



(10) 授权公告号 CN 109963947 B

(45) 授权公告日 2023. 07. 04

(21) 申请号 201780066783.5

(22) 申请日 2017.10.26

(65) 同一申请的已公布的文献号
申请公布号 CN 109963947 A

(43) 申请公布日 2019.07.02

(30) 优先权数据
62/413,044 2016.10.26 US
62/417,609 2016.11.04 US

(85) PCT国际申请进入国家阶段日
2019.04.26

(86) PCT国际申请的申请数据
PCT/JP2017/038798 2017.10.26

(87) PCT国际申请的公布数据
W02018/079686 EN 2018.05.03

(73) 专利权人 味之素株式会社
地址 日本东京都

(72) 发明人 B.米吉茨 C.罗切 P.凯利
朝里彩香 户矢崎未来 福井启太

(74) 专利代理机构 北京市柳沈律师事务所
11105

专利代理师 张文辉

(51) Int.Cl.
C12N 15/77 (2006.01)
C12P 13/12 (2006.01)
C12P 17/10 (2006.01)
C12P 7/22 (2006.01)
C12P 7/24 (2006.01)
C12P 7/42 (2006.01)
C12P 13/00 (2006.01)
C12P 19/40 (2006.01)

(56) 对比文件
CN 104411821 A,2015.03.11
US 2009281353 A1,2009.11.12
IKEDA,M.等.“Methionine synthase II
(Cobalamin-independent),EC=2.1.1.14,
Database accession no. Q8NNQ9.《DATABASE
UniProt》.2002,Submitted name,Sequence,
family and domain databases部分.

审查员 袁一方

权利要求书2页 说明书46页
序列表68页

(54) 发明名称

用于生产L-甲硫氨酸或需要S-腺苷甲硫氨酸进行合成的代谢物的方法

(57) 摘要

提供了用于生产目标物质,诸如香草醛和香草酸的方法。通过使用具有目标物质生产能力的微生物从碳源或目标物质的前体生产目标物质,其中微生物已经以NCg12048蛋白的活性得以降低的方式进行了修饰。

1. 一种生产目标物质的方法,所述方法包括以下步骤:
通过使用具有生产所述目标物质的能力的微生物来生产所述目标物质,
其中所述微生物是谷氨酸棒杆菌,
其中所述微生物已经以与未修饰的菌株相比由NCg12048基因编码的蛋白质的活性得以降低的方式进行了修饰,
其中通过替换基因启动子来弱化所述NCg12048基因的表达降低由所述NCg12048基因编码的蛋白质的活性,或通过破坏所述NCg12048基因来降低由所述NCg12048基因编码的蛋白质的活性,
其中所述NCg12048基因编码的蛋白质包含SEQ ID NO:93的氨基酸序列,且
其中所述目标物质是香草酸。
2. 根据权利要求1所述的方法,其中所述生产包括:
在含有碳源的培养基中培养所述微生物以在所述培养基中生产和积累所述目标物质。
3. 根据权利要求1所述的方法,其中所述生产包括:
通过使用所述微生物来将所述目标物质的前体转化成所述目标物质。
4. 根据权利要求3所述的方法,其中所述转化包括:
在含有所述前体的培养基中培养所述微生物以在所述培养基中生产和积累所述目标物质。
5. 根据权利要求3所述的方法,其中所述转化包括:
在反应混合物中允许所述微生物的细胞作用于所述前体以在所述反应混合物中生产和积累所述目标物质。
6. 根据权利要求5所述的方法,其中所述细胞是存在于所述微生物的培养液中的细胞、从所述培养液中收集的细胞、存在于所述培养液的加工产物中的细胞、存在于收集的细胞的加工产物中的细胞,或这些的组合。
7. 根据权利要求3至6中任一项所述的方法,其中所述前体是原儿茶酸。
8. 根据权利要求1至7中任一项所述的方法,所述方法还包括收集所述目标物质。
9. 根据权利要求1-8中任一项所述的方法,其中所述微生物已经以与未修饰的菌株相比参与所述目标物质的生物合成的酶的活性得以增加的方式进行了进一步修饰。
10. 根据权利要求9所述的方法,其中所述参与目标物质的生物合成的酶选自以下组成的组:3-脱氧-D-阿拉伯糖-庚酮糖酸7-磷酸合酶、3-脱氢奎尼酸合酶、3-脱氢奎尼酸脱水酶、3-脱氢莽草酸脱水酶、O-甲基转移酶、芳族醛氧化还原酶和它们的组合。
11. 根据权利要求1-10中任一项所述的方法,其中所述微生物已经以与未修饰的菌株相比磷酸泛酰巯基乙胺基转移酶的活性得以增加的方式进行了进一步的修饰。
12. 根据权利要求1至11中任一项所述的方法,其中所述微生物已经以与未修饰的菌株相比参与除所述目标物质以外的物质的副生产的酶的活性得以降低的方式进行了进一步的修饰。
13. 根据权利要求12所述的方法,其中所述参与除目标物质以外的物质的副生产的酶选自以下组成的组:香草酸脱甲基酶、原儿茶酸3,4-双加氧酶、醇脱氢酶、莽草酸脱氢酶和它们的组合。
14. 一种生产香草醛的方法,所述方法包括:

