}:N-N-N-三甲基甘氨酸

[0252] 将种子培养所得的菌体60mL注入装有上述组成的培养基240mL的1L容量的微型罐中,温度34℃、36℃或38℃,pH4.9进行培养。在培养基中的葡萄糖全部耗尽的时间点上结束培养。L-谷氨酸浓度通过将培养液上清用水稀释适当倍率后、使用BIOTECHANALYSER(生物技术分析仪)(AS-210SAKURA精机)来测定。

[0253] 结果如表1所示。可以判定:与比较对照株NA1株相比,gcd缺损株NA1::Δgcd株的L-谷氨酸蓄积提高。

[0254] [表1]

[0255] 表1生成的L-谷氨酸(g/罐)

[0256]

菌株	34℃	36℃	38℃

NA1	13.0	12.4	13.3
NA1::Δgcd	15.4	16.0	16.7

[0257] (实施例 2) 基于 gcd 基因缺损株的 L- 半胱氨酸生产

[0258] (1) L- 半胱氨酸生产菌的构建

[0259] 为了考察在 *P. Ananatis* 中 gcd 基因缺损对 L- 半胱氨酸生产的效果, 构建了 L- 半胱氨酸生产菌。

[0260] (1-1) yeaS 基因表达质粒的构建

[0261] 首先, 构建了用于构建上述菌株的质粒。方法如下。

[0262] 以大肠杆菌 MG1655 (ATCC No. 47076) 的染色体 DNA 为模板, 使用 P11 (agctgagtcg acccccagga aaaattgggtt aataac; SEQ ID NO: 51) 和 P12 (agctgagcatgcttccaact gcgctaata ga cgc; SEQ ID NO: 52) 作为引物, 通过 PCR 获得了含有 nlpD 基因启动子区域 (以下将野生型 nlpD 基因启动子记作“Pnlp0”) 约 300bp 的 DNA 片段。在上述引物的 5' 末端、3' 末端分别设计了限制性内切酶 SalI 以及 PaeI 的位点。PCR 循环如下所述: 95°C 3 分钟后, 95°C 60 秒、50°C 30 秒、72°C 40 秒循环 2 次, 94°C 20 秒、55°C 20 秒、72°C 15 秒循环 25 次, 最后 72°C、5 分钟。用 SalI 以及 PaeI 处理得到的片段, 并将其插入到 pMIV-5JS (特开 2008-99668) 的 SalI-PaeI 位点中, 得到 pMIV-Pnlp0 质粒。在该质粒 pMIV-Pnlp0 中插入的 Pnlp0 启动子的 PaeI-SalI 片段的碱基序列如 SEQ ID NO: 41 所示。

[0263] 随后, 以 MG1655 的染色体 DNA 为模板, 使用 P13 (agctgatcta gaaaacagaatttgccctggc ggc; SEQ ID NO: 53) 和 P14 (agctgaggat ccaggaagag tttgtagaaa cgc; SEQ ID NO: 54) 作为引物, 通过 PCR 获得了含有 rrnB 基因的终止子区域约 300bp 的 DNA 片段。在上述引物的 5' 末端分别设计了限制性内切酶 XbaI 以及 BamHI 的位点。PCR 循环如下所述: 95°C 3 分钟后, 95°C 60 秒、50°C 30 秒、72°C 40 秒循环 2 次, 94°C 20 秒、59°C 20 秒、72°C 15 秒循环 25 次, 最后在 72°C 下 5 分钟。用 XbaI 以及 BamHI 处理得到的片段, 并将其插入到 pMIV-Pnlp0 的 XbaI-BamHI 位点处从而得到 pMIV-Pnlp0-ter 质粒。

[0264] 接下来, 以 MG1655 的染色体 DNA 为模板, 使用 P15 (agctgagtcgacgtgttcgc tgaatacggg gt; SEQ ID NO: 55) 和 P16 (agctgatcta gagaaagcat caggattgcagc; SEQ ID NO: 56) 作为引物, 通过 PCR 获得了含有 yeaS 基因的约 700bp 的 DNA 片段。在上述引物的 5' 末端分别设计了限制性内切酶 SalI 以及 XbaI 的位点。PCR 循环如下所述: 95°C 3 分钟后, 95°C 60 秒、50°C 30 秒、72°C 40 秒循环 2 次, 94°C 20 秒、55°C 20 秒、72°C 15 秒循环 25 次, 最后在 72°C 下 5 分钟。用 SalI 以及 XbaI 处理得到的片段, 并将其插入到 pMIV-Pnlp0-ter 的 SalI-XbaI 位点处从而得到 pMIV-Pnlp0-YeaS3 质粒。由此, 在 pMIV-5JS 载体上构建了按照 nlpD 启动子、yeaS 基因以及 rrnB 终止子的顺序连接的 yeaS 的表达单元。

[0265] 为了通过修饰 nlpD 启动子的 -10 区域来得到更强的启动子, 采用下述方法进行了 -10 区域的随机化。在 nlpD 启动子区域 (图 3) 中存在两个推测为起启动子功能的区域, 在图中分别由 pnlp1 和 pnlp2 表示。以质粒 pMIV-Pnlp0 为模板, 使用 P11 以及 P17 (atcgtgaaga tcttttccag tgtnnannag ggtgccttgccacggtnatna ngtcactgg (“n”表示可以为 a, t, g, c 中的任意一种), SEQ ID NO: 57) 作为引物, 通过 PCR 获得了将包含于 nlpD 启动子的 3' 末端侧的 -10 区域 (记为 -10 (Pnlp1)) 随机化而得的 DNA 片段 (图 3)。PCR 循

环如下所述:95℃ 3 分钟后,95℃ 60 秒、50℃ 30 秒、72℃ 40 秒循环 2 次,94℃ 20 秒、60℃ 20 秒、72℃ 15 秒循环 25 次,最后在 72℃ 下 5 分钟。

[0266] 另一方面,同样地以质粒 pMIV-Pnlp0 为模板,使用 P12 以及 P18(tggaaaagatcttcannnnn cgctgacctg cg(“n”表示可以为 a,t,g,c 中的任意一种),SEQ ID NO:58) 作为引物,通过 PCR 获得了将包含于 nlpD 启动子的 5' 末端侧的 -10 区域(记为 -10(Pnlp2)) 随机化而得的 DNA 片段(图 3)。PCR 循环如下所述:95℃ 3 分钟后,95℃ 60 秒、50℃ 30 秒、72℃ 40 秒循环 2 次,94℃ 20 秒、60℃ 20 秒、72℃ 15 秒循环 25 次,最后在 72℃ 下 5 分钟。

[0267] 得到的 3' 末端侧和 5' 末端侧的片段能够通过引物 P17 和 P18 上设计的 BglII 位点相连接,从而能够构建 2 个 -10 领域被随机化的完整 nlpD 启动子。以该片段为模板,使用 P11 以及 P12 作为引物,通过 PCR 获得了修饰型完整 nlpD 启动子的 DNA 片段。PCR 循环如下所述:95℃ 3 分钟后,95℃ 60 秒、50℃ 30 秒、72℃ 40 秒循环 2 次,94℃ 20 秒、60℃ 20 秒、72℃ 15 秒循环 12 次,最后在 72℃ 下 5 分钟。

[0268] 用在引物的 5' 末端上设计的限制性内切酶 SalI 以及 PaeI 处理扩增的片段,并通过将其插入到同样地用 SalI 以及 PaeI 处理的质粒 pMIV-Pnlp0-YeaS3 中,将质粒上的野生型 nlpD 启动子位点(Pnlp0) 替换成突变型 Pnlp。从其中选择了具有如图 3 所示的启动子序列(Pnlp8) 的质粒,并将其命名为 pMIV-Pnlp8-YeaS7。插入在该质粒中的 Pnlp8 启动子的 PaeI-SalI 片段的碱基序列如 SEQ ID NO:42 所示。

[0269] (1-2) 突变型 cysE 表达质粒的构建

[0270] 随后,从 pMW-Pomp-cysE5(W02005007841) 中,用 PaeI、SacI 将 Pomp-cysE5 盒(cassette) 的部分切出,插入到 pMIV-5JS 的相同位点处,从而构建了 pMIV-Pomp-CysE5。pMW-Pomp-cysE5 为将连接有 ompC 基因启动子的编码突变型 SAT 的基因 cysE5 插入到 pMW118 中而得到的质粒。从 pACYC184(GenBank/EMBL 登录号 X06403、可从 NIIPONGENE 公司获得) 上,用 XbaI、Eco88I 将四环素抗性基因切出,将该基因片段通过克列诺片段(Klenow fragment) 处理后,插入到 pMIV-Pomp-CysE5 的 PvuI 位点处,从而构建了 pMT-Pomp-CysE5。接下来,用 HindIII 酶切 pMIV-Pnlp8-YeaS7,并通过克列诺片段进行平滑末端化后,用 NcoI 进行酶切,将含有 Pnlp8-YeaS-rrnB 终止子的盒和氯霉素抗性标记的片段切出。将该片段和同样以 pMIV-5JS 为骨架(backbone) 的 pMT-Pomp-CysE5 的 SmaI、NcoI 酶切片段分别连接,从而构建了 pMT-EY2。pMT-EY2 为在同一个质粒上具有 Pnlp8-YeaS-rrnB 终止子盒和 Pomp-CysE5 盒的质粒。

[0271] (1-3) 将 cysE5、yeaS 导入 P. ananatis SC17 株

[0272] 上述的 pMT-EY2 具备来源于 pMIV-5JS(特开 2008-99668) 的 Mu 噬菌体附着位点(attachment site)。通过使质粒与具有 Mu 转座酶(Mu transposase) 的辅助质粒 pMH10(Zimenkov D. et al., Biotechnologiya(俄语), 6, 1-22(2004)) 在同一细胞内共存,可以将该 pMT-EY2 质粒上以被 Mu 噬菌体的附着位点包夹的形式存在的含有氯霉素抗性标记的 PompC-cysE5-Pnlp8-YeaS-rrnBterminator(终止子) 盒插入到 P. ananatis SC17(美国专利 6596517) 菌株的染色体上。而且,由于在质粒 pMT-EY2 上存在的氯霉素抗性标记具有夹在 2 个 λ 噬菌体的辅助位点(λ attR 和 λ attL) 之间的结构,可以通过后述的方法将氯霉素抗性标记除去。

[0273] 首先,将通过电穿孔向 SC17 菌株中导入了 pMH10 而得到的菌株在含有 20mg/L 卡

那霉素的 LB 琼脂培养基上 30℃ 过夜培养来进行选择。将得到的转化菌株在 30℃ 下培养,并进一步通过电穿孔向该菌株中导入 pMT-EY2。将这个用 pMH10 和 pMT-EY2 两者转化的菌株在 42℃, 20 分钟的条件下进行热冲击,然后在含有 20mg/L 氯霉素的 LB 琼脂培养基上选择氯霉素抗性菌株的菌落。此时培养温度为 39℃。通过上述方法,获得约 50 个克隆,通过分别在 LB 琼脂培养基上在 39℃ 下培养 48 小时,进行 pMH10 以及 pMT-EY2 的消除 (curing)。获得了因在染色体上插入了盒而显示出氯霉素抗性,且作为消除了两个质粒的结果对卡那霉素以及氨苄青霉素显示敏感性的菌株。随后,以该菌株的染色体 DNA 为模板,使用 P11 和 P16 作为引物,通过 PCR 确认了得到的菌株的染色体上目的盒的插入。得到的全部克隆分别命名为 EY01 ~ EY50,进行 EY01 ~ EY50 菌株的 L- 半胱氨酸生产培养。培养通过下述方法进行。结果,筛选出 L- 半胱氨酸生产最多的克隆 EY19 菌株。

[0274] 通过来源于 λ 噬菌体的切出系统除去 EY19 菌株中导入的氯霉素抗性标记。具体来说,用携带了 λ 噬菌体的 Int-Xis 基因的 pMT-Int-Xis2(W02005/010175) 转化 EY19 菌株,从得到的转化菌株中获得了显示氯霉素敏感性的 EY19(s) 菌株。

[0275] (1-4) 由 EY19(s) 菌株制作 cysPTWA 基因表达强化菌株

[0276] 接下来,为了强化 cysPTWA 基因的表达,将染色体上的在 cysPTWA 基因簇上游存在的启动子替换为上述的强启动子 Pnlp8。首先以 pMIV-Pnlp8-YeaS7 为模板,使用 P11 以及 P12 通过 PCR 获得含有 nlp8 启动子约 300bp 的 DNA 片段。PCR 循环如下所述:95℃ 3 分钟后,95℃ 60 秒、50℃ 30 秒、72℃ 40 秒循环 2 次,94℃ 20 秒、59℃ 20 秒、72℃ 15 秒循环 20 次,最后在 72℃ 下 5 分钟。

[0277] 将扩增得到的含有 nlp8 启动子的 DNA 片段利用克列诺片段处理后插入到用 XbaI 酶切后通过克列诺片段处理的质粒 pMW118-(λ att L-KmR- λ attR) (W02006/093322A2) 中,从而得到了质粒 pMW-Km-Pnlp8。将 pMW-Km-Pnlp8 作为模板使用,并使用引物 P19(tccgctcacg atttttttca tcgctggttaaggtcatttat cccccaggaa aaattgggtta, SEQ ID NO:59) 和 P20(tttcacaccg ctcaaccgcagggcataacc ggcccttgaa gcctgctttt ttatactaag ttg, SEQ ID NO:60) 进行 PCR,扩增得到含有 Km-Pnlp8 盒的约 1.6kb 的 DNA 片段。此时的 PCR 循环如下所述:95℃ 3 分钟后,95℃ 60 秒、50℃ 30 秒、72℃ 40 秒循环 2 次,94℃ 20 秒、54℃ 20 秒、72℃ 90 秒循环 30 次,最后在 72℃ 下 5 分钟。在两个引物上设计有作为染色体上的靶标用于通过 λ 依赖性整合(被称为“Red 驱动整合”)的方法(Proc. Natl. Acad. Sci. USA, 2000, vol. 97, No. 12, p6640-6645)) 插入目的片段的序列(这种情况下为 cysPTWA 启动子附近的序列)。因此,在将得到的 DNA 片段通过该 λ 依赖性整合插入到目的菌株中时,在染色体上的 cysPTWA 基因的紧邻前方插入了 Km-Pnlp8,从而形成了 nlp8 启动子连接 cysPTWA 基因的结构。在 SEQ ID NO:43 中显示了 cysPTWA 基因簇的碱基序列,在 SEQ ID NO:44 ~ 46 中显示了 cysP、cysT、cysW 各基因编码的氨基酸序列。SEQ ID NO:47、48 分别显示了 cysA 基因的碱基序列,以及该基因编码的氨基酸序列。

[0278] *P. ananatis* SC17(0)/RSF-Red-TER 株是用于高效进行 λ 依赖性整合的宿主菌株,是在对 λ Red 基因产物抗性的 *P. ananatis* 菌株 SC17(0) 株中导入了表达 λ 的 gam、bet 和 exo 各基因(以下称作“ λ Red 基因”)的辅助质粒 RSF-Red-TER 而得到的菌株(W02008/075483)。SC17(0) 株于 2005 年 9 月 21 日保藏在俄罗斯国立工业微生物保藏中心(Russian National Collection of Industrial Microorganisms(VKPM),GNII Genetika)

(地址:Russia,117545Moscow,1 Dorozhny proezd.1),保藏号VKPM B-9246。此外,该RSF-Red-TER质粒的构建方法在W02008/075483中有详细描述。

[0279] 在添加IPTG以诱导 λ Red基因的表达的条件下培养上述SC17(0)/RSF-Red-TER菌株,制备了电穿孔用的细胞。通过电穿孔向这些细胞中导入了上述的目的DNA片段,以卡那霉素抗性为指标,获得了通过 λ 依赖性整合而在cysPTWA基因上游插入了nlp8启动子的重组菌株。以获得的菌株的染色体DNA为模板,使用P21(ctttgtccct ttagtgaagg:SEQ ID NO:61)、P22(agctgatcta gaagctgact cgagttaatg gcctcccaga cgac:SEQ ID NO:62)作为引物,通过PCR反应,确认形成了目标Km-Pnlp8-cysPTWA结构,并将该菌株命名为SC17(0)-Pnlp8-PTWA菌株。