通过权利要求1至13中任一项所述的方法来生产香草酸;和
将所述香草酸转化成香草醛。

用于生产L-甲硫氨酸或需要S-腺苷甲硫氨酸进行合成的代谢物的方法

发明领域

[0001] 本发明涉及通过使用微生物生产目标物质,诸如香草醛和香草酸的方法。

[0002] 发明背景

[0003] 香草醛是提供香草气味的主要成分,并且用作食品、饮料、香水等中的芳香物。香草醛通常通过从天然产物中提取或通过化学合成来生产。

[0004] 已经尝试在生产香草醛的方法中尝试生物工程技术,例如通过使用各种微生物与丁香油酚、异丁香油酚、阿魏酸、葡萄糖、香草酸、椰子壳等(Kaur B.and Chakraborty D., Biotechnological and molecular approaches for vanillin production:a review.Appl Biochem Biotechnol.2013 Feb;169(4):1353-72)。另外,使用生物工程技术生产香草醛的其他方法包括以糖苷生产(WO2013/022881和WO2004/111254),使用香草醛合酶从阿魏酸生产香草醛(JP2015-535181),通过大肠杆菌发酵来生产香草酸,然后将香草酸酶促转化为香草醛(美国专利No.6,372,461)。

[0005] 谷氨酸棒杆菌(*Corynebacterium glutamicum*)的NCgl2048基因编码与MetE和MetH蛋白两者同源的蛋白质,所述MetE和MetH蛋白分别由metE和metH基因编码。虽然由NCgl2048基因编码的蛋白质在一些数据库中标注为甲硫氨酸合酶,但是其实际功能尚未得到鉴定。

[0006] 发明概述

[0007] 本发明描述了用于改善目标物质,诸如香草醛和香草酸的生产的新技术,从而提供了用于有效生产目标物质的方法。

[0008] 本发明的一方面是通过以NCgl2048基因的表达得以弱化的方式修饰微生物,微生物可以以显著改善的方式生产目标物质,诸如香草酸。

[0009] 本发明的一方面提供了生产目标物质的方法,所述方法包括以下步骤:通过使用具有生产所述目标物质的能力的微生物来生产所述目标物质,其中所述微生物已经以与未修饰的菌株相比由NCgl2048基因编码的蛋白质的活性得以降低的方式进行了修饰,且

[0010] 且其中所述目标物质选自以下组成的组:L-甲硫氨酸、需要S-腺苷甲硫氨酸来进行生物合成的代谢物和它们的组合。

[0011] 本发明的另一方面提供了如上文描述的方法,其中所述生产包括在含有碳源的培养基中培养所述微生物以在所述培养基中生产和积累所述目标物质。

[0012] 本发明的另一方面提供了如上文描述的方法,其中所述生产包括通过使用所述微生物来将所述目标物质的前体转化成所述目标物质。

[0013] 本发明的另一方面提供了如上文描述的方法,其中所述转化包括在含有所述前体的培养基中培养所述微生物以在所述培养基中生产和积累所述目标物质。

[0014] 本发明的另一方面提供了如上文描述的方法,其中所述转化包括在反应混合物中允许所述微生物的细胞作用于所述前体以在所述反应混合物中生产和积累所述目标物质。

[0015] 本发明的另一方面提供了如上文描述的方法,其中所述细胞是存在于所述微生物

的培养液中的细胞、从所述培养液中收集的细胞、存在于所述培养液的加工产物中的细胞、存在于收集的细胞的加工产物中的细胞,或这些的组合。

[0016] 本发明的另一方面提供了如上文描述的方法,其中所述前体选自由以下组成的组:原儿茶酸、原儿茶醛、L-色氨酸、L-组氨酸、L-苯丙氨酸、L-酪氨酸、L-精氨酸、L-鸟氨酸、甘氨酸和它们的组合。

[0017] 本发明的另一方面提供了如上文描述的方法,所述方法还包括收集所述目标物质。

[0018] 本发明的另一方面提供了如上文描述的方法,其中所述NCg12048基因编码选自由以下组成的组的蛋白质:

[0019] (a)包含SEQ ID NO:93的氨基酸序列的蛋白质,

[0020] (b)包含SEQ ID NO:93的氨基酸序列,且该氨基酸序列包含1至10个氨基酸残基的取代、缺失、插入和/或添加的蛋白质,且其中所述蛋白质具有微生物中所述蛋白质的活性降低导致目标物质的生产增加的性质,和

[0021] (c)包含与SEQ ID NO:93的氨基酸序列具有90%或更高的同一性的氨基酸序列的蛋白质,且其中所述蛋白质具有微生物中所述蛋白质的活性降低导致目标物质的生产增加的性质。

[0022] 本发明的另一方面提供了如上文描述的方法,其中通过弱化所述NCg12048基因的表达或通过破坏所述NCg12048基因来降低由所述NCg12048基因编码的蛋白质的活性。

[0023] 本发明的另一方面提供了如上文描述的方法,其中通过对所述NCg12048基因的表达控制序列进行修饰来弱化所述NCg12048基因的表达。

[0024] 本发明的另一方面提供了如上文描述的方法,其中所述微生物是属于肠杆菌科(Enterobacteriaceae)的细菌、棒状杆菌细菌或酵母。

[0025] 本发明的另一方面提供了如上文描述的方法,其中所述微生物是属于棒状杆菌属(Corynebacterium)的细菌。

[0026] 本发明的另一方面提供了如上文描述的方法,其中所述微生物是谷氨酸棒杆菌(Corynebacterium glutamicum)。

[0027] 本发明的另一方面提供了如上文描述的方法,其中所述微生物是属于埃希氏菌属(Escherichia)的细菌。

[0028] 本发明的另一方面提供了如上文描述的方法,其中所述微生物是大肠杆菌(Escherichia coli)。

[0029] 本发明的另一方面提供了如上文描述的方法,其中所述代谢物选自由以下组成的组:香草醛、香草酸、褪黑激素、麦角硫因、麦根酸、阿魏酸、多胺、愈创木酚、4-乙烯基愈创木酚、4-乙基愈创木酚和肌酸。

[0030] 本发明的另一方面提供了如上文描述的方法,其中所述微生物已经以与未修饰的菌株相比参与所述目标物质的生物合成的酶的活性得以增加的方式进行了进一步修饰。

[0031] 本发明的另一方面提供了如上文描述的方法,其中所述参与目标物质的生物合成的酶选自由以下组成的组:3-脱氧-D-阿拉伯糖-庚酮糖酸7-磷酸合酶、3-脱氢奎尼酸合酶、3-脱氢奎尼酸脱水酶、3-脱氢莽草酸脱水酶、O-甲基转移酶、芳族醛氧化还原酶和它们的组合。

[0032] 本发明的另一方面提供了如上文描述的方法,其中所述微生物已经以与未修饰的菌株相比磷酸泛酰巯基乙胺基转移酶的活性得以增加的方式进行了进一步的修饰。

[0033] 本发明的另一方面提供了如上文描述的方法,其中所述微生物已经以与未修饰的菌株相比参与除所述目标物质以外的物质的副生产的酶的活性得以降低的方式进行了进一步的修饰。

[0034] 本发明的另一方面提供了如上文描述的方法,其中所述参与除目标物质以外的物质的副生产的酶选自以下组成的组:香草酸脱甲基酶、原儿茶酸3,4-双加氧酶、醇脱氢酶、莽草酸脱氢酶和它们的组合。

[0035] 本发明的另一方面提供了用于生产香草醛的方法,所述方法包括通过如上文所述的方法生产香草酸;并且将所述香草酸转化成香草醛。

[0036] 发明详述

[0037] <1>微生物

[0038] 如本文所述的微生物是具有生产目标物质的能力的微生物,所述微生物已经以NCg12048蛋白的活性得以降低的方式进行了修饰,所述NCg12048蛋白是由NCg12048基因编码的蛋白质。生产目标物质的能力也可以称为“目标物质生产能力”。