[0280] 接下来,纯化SC17(0)-Pnlp8-PTWA菌株的染色体DNA,将10 μ g该染色体DNA通过电穿孔法导入到EY19(s)菌株中,获得卡那霉素抗性菌株。以得到的菌株的染色体DNA为模板,使用P21、P22作为引物利用PCR进行扩增,确认了在EY19(s)菌株的染色体上导入了Km-Pnlp8-cysPTWA结构。将通过上述方法获得的菌株命名为EYP197菌株。进一步,使用上述的pMT-Int-Xis2将卡那霉素抗性标记从染色体上除去,将变为卡那霉素敏感性的菌株命名为EYP197(s)菌株。

[0281] (1-5)由EYP197(s)菌株制作携带了突变型3-磷酸甘油酸脱氢酶(serA348)基因的菌株

[0282] 通过下述的方法构建了serA348基因,该基因编码来源于Pantoea ananatis的3-磷酸甘油酸脱氢酶作为向L-半胱氨酸生产菌导入的3-磷酸甘油酸脱氢酶,且其编码的是用丙氨酸取代了第348位的天冬酰胺残基的突变型酶(N348A)(J. Biol. Chem. 1996; 271(38):23235-8)。

[0283] Pantoea ananatis来源的野生型serA基因的序列如SEQ ID NO:49所示,该基因编码的氨基酸序列如SEQ ID NO:50所示。为了得到导入了上述突变的serA基因的3'侧的DNA片段,进行下述操作,以SC17菌株染色体DNA为模板,使用P23(agctgagtcg acatggcaaa ggtatcactg gaa:SEQ ID NO:63)以及P24(gagaacgccc gggcgggctt cgtgaatatg cagc:SEQ ID NO:64)作为引物进行PCR(95 $^{\circ}$ C 3分钟后,95 $^{\circ}$ C 60秒、50 $^{\circ}$ C 30秒、72 $^{\circ}$ C 40秒循环2次,94 $^{\circ}$ C 20秒、60 $^{\circ}$ C 20秒、72 $^{\circ}$ C 60秒循环25次,最后在72 $^{\circ}$ C下5分钟)。接下来,为了同样地获得导入了突变的5'侧的DNA片段,以SC17菌株染色体DNA为模板,使用P25(agctgatcta gacgtgggat cagtaaagca gg:SEQ ID NO:65)和P26(aaaaccgccc gggcgttctc ac:SEQ ID NO:66)作为引物进行PCR(95 $^{\circ}$ C 3分钟后,95 $^{\circ}$ C 60秒、50 $^{\circ}$ C 30秒、72 $^{\circ}$ C 40秒循环2次,94 $^{\circ}$ C 20秒、60 $^{\circ}$ C 20秒、72 $^{\circ}$ C 20秒循环20次,最后在72 $^{\circ}$ C下5分钟)。通过限制性内切酶SmaI处理得到的两个PCR片段之后,通过DNA连接酶进行连接,得到含有目的突变(N348A)的突变型serA基因全长的DNA片段。以该DNA片段为模板,使用P23和P25作为引物,进行PCR扩增(95 $^{\circ}$ C 3分钟后,95 $^{\circ}$ C 60秒、50 $^{\circ}$ C 30秒、72 $^{\circ}$ C 40秒循环2次,94 $^{\circ}$ C 20秒、60 $^{\circ}$ C 20秒、72 $^{\circ}$ C 75秒循环15次,最后在72 $^{\circ}$ C下5分钟)。利用SalI、XbaI处理在P23和P25引物上设计的SalI、XbaI限制性内切酶位点之后,将其插入到同样用SalI、XbaI处理的pMIV-Pnlp8-ter中,制得pMIV-Pnlp8-serA348。

[0284] 构建的pMIV-Pnlp8-serA348上携带了来源于pMIV-5JS(特开2008-99668)的Mu的附着位点。如果使用该质粒,则如上所述,通过使用具有Mu转座酶的辅助质粒pMH10,可

以将含有氯霉素抗性标记的 Pnlp8-serA348-rrnBterminator 盒插入到 *P. ananatis* SC17 菌株的染色体上。在 SC17(0) 菌株中导入了质粒 pMIV-Pnlp8-serA348 以及 pMH10, 获得了在染色体上插入了 Pnlp8-serA348-rrnB terminator 盒的菌株。使用引物 P11、P25, 通过 PCR 确认了目的盒存在于细胞中。针对得到的 50 个克隆, 测定细胞提取液中的 3-磷酸甘油酸脱氢酶的活性, 选择出活性最高的菌株, 将其命名为 SC17int-serA348 菌株。接下来, 将 10 μg 的 SC17int-serA348 株的染色体 DNA 通过电穿孔导入到 EYP197(s) 菌株中, 获得了氯霉素抗性菌株, 使用引物 P11、P25, 通过 PCR 确认了在 EYP197(s) 菌株的染色体上与氯霉素抗性标记一起导入了 Pnlp8-serA348 的结构。将这样得到的菌株命名为 EYPS1976 株。

[0285] 通过上述的利用 pMT-Int-Xis2 的标记清除方法, 将氯霉素抗性标记除去, 将获得的氯霉素敏感性菌株命名为 EYPS1976(s) 株。

[0286] (1-6) 从 EYPS1976(s) 株出发制作 gcd 基因缺损株

[0287] 从实施例 1 所述的 SC17(0): Δgcd 株制备基因组 DNA, 通过电穿孔将其导入 EYPS1976(s) 株, 以卡那霉素抗性为指标, 从 EYPS1976(s) 株获得了 gcd 缺损株 (EYPS1976 Δgcd 株)。

[0288] (2) L-半胱氨酸生产菌 EYPS1976(s) 株与 EYPS1976 Δgcd 株的培养

[0289] 为了考察 gcd 基因缺损对 L-半胱氨酸以及 L-半胱氨酸前体 O-乙酰丝氨酸的发酵生产的效果, 使用 L-半胱氨酸生产菌 EYPS1976(s) 株及由该菌株衍生的 gcd 缺损株 EYPS1976 Δgcd 株进行了发酵生产培养, 并对生产的 L-半胱氨酸和 O-乙酰丝氨酸的量进行了比较。培养使用了下述组成的 L-半胱氨酸生产培养基。

[0290]

〔L-半胱氨酸生产培养基〕(各成分的浓度为最终浓度)

成分 1:

(NH ₄) ₂ SO ₄	15g/L
KH ₂ PO ₄	1.5g/L
MgSO ₄ ·7H ₂ O	1g/L
硫胺素盐酸盐	0.1mg/L

成分 2:

FeSO ₄ ·7H ₂ O	1.7mg/L
Na ₂ MoO ₄ ·2H ₂ O	0.15mg/L

[0291]

CoCl ₂ ·6H ₂ O	0.7mg/L
MnCl·4H ₂ O	1.6mg/L
ZnSO ₄ · 7H ₂ O	0.3mg/L
CuSO ₄ ·5H ₂ O	0.25mg/L
成分 3:	
胰蛋白胨	0.6g/L
酵母提取物	0.3g/L
氯化钠	0.6g/L
成分 4:	
碳酸钙	20g/L
成分 5:	
一水合 L-组氨酸盐酸盐	135mg/L
成分 6:	
硫代硫酸钠	6g/L
成分 7:	
吡哆醇盐酸盐	2mg/L
成分 8:	
葡萄糖	40g/L

[0292] 对于各成分,分别预先制备 10 倍(成分 1)、1000 倍(成分 2)、100/6 倍(成分 3)、100 倍(成分 5)、350g/L(成分 6)、1000 倍(成分 7)、10 倍(成分 8)的保存溶液,并在使用时混合,并用灭菌水调整至规定量,从而制备成最终浓度。灭菌分别采用 110℃、30 分钟的高压灭菌(成分 1、2、3、5、8)、180℃、5 小时以上的干热灭菌(成分 4)以及过滤灭菌(成分 6、7)。

[0293] L-半胱氨酸生产培养按照下述规程进行。将 EYPS1976(s) 株和 EYPS1976 Δgcd 株在 LB 琼脂培养基上涂布开,在 34℃ 下进行过夜预培养,然后用 10 μL 大小的接种环(NUNC 公司 Blue loop)从平板上 2 次挑取相当于约 7cm 的菌体,并接种到已经放入大试管(内径 23mm、长度 20cm)的 2ml 的 L-半胱氨酸生产培养基中,使得培养开始时刻的菌体量基本相同。

[0294] 分别在 34℃ 和 38℃ 下进行振荡培养,24 小时后结束培养。确认了此时作为碳源的葡萄糖被完全消耗。按照 Gaitonde, M. K. (Biochem J. 1967Aug ;104(2) :627-33.) 中记载的方法对培养基中产生的 L-半胱氨酸进行定量。此外,利用 HPLC 对培养基中产生的 OAS(O-乙酰丝氨酸)进行了定量。此时,采用了将样品用 200mM 的 Tris-HCl (pH9.0) 稀释,使 OAS 转化为更稳定的 NAS(N-acetylserine, N-乙酰丝氨酸),再进行检测的方法。HPLC 的条件如下。

[0295] 柱:Inertsil ODS-3(疏水性柱/GL science 公司制造)

[0296] 缓冲液流速:1.0mL/min

[0297] 柱温度:40℃

[0298] 检测器:UV210nm

[0299] 上样量:10mL

[0300] 缓冲液:0.1M $\text{KH}_2\text{PO}_4 \cdot \text{H}_3\text{PO}_4$ (pH2.2)、5mM 1-辛磺酸钠。

[0301] 各菌株均进行了6组平行实验,各平均值和标准偏差如表2所示。由表2可知,gcd 缺损在培养温度 34℃及 38℃均具有增加 L- 半胱氨酸和 O- 乙酰丝氨酸的效果。增加的幅度在 38℃更为显著,可知,在高温培养中具有特别好的效果。此外,还可知,在 38℃具有增加菌量 (OD) 的效果。

[0302] [表2]

[0303] 表2gcd 缺损株的 L- 半胱氨酸和 NAS 生产能力

[0304]

		Cys(g/L)	NAS(g/L)	OD600
34℃	EYPS1976(s)	1.02±0.16	0.81±0.10	15.1±0.50
	EYPS1976 Δgcd	1.29±0.16	1.01±0.10	15.0±0.24
38℃	EYPS1976(s)	0.31±0.12	0.54±0.11	11.1±1.68
	EYPS1976 Δgcd	0.95±0.21	0.87±0.10	19.2±0.57

[0305] (序列表说明)

[0306] SEQ ID NO:1:Pantoea ananatis 的 gcd 基因的碱基序列

[0307] SEQ ID NO:2:Pantoea ananatis 的 GCD 的氨基酸序列

[0308] SEQ ID NO:3:Pantoea ananatis 的 hisD 基因的碱基序列

[0309] SEQ ID NO:4:用于将 Km^r 基因组入 hisD 基因的片段的扩增用引物

[0310] SEQ ID NO:5:用于将 Km^r 基因组入 hisD 基因的片段的扩增用引物

[0311] SEQ ID NO:6:cat 基因扩增用引物

[0312] SEQ ID NO:7:cat 基因扩增用引物

[0313] SEQ ID NO:8:sacB 基因扩增用引物

[0314] SEQ ID NO:9:sacB 基因扩增用引物

[0315] SEQ ID NO:10:含 PlacUV5 启动子的 DNA 片段扩增用引物

[0316] SEQ ID NO:11:含 PlacUV5 启动子的 DNA 片段扩增用引物

[0317] SEQ ID NO:12:含 λ Red α β γ 基因以及 tL3 的 DNA 片段扩增用引物

[0318] SEQ ID NO:13:含 λ Red α β γ 基因以及 tL3 的 DNA 片段扩增用引物

[0319] SEQ ID NO:14:含 PlacUV5 启动子和 TrrnB 的 DNA 片段扩增用引物

[0320] SEQ ID NO:15:含 PlacUV5 启动子和 TrrnB 的 DNA 片段扩增用引物

[0321] SEQ ID NO:16:attL 扩增用引物

- [0322] SEQ ID NO:17:attL 扩增用引物
- [0323] SEQ ID NO:18:attL 的碱基序列
- [0324] SEQ ID NO:19:attR 扩增用引物
- [0325] SEQ ID NO:20:attR 扩增用引物
- [0326] SEQ ID NO:21:attR 的碱基序列
- [0327] SEQ ID NO:22:含 bla 基因的 DNA 片段扩增用引物
- [0328] SEQ ID NO:23:含 bla 基因的 DNA 片段扩增用引物
- [0329] SEQ ID NO:24:含 ter_rrnB 的 DNA 片段扩增用引物
- [0330] SEQ ID NO:25:含 ter_rrnB 的 DNA 片段扩增用引物
- [0331] SEQ ID NO:26:含 ter_thrL 终止子的 DNA 片段的碱基序列
- [0332] SEQ ID NO:27:含 ter_thrL 终止子的 DNA 片段扩增用引物
- [0333] SEQ ID NO:28:含 ter_thrL 终止子的 DNA 片段扩增用引物
- [0334] SEQ ID NO:29:用于扩增 gltA 基因的 ORF 以外的部分的引物
- [0335] SEQ ID NO:30:用于扩增 gltA 基因的 ORF 以外的部分的引物
- [0336] SEQ ID NO:31:prpC 基因扩增用引物
- [0337] SEQ ID NO:32:prpC 基因扩增用引物
- [0338] SEQ ID NO:33:gcd 缺损用引物
- [0339] SEQ ID NO:34:gcd 缺损用引物
- [0340] SEQ ID NO:35:gcd 缺损确认用引物
- [0341] SEQ ID NO:36:gcd 缺损确认用引物
- [0342] SEQ ID NO:37:野生型 cysE 基因的碱基序列
- [0343] SEQ ID NO:38:野生型 cysE 编码的丝氨酸乙酰转移酶的氨基酸序列
- [0344] SEQ ID NO:39:野生型 yeaS 基因的碱基序列
- [0345] SEQ ID NO:40:野生型 YeaS 的氨基酸序列
- [0346] SEQ ID NO:41:Pnlp0 的碱基序列
- [0347] SEQ ID NO:42:Pnlp8 的碱基序列
- [0348] SEQ ID NO:43:cysPTWA 基因簇的碱基序列
- [0349] SEQ ID NO:44:cysP 基因编码的氨基酸序列
- [0350] SEQ ID NO:45:cysT 基因编码的氨基酸序列
- [0351] SEQ ID NO:46:cysW 基因编码的氨基酸序列
- [0352] SEQ ID NO:47:cysA 基因的碱基序列
- [0353] SEQ ID NO:48:cysA 基因编码的氨基酸序列
- [0354] SEQ ID NO:49:Pantoea ananatis 野生型 serA 基因的碱基序列
- [0355] SEQ ID NO:50:Pantoea ananatis 野生型 serA 基因编码的氨基酸序列
- [0356] SEQ ID NO:51 ~ 66:引物 P11 ~ P26

[0001]