[0039] <1-1>具有目标物质生产能力的微生物

[0040] 短语“具有目标物质生产能力的微生物”可以指能够生产目标物质的微生物。

[0041] 短语“具有目标物质生产能力的微生物”可以指在微生物用于发酵方法的情况下能够通过发酵生产目标物质的微生物。也就是说,短语“具有目标物质生产能力的微生物”可以指能够从碳源生产目标物质的微生物。具体而言,短语“具有目标物质生产能力的微生物”可以指当在含有碳源的培养基等培养基中培养时能够在培养基中以一定程度生产并积累目标物质,从而可以从培养基中收集目标物质的微生物。

[0042] 短语“具有目标物质生产能力的微生物”还可以指在将微生物用于生物转化方法的情况下能够通过生物转化生产目标物质的微生物。也就是说,短语“具有目标物质生产能力的微生物”可以指能够从目标物质的前体生产目标物质的微生物。具体而言,短语“具有目标物质生产能力的微生物”可以指当在含有目标物质的前体的培养基中培养时,能够在培养基中以一定程度生产并积累目标物质,从而可以从培养基中收集目标物质的微生物。具体而言,短语“具有目标物质生产能力的微生物”还可以指能够在被允许作用于反应混合物中的目标物质的前体时在反应混合物中以一定程度生产和积累目标物质,从而可以从反应混合物中收集目标物质的微生物。

[0043] 具有目标物质生产能力的微生物可以是能够在培养基或反应混合物中以比用非修饰菌株可以获得的量更大的量生产和积累目标物质的微生物。非修饰的菌株也可以称为“非修饰的微生物的菌株”。短语“非修饰微生物的菌株”或“非修饰的菌株”可以指尚未以使NCg12048蛋白的活性得以降低的方式进行修饰的对照菌株。具有目标物质生产能力的微生物可以能够在培养基或反应混合物中以例如0.01g/L以上、0.05g/L以上或0.09g/L以上的量积累目标物质。

[0044] 目标物质可以选自L-甲硫氨酸和需要S-腺苷甲硫氨酸(SAM)进行生物合成的代谢物。需要SAM进行生物合成的代谢物的实例可以包括例如香草醛、香草酸、褪黑激素、麦角硫因、麦根酸、阿魏酸、多胺、愈创木酚、4-乙烯基愈创木酚、4-乙基愈创木酚和肌酸。微生物可

以能够仅产生一种目标物质的能力,或者可以能够产生两种或更多种目标物质的能力。此外,微生物可以能够从目标物质的一种前体或目标物质的两种或更多种前体产生目标物质的能力。

[0045] 当目标物质是能够形成盐的化合物时,目标物质可以作为游离化合物,其盐或它们的混合物获得。也就是说,除非另有说明,术语“目标物质”可以指游离形式的目标物质、其盐或其混合物。盐的实例可以包括例如硫酸盐,盐酸盐,碳酸盐,铵盐,钠盐和钾盐。作为目标物质的盐,可以使用一种盐,或者可以组合使用两种或更多种盐。

[0046] 可以用作亲本菌株以构建如本文中描述的微生物的微生物没有特别限制。微生物的实例可以包括细菌和酵母。

[0047] 细菌的实例可以包括属于肠杆菌科的细菌和棒状杆菌细菌。

[0048] 属于肠杆菌科的细菌的实例可以包括属于埃希氏菌属、肠杆菌属(*Enterobacter*)、泛菌属(*Pantoea*)、克雷伯氏菌属(*Klebsiella*)、沙雷菌属(*Serratia*)、欧文氏菌属(*Erwinia*)、发光杆菌属(*Photobacterium*)、普罗威登斯菌属(*Providencia*)、沙门氏菌属(*Salmonella*)、摩根菌属(*Morganella*)的细菌。具体地,可以使用根据用于NCBI(National Center for Biotechnology Information)数据库的分类学(ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=91347)分类为肠杆菌科的细菌。

[0049] 埃希氏菌属细菌没有特别限制,并且其实例可以包括根据微生物学领域技术人员已知的分类学分类为埃希氏菌属的细菌。埃希氏菌属细菌的实例可以包括例如记载于Neidhardt et al.(Backmann B.J.,1996,Derivations and Genotypes of some mutant derivatives of *Escherichia coli* K-12,pp.2460-2488,Table 1,In F.D.Neidhardt(编),*Escherichia coli* and *Salmonella* Cellular and Molecular Biology/第二版,American Society for Microbiology Press,Washington,D.C.)工作中的那些细菌。埃希氏菌属细菌的实例可以包括例如大肠杆菌。大肠杆菌的具体实例可以包括大肠杆菌K-12菌株,诸如W3110菌株(ATCC 27325)和MG1655菌株(ATCC 47076);大肠杆菌K5菌株(ATCC 23506);大肠杆菌B菌株,诸如BL21(DE3)菌株;以及其衍生菌株。

[0050] 肠杆菌属细菌没有特别限制,并且其实例可以包括根据微生物学领域技术人员已知的分类学分类为肠杆菌属的细菌。肠杆菌属细菌的实例可以包括例如聚团肠杆菌(*Enterobacter agglomerans*)和产气肠杆菌(*Enterobacter aerogenes*)。聚团肠杆菌的具体实例可以包括例如聚团肠杆菌ATCC 12287菌株。产气肠杆菌的具体实例可以包括例如产气肠杆菌ATCC 13048菌株、NBRC 12010菌株(Biotechnol.Bioeng.,2007,Mar.27;98(2):340-348)和AJ110637菌株(FERM BP-10955)。肠杆菌属细菌的实例还可以包括例如记载于欧洲专利申请公开(EP-A)No.0952221中的菌株。另外,聚团肠杆菌还可以包括分类为成团泛菌(*Pantoea agglomerans*)的一些菌株。

[0051] 泛菌属细菌没有特别限制,并且其实例可以包括根据微生物学领域技术人员已知的分类学分类为泛菌属的细菌。泛菌属细菌的实例可以包括例如菠萝泛菌(*Pantoea ananatis*)、斯氏泛菌(*Pantoea stewartii*)、成团泛菌和柠檬泛菌(*Pantoea citrea*)。菠萝泛菌的具体实例可以包括例如菠萝泛菌LMG20103菌株、AJ13355菌株(FERM BP-6614)、AJ13356菌株(FERM BP-6615)、AJ13601菌株(FERM BP-7207)、SC17菌株(FERM BP-11091)、SC17(0)菌株(VKPM B-9246)和SC17sucA菌株(FERM BP-8646)。一些肠杆菌属细菌和欧文氏

菌属细菌重新分类为泛菌属(Int.J.Syst.Bacteriol.,39,337-345(1989); Int.J.Syst.Bacteriol.,43,162-173(1993))。例如,基于16S rRNA等的核苷酸序列分析,一些聚团肠杆菌菌株最近重新分类为成团泛菌、菠萝泛菌、斯氏泛菌等(Int.J.Syst.Bacteriol.,39,337-345(1989))。泛菌属细菌可以包括如上文描述的重新分类为泛菌属的那些细菌。

[0052] 欧文氏菌属细菌的实例可以包括解淀粉欧文氏菌(*Erwinia amylovora*)和胡萝卜软腐欧文氏菌(*Erwinia carotovora*)。克雷伯氏菌属细菌的实例可以包括植物克雷伯氏菌(*Klebsiella planticola*)。

[0053] 棒状杆菌细菌的实例可以包括属于棒杆菌属、短杆菌属(*Brevibacterium*)、细杆菌属(*Microbacterium*)等的细菌。

[0054] 此类棒状杆菌细菌的具体实例可以包括以下物种。

[0055] 嗜乙酰乙酸棒状杆菌(*Corynebacterium acetoacidophilum*)

[0056] 醋谷氨酸棒杆菌(*Corynebacterium acetoglutamicum*)

[0057] 解烷棒状杆菌(*Corynebacterium alkanolyticum*)

[0058] 帚石南棒状杆菌(*Corynebacterium callunae*)

[0059] 钝齿棒状杆菌(*Corynebacterium crenatum*)

[0060] 谷氨酸棒杆菌(*Corynebacterium glutamicum*)

[0061] 百合棒杆菌(*Corynebacterium lilium*)

[0062] 栖糖蜜棒杆菌(*Corynebacterium melassecola*)

[0063] 热产氨棒状杆菌(*Corynebacterium thermoaminogenes*) (有效棒杆菌(*Corynebacterium efficiens*))