序列表

<110> 味之素株式会社(Ajinomoto Co., Inc.)
 <120> 生产 L-氨基酸的微生物和 L-氨基酸的生产方法
 <130> D095-C9293
 <150> JP2008-229736
 <151> 2008-09-08
 <150> JP2009-032839
 <151> 2009-02-16
 <160> 66
 <170> PatentIn version 3.5
 <210> 1
 <211> 2931
 <212> DNA
 <213> Pantoea anantis
 <220>
 <221> CDS
 <222> (301).. (2691)
 <400> 1
 gcgtctggcg atctcattcc ctttacgctt taacctttct tatctctctgg cactattgaa 60
 tcatacggcg aaggccccat taacaccctc taaaglatig ccttagatgc gtigttaaga 120
 gctttaagcc aaaaaaatca aaactcatac caattttgct ataacattta acagatatgg 180
 atcatcacgc ttcagggaag agcgaaataa gaatgccttt catcttttgt actttatgac 240
 tgacaacaat ctatctgatt gttttcctga gttttctggc aacgaaatga ggicaacatt 300
 atg ggg aaa aac tcc tca tcc ttt agc gtc gtc cgt ttt tta acg gtc 348
 Met Gly Lys Asn Ser Ser Phe Ser Val Val Arg Phe Leu Thr Val
 1 5 10 15
 ctg ttc gcc gtg cta acg ggt gcg ttc atg tta att ggt ggt atc tgg 396
 Leu Phe Ala Val Leu Thr Gly Ala Phe Met Leu Ile Gly Gly Ile Trp
 20 25 30
 ctg gcc acg atc ggt ggt tcc tgg tac tac atc atc ggc ggt gca gcc 444
 Leu Ala Thr Ile Gly Gly Ser Trp Tyr Tyr Ile Ile Gly Gly Ala Ala
 35 40 45
 atg ctg ctt acc gct ttc ctg ctg tgg cga cgt aac agc gct gcc ctg 492
 Met Leu Leu Thr Ala Phe Leu Leu Trp Arg Arg Asn Ser Ala Ala Leu
 50 55 60
 gtt gtc tat gcg ctc tta ctg ctg gct acg ctg gcc tgg ggc gtt tgg 540
 Val Val Tyr Ala Leu Leu Leu Ala Thr Leu Ala Trp Gly Val Trp
 65 70 75 80
 gaa gtc ggc acc gac ttc tgg gca ctg gca ccg cgt acc gac gta ctg 588
 Glu Val Gly Thr Asp Phe Trp Ala Leu Ala Pro Arg Thr Asp Val Leu
 85 90 95
 gtg atc ttt ggc gtc tgg ctg gtg ttg ccc ttt gtc tat cgc ggc tta 636
 Val Ile Phe Gly Val Trp Leu Val Leu Pro Phe Val Tyr Arg Gly Leu
 100 105 110
 tac cag ccg ggt aaa ggc gca ctg ggt gcc atg ggc gta gcg ctg gtt 684
 Tyr Gln Pro Gly Lys Gly Ala Leu Gly Ala Met Gly Val Ala Leu Val
 115 120 125
 gcc agt gca gcg gtg tta acc tat tcc gtc ttt aat gat ccg caa gtg 732
 Ala Ser Ala Ala Val Leu Thr Tyr Ser Val Phe Asn Asp Pro Gln Val
 130 135 140
 gtt aac ggt gca tta ccg gca aca gcg gat aat gcg cct cag gca cag 780
 Val Asn Gly Ala Leu Pro Ala Thr Ala Asp Asn Ala Pro Gln Ala Gln
 145 150 155 160
 ccg ttg agc aat att gct gat ggt gac tgg ccg gcc tat gcg cgc gat 828
 Pro Leu Ser Asn Ile Ala Asp Gly Asp Trp Pro Ala Tyr Ala Arg Asp
 165 170 175
 cag caa ggg acg cgc ttc tgc ccg ctc aag cag atc aac cac gac aat 876
 Gln Gln Gly Thr Arg Phe Ser Pro Leu Lys Gln Ile Asn His Asp Asn
 180 185 190
 gtg aaa gaa ctg cag gtt gcc tgg caa ttc cag acc ggt gat atg aaa 924
 Val Lys Glu Leu Gln Val Ala Trp Gln Phe Gln Thr Gly Asp Met Lys
 195 200 205
 cgc cca agc gat ccg ggc gaa att acc gat gaa gtg acg cca atc aag 972
 Arg Pro Ser Asp Pro Gly Glu Ile Thr Asp Glu Val Thr Pro Ile Lys
 210 215 220
 att cgc gac acg ctg tat ctt tgc acg cca cat cag att tta ttt gct 1020
 Ile Arg Asp Thr Leu Tyr Leu Cys Thr Pro His Gln Ile Leu Phe Ala
 225 230 235 240
 ctg gat gcg gcc acc ggc aag caa aag tgg aag ttt gat ccc ggc ctg 1068

[0002]

Leu	Asp	Ala	Ala	Thr	Gly	Lys	Gln	Lys	Trp	Lys	Phe	Asp	Pro	Gly	Leu	
aaa	acc	aac	cca	acc	ttc	cag	cac	gtg	acc	tgt	cgt	ggt	gtg	tca	tac	1116
Lys	Thr	Asn	Pro	Thr	Phe	Gln	His	Val	Thr	Cys	Arg	Gly	Val	Ser	Tyr	
			260					265				270				
cac	gaa	ttc	cct	gca	gcg	aag	gat	gcg	tcc	aat	acc	cag	cct	gcg	ctg	1164
His	Glu	Phe	Pro	Ala	Ala	Lys	Asp	Ala	Ser	Asn	Thr	Gln	Pro	Ala	Leu	
			275					280				285				
tgc	tcg	cgt	cgt	atc	tac	ctg	cca	gtc	aat	gac	ggg	cgt	ttg	ttc	gcg	1212
Cys	Ser	Arg	Arg	Ile	Tyr	Lcu	Pro	Val	Asn	Asp	Gly	Arg	Lcu	Phe	Ala	
			290					295				300				
ctg	gat	gcg	gaa	acc	ggg	gaa	cgc	tgc	ccg	gcc	ttt	ggt	aac	aac	ggt	1260
Leu	Asp	Ala	Glu	Thr	Gly	Glu	Arg	Cys	Pro	Ala	Phe	Gly	Asn	Asn	Gly	
			305			310					315				320	
gag	ctg	gat	ctg	cag	cac	aag	cag	ccg	gtc	aca	acg	cca	ggc	atg	tat	1308
Glu	Leu	Asp	Leu	Gln	His	Lys	Gln	Pro	Val	Thr	Thr	Pro	Gly	Met	Tyr	
			325					330					335			
gag	cca	acc	tcg	cca	ccg	gtg	att	act	gac	acc	acc	att	gtg	atg	gct	1356
Glu	Pro	Thr	Ser	Pro	Pro	Val	Ile	Thr	Asp	Thr	Thr	Ile	Val	Met	Ala	
			340					345					350			
ggc	gcg	gta	acc	gat	aac	ttt	tca	acc	cgt	gaa	cct	tca	ggc	gcc	atc	1404
Gly	Ala	Val	Thr	Asp	Asn	Phe	Ser	Thr	Arg	Glu	Pro	Ser	Gly	Ala	Ile	
			355					360				365				
cgt	ggc	ttt	gat	gtg	aac	acc	ggg	aag	ctg	ttg	tgg	gtg	ttc	gat	ccg	1452
Arg	Gly	Phe	Asp	Val	Asn	Thr	Gly	Lys	Leu	Leu	Trp	Val	Phe	Asp	Pro	
			370				375					380				
ggc	gcg	aaa	gat	cct	aac	gcg	att	ccg	gcg	gat	gaa	cac	acg	ttc	acc	1500
Gly	Ala	Lys	Asp	Pro	Asn	Ala	Ile	Pro	Ala	Asp	Glu	His	Thr	Phe	Thr	
			385			390					395				400	
atg	aac	tcc	cct	aac	tcg	tgg	gca	cct	gcg	gtt	tac	gat	ccg	aag	ctg	1548
Met	Asn	Ser	Pro	Asn	Ser	Trp	Ala	Pro	Ala	Val	Tyr	Asp	Pro	Lys	Leu	
			405										410			
gat	atc	gtt	tac	ctg	cca	atg	ggg	gtg	acc	acg	ccg	gat	atc	tgg	ggc	1596
Asp	Ile	Val	Tyr	Leu	Pro	Met	Gly	Val	Thr	Thr	Pro	Asp	Ile	Trp	Gly	
			420					425					430			
ggc	aac	cgc	aca	cct	gag	cag	gaa	cgt	tat	gcc	agc	agc	gtg	ctg	gcg	1644
Gly	Asn	Arg	Thr	Pro	Glu	Gln	Glu	Arg	Tyr	Ala	Ser	Ser	Val	Leu	Ala	
			435					440					445			
ctg	aac	gcg	acg	acc	ggg	aag	ctg	gtg	tgg	tca	tat	cag	act	gtg	cat	1692
Leu	Asn	Ala	Thr	Thr	Gly	Lys	Leu	Val	Trp	Ser	Tyr	Gln	Thr	Val	His	
			450					455				460				
cac	gat	ctg	tgg	gat	atg	gac	ctg	cct	tcg	cag	ccg	acg	ctg	gcg	gat	1740
His	Asp	Leu	Trp	Asp	Met	Asp	Leu	Pro	Ser	Gln	Pro	Thr	Leu	Ala	Asp	
			465			470					475				480	
att	acc	gat	aaa	gac	ggg	aat	acc	gtg	ccg	gtt	atc	tat	gcc	cct	gcc	1788
Ile	Thr	Asp	Lys	Asp	Gly	Asn	Thr	Val	Pro	Val	Ile	Tyr	Ala	Pro	Ala	
			485										490		495	
aaa	acc	ggg	aac	atc	ttt	gtt	ctg	gal	cgc	cgc	aca	ggg	aaa	act	gtg	1836
Lys	Thr	Gly	Asn	Ile	Phe	Val	Leu	Asp	Arg	Arg	Thr	Gly	Lys	Thr	Val	
			500					505					510			
gtt	ccg	gcc	ccg	gaa	acc	cct	gtt	ccg	cag	ggc	gca	gct	aag	ggc	gac	1884
Val	Pro	Ala	Pro	Glu	Thr	Pro	Val	Pro	Gln	Gly	Ala	Ala	Lys	Gly	Asp	
			515					520					525			
cat	gtc	tca	gct	aca	cag	cct	tac	tct	gaa	ctg	acc	ttc	cgt	ccg	aaa	1932
His	Val	Ser	Ala	Thr	Gln	Pro	Tyr	Ser	Glu	Lcu	Thr	Phe	Arg	Pro	Lys	
			530				535					540				
cag	aac	ctg	acg	gat	aag	gac	atg	tgg	ggc	gcg	acg	atg	tat	gac	cag	1980
Gln	Asn	Leu	Thr	Asp	Lys	Asp	Met	Trp	Gly	Ala	Thr	Met	Tyr	Asp	Gln	
			545			550					555				560	
ctg	gtg	tgc	cgc	gtg	att	ttc	aaa	cgt	ctg	cgc	tac	gaa	ggg	ccg	ttc	2028
Leu	Val	Cys	Arg	Val	Ile	Phe	Lys	Arg	Leu	Arg	Tyr	Glu	Gly	Pro	Phe	
			565										570			
acg	cca	cct	tct	gag	cag	ggc	acc	ctg	gtc	ttc	ccg	ggc	aac	ctg	ggc	2076
Thr	Pro	Pro	Ser	Glu	Gln	Gly	Thr	Leu	Val	Phe	Pro	Gly	Asn	Leu	Gly	
			580					585					590			
atg	ttt	gaa	tgg	ggc	ggc	att	tcc	gtt	gat	ccg	cat	cgt	cag	att	gcg	2124
Met	Phe	Glu	Trp	Gly	Gly	Ile	Ser	Val	Asp	Pro	His	Arg	Gln	Ile	Ala	
			595				600						605			
att	gct	aac	cca	atg	gcg	ctg	ccg	ttc	gtg	tct	aag	ctg	atc	cca	cgc	2172
Ile	Ala	Asn	Pro	Met	Ala	Leu	Pro	Phe	Val	Ser	Lys	Leu	Ile	Pro	Arg	
			610				615						620			
ggg	ccg	ggg	aat	ccg	gaa	gag	cca	cca	aaa	ggc	gca	acg	ggc	ggg	tca	2220
Gly	Pro	Gly	Asn	Pro	Glu	Glu	Pro	Pro	Lys	Gly	Ala	Thr	Gly	Gly	Ser	
			625			630					635				640	
ggg	act	gaa	acc	ggg	att	cag	ccg	cag	tac	ggg	gtg	cca	tat	ggc	gtt	2268
Gly	Thr	Glu	Thr	Gly	Ile	Gln	Pro	Gln	Tyr	Gly	Val	Pro	Tyr	Gly	Val	
			645						650					655		
gaa	ctg	aat	ccg	ttc	ctg	tca	cct	ttt	ggg	ctg	ccg	tgt	aaa	caa	cct	2316
Glu	Leu	Asn	Pro	Phe	Leu	Ser	Pro	Phe	Gly	Leu	Pro	Cys	Lys	Gln	Pro	
			660					665					670			

[0003]

```

gca tgg ggt tat gtt tct gct gtt gac ctg aaa acc aac gaa gtg gtg 2364
Ala Trp Gly Tyr Val Ser Ala Val Asp Leu Lys Thr Asn Glu Val Val
675 680 685
tgg aaa caa cgt att ggt acc gtt cgt gac agc tca cct gta ccg ctg 2412
Trp Lys Gln Arg Ile Gly Thr Val Arg Asp Ser Ser Pro Val Pro Leu
690 695 700
ccc ttt aaa atg ggt atg cca atg ctg ggc gga ccg gtt gcc acc gca 2460
Pro Phe Lys Met Gly Met Pro Met Leu Gly Gly Pro Val Ala Thr Ala
705 710 715 720
ggc aaa gtg ttc ttt att ggc gca acg gct gat aac tac ctg cgc gct 2508
Gly Lys Val Phe Phe Ile Gly Ala Thr Ala Asp Asn Tyr Leu Arg Ala
725 730 735
ttc agc acc gac acc ggt gaa ctc ttg tgg cag gcg cgc ctg cca gcc 2556
Phe Ser Thr Asp Thr Gly Glu Leu Leu Trp Gln Ala Arg Leu Pro Ala
740 745 750
ggt ggt cag gca acg cca atg acc tal gaa gtt aac ggc aag caa tac 2604
Gly Gly Gln Ala Thr Pro Met Thr Tyr Glu Val Asn Gly Lys Gln Tyr
755 760 765
gtt gtg att gct gcc ggt ggc cat ggt tca ttc ggc acc aag ctg ggc 2652
Val Val Ile Ala Ala Gly Gly His Gly Ser Phe Gly Thr Lys Leu Gly
770 775 780
gat tac gtg att gcc tat gcg ctg ccc gac cag aag taa ttaacacctg 2701
Asp Tyr Val Ile Ala Tyr Ala Leu Pro Asp Gln Lys
785 790 795
aacagaggcgc gactccggtc cgcctttttt tatgcctgct atctgccctg tgcttttgcg 2761
cgiggggagc gccagcttaa ccaggcgcac agcccatga ccatgcaggt gccagaaat 2821
acggcccgca ttcccaacgc gccaccaatc accccaccaa aaaggggacc acigacctgg 2881
ccgatatact gcgccgacgt cgaatagccg agcatgcgcc ccacctgcac 2931

```

<210> 2
 <211> 796
 <212> PRT
 <213> Pantoea anantis

<400> 2
 Met Gly Lys Asn Ser Ser Phe Ser Val Val Arg Phe Leu Thr Val
 1 5 10 15
 Leu Phe Ala Val Leu Thr Gly Ala Phe Met Leu Ile Gly Gly Ile Trp
 20 25 30
 Leu Ala Thr Ile Gly Gly Ser Trp Tyr Tyr Ile Ile Gly Gly Ala Ala
 35 40 45
 Met Leu Leu Thr Ala Phe Leu Leu Trp Arg Arg Asn Ser Ala Ala Leu
 50 55 60
 Val Val Tyr Ala Leu Leu Leu Ala Thr Leu Ala Trp Gly Val Trp
 65 70 75 80
 Glu Val Gly Thr Asp Phe Trp Ala Leu Ala Pro Arg Thr Asp Val Leu
 85 90 95
 Val Ile Phe Gly Val Trp Leu Val Leu Pro Phe Val Tyr Arg Gly Leu
 100 105 110
 Tyr Gln Pro Gly Lys Gly Ala Leu Gly Ala Met Gly Val Ala Leu Val
 115 120 125
 Ala Ser Ala Ala Val Leu Thr Tyr Ser Val Phe Asn Asp Pro Gln Val
 130 135 140
 Val Asn Gly Ala Leu Pro Ala Thr Ala Asp Asn Ala Pro Gln Ala Gln
 145 150 155 160
 Pro Leu Ser Asn Ile Ala Asp Gly Asp Trp Pro Ala Tyr Ala Arg Asp
 165 170 175
 Gln Gln Gly Thr Arg Phe Ser Pro Leu Lys Gln Ile Asn His Asp Asn
 180 185 190
 Val Lys Glu Leu Gln Val Ala Trp Gln Phe Gln Thr Gly Asp Met Lys
 195 200 205
 Arg Pro Ser Asp Pro Gly Glu Ile Thr Asp Glu Val Thr Pro Ile Lys
 210 215 220
 Ile Arg Asp Thr Leu Tyr Leu Cys Thr Pro His Gln Ile Leu Phe Ala
 225 230 235 240
 Leu Asp Ala Ala Thr Gly Lys Gln Lys Trp Lys Phe Asp Pro Gly Leu
 245 250 255
 Lys Thr Asn Pro Thr Phe Gln His Val Thr Cys Arg Gly Val Ser Tyr
 260 265 270
 His Glu Phe Pro Ala Ala Lys Asp Ala Ser Asn Thr Gln Pro Ala Leu
 275 280 285
 Cys Ser Arg Arg Ile Tyr Leu Pro Val Asn Asp Gly Arg Leu Phe Ala
 290 295 300
 Leu Asp Ala Glu Thr Gly Glu Arg Cys Pro Ala Phe Gly Asn Asn Gly
 305 310 315 320
 Glu Leu Asp Leu Gln His Lys Gln Pro Val Thr Thr Pro Gly Met Tyr
 325 330 335
 Glu Pro Thr Ser Pro Pro Val Ile Thr Asp Thr Thr Ile Val Met Ala
 340 345 350