[0064] 力士棒杆菌(*Corynebacterium herculis*)

[0065] 二歧短杆菌(*Brevibacterium divaricatum*) (谷氨酸棒杆菌)

[0066] 黄色短杆菌(*Brevibacterium flavum*) (谷氨酸棒杆菌)

[0067] 未成熟短杆菌(*Brevibacterium immariophilum*)

[0068] 乳糖发酵短杆菌(*Brevibacterium lactofermentum*) (谷氨酸棒杆菌)

[0069] 玫瑰色短杆菌(*Brevibacterium roseum*)

[0070] 解糖短杆菌(*Brevibacterium saccharolyticum*)

[0071] 硫殖短杆菌(*Brevibacterium thiogenitalis*)

[0072] 产氨短杆菌(*Corynebacterium ammoniagenes*) (*Corynebacterium stationis*)

[0073] 白色短杆菌(*Brevibacterium album*)

[0074] 蜡状短杆菌(*Brevibacterium cerinum*)

[0075] 嗜氨细杆菌(*Microbacterium ammoniophilum*)

[0076] 棒状杆菌细菌的具体实例可以包括以下菌株。

[0077] 嗜乙酰乙酸棒杆菌(*Corynebacterium acetoacidophilum*) ATCC 13870

[0078] 醋谷氨酸棒杆菌(*Corynebacterium acetoglutamicum*) ATCC 15806

[0079] 解烷棒杆菌(*Corynebacterium alkanolyticum*) ATCC 21511

[0080] 帚石南棒杆菌(*Corynebacterium callunae*) ATCC 15991

[0081] 钝齿棒杆菌(*Corynebacterium crenatum*) AS1.542

- [0082] 谷氨酸棒杆菌ATCC 13020,ATCC 13032,ATCC 13060,ATCC 13869,FERM BP-734
- [0083] 百合棒杆菌(*Corynebacterium lilium*)ATCC 15990
- [0084] 栖糖蜜棒杆菌(*Corynebacterium melassecola*)ATCC 17965
- [0085] 有效棒杆菌(*Corynebacterium efficiens*) (热产氨棒状杆菌(*Corynebacterium thermoaminogenes*))AJ12340(FERM BP-1539)
- [0086] 力士棒杆菌(*Corynebacterium herculis*)ATCC 13868
- [0087] 二歧短杆菌(*Brevibacterium divaricatum*) (谷氨酸棒杆菌)ATCC 14020
- [0088] 黄色短杆菌(*Brevibacterium flavum*) (谷氨酸棒杆菌)ATCC 13826,ATCC 14067, AJ12418(FERM BP-2205)
- [0089] 未成熟短杆菌(*Brevibacterium immariophilum*)ATCC 14068
- [0090] 乳糖发酵短杆菌(*Brevibacterium lactofermentum*) (谷氨酸棒杆菌)ATCC 13869
- 玫瑰色短杆菌(*Brevibacterium roseum*)ATCC 13825
- [0091] 解糖短杆菌(*Brevibacterium saccharolyticum*)ATCC 14066
- [0092] 硫殖短杆菌(*Brevibacterium thiogenitalis*)ATCC 19240
- [0093] 产氨短杆菌(*Corynebacterium ammoniagenes*) (停滞棒杆菌(*Corynebacterium stationis*))ATCC 6871,ATCC 6872
- [0094] 白色短杆菌(*Brevibacterium album*)ATCC 15111
- [0095] 蜡状短杆菌(*Brevibacterium cerinum*)ATCC 15112
- [0096] 嗜氨细杆菌(*Microbacterium ammoniophilum*)ATCC 15354
- [0097] 棒状杆菌细菌可以包括先前被分类为短杆菌属但目前并入棒状杆菌属中的细菌(Int.J.Syst.Bacteriol.,41,255(1991))。此外,停滞棒杆菌可以包括先前已被分类为产氨棒状杆菌,但目前基于16S rRNA的核苷酸序列分析等重新分类为停滞棒杆菌的细菌(Int.J.Syst.Evol.Microbiol.,60,874-879(2010))。
- [0098] 酵母可以是芽殖酵母或裂殖酵母。酵母可以是单倍体或二倍体以上的多倍体酵母。酵母的实例可包括属于酵母属(*Saccharomyces*)的酵母,例如酿酒酵母(*Saccharomyces cerevisiae*),属于毕赤酵母属(*Pichia*) (其也可以称为Wickerhamomyces属)的酵母,如西弗毕赤酵母(*Pichia ciferrii*),赛道威毕赤酵母(*Pichia sydowiorum*)和巴斯德毕赤酵母(*Pichia pastoris*),属于假丝酵母属(*Candida*)的酵母,诸如实用假性酵母(*Candida utilis*),属于汉逊酵母属(*Hansenula*)的酵母,如多形汉逊酵母(*Hansenula polymorpha*)和属于裂殖酵母(*Schizosaccharomyces*)的酵母,如粟酒裂殖酵母(*Schizosaccharomyces pombe*)。
- [0099] 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;或atcc.org)获得。也就是说,注册号被赋予各自的菌株,并且菌株可以通过使用这些注册号订购(参见atcc.org)。美国典型培养物保藏中心的目录中列出了菌株的注册号。这些菌株也可以从例如菌株保藏的保藏所获得。
- [0100] 微生物固有具有目标物质生产能力,或者可以已经以使它具有目标物质生产能力的方式进行了修饰。具有目标物质生产能力的微生物可以通过赋予如上所述的此类微生物以目标物质生产能力或增强如上提及的此类微生物的目标物质生产能力来获得。
- [0101] 以下,说明赋予或提高目标物质生成能力的方法的具体实例。为了赋予或提高目