[0004]

Gly Ala Val Thr Asp Asn Phe Ser Thr Arg Glu Pro Ser Gly Ala Ile
 355 360 365
 Arg Gly Phe Asp Val Asn Thr Gly Lys Leu Leu Trp Val Phe Asp Pro
 370 375 380
 Gly Ala Lys Asp Pro Asn Ala Ile Pro Ala Asp Glu His Thr Phe Thr
 385 390 400
 Met Asn Ser Pro Asn Ser Trp Ala Pro Ala Val Tyr Asp Pro Lys Leu
 405 410 415
 Asp Ile Val Tyr Leu Pro Met Gly Val Thr Thr Pro Asp Ile Trp Gly
 420 425 430
 Gly Asn Arg Thr Pro Glu Gln Glu Arg Tyr Ala Ser Ser Val Leu Ala
 435 440 445
 Leu Asn Ala Thr Thr Gly Lys Leu Val Trp Ser Tyr Gln Thr Val His
 450 455 460
 His Asp Leu Trp Asp Met Asp Leu Pro Ser Gln Pro Thr Leu Ala Asp
 465 470 475 480
 Ile Thr Asp Lys Asp Gly Asn Thr Val Pro Val Ile Tyr Ala Pro Ala
 485 490 495
 Lys Thr Gly Asn Ile Phe Val Leu Asp Arg Arg Thr Gly Lys Thr Val
 500 505 510
 Val Pro Ala Pro Glu Thr Pro Val Pro Gln Gly Ala Ala Lys Gly Asp
 515 520 525
 His Val Ser Ala Thr Gln Pro Tyr Ser Glu Leu Thr Phe Arg Pro Lys
 530 535 540
 Gln Asn Leu Thr Asp Lys Asp Met Trp Gly Ala Thr Met Tyr Asp Gln
 545 550 555 560
 Leu Val Cys Arg Val Ile Phe Lys Arg Leu Arg Tyr Glu Gly Pro Phe
 565 570 575
 Thr Pro Pro Ser Glu Gln Gly Thr Leu Val Phe Pro Gly Asn Leu Gly
 580 585 590
 Met Phe Glu Trp Gly Gly Ile Ser Val Asp Pro His Arg Gln Ile Ala
 595 600 605
 Ile Ala Asn Pro Met Ala Leu Pro Phe Val Ser Lys Leu Ile Pro Arg
 610 615 620
 Gly Pro Gly Asn Pro Glu Glu Pro Pro Lys Gly Ala Thr Gly Gly Ser
 625 630 635 640
 Gly Thr Glu Thr Gly Ile Gln Pro Gln Tyr Gly Val Pro Tyr Gly Val
 645 650 655
 Glu Leu Asn Pro Phe Leu Ser Pro Phe Gly Leu Pro Cys Lys Gln Pro
 660 665 670
 Ala Trp Gly Tyr Val Ser Ala Val Asp Leu Lys Thr Asn Glu Val Val
 675 680 685
 Trp Lys Gln Arg Ile Gly Thr Val Arg Asp Ser Ser Pro Val Pro Leu
 690 695 700
 Pro Phe Lys Met Gly Met Pro Met Leu Gly Gly Pro Val Ala Thr Ala
 705 710 715 720
 Gly Lys Val Phe Phe Ile Gly Ala Thr Ala Asp Asn Tyr Leu Arg Ala
 725 730 735
 Phe Ser Thr Asp Thr Gly Glu Leu Leu Trp Gln Ala Arg Leu Pro Ala
 740 745 750
 Gly Gly Gln Ala Thr Pro Met Thr Tyr Glu Val Asn Gly Lys Gln Tyr
 755 760 765
 Val Val Ile Ala Ala Gly Gly His Gly Ser Phe Gly Thr Lys Leu Gly
 770 775 780
 Asp Tyr Val Ile Ala Tyr Ala Leu Pro Asp Gln Lys
 785 790 795

<210> 3
 <211> 1308
 <212> DNA
 <213> *Pantoea ananatis*

<400> 3
 atgagcagaa tcatgacgcc cgtgaactgg gaagccatgca gcagcgaggc gcagcaggcg 60
 ctgtttggcac gccctgcgct cgcctcgtct gacagcatca gccagatcgt gcgcgatgtg 120
 ttggtcagag tgaagagga aggcgatgcg gctttacgag aattcagcgc gcgcittgac 180
 aaggttgaana cagacgacct gcgcgttacg ccacagcaga tgcaggcggc cagcgatcgc 240
 ctgttgtagc agctgaaaca ggcatggcc gtggccattg gcaatattga aaccittcac 300
 cglgcgcaga tcccgccgcc ggtggatlg gaaacgcagc ccggcgtlcc cgtlcagcaa 360
 attacgcgcc cgatgaaate ggtgggcttg tataatccgg gcggttctgc cccgctgttt 420
 tctaccgttc tgatgctggc taccocggcg cggattgcgg gctgtggtcg cgtgtgtctg 480
 tgcctgcccc cgcgattgc tgaigaaatt cctacgcgg ccaaacittg cgtgtggaa 540
 gaagtgttcc aggtgggtgg atcacaggcg attgccgcc tggcttttg caccgaaagc 600
 atccctaagg tagataaaat ttttggctcg ggcaacgcgt gggttacga agccaaacgt 660
 caggtcagcc agcccttga tggcgcggcg attgatatgc ccgctggccc gtcggaagt 720
 ctggtgatig ccgatgaagg tgcacaccg gccttcgttg cctctgatct gctgtcgag 780
 gcggaacacg gccctgacgc gcaggtgatt ttactgacgc cttegtgtgg cgtggccgag 840
 cgcgtgcggc aggcgggtga ggaacagctg gccagttgc cactgtcgcc gacagccgcg 900
 caggcactgg aaagcagccg cctgatcgtc gcccgggata tgcagcaatg cattgcgatc 960
 tccaaccgct atgttccgga gcacctgatt ctgcaaaccc gcacgccacg ggatctgggt 1020

[0005]

```

gaacagatta ccagcgccgg ttccggtttt ctgggcgact ggtcacccga atccgcagga 1080
gattatgctt cgggcaccaa ccacgtgctg ccgacctacg gctataccgc gacatgctcc 1140
agccctgggcc tggccgactt tcagaaacgc atgacggtac aggagctgac gccgcagggc 1200
ttccctgaacc tggcggcgac catcgaacc ctggcggccg ctgaacagct gcacgccac 1260
aaaaatgccg tcacgttgcg cgttgcgcga ctcaaggagc aagcatga 1308

```

<210> 4
 <211> 68
 <212> DNA
 <213> 人工序列

<220>
 <223> 引物

```

<400> 4
ccatagcggg tggagatgc aatgcattgc tgcataatcc tgaagccigc tttttatac 60
taagttagg 68

```

<210> 5
 <211> 68
 <212> DNA
 <213> 人工序列

<220>
 <223> 引物

```

<400> 5
gccccccagg cacitgaaag cagccgccig atcgtcgccc cgctcaagtt agtataaaaa 60
agctgaac 68

```

<210> 6
 <211> 33
 <212> DNA
 <213> 人工序列

<220>
 <223> 引物

```

<400> 6
tagcgagatc tctgatgtcc ggccggtgctt ttg 33

```

<210> 7
 <211> 32
 <212> DNA
 <213> 人工序列

<220>
 <223> 引物

```

<400> 7
aaaaagagct cttacgcccc gccctgccac tc 32

```

<210> 8
 <211> 34
 <212> DNA
 <213> 人工序列

<220>
 <223> 引物

```

<400> 8
caggatctag aaggagacat gaacgatgaa catc 34

```

<210> 9
 <211> 36
 <212> DNA
 <213> 人工序列

<220>
 <223> 引物

```

<400> 9
gataaggatc cgaataaaaa gaaaatgcc aatagga 36

```

[0006]

<210>	10	
<211>	31	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	引物	
<400>	10	
	cctttgagct cgcgggcagt gagcgcaacg c	31
<210>	11	
<211>	48	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	引物	
<400>	11	
	ctagagcggc cgccgatcgg gatcctcctg tgtgaaattg ttatccgc	48
<210>	12	
<211>	42	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	引物	
<400>	12	
	ctctacgac gaggaggita taaaaaatgg atattaatac tg	42
<210>	13	
<211>	36	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	引物	
<400>	13	
	tcaaagcggc cgcttcttcg tctgtttcta ctggtta	36
<210>	14	
<211>	31	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	引物	
<400>	14	
	cctttggtac cgcgggcagt gagcgcaacg c	31
<210>	15	
<211>	34	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	引物	
<400>	15	
	aacaggaatt ctttgcttgg cggcagtagc gcgg	34
<210>	16	
<211>	40	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	引物 P1	
<400>	16	

[0007]

ctagtaagat ctigaagcct gcttttttat actaagttgg	40
<210> 17	
<211> 41	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引物 P2	
<400> 17	
atgatacgaat tcgaaatcaa ataatagattt tattttgact g	41
<210> 18	
<211> 120	
<212> DNA	
<213> 人工序列	
<220>	
<223> 包含 attL 的 DNA 片段	
<400> 18	
agatcttgaa gcctgccttt ttataactaag tiggcattat aaaaaagcat tgcttatcaa	60
tttggttgcaa cgaacaggct actatcagtc aaaataaaat cattatttga tticgaattc	120
<210> 19	
<211> 41	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引物 P3	
<400> 19	
atgccactgc agtcctgllac agglcactaa taccatctaa g	41
<210> 20	
<211> 46	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引物 P4	
<400> 20	
accgttaagc ttcttagacg ctcaagitag tataaaaaag ctgaac	46
<210> 21	
<211> 184	
<212> DNA	
<213> 人工序列	
<220>	
<223> 包含 attR 的 DNA 片段	
<400> 21	
ctgcagtcig ttacaggica ctaataccat ctaagtagtt gattcatagt gactgcatat	60
gttggtgttt acagtattat gtagtcgtt ttttatgcaa aatctaatit aatatattga	120
tatttataic attttacgtt tctcgttcag cttttttata ctaacttagc cgtctagaaa	180
gctt	184
<210> 22	
<211> 38	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引物 P5	
<400> 22	
ttcttagacg tcaggtaggca cttttcgggg aaatgtgc	38
<210> 23	

[0008]

<211>	37		
<212>	DNA		
<213>	人工序列		
<220>			
<223>	引物 P6		
<400>	23		
taacagagat	ctcgcgcaga	aaaaaaggat ctcaaga	37
<210>	24		
<211>	46		
<212>	DNA		
<213>	人工序列		
<220>			
<223>	引物 P7		
<400>	24		
aacagagatc	taagcttaga	tccittgcct ggcggcagta gcgcgg	46
<210>	25		
<211>	35		
<212>	DNA		
<213>	人工序列		
<220>			
<223>	引物 P8		
<400>	25		
ataaactgca	gcaaaaagag	ttttagaaaa cgcaa	35
<210>	26		
<211>	1388		
<212>	DNA		
<213>	人工序列		
<220>			
<223>	包含 Tc 基因 和 ter_thrl 的 DNA 片段		
<400>	26		
gaatttctcat	gittgacagc	ttatcatcga taagctttaa tgcggtagtt taicacagtt	60
aaatttgcata	cgcagtcagg	caccgtgtat gaaatctaac aatgcgcctca tegtatectt	120
cggcaccgtc	accctggatg	ctgtaggcat agccttgggt atgccggtac tgcggggcct	180
cttgcgggat	atcgctccatt	ccgacagcat cgcacagtcac tatggcgtgc tgcctagcgt	240
atalgcgllg	algcaatlcc	talgcgcacc cgtlctcgga gcacgtlccg accgcctllg	300
ccgcgcccca	gtcctgtctc	cttcgtact tggagccact atcgactacg cgaatcagtc	360
gaccacaccc	gtcctgttga	lccctctacgc cggacgcate gtggccggca tcaccggcgc	420
cacaggttgc	gttgcctggc	ccatatatgc cgacatcacc gatggggaag atcgggctcg	480
ccacttcggg	ctcatgagcg	cttgtttcgg cgtgggtatg gtggcaggcc cgttggccgg	540
gggacgtllg	ggcgccaatc	ccctlgtatgc accatlcctt gcggcggcgg lgtlcaacgg	600
ccctcaacct	ctactgggct	gtcttctaata gcaggagtcg cataaggagg agcgtcgacc	660
gatgcccttg	agagccctta	accacgtcag ctccctccgg tgggcgcggg gcatgactat	720
cgtcgcgcga	cttatgactg	tcttctttat catgcaactc gtaggacagg tgcgggcagc	780
gtctctgggt	atcttcggcg	aggaccgctt tgcctggagc gcgacgatga tgggctgtc	840
gcttgcggta	ttcggaaatc	lgtcacgccc cgtlcaagcc ttgctacatg gtccccccac	900
caaacgtttc	ggcgagaagc	aggccattat cgcgggcatg gcggccgacg cgttgggcta	960
cgtcttgcgt	gcgttcgcga	cgcgaggctg gatggccttc cccattatga ttcttctcgc	1020
ttccggcggc	atcgggatgc	ccgcgttgca ggccatgctg tccaggcagg tagatgacga	1080
ccatcaggga	cagcttcaag	gatcgtctcg ggctcttacc agcctaactt cgaatcactg	1140
accgtgatac	gtcacggcga	tttatgccgc ctccggcagc acatggaaac ggittggcatg	1200
gattgtaggc	gcgcgccctat	accttgtctg cctccccgcg ttgcgtcgcg gtgcattggag	1260
ccggggcacc	tgcacttgaa	tggaaagccg cggcacctcg ctaacggatt caccactcca	1320
actagaaagc	ttaacacaga	aaaaagcccg caccctgacag tgcgggcttt ttttttcgac	1380
cactgcag			1388
<210>	27		
<211>	36		
<212>	DNA		
<213>	人工序列		
<220>			
<223>	引物 P9		
<400>	27		
aglaattcta	gaaagcttaa	cacagaaaa agccccg	36

[0009]

<210>	28	
<211>	43	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	引物 P10	
<400>	28	
	ctagtaggat ccctgcagtg gtcgaaaaaa aaagcccgca ctg	43
<210>	29	
<211>	33	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	引物 1	
<400>	29	
	ggaagatcta ttgacctcg cacatcaacc tgg	33
<210>	30	
<211>	30	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	引物 2	
<400>	30	
	cggggtacct tgtaaatatt ttaacccgcc	30
<210>	31	
<211>	56	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	引物 3	
<400>	31	
	ggaagatcta aggagacctt aaatgagcga cacaacgatc ctgcaaaaca gtaccc	56
<210>	32	
<211>	39	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	引物 4	
<400>	32	
	cggggtacct cgtagagggt tactggcgct tatccagcg	39
<210>	33	
<211>	80	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	引物	
<400>	33	
	ctatctgatt gttttcctga gttttctggc aacgaaatga ggtcaacatt tgaagcctgc	60
	ttttttatac taagttggca	80
<210>	34	
<211>	80	
<212>	DNA	
<213>	人工序列	

[0010]

<220>		
<223>	引物	
<400>	34	
agcaggcata aaaaaaggcg gaccggagtc cgcctctgtt cagggtgttaa cgctcaagtt		60
agtataaaaa agctgaacga		80
<210>	35	
<211>	21	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	引物	
<400>	35	
gatatggatc atcacgcttc a		21
<210>	36	
<211>	18	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	引物	
<400>	36	
catgctcggc tattegac		18
<210>	37	
<211>	1422	
<212>	DNA	
<213>	大肠杆菌 (Escherichia coli)	
<220>		
<221>	CDS	
<222>	(301).. (1119)	
<400>	37	
tgggtcaggg catggatgta caaagcgcgc aggagaagat tggtcagggtg gtggaaggct		60
accgcaatac gaaagaagtc cgcgaactgg cgcacgcctt cggcgttgaa atgccataaa		120
ccgaggaaat ttatcaagta ttatattgcg gaaaaaacgc gcgcgaggca gcattgacct		180
tactaggtcg tgcacgcaag gacgagcgca gcagccacta accccaggga acctttgita		240
ccgctatgac ccggcccgcg cagaacgggc cggtcattat ctcatcggtt ggagtaagca		300
atg tcg igt gaa gaa ctg gaa att gtc tgg aac aat att aaa gcc gaa		348
Met Ser Cys Glu Glu Leu Glu Ile Val Trp Asn Asn Ile Lys Ala Glu		
1 5 10 15		
gcc aga acg ctg gcg gac tgt gag cca atg ctg gcc agt ttt tac cac		396
Ala Arg Thr Leu Ala Asp Cys Glu Pro Met Leu Ala Ser Phe Tyr His		
20 25 30		
gcg acg cta ctc aag cac gaa aac ctt ggc agt gca ctg agc tac atg		444
Ala Thr Leu Leu Lys His Glu Asn Leu Gly Ser Ala Leu Ser Tyr Met		
35 40 45		
ctg gcg aac aag ctg tca tgc cca att atg cct gct att gct atc cgt		492
Leu Ala Asn Lys Leu Ser Ser Pro Ile Met Pro Ala Ile Ala Ile Arg		
50 55 60		
gaa gtg gtg gaa gaa gcc tac gcc gct gac ccg gaa atg atc gcc tct		540
Glu Val Val Glu Glu Ala Tyr Ala Ala Asp Pro Glu Met Ile Ala Ser		
65 70 75 80		
gcg gcc tgt gat att cag gcg gtg cgt acc cgc gac ccg gca gtc gat		588
Ala Ala Cys Asp Ile Gln Ala Val Arg Thr Arg Asp Pro Ala Val Asp		
85 90 95		
aaa tac tca acc ccg ttg tta tac ctg aag ggt ttt cat gcc ttg cag		636
Lys Tyr Ser Thr Pro Leu Leu Tyr Leu Lys Gly Phe His Ala Leu Gln		
100 105 110		
gcc tat cgc atc ggt cac tgg ttg tgg aat cag ggg cgt cgc gca ctg		684
Ala Tyr Arg Ile Gly His Trp Leu Trp Asn Gln Gly Arg Arg Ala Leu		
115 120 125		
gca atc ttt ctg caa aac cag gtt tct gtg acg ttc cag gtc gat att		732
Ala Ile Phe Leu Gln Asn Gln Val Ser Val Thr Phe Gln Val Asp Ile		
130 135 140		
cac ccg gca gca aaa att ggt cgc ggt atc atg ctt gac cac gcg aca		780
His Pro Ala Ala Lys Ile Gly Arg Gly Ile Met Leu Asp His Ala Thr		
145 150 155 160		
ggc atc gtc gtt ggt gaa acg gcg gtg att gaa aac gac gta tgc att		828
Gly Ile Val Val Gly Glu Thr Ala Val Ile Glu Asn Asp Val Ser Ile		

[0011]

```

      165      170      175
ctg caa tct gtg acg ctt ggc ggt acg ggt aaa tct ggt ggt gac cgt 876
Leu Gln Ser Val Thr Leu Gly Gly Thr Gly Lys Ser Gly Gly Asp Arg
      180      185      190
cac ccg aaa att cgt gaa ggt gtg atg att ggc gcg ggc gcg aaa atc 924
His Pro Lys Ile Arg Glu Gly Val Met Ile Gly Ala Gly Ala Lys Ile
      195      200      205
ctc ggc aat att gaa gtt ggg cgc ggc gcg aag att ggc gca ggt tcc 972
Leu Gly Asn Ile Glu Val Gly Arg Gly Ala Lys Ile Gly Ala Gly Ser
      210      215      220
gtg gtg ctg caa ccg gtg ccg ccg cat acc acc gcc gct ggc gtt ccg 1020
Val Val Leu Gln Pro Val Pro Pro His Thr Thr Ala Ala Gly Val Pro
      225      230      235      240
gct cgt att gtc ggt aaa cca gac agc gat aag cca tca atg gat atg 1068
Ala Arg Ile Val Gly Lys Pro Asp Ser Asp Lys Pro Ser Met Asp Met
      245      250      255
gac cag cat ttc aac ggt att aac cat aca ttt gag tat ggg gat ggg 1116
Asp Gln His Phe Asn Gly Ile Asn His Thr Phe Glu Tyr Gly Asp Gly
      260      265      270
atc taatgtccig tgatcgtgcc ggatgcgaig taatcatcta tccggcctac 1169
Ile

agtaactaat ctctcaatac cgtccccgga taccccaact gccgccaggc ttcatacacc 1229
actaccgaca ccgcatigga cagattcatg ctgcggctgt ccggcaccat cggaaatgcga 1289
attttttgtt cagcgggcag ggcatcaaga atgctcgcig gcagcccgcg tgtttccggg 1349
ccgaacatca gataatcgcc atctgatag ctacggcgcg tgtgagcagg tgiaccttic 1409
gtggtgaggg cga 1422

```

<210> 38
 <211> 273
 <212> PRT
 <213> 大肠杆菌 (Escherichia coli)

```

<400> 38
Met Ser Cys Glu Glu Leu Glu Ile Val Trp Asn Asn Ile Lys Ala Glu
1      5      10      15
Ala Arg Thr Leu Ala Asp Cys Glu Pro Met Leu Ala Ser Phe Tyr His
20     25     30
Ala Thr Leu Leu Lys His Glu Asn Leu Gly Ser Ala Leu Ser Tyr Met
35     40     45
Leu Ala Asn Lys Leu Ser Ser Pro Ile Met Pro Ala Ile Ala Ile Arg
50     55     60
Glu Val Val Glu Glu Ala Tyr Ala Ala Asp Pro Glu Met Ile Ala Ser
65     70     75     80
Ala Ala Cys Asp Ile Gln Ala Val Arg Thr Arg Asp Pro Ala Val Asp
85     90     95
Lys Tyr Ser Thr Pro Leu Leu Tyr Leu Lys Gly Phe His Ala Leu Gln
100    105    110
Ala Tyr Arg Ile Gly His Trp Leu Trp Asn Gln Gly Arg Ala Leu
115    120    125
Ala Ile Phe Leu Gln Asn Gln Val Ser Val Thr Phe Gln Val Asp Ile
130    135    140
His Pro Ala Ala Lys Ile Gly Arg Gly Ile Met Leu Asp His Ala Thr
145    150    155    160
Gly Ile Val Val Gly Glu Thr Ala Val Ile Glu Asn Asp Val Ser Ile
165    170    175
Leu Gln Ser Val Thr Leu Gly Gly Thr Gly Lys Ser Gly Gly Asp Arg
180    185    190
His Pro Lys Ile Arg Glu Gly Val Met Ile Gly Ala Gly Ala Lys Ile
195    200    205
Leu Gly Asn Ile Glu Val Gly Arg Gly Ala Lys Ile Gly Ala Gly Ser
210    215    220
Val Val Leu Gln Pro Val Pro Pro His Thr Thr Ala Ala Gly Val Pro
225    230    235    240
Ala Arg Ile Val Gly Lys Pro Asp Ser Asp Lys Pro Ser Met Asp Met
245    250    255
Asp Gln His Phe Asn Gly Ile Asn His Thr Phe Glu Tyr Gly Asp Gly
260    265    270
Ile

```

<210> 39
 <211> 1039
 <212> DNA
 <213> 大肠杆菌 (Escherichia coli)

<220>
 <221> CDS
 <222> (201).. (839)

[0012]

```

<400> 39
tgcggataac ggtagaattt ttacgccagt attttgccga gcactaccgg aatttttcac 60
tggagcatgc ctgattaatg attcaattat cgggttgata tcagggttaa acctgatttt 120
ctccitttcta agcgcctaca gatgggttag cataattacc tttaatcgcg catgatcgaa 180
agataattaa agagggttaat gtg ttc gct gaa tac ggg gtt ctg aat tac tgg 233
          Val Phe Ala Glu Tyr Gly Val Leu Asn Tyr Trp
          1          5          10
acc tat ctg gtt ggg gcc att ttt att gtg ttg gtg cca ggg cca aat 281
Thr Tyr Leu Val Gly Ala Ile Phe Ile Val Leu Val Pro Gly Pro Asn
          15          20          25
acc ctg ttt gta etc aaa aat agc gtc agt agc ggt atg aaa ggc ggt 329
Thr Leu Phe Val Leu Lys Asn Ser Val Ser Ser Gly Met Lys Gly Gly
          30          35          40
tat ctt gcg gcc tgc ggt gta ttt att ggc gat gcg gta ttg atg ttt 377
Tyr Leu Ala Ala Cys Gly Val Phe Ile Gly Asp Ala Val Leu Met Phe
          45          50          55
ctg gca tgg gct gga gtg gcg aca tta att aag acc acc ccg ata tta 425
Leu Ala Trp Ala Gly Val Ala Thr Leu Ile Lys Thr Thr Pro Ile Leu
          60          65          70          75
ttc aac att gta cgt tat ctt ggt gcg ttt tat ttg etc tat ctg ggg 473
Phe Asn Ile Val Arg Tyr Leu Gly Ala Phe Tyr Leu Leu Tyr Leu Gly
          80          85          90
agt aaa att ctt tac gcg acc ctg aag ggt aaa aat agc gag gcc aaa 521
Ser Lys Ile Leu Tyr Ala Thr Leu Lys Gly Lys Asn Ser Glu Ala Lys
          95          100          105
tcc gat gag ccc caa tac ggt gct att ttt aaa cgc gcg tta att ttg 569
Ser Asp Glu Pro Gln Tyr Gly Ala Ile Phe Lys Arg Ala Leu Ile Leu
          110          115          120
agc ctg act aat ccg aaa gcc att ttg ttc tat gtg tgg ttt ttc gta 617
Ser Leu Thr Asn Pro Lys Ala Ile Leu Phe Tyr Val Ser Phe Phe Val
          125          130          135
cag ttt atc gat gtt aat gcc cca cat acg gga att tca ttc ttt att 665
Gln Phe Ile Asp Val Asn Ala Pro His Thr Gly Ile Ser Phe Phe Ile
          140          145          150          155
ctg gcg gcg acg ctg gaa ctg gtg agt ttc tgc tat ttg agc ttc ctg 713
Leu Ala Ala Thr Leu Glu Leu Val Ser Phe Cys Tyr Leu Ser Phe Leu
          160          165          170
att ata tct ggt gct ttt gtc acg cag tac ata cgt acc aaa aag aaa 761
Ile Ile Ser Gly Ala Phe Val Thr Gln Tyr Ile Arg Thr Lys Lys Lys
          175          180          185
ctg gct aaa gtt ggc aac tca ctg att ggt ttg atg ttc gtg ggt ttc 809
Leu Ala Lys Val Gly Asn Ser Leu Ile Gly Leu Met Phe Val Gly Phe
          190          195          200
gct gcc cga ctg gcg acg ctg caa tcc tga tgccttcagc ccgcgttgtc 859
Ala Ala Arg Leu Ala Thr Leu Gln Ser
          205          210
gggggttccc catctataat cctcccctgat tcttcgctga tatgggtgcta aaaagtaacc 919
aataaatggt atttaaaatg caaattatca ggcgtaccct gaaacggctg gaataaacgg 979
ttttcagcgc attcaccgaa ggagggaataa ggatgcttca aatcccacag aattatattc 1039

```

```

<210> 40
<211> 212
<212> PRT
<213> 大肠杆菌(Escherichia coli)

```

```

<400> 40
Val Phe Ala Glu Tyr Gly Val Leu Asn Tyr Trp Thr Tyr Leu Val Gly
1          5          10          15
Ala Ile Phe Ile Val Leu Val Pro Gly Pro Asn Thr Leu Phe Val Leu
          20          25          30
Lys Asn Ser Val Ser Ser Gly Met Lys Gly Gly Tyr Leu Ala Ala Cys
          35          40          45
Gly Val Phe Ile Gly Asp Ala Val Leu Met Phe Leu Ala Trp Ala Gly
          50          55          60
Val Ala Thr Leu Ile Lys Thr Thr Pro Ile Leu Phe Asn Ile Val Arg
          65          70          75          80
Tyr Leu Gly Ala Phe Tyr Leu Leu Tyr Leu Gly Ser Lys Ile Leu Tyr
          85          90          95
Ala Thr Leu Lys Gly Lys Asn Ser Glu Ala Lys Ser Asp Glu Pro Gln
          100          105          110
Tyr Gly Ala Ile Phe Lys Arg Ala Leu Ile Leu Ser Leu Thr Asn Pro
          115          120          125
Lys Ala Ile Leu Phe Tyr Val Ser Phe Phe Val Gln Phe Ile Asp Val
          130          135          140
Asn Ala Pro His Thr Gly Ile Ser Phe Phe Ile Leu Ala Ala Thr Leu
          145          150          155          160
Glu Leu Val Ser Phe Cys Tyr Leu Ser Phe Leu Ile Ile Ser Gly Ala
          165          170          175

```

[0013]

Phe Val Thr Gln Tyr Ile Arg Thr Lys Lys Lys Leu Ala Lys Val Gly	
180 185 190	
Asn Ser Leu Ile Gly Leu Met Phe Val Gly Phe Ala Ala Arg Leu Ala	
195 200 205	
Thr Leu Gln Ser	
210	
<210> 41	
<211> 313	
<212> DNA	
<213> 大肠杆菌 (Escherichia coli)	
<400> 41	
gcatgcttcc aactgcgcta atgacgcagc tggacgaagg cgggattctc gtcttaccgc	60
taggggagga gcaccagtat ttgaaacggg tgcgtcgtcg gggaggcgaa ttattatcg	120
ataccgtgga ggcggtgcgc ttgtccctt tagtgaaggg tgagctggct taaaacgtga	180
ggaaatacct ggatttttcc tggttatttt gccgcaggic agcgtatcgt gaacatcttt	240
tccagtgttc agtaggggtgc ctigcacggt aattatgtca ctggtattata accaattttt	300
cttgggggtc gac	313
<210> 42	
<211> 312	
<212> DNA	
<213> 人工序列	
<220>	
<223> 突变 nlpD 启动子	
<400> 42	
gcatgcttcc aactgcgcta atgacgcagc tggacgaagg cgggattctc gtcttaccgc	60
taggggagga gcaccagtat ttgaaacggg tgcgtcgtcg gggaggcgaa ttattatcg	120
ataccgtgga ggcggtgcgc ttgtccctt tagtgaaggg tgagctggct taaaacgtga	180
ggaaatacct ggatttttcc tggttatttt gccgcaggic agcgtataat gaagatcttt	240
tccagtgttg acaagggtcc ttgcacggtt ataattgtcac tggttattaa ccaatttttc	300
ctgggggtcg ac	312
<210> 43	
<211> 4403	
<212> DNA	
<213> Pantoea ananatis	
<220>	
<221> CDS	
<222> (301).. (1311)	
<220>	
<221> CDS	
<222> (1317).. (2147)	
<220>	
<221> CDS	
<222> (2150).. (3022)	
<400> 43	
tacagcggaa cctggcacgg gccagaaggg ttgatgccgt cggatgacac actcaagagc	60
tggacgctca gcaaaattgt ctggcagcgc taagtctttt ticacaccgc tcaaccgcag	120
ggcataacgc gccctgcgcg tccaattctg tttttcgtct gtcttttccc gccgccttat	180
gcctttttcg actttgaaat cagcaaacga tatataaaac cgttacgggt ttacgcigag	240
ttataaataa actgctgtat ctgcagatga gatctgcac aaatttcttc aggggtgaacc	300
atg acc tta cca gcg atg aaa aaa atc gtg agc gga ctc gca ctg tcg	348
Met Thr Leu Pro Ala Met Lys Lys Ile Val Ser Gly Leu Ala Leu Ser	
1 5 10 15	
ctg agt ctg gcc ggt gcc gca aac gcg acc gag ctg ttg aac agc tct	396
Leu Ser Leu Ala Gly Ala Ala Asn Ala Thr Glu Leu Leu Asn Ser Ser	
20 25 30	
tac gat gtc gca cgt gaa tta ttt gtc gcc ctg aat gcg cct ttt gtc	444
Tyr Asp Val Ala Arg Glu Leu Phe Val Ala Leu Asn Ala Pro Phe Val	
35 40 45	
agc cag tgg gat gcc agc cat cct gac gac aag ctg acc att aag atg	492
Ser Gln Trp Asp Ala Ser His Pro Asp Asp Lys Leu Thr Ile Lys Met	
50 55 60	
tcc cat gcc ggg tca tcc aaa cag gcg ctg gcg atc ctg caa ggc ctg	540
Ser His Ala Gly Ser Ser Lys Gln Ala Leu Ala Ile Leu Gln Gly Leu	
65 70 75 80	
cgt gcc gat gtg gtg acc tat aac cag gtc acc gat gtg cag gtg ctg	588
Arg Ala Asp Val Val Thr Tyr Asn Gln Val Thr Asp Val Gln Val Leu	
85 90 95	