标物质生产能力,下面举例说明的此类修饰可以独立使用,或者以适当组合使用。

[0102] 目标物质可以通过参与目标物质生物合成的酶的作用产生。此类酶也可以称为“目标物质生物合成酶”。因此,微生物可以具有目标物质生物合成酶。换言之,微生物可以具有编码目标物质生物合成酶的基因。此类基因也可以称为“目标物质生物合成基因”。微生物可以固有具有目标物质生物合成基因,或者可以已经引入了目标物质生物合成基因。本文解释了用于引入基因的方法。

[0103] 另外,通过提高目标物质生物合成酶的活性,可以改善微生物的目标物质生产能力。也就是说,赋予或提高目标物质生成能力的方法的实例可以包括提高目标物质生物合成酶的活性的方法。也就是说,微生物可以经修饰而使得增加了目标物质生物合成酶的活性。一种目标物质生物合成酶的活性可以得以增加,或者两种或更多种目标物质生物合成酶的活性可以得以增加。本文描述了用于增加蛋白质(诸如酶等)的活性的方法。通过例如增加编码蛋白质的基因的表达,可以提高蛋白质(诸如酶等)的活性。

[0104] 目标物质可以从例如碳源和/或目标物质的前体产生。因此,目标物质生物合成酶的实例可以包括例如催化从碳源和/或前体向目标物质转化的酶。例如,3-脱氢莽草酸可以通过部分莽草酸途径产生,其可以包括由3-脱氧-D-阿拉伯糖-庚酮糖酸7-磷酸合酶(DAHP合酶),3-脱氢奎尼酸合酶和3-脱氢奎尼酸脱水酶催化的步骤;可以通过3-脱氢莽草酸脱水酶(DHSD)的作用将3-脱氢莽草酸转化为原儿茶酸;可以通过O-甲基转移酶(OMT)或芳香醛氧化还原酶,诸如芳香羧酸还原酶,ACAR的作用将原儿茶酸分别转化为香草酸或原儿茶醛;以及可以通过ACAR或OMT的作用将香草酸或原儿茶醛转化为香草醛。也就是说,目标物质生物合成酶的具体实例可以包括例如DAHP合酶,3-脱氢奎尼酸合酶,3-脱氢奎尼酸脱水酶,DHSD,OMT和ACAR。

[0105] 术语“3-脱氧-D-阿拉伯糖-庚酮糖酸7-磷酸合酶(DAHP合酶)”可以指具有催化将D-赤藓糖-4-磷酸和磷酸烯醇丙酮酸转化成3-脱氧-D-阿拉伯糖-庚酮糖酸7-磷酸(DAHP)和磷酸的反应的活性的蛋白质(EC 2.5.1.54)。编码DAHP合酶的基因也可以称为“DAHP合酶基因”。DAHP合酶的实例可以包括分别由aroF、aroG和aroH基因编码的AroF、AroG和AroH蛋白。在这些中,AroG可以作为主要的DAHP合酶发挥功能。DAHP合酶诸如AroF、AroG和AroH蛋白的实例包括各种生物体如肠杆菌科(Enterobacteriaceae)细菌和棒状杆菌细菌的那些。DAHP合酶的具体实例可以包括大肠杆菌天然的AroF、AroG和AroH蛋白。大肠杆菌K-12MG1655菌株天然的aroG基因的核苷酸序列如SEQ ID NO:1所示,并且由该基因编码的AroG蛋白的氨基酸序列如SEQ ID NO:2所示。