[0014]

cac gat aaa ggc aaa ctg atc cct gcc gac tgg caa acc cgc ctg ccg His Asp Lys Gly Lys Leu Ile Pro Ala Asp Trp Gln Thr Arg Leu Pro 100 105	636
aat aac agt tgc ccg ttt tac tcc acc atg gcg ttc ctg gtg cgc aag Asn Asn Ser Ser Pro Phe Tyr Ser Thr Met Ala Phe Leu Val Arg Lys 115 120 125	684
gga aac cca aag cag att cac gac tgg tcc gat tta acc cgt gac gat Gly Asn Pro Lys Gln Ile His Asp Trp Ser Asp Leu Thr Arg Asp Asp 130 135 140	732
gtg aag ctg att ttt cct aat ccc aaa acc tgc ggc aac gga cgt tat Val Lys Leu Ile Phe Pro Asn Pro Lys Thr Ser Gly Asn Gly Arg Tyr 145 150 155 160	780
acc tat ctt gct gcc tgg ggc gcc agc aac act gac ggg ggc gal Thr Tyr Leu Ala Ala Trp Gly Ala Ala Ser Asn Thr Asp Gly Gly Asp 165 170 175	828
cag gct aaa acc cgc gct ttt atg aca aaa ttt ctg aaa aat gtt gaa Gln Ala Lys Thr Arg Ala Phe Met Thr Lys Phe Leu Lys Asn Val Glu 180 185 190	876
gtc ttc gat acc ggt ggc cga ggt gct acg acc acc ttt gct gaa cgc Val Phe Asp Thr Gly Gly Arg Gly Ala Thr Thr Thr Phe Ala Glu Arg 195 200 205	924
ggt ctg ggc gat gtg ttg atc agt ttt gag tct gaa gtg aat aac atc Gly Leu Gly Asp Val Leu Ile Ser Phe Glu Ser Glu Val Asn Asn Ile 210 215 220	972
cgc aac cag tac ggc aaa gac gac tac gaa gtc gtg gtg cct aaa acc Arg Asn Gln Tyr Gly Lys Asp Asp Tyr Glu Val Val Val Pro Lys Thr 225 230 235 240	1020
gal att ctg gcg gag ttt ccc gtt gcc tgg gla gat aaa aac gtc gag Asp Ile Leu Ala Glu Phe Pro Val Ala Trp Val Asp Lys Asn Val Glu 245 250 255	1068
cag aat aaa aca gcc gat gca gcg aaa gcc tat ctg acc tgg ctg tat Gln Asn Lys Thr Ala Asp Ala Ala Lys Ala Tyr Leu Thr Trp Leu Tyr 260 265 270	1116
tct cct gcg gcg cag aaa att att acg gat ttc tat tac cgc gtg aac Ser Pro Ala Ala Gln Lys Ile Ile Thr Asp Phe Tyr Tyr Arg Val Asn 275 280 285	1164
aat ccg cag tta atg gcg cag caa aaa gcc cgt ttt cct gcc acg aac Asn Pro Gln Leu Met Ala Gln Gln Lys Ala Arg Phe Pro Ala Thr Asn 290 295 300	1212
ctg ttt cgt gtt gaa gac att ttt ggc ggc tgg gat aac gtg atg aaa Leu Phe Arg Val Glu Asp Ile Phe Gly Gly Trp Asp Asn Val Met Lys 305 310 315 320	1260
acc cat ttc gcc agc ggt ggc gag cta gac cag tta tta gcg gcg ggg Thr His Phe Ala Ser Gly Gly Glu Leu Asp Gln Leu Leu Ala Ala Gly 325 330 335	1308
cgg tgatc atg ttt gca gcc agc caa aaa cgc gtc ctg ccc ggt ttc ggt Arg Met Phe Ala Ala Ser Gln Lys Arg Val Leu Pro Gly Phe Gly 340 345 350	1358
ctc agc ctg ggc acc agc ctg ctg ttt acc tgt ctg gtc ctg ctg ctg Leu Ser Leu Gly Thr Ser Leu Leu Phe Thr Cys Leu Val Leu Leu Leu 355 360 365	1406
cca atc agc gca ctg att atg cag ctg tgc cag atg acg ttg cag caa Pro Ile Ser Ala Leu Ile Met Gln Leu Ser Gln Met Thr Leu Gln Gln 370 375 380	1454
tac tgg gac gtg gtc acc aat ccg cag ctg atc gcg gcc tat aag gtc Tyr Trp Asp Val Val Thr Asn Pro Gln Leu Ile Ala Ala Tyr Lys Val 385 390 395	1502
acg ctg ctg tgc gcc ggt gtg gcc tca ctg ttt aat gcc gta ttc ggc Thr Leu Leu Ser Ala Gly Val Ala Ser Leu Phe Asn Ala Val Phe Gly 400 405 410 415	1550
atg tta atg gcg tgg atc tta acg cgt tac cgt ttt ccg ggc cgc acg Met Leu Met Ala Trp Ile Leu Thr Arg Tyr Arg Phe Pro Gly Arg Thr 420 425 430	1598
ctg ctg gat ggt ctg atg gat ctg ccg ttt gcg ctg ccg acc gcg gtt Leu Leu Asp Gly Leu Met Asp Leu Pro Phe Ala Leu Pro Thr Ala Val 435 440 445	1646
gct ggc ctg acg ctg gcc ggt ctg ttt tcc gtg aac ggc tgg tac gga Ala Gly Leu Thr Leu Ala Gly Leu Phe Ser Val Asn Gly Trp Tyr Gly 450 455 460	1694
caa tgg ttc gcg cat ttt gat atc aag atc tcc tat acc tgg atc ggt Gln Trp Phe Ala His Phe Asp Ile Lys Ile Ser Tyr Thr Trp Ile Gly 465 470 475	1742
atc gcg ctg gcg atg gcc ttc acc agt att ccg ttt gtg gtg cgt acc Ile Ala Leu Ala Met Ala Phe Thr Ser Ile Pro Phe Val Val Arg Thr 480 485 490 495	1790
gtg cag ccg gtg ctg gaa gag ctg ggg cct gaa tat gag gaa gcg gct Val Gln Pro Val Leu Glu Glu Leu Gly Pro Glu Tyr Glu Glu Ala Ala 500 505 510	1838
caa acg ctg ggc gcc acg ccc tgg cag agc ttc cgc ccg gtc gtt ctg Gln Thr Leu Gly Ala Thr Pro Trp Gln Ser Phe Arg Arg Val Val Leu 515 520 525	1886

[0015]

cct gaa gtg gca ccg gcc tta ctt gcg ggc acc gcg ctg tcg ttt acc	515	520	525	1934
Pro Glu Val Ala Pro Ala Leu Leu Ala Gly Thr Ala Leu Ser Phe Thr	530	535	540	
cgc agc ctg ggc gag ttt ggt gcg gta atc ttt att gcc ggc aac atc	545	550	555	1982
Arg Ser Leu Gly Glu Phe Val Thr Ser Leu Met Ile Phe Val Arg Leu Gln	560	565	570	2030
gag ttt gac tat ccg gca gcc agc gcc att gcc tcg gtc att ctg gcg	580	585	590	2078
Glu Phe Asp Tyr Pro Ala Ala Ser Ala Ile Ala Ser Val Ile Leu Ala	595	600	605	2126
gca tca ctg ctg tta ctt ttc gct atc aat acc tta caa agc cgc ttt	610	615	620	2173
Gly Arg Arg Leu Gly Gly His Met Ala Glu Ile Ser Gln Leu Asn	625	630	635	2221
cat gcc gac cgc cag cct gtt aac tgg gcc aag tgg ctg ctt att ggt	640	645	650	2269
His Ala Asp Arg Gln Pro Val Asn Trp Ala Lys Trp Leu Leu Ile Gly	655	660	665	2317
att ggt gcg ctg ata tcc ttg ctg ctg ctg gtc gtg ccg atg gtg tcc	670	675	680	2365
Ile Gly Ala Leu Ile Ser Leu Leu Leu Val Val Pro Met Val Ser	685	690	695	2413
atc ttc tgg gag gcc ctg cat aaa gga ctg gcc gtc acc tta agt aat	700	705	710	2461
Ile Phe Trp Glu Ala Leu His Lys Gly Leu Gly Val Thr Leu Ser Asn	715	720	725	2509
ctg acc gac agc gac atg ctg cat gcc ata tgg ctg acg gtg ctg gtc	730	735	740	2557
Leu Thr Asp Ser Asp Met Leu His Ala Ile Trp Leu Thr Val Leu Val	745	750	755	2605
gca ttg att acc gtg ccg gtg aat tta gtg ttc gcc acg ctg ctg gcc	760	765	770	2653
Ala Leu Ile Thr Val Pro Val Asn Leu Val Phe Gly Thr Leu Leu Ala	775	780	785	2701
tgg ctg gtg aca cgc ttt acc ttt ccg gga cgt cag ctg ctt ttg acg	790	795	800	2749
Trp Leu Val Thr Arg Phe Thr Phe Pro Gly Arg Gln Leu Leu Leu Thr	805	810	815	2797
ctg ttc gat att ccc ttt gcg gta tcg cct gtg gtc gcc ggt ctg atg	820	825	830	2845
Leu Phe Asp Ile Pro Phe Ala Val Ser Pro Val Val Ala Gly Leu Met	835	840	845	2893
tat ctg ctg ttc tgg gcc att aac ggc ccg gcg gcc ggc tgg ctg gat	850	855	860	2941
Tyr Leu Leu Phe Trp Gly Ile Asn Gly Pro Ala Gly Gly Trp Leu Asp	865	870	875	2989
gcc cat aat att cag gtg atg ttc tcc tgg cct gcc atg gtg ctg gtc	880	885	890	3042
Ala His Asn Ile Gln Val Met Phe Ser Trp Pro Gly Met Val Leu Val	895	900	905	3102
acc gtc ttc gtt acc tgt ccg ttt gtg gtg cgc gaa ctg gtg ccg gtg	910	915	920	3162
Thr Val Phe Val Thr Cys Pro Phe Val Val Arg Glu Leu Val Pro Val	925	930	935	3222
atg ctg agc cag gcc agt cat gaa gat gaa gcc gcg gtg ctg tta ggt	940	945	950	3282
Met Leu Ser Gln Gly Ser His Glu Asp Glu Ala Ala Val Leu Leu Gly	955	960	965	3342
gcc tcg gcc tgg cag atg ttc cgt cgc gtg acg ctg ccg aat att cgc	970	975	980	3402
Ala Ser Gly Trp Gln Met Phe Arg Arg Val Thr Leu Pro Asn Ile Arg	985	990	995	3462
tgg gcc atg ctg tat gcc gtc gtg ctg acc aac gcc cgc gcg att ggt				
Trp Ala Met Leu Tyr Gly Val Val Leu Thr Asn Ala Arg Ala Ile Gly				
gag ttt gcc gcg gtt tcc gtg gtt tcg ggt tct att cgc ggt gaa acc				
Glu Phe Gly Ala Val Ser Val Val Ser Gly Ser Ile Arg Gly Glu Thr				
tac act tta ccg ctt cag gtt gaa tta ctg cat cag gat tac aac acg				
Tyr Thr Leu Pro Leu Gln Val Glu Leu His Gln Asp Tyr Asn Thr				
gtg gcc gcc ttt act gcc gca gcc tta ctg acc gtg atg gca atc gtg				
Val Gly Ala Phe Thr Ala Ala Ala Leu Leu Thr Val Met Ala Ile Val				
acg ctg ttt ctg aaa agc att gtg caa tgg cgt tta gag caa cag cac				
Thr Leu Phe Leu Lys Ser Ile Val Gln Trp Arg Leu Glu Gln Gln His				
aaa cgc ctg caa ctg gag gac aat cat gag cat tgagattaac cagatcaaca				
Lys Arg Leu Gln Leu Glu Asp Asn His Glu His				
aatccttttgg tcgcacagcg gtgctgaacg atatctcact ggatattcct tcigggccaga				
tgggtggccctt actggggccg tccggttccg gtataaacac gctgctgcgc atcatgtcg				
gactggaaaca tcagaaacagc ggtcagattc gttttcacga ccacgatgtc agccgcctgc				
acgcccgcga tcgccaggtc ggatttgtct tccagcacta tgcgctgttc cgtcatatga				
cggctcttga caatattgcc ttggcctga ccgtgctgcc gcgccgtgag cgtccgtcca				
gtgcggaaat taaaaaacgc gtcacgcgc tgctggagat ggtgcagctt tccatctgg				
cgaacgglll ccgggccag ctllcgggag ggcagaagca gcgcgtcgcg ctggcaagag				

[0016]

```

ccctggccgt ggaaccgcaa atccgtgtgc tggatgagcc ctttgggtgc ctggacgctc 3522
aggtgcgtaa agagctgcgc cgttgggtac gtcagctgca cgaagaattg aagttcacca 3582
gctgttctgt caccacgat caggaagagg cgatggaagt ggccgatcgc gtggtggtga 3642
tgagccaggg cagcatcgaa cagggtggga cgcggatga agtctggcgc gatcccgcca 3702
cgcgttctgt gctggaattc ctgggtgagg ttaaccgctt cgacggtaga gtcattggtt 3762
ctcagttcca tctcggggcg caccactggc cgttaggcta tactcttgca catcaggcgc 3822
cggctgatat gttcctgcgc cgttgggaaa tcgacgttcc gcgcagaagt agcctggaaa 3882
cgcctctgcc cgttcaggtc ttagaagtga gtccctctgg tcacttctgg cagctggttg 3942
tgcagccaac gggatggcag agcagaccct tctcgttggc ctttgacggt gaacagaccg 4002
cgccgttgcg cggcgagcgc ctgttctgtg ggctgcagca ggccagactg taccaggcgc 4062
cgacaccgtt acgggcggtt gcctttgcac acagcgctcg ataggttagg tgaatgttaa 4122
acgcccgag gccttcccg cgatccgggc tttttaatgg caaggtttgt aacctgtaga 4182
cttgataaga cgcgcaagcg tcgcatcagg caacaccacg tatggataga gatcgtgagt 4242
acattagaac aaacaatagg caatacgccct ctggtgaagt tgcagcgaat ggggcgggat 4302
aacggcagtg aagtgtggtt aaaactggaa ggcaataacc cggcagggtc ggtgaaagat 4362
cgtcggcgac ttctgatgat cgtcagggcg gaaaagcgcg g 4403

```

<210> 44
 <211> 337
 <212> PRT
 <213> Pantoca ananatis

<400> 44
 Met Thr Leu Pro Ala Met Lys Lys Ile Val Ser Gly Leu Ala Leu Ser
 1 5 10 15
 Leu Ser Leu Ala Gly Ala Ala Asn Ala Thr Glu Leu Leu Asn Ser Ser
 20 25 30
 Tyr Asp Val Ala Arg Glu Leu Phe Val Ala Leu Asn Ala Pro Phe Val
 35 40 45
 Ser Gln Trp Asp Ala Ser His Pro Asp Asp Lys Leu Thr Ile Lys Met
 50 55 60
 Ser His Ala Gly Ser Ser Lys Gln Ala Leu Ala Ile Leu Gln Gly Leu
 65 70 75 80
 Arg Ala Asp Val Val Thr Tyr Asn Gln Val Thr Asp Val Gln Val Leu
 85 90 95
 His Asp Lys Gly Lys Leu Ile Pro Ala Asp Trp Gln Thr Arg Leu Pro
 100 105 110
 Asn Asn Ser Ser Pro Phe Tyr Ser Thr Met Ala Phe Leu Val Arg Lys
 115 120 125
 Gly Asn Pro Lys Gln Ile His Asp Trp Ser Asp Leu Thr Arg Asp Asp
 130 135 140
 Val Lys Leu Ile Phe Pro Asn Pro Lys Thr Ser Gly Asn Gly Arg Tyr
 145 150 155 160
 Thr Tyr Leu Ala Ala Trp Gly Ala Ala Ser Asn Thr Asp Gly Gly Asp
 165 170 175
 Gln Ala Lys Thr Arg Ala Phe Met Thr Lys Phe Leu Lys Asn Val Glu
 180 185 190
 Val Phe Asp Thr Gly Gly Arg Gly Ala Thr Thr Thr Phe Ala Glu Arg
 195 200 205
 Gly Leu Gly Asp Val Leu Ile Ser Phe Glu Ser Glu Val Asn Asn Ile
 210 215 220
 Arg Asn Gln Tyr Gly Lys Asp Asp Tyr Glu Val Val Val Pro Lys Thr
 225 230 235 240
 Asp Ile Leu Ala Glu Phe Pro Val Ala Trp Val Asp Lys Asn Val Glu
 245 250 255
 Gln Asn Lys Thr Ala Asp Ala Ala Lys Ala Tyr Leu Thr Trp Leu Tyr
 260 265 270
 Ser Pro Ala Ala Gln Lys Ile Ile Thr Asp Phe Tyr Tyr Arg Val Asn
 275 280 285
 Asn Pro Gln Leu Met Ala Gln Gln Lys Ala Arg Phe Pro Ala Thr Asn
 290 295 300
 Leu Phe Arg Val Glu Asp Ile Phe Gly Gly Trp Asp Asn Val Met Lys
 305 310 315 320
 Thr His Phe Ala Ser Gly Gly Glu Leu Asp Gln Leu Leu Ala Ala Gly
 325 330 335
 Arg

<210> 45
 <211> 277
 <212> PRT
 <213> Pantoea ananatis

<400> 45
 Met Phe Ala Ala Ser Gln Lys Arg Val Leu Pro Gly Phe Gly Leu Ser
 1 5 10 15
 Leu Gly Thr Ser Leu Leu Phe Thr Cys Leu Val Leu Leu Leu Pro Ile
 20 25 30

[0017]

```

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

```

```

<210> 46
<211> 291
<212> PRT
<213> Pantoea ananatis

```

```

<400> 46
Met Ala Glu Ile Ser Gln Leu Asn His Ala Asp Arg Gln Pro Val Asn
1           5           10           15
Trp Ala Lys Trp Leu Leu Ile Gly Ile Gly Ala Leu Ile Ser Leu Leu
           20           25           30
Leu Leu Val Val Pro Met Val Ser Ile Phe Trp Glu Ala Leu His Lys
           35           40           45
Gly Leu Gly Val Thr Leu Ser Asn Leu Thr Asp Ser Asp Met Leu His
           50           55           60
Ala Ile Trp Leu Thr Val Leu Val Ala Leu Ile Thr Val Pro Val Asn
           65           70           75           80
Leu Val Phe Gly Thr Leu Leu Ala Trp Leu Val Thr Arg Phe Thr Phe
           85           90           95
Pro Gly Arg Gln Leu Leu Leu Thr Leu Phe Asp Ile Pro Phe Ala Val
           100          105          110
Ser Pro Val Val Ala Gly Leu Met Tyr Leu Leu Phe Trp Gly Ile Asn
           115          120          125
Gly Pro Ala Gly Gly Trp Leu Asp Ala His Asn Ile Gln Val Met Phe
           130          135          140
Ser Trp Pro Gly Met Val Leu Val Thr Val Phe Val Thr Cys Pro Phe
           145          150          155          160
Val Val Arg Glu Leu Val Pro Val Met Leu Ser Gln Gly Ser His Glu
           165          170          175
Asp Glu Ala Ala Val Leu Leu Gly Ala Ser Gly Trp Gln Met Phe Arg
           180          185          190
Arg Val Thr Leu Pro Asn Ile Arg Trp Ala Met Leu Tyr Gly Val Val
           195          200          205
Leu Thr Asn Ala Arg Ala Ile Gly Glu Phe Gly Ala Val Ser Val Val
           210          215          220
Ser Gly Ser Ile Arg Gly Glu Thr Tyr Thr Leu Pro Leu Gln Val Glu
           225          230          235          240
Leu Leu His Gln Asp Tyr Asn Thr Val Gly Ala Phe Thr Ala Ala Ala
           245          250          255
Leu Leu Thr Val Met Ala Ile Val Thr Leu Phe Leu Lys Ser Ile Val
           260          265          270
Gln Trp Arg Leu Glu Gln Gln His Lys Arg Leu Gln Leu Glu Asp Asn
           275          280          285
His Glu His
           290

```

[0018]

```

<210> 47
<211> 1403
<212> DNA
<213> Pantoea ananatis

<220>
<221> CDS
<222> (15)..(1103)

<400> 47
actggaggac aatc atg agc att gag att aac cag atc aac aaa tcc ttt 50
          Met Ser Ile Glu Ile Asn Gln Ile Asn Lys Ser Phe
          1          5          10
ggt cgc aca gcg gtg ctg aac gat atc tca ctg gat att cct tct ggc 98
Gly Arg Thr Ala Val Leu Asn Asp Ile Ser Leu Asp Ile Pro Ser Gly
          15          20          25
cag atg gtg gcc tta ctg ggg ccg tcc ggt tcc ggt aaa acc acg ctg 146
Gln Met Val Ala Leu Leu Gly Pro Ser Gly Ser Gly Lys Thr Thr Leu
          30          35          40
ctg cgc atc att gct gga ctg gaa cat cag aac agc ggt cag att cgt 194
Leu Arg Ile Ile Ala Gly Leu Glu His Gln Asn Ser Gly Gln Ile Arg
          45          50          55
ttt cac gac cac gat gtc agc cgc ctg cac gcc cgc gat cgc cag gtc 242
Phe His Asp His Asp Val Ser Arg Leu His Ala Arg Asp Arg Gln Val
          60          65          70
gga ttt gtc ttc cag cac tat gcg ctg ttc cgt cat atg acg gtc ttc 290
Gly Phe Val Phe Gln His Tyr Ala Leu Phe Arg His Met Thr Val Phe
          80          85          90
gac aat att gcc ttt ggc ctg acc gtg ctg ccg cgc cgt gag cgt ccg 338
Asp Asn Ile Ala Phe Gly Leu Thr Val Leu Pro Arg Arg Glu Arg Pro
          95          100          105
tcc agt gcg gaa att aaa aaa cgc gtc acg cgc ctg ctg gag atg gtg 386
Ser Ser Ala Glu Ile Lys Lys Arg Val Thr Arg Leu Leu Glu Met Val
          110          115          120
cag ctt tcc cat ctg gcg aac cgt ttc ccg gcc cag ctt tcc gga ggg 434
Gln Leu Ser His Leu Ala Asn Arg Phe Pro Ala Gln Leu Ser Gly Gly
          125          130          135
cag aag cag cgc gtc gcg ctg gca aga gcc ctg gcc gtg gaa ccg caa 482
Gln Lys Gln Arg Val Ala Leu Ala Arg Ala Leu Ala Val Glu Pro Gln
          140          145          150
atc ctg ttg ctg gat gag ccc ttt ggt gcg ctg gac gct cag gtc cgt 530
Ile Leu Leu Leu Asp Glu Pro Phe Gly Ala Leu Asp Ala Gln Val Arg
          160          165          170
aaa gag ctg cgc cgt tgg tta cgt cag ctg cac gaa gaa ttg aag ttc 578
Lys Glu Leu Arg Arg Trp Leu Arg Gln Leu His Glu Glu Leu Lys Phe
          175          180          185
acc agc gtg ttc gtc acc cac gat cag gaa gag gcg atg gaa gtg gcc 626
Thr Ser Val Phe Val Thr His Asp Gln Glu Glu Ala Met Glu Val Ala
          190          195          200
gat cgc gtg gtg gtg atg agc cag gcc agc atc gaa cag gtg ggg acg 674
Asp Arg Val Val Val Met Ser Gln Gly Ser Ile Glu Gln Val Gly Thr
          205          210          215
ccg gat gaa gtc tgg cgc gat ccc gcc acg cgc ttc gtg ctg gaa ttc 722
Pro Asp Glu Val Trp Arg Asp Pro Ala Thr Arg Phe Val Leu Glu Phe
          220          225          230
ctg ggt gag gtt aac cgc ttc gac ggt gaa gtg cat ggt tct cag ttc 770
Leu Gly Glu Val Asn Arg Phe Asp Gly Glu Val His Gly Ser Gln Phe
          240          245          250
cat gtc ggg gcg cac cac tgg ccg tta gcc tat acc tct gca cat cag 818
His Val Gly Ala His His Trp Pro Leu Gly Tyr Thr Ser Ala His Gln
          255          260          265
ggc gcg gtc gat ctg ttc ctg cgc ccg tgg gaa atc gac gtt tcc cgc 866
Gly Ala Val Asp Leu Phe Leu Arg Pro Trp Glu Ile Asp Val Ser Arg
          270          275          280
aga agt agc ctg gaa acg ccg ctg ccc gtt cag gtc tta gaa gtg agt 914
Arg Ser Ser Leu Glu Thr Pro Leu Pro Val Gln Val Leu Glu Val Ser
          285          290          295
cct cgt ggt cac ttc tgg cag ctg gtg gtg cag cca acg gga tgg cag 962
Pro Arg Gly His Phe Trp Gln Leu Val Val Gln Pro Thr Gly Trp Gln
          300          305          310
agc gag ccc ttc tcc ctg gtc ttt gac ggt gaa cag acc gcg ccg ttg 1010
Ser Glu Pro Phe Ser Leu Val Phe Asp Gly Glu Gln Thr Ala Pro Leu
          315          320          325
cgc gcc gag cgc ctg ttc gtg ggg ctg cag cag gcc aga ctg tac cag 1058
Arg Gly Glu Arg Leu Phe Val Gly Leu Gln Gln Ala Arg Leu Tyr Gln
          330          335          340
ggc gcg aca ccg tta cgg gcg gtt gcc ttt gca cac agc gcc tga 1103
Gly Ala Thr Pro Leu Arg Ala Val Ala Phe Ala His Ser Ala
          345          350          355
taggttagt gaatgtaaa cgcgcggagg cgttccgcg gatccgggtc tttaaatggc 1163
          360

```

[0019]

```

aaggtttgta acctgttagac ctgataagac ggcgaagcgt cgcacacaggc aacaccacgt 1223
atggatagag atcgttagta cattagaaca aacaataggc aatacgccctc tggtagaagt 1283
gcagcgaaatg gggccggata acggcagtga agtgtggta aaactggaag gcaataaccc 1343
ggcaggttcg gtgaaagatc gtgcggcact ttcatgatc gtccaggcgg aaaagcgcgg 1403

```

<210> 48
 <211> 362
 <212> PRT
 <213> Pantoea ananatis

<400> 48
 Met Ser Ile Glu Ile Asn Gln Ile Asn Lys Ser Phe Gly Arg Thr Ala
 1 5 10 15
 Val Leu Asn Asp Ile Ser Leu Asp Ile Pro Ser Gly Gln Met Val Ala
 20 25 30
 Leu Leu Gly Pro Ser Gly Ser Gly Lys Thr Thr Leu Leu Arg Ile Ile
 35 40 45
 Ala Gly Leu Glu His Gln Asn Ser Gly Gln Ile Arg Phe His Asp His
 50 55 60
 Asp Val Ser Arg Leu His Ala Arg Asp Arg Gln Val Gly Phe Val Phe
 65 70 75 80
 Gln His Tyr Ala Leu Phe Arg His Met Thr Val Phe Asp Asn Ile Ala
 85 90 95
 Phe Gly Leu Thr Val Leu Pro Arg Arg Glu Arg Pro Ser Ser Ala Glu
 100 105 110
 Ile Lys Lys Arg Val Thr Arg Leu Leu Glu Met Val Gln Leu Ser His
 115 120 125
 Leu Ala Asn Arg Phe Pro Ala Gln Leu Ser Gly Gly Gln Lys Gln Arg
 130 135 140
 Val Ala Leu Ala Arg Ala Leu Ala Val Glu Pro Gln Ile Leu Leu Leu
 145 150 155 160
 Asp Glu Pro Phe Gly Ala Leu Asp Ala Gln Val Arg Lys Glu Leu Arg
 165 170 175
 Arg Trp Leu Arg Gln Leu His Glu Glu Leu Lys Phe Thr Ser Val Phe
 180 185 190
 Val Thr His Asp Gln Glu Glu Ala Met Glu Val Ala Asp Arg Val Val
 195 200 205
 Val Met Ser Gln Gly Ser Ile Glu Gln Val Gly Thr Pro Asp Glu Val
 210 215 220
 Trp Arg Asp Pro Ala Thr Arg Phe Val Leu Glu Phe Leu Gly Glu Val
 225 230 235 240
 Asn Arg Phe Asp Gly Glu Val His Gly Ser Gln Phe His Val Gly Ala
 245 250 255
 His His Trp Pro Leu Gly Tyr Thr Ser Ala His Gln Gly Ala Val Asp
 260 265 270
 Leu Phe Leu Arg Pro Trp Glu Ile Asp Val Ser Arg Arg Ser Ser Leu
 275 280 285
 Glu Thr Pro Leu Pro Val Gln Val Leu Glu Val Ser Pro Arg Gly His
 290 295 300
 Phe Trp Gln Leu Val Val Gln Pro Thr Gly Trp Gln Ser Glu Pro Phe
 305 310 315 320
 Ser Leu Val Phe Asp Gly Glu Gln Thr Ala Pro Leu Arg Gly Glu Arg
 325 330 335
 Leu Phe Val Gly Leu Gln Gln Ala Arg Leu Tyr Gln Gly Ala Thr Pro
 340 345 350
 Leu Arg Ala Val Ala Phe Ala His Ser Ala
 355 360

<210> 49
 <211> 1779
 <212> DNA
 <213> Pantoea ananatis

<220>
 <221> CDS
 <222> (301)..(1536)

<400> 49
 gtcaaaaccc tcaaaaaata aagcaccggg cgcacaacag ggcccccgt tttttgatt 60
 taaaaaaact ttctcacca ggcigaaatt tggtagacta tgtcacataa ccgtcatcgg 120
 cagcgggttc gttcttctcg atgcggccaa cccacgattt tgtctggcaa agtacgtcct 180
 ctgagccctg ccatgtctgc ggtcaggcaa tcttttgtat tgcgcaggc gatttttttg 240
 atattttgac agacggctga ctgcgttcag tctcgttga attctgaata gggttgggaa 300
 atg gca aag gta tca ctg gaa aaa gac aaa att aag ttc ctg ctg gtg 348
 Met Ala Lys Val Ser Leu Glu Lys Asp Lys Ile Lys Phe Leu Leu Val
 1 5 10 15
 gaa ggt gtc cat cag agc gcg ctg gaa aat ctt cgt gct gca ggt tac 396
 Glu Gly Val His Gln Ser Ala Leu Glu Asn Leu Arg Ala Ala Gly Tyr

[0020]

acc aat att gaa ttc cac aaa ggc gca ctg gat gcc gag gcg tta aaa	444
Thr Asn Ile Glu Phe His Lys Gly Ala Leu Asp Ala Glu Ala Leu Lys	
35 40 45	
gct tcc gct cgc gat gcg cat ttt atc ggt atc cgt tcc cgt tcc caa	492
Ala Ser Ala Arg Asp Ala His Phe Ile Gly Ile Arg Ser Arg Ser Gln	
50 55 60	
ctg acc gaa gag att ttt gcc gct gca gaa aaa ctg gta gcg gtg ggc	540
Leu Thr Glu Glu Ile Phe Ala Ala Ala Glu Lys Leu Val Ala Val Gly	
65 70 75 80	
tgt ttc tgt atc gga acg aat cag gtt gat tta aat gcc gca gcg aaa	588
Cys Phe Cys Ile Gly Thr Asn Gln Val Asp Leu Asn Ala Ala Ala Lys	
85 90 95	
cgc ggt atc ccg gtt ttt aac gca cct ttc tca aat acg cgc tct gtg	636
Arg Gly Ile Pro Val Phe Asn Ala Pro Phe Ser Asn Thr Arg Ser Val	
100 105 110	
gcc gag ctg gtt att ggc gag atg ctg ctg atg ctg cgc ggt gtt ccg	684
Ala Glu Leu Val Ile Gly Glu Met Leu Leu Met Leu Arg Gly Val Pro	
115 120 125	
gaa gcg aat gcc aaa gcg cac cgt ggt atc tgg aat aaa atc gcc aaa	732
Glu Ala Asn Ala Lys Ala His Arg Gly Ile Trp Asn Lys Ile Ala Lys	
130 135 140	
ggc tct ttt gaa gcg cgc ggt aaa aag ctg ggt atc att ggc tat ggc	780
Gly Ser Phe Glu Ala Arg Gly Lys Lys Leu Gly Ile Ile Gly Tyr Gly	
145 150 155 160	
cat atc ggt atg caa ctg ggc gtg ctg gca gaa agt ctg ggc atg cac	828
His Ile Gly Met Gln Leu Gly Val Leu Ala Glu Ser Leu Gly Met His	
165 170 175	
ggt tac ttc tat gac atc gaa aac aag ctg ccg ttg ggc aac gca tca	876
Val Tyr Phe Tyr Asp Ile Glu Asn Lys Leu Pro Leu Gly Asn Ala Ser	
180 185 190	
cag gtt cgt agc ctg acg cag ttg cta aat atg agt gac gtt gtc agc	924
Gln Val Arg Ser Leu Thr Gln Leu Leu Asn Met Ser Asp Val Val Ser	
195 200 205	
ctg cat gtc ccg gaa acc gcc tct acg caa aat atg att tct gcc aat	972
Leu His Val Pro Glu Thr Ala Ser Thr Gln Asn Met Ile Ser Ala Asn	
210 215 220	
gag ctg gct cag atg aag cct ggc ggc ctg ctg ala aat gcc tca cgc	1020
Glu Leu Ala Gln Met Lys Pro Gly Gly Leu Leu Ile Asn Ala Ser Arg	
225 230 235 240	
ggc acc gtg gta gat att cct gct ttg tgc gaa gcg ctg gcc agc aag	1068
Gly Thr Val Val Asp Ile Pro Ala Leu Cys Glu Ala Leu Ala Ser Lys	
245 250 255	
cag gtt ggt ggc gct gcg att gat gtg ttc cct gta gag ccg gcg acc	1116
Gln Val Gly Gly Ala Ala Ile Asp Val Phe Pro Val Glu Pro Ala Thr	
260 265 270	
aac agc gat ccg ttt gtt tcc cca ctg agc gaa ttc gac aac gtt atc	1164
Asn Ser Asp Pro Phe Val Ser Pro Leu Ser Glu Phe Asp Asn Val Ile	
275 280 285	
ctg acg ccg cac atc ggg gga tgc acg gaa gaa gct cag gag aat atc	1212
Leu Thr Pro His Ile Gly Gly Ser Thr Glu Glu Ala Gln Glu Asn Ile	
290 295 300	
ggg att gaa gtc gcg ggc aag ctg gcg aaa tat tgc gat aac ggt tca	1260
Gly Ile Glu Val Ala Gly Lys Leu Ala Lys Tyr Ser Asp Asn Gly Ser	
305 310 315 320	
acg ctg tcc gcc gtc aat ttc ccg gaa gtg tca ttg ccg atg cac ggc	1308
Thr Leu Ser Ala Val Asn Phe Pro Glu Val Ser Leu Pro Met His Gly	
325 330 335	
att agc gcc agt cgt ctg ctg cat att cac gaa aac cgt ccg ggc gtt	1356
Ile Ser Ala Ser Arg Leu Leu His Ile His Glu Asn Arg Pro Gly Val	
340 345 350	
ctc acc gcg atc aac cag att ttc gct gaa caa ggc atc aac att gcc	1404
Leu Thr Ala Ile Asn Gln Ile Phe Ala Glu Gln Gly Ile Asn Ile Ala	
355 360 365	
gct cag tac ctg caa acc tct ccg atg atg ggt tat gtg gtc atc gac	1452
Ala Gln Tyr Leu Gln Thr Ser Pro Met Met Gly Tyr Val Val Ile Asp	
370 375 380	
att gat gct gag cac gaa ctg gca gag aaa gct ctg caa ctg atg aag	1500
Ile Asp Ala Glu His Glu Leu Ala Glu Lys Ala Leu Gln Leu Met Lys	
385 390 395 400	
gcg att ccg gga acg att cgc gcc cgc ctg ctt tac tgatccacg	1546
Ala Ile Pro Gly Thr Ile Arg Ala Arg Leu Leu Tyr	
405 410	
ctgtcaccta cccgggcaca caagcatgcc cgggtttatt catcccatag ccacagtttt	1606
gatggcgtca gcacggccgg caaaggaatg tcccacgcc ctgtaggcag cgcgtcaacc	1666
cgcgtacagt catgagcgtat gccacccggt aaaaacccat gctgtttcca gtctgttaag	1726
gtgcgacgtt agaagccgcc cccatttctt aaacgtgttc cggcgcgac gaa	1779

<210> 50

[0021]

```

<211> 412
<212> PRT
<213> Pantoea ananatis

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

```

```

<210> 51
<211> 36
<212> DNA
<213> 人工序列

```

```

<220>
<223> 引物 P11

```

```

<400> 51
agctgagtcg acccccagga aaaattgggtt aataac

```

36

```

<210> 52
<211> 33
<212> DNA
<213> 人工序列

```

```

<220>
<223> 引物 P12

```

```

<400> 52

```

[0022]

agctgagcat gcttccaact gcgctaata ga cgc	33
<210> 53	
<211> 33	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引物 P13	
<400> 53	
agctgatc t a g a a a c a g a a t t t g c c t g g c g g c	33
<210> 54	
<211> 33	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引物 P14	
<400> 54	
agctgaggat ccaggaagag tttgtagaaa cgc	33
<210> 55	
<211> 32	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引物 P15	
<400> 55	
agctgagtcg acgtgttcgc tgaatacggg gt	32
<210> 56	
<211> 32	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引物 P16	
<400> 56	
agctgatc t a g a g a a a g c a t c a g g a t t g c a g c	32
<210> 57	
<211> 59	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引物 P17	
<220>	
<221> misc-feature	
<222> (25)..(25)	
<223> n 是 a, c, g, 或 t	
<220>	
<221> misc-feature	
<222> (27)..(28)	
<223> n 是 a, c, g, 或 t	
<220>	
<221> misc-feature	
<222> (46)..(46)	
<223> n 是 a, c, g, 或 t	
<220>	
<221> misc-feature	
<222> (49)..(49)	
<223> n 是 a, c, g, 或 t	
<220>	

[0023]

<221> misc.feature
 <222> (51)..(51)
 <223> n是a, c, g, 或 t

 <400> 57
 atcgtgaaga tcttttccag tgttnannag ggtgccttgc acggtinatna ngtcactgg 59

 <210> 58
 <211> 32
 <212> DNA
 <213> 人工序列

 <220>
 <223> 引物 P18

 <220>
 <221> misc.feature
 <222> (16)..(20)
 <223> n是a, c, g, 或 t

 <400> 58
 tggaaaagat cttcannnnn cgtgacctg cg 32

 <210> 59
 <211> 60
 <212> DNA
 <213> 人工序列

 <220>
 <223> 引物 P19

 <400> 59
 tccgtcacg atttttttca tcgttggtta ggtcatttat cccccaggaa aaattggtta 60

 <210> 60
 <211> 63
 <212> DNA
 <213> 人工序列

 <220>
 <223> 引物 P20

 <400> 60
 tttaacacg ctcaaccgca gggcataacc ggcccttgaa gccctgcttt ttatactaag 60
 ttg 63

 <210> 61
 <211> 20
 <212> DNA
 <213> 人工序列

 <220>
 <223> 引物 P21

 <400> 61
 ctttgtccct ttagtgaagg 20

 <210> 62
 <211> 44
 <212> DNA
 <213> 人工序列

 <220>
 <223> 引物 P22

 <400> 62
 agctgatcia gaagctgact cgagttaatg gcctcccaga cgac 44

 <210> 63
 <211> 33
 <212> DNA
 <213> 人工序列

 <220>

[0024]

<223> 引物 P23

<400> 63
agctgagtcg acatggcaaa ggtatcactg gaa 33

<210> 64
<211> 34
<212> DNA
<213> 人工序列

<220>
<223> 引物 P24

<400> 64
gagaacgccc gggcgggctt cgtgaatatg cagc 34

<210> 65
<211> 32
<212> DNA
<213> 人工序列

<220>
<223> 引物 P25

<400> 65
agctgatcta gacgtgggat cagtaaagca gg 32

<210> 66
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 引物 P26

<400> 66
aaaaccgccc gggcgttctc ac 22

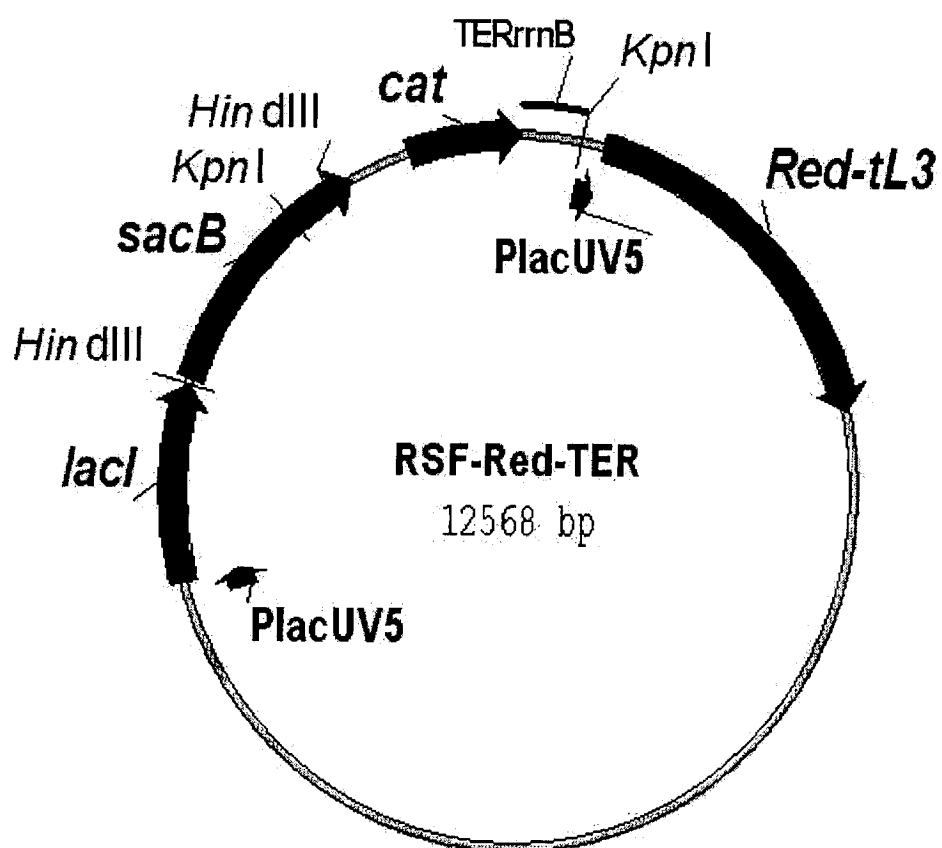


图 1

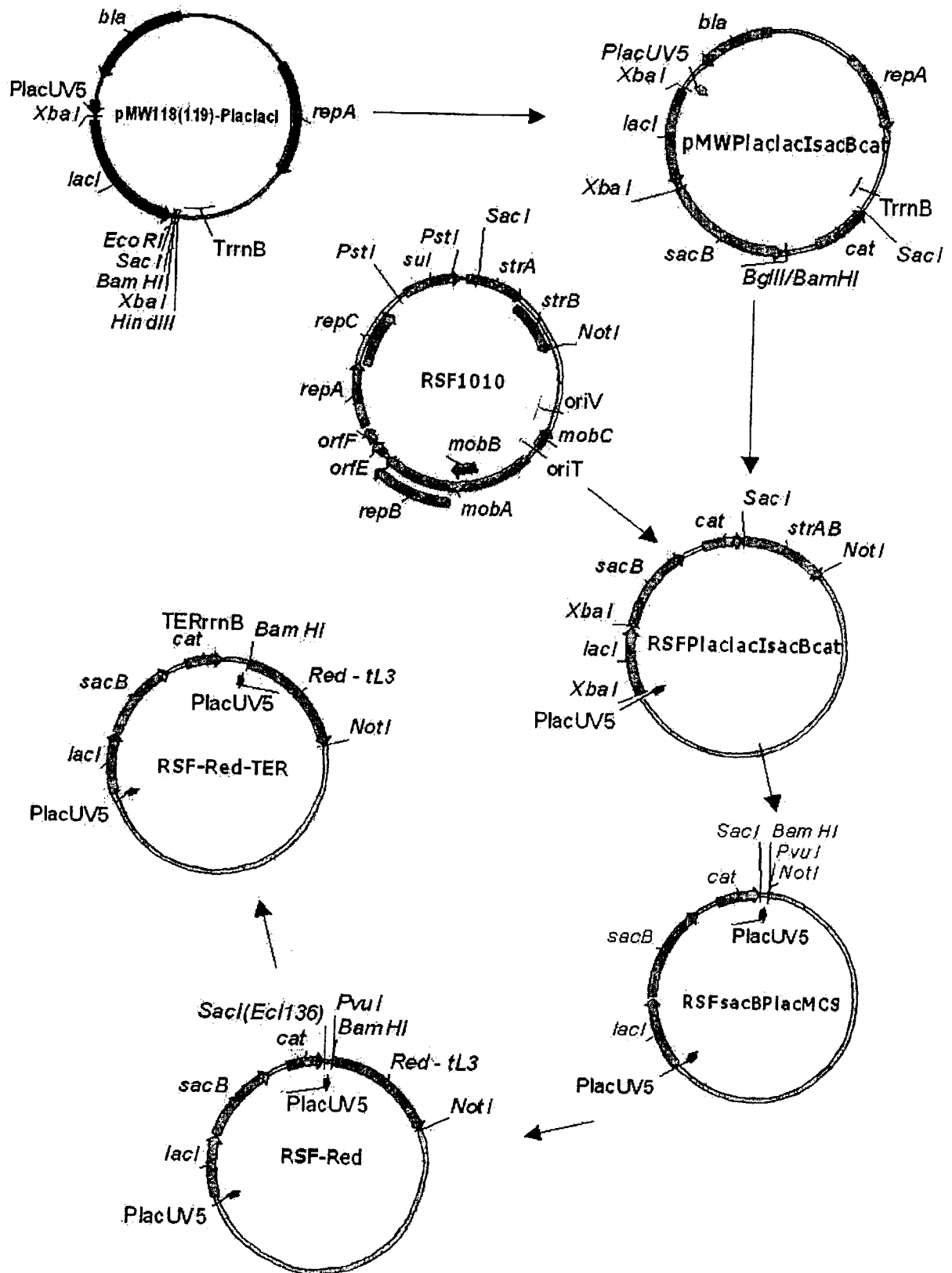


图 2

天然启动子 Pnlp 的序列

-35(Pnlp2) -10(Pnlp2)
 aaaacgtgaggaaatacctggatttttcc**tggtta**ttttgccgcaggtcagcgt**tatcgtg**
 -35(Pnlp1) -10(Pnlp1) 转录起始点
 aagatcttttccagtg**ttcagtg**agggtgccttgcacgggt**aattat**gtcact**tggtt**attaa
 M S
 ccaatttttctgggggataa**atg**agc

图 3