[0106] DAHP合酶活性可以通过例如将酶与底物,诸如D-赤藓糖-4-磷酸和磷酸烯醇丙酮酸温育并测量DAHP的酶和底物依赖性生成来测量。

[0107] 术语“3-脱氢奎尼酸合酶”可以指具有催化使DAHP脱磷酸化以产生3-脱氢奎宁酸的反应的活性的蛋白质(EC 4.2.3.4)。编码3-脱氢奎尼酸合酶的基因也可以称为“3-脱氢奎尼酸合酶基因”。3-脱氢奎尼酸合酶的实例可以包括由aroB基因编码的AroB蛋白。3-脱氢奎尼酸合酶如AroB蛋白的实例可以包括各种生物体如肠杆菌科细菌和棒状杆菌细菌天然的那些。3-脱氢奎尼酸合酶的具体实例可以包括大肠杆菌天然的AroB。大肠杆菌K-12MG1655菌株天然的aroB基因的核苷酸序列如SEQ ID NO:3所示,并且由该基因编码的AroB蛋白的氨基酸序列如SEQ ID NO:4所示。

[0108] 3-脱氢奎尼酸合酶活性可以通过例如将酶与底物,诸如DAHP温育并测量3-脱氢奎宁酸的酶和底物依赖性产生来测量。

[0109] 术语“3-脱氢奎尼酸脱水酶”可以指具有催化使3-脱氢奎宁酸脱水以生成3-脱氢莽草酸的反应的活性的蛋白质(EC 4.2.1.10)。编码3-脱氢奎尼酸脱水酶的基因也可以称为“3-脱氢奎尼酸脱水酶基因”。3-脱氢奎尼酸脱水酶的实例可以包括由aroD基因编码的AroD蛋白。3-脱氢奎尼酸脱水酶诸如AroD蛋白的实例可以包括各种生物体如肠杆菌科细菌和棒杆菌的那些。3-脱氢奎尼酸脱水酶的具体实例可以包括大肠杆菌天然的AroD蛋白。大肠杆菌K-12MG1655菌株天然的aroD基因的核苷酸序列如SEQ ID NO:5所示,并且由该基因编码的AroD蛋白的氨基酸序列如SEQ ID NO:6所示。

[0110] 3-脱氢奎尼酸脱水酶活性可以通过例如将酶与底物,诸如3-脱氢奎宁酸温育并测量3-脱氢莽草酸的酶和底物依赖性产生来测量。

[0111] 术语“3-脱氢草酸脱水酶(DHSD)”可以指具有催化使3-脱氢莽草酸脱水以生成原儿茶酸的反应的活性的蛋白质(EC 4.2.1.118)。编码DHSD的基因也可以称为“DHSD基因”。DHSD的实例可以包括由asbF基因编码的AsbF蛋白。DHSD例如AsbF蛋白的实例包括各种生物体如苏云金芽孢杆菌(*Bacillus thuringiensis*),粗糙脉孢菌(*Neurospora crassa*)和*Podospora pauciseta*天然的那些。苏云金芽孢杆菌BMB171菌株天然的asbF基因的核苷酸序列如SEQ ID NO:7所示,并且由该基因编码的AsbF蛋白的氨基酸序列如SEQ ID NO:8所示。

[0112] DHSD活性可以通过例如将酶与底物,诸如3-脱氢莽草酸温育并测量原儿茶酸的酶和底物依赖性产生来测量。

[0113] 编码莽草酸途径的酶如DAHP合酶,3-脱氢奎尼酸合酶和3-脱氢奎尼酸脱水酶的基因的表达受到酪氨酸阻遏物TyrR阻遏,所述TyrR由tyrR基因编码。因此,莽草酸途径酶的活性也可以通过降低酪氨酸阻遏物TyrR的活性而增加。大肠杆菌K-12MG1655菌株天然的tyrR基因的核苷酸序列如SEQ ID NO:9所示,并且由该基因编码的TyrR蛋白的氨基酸序列如SEQ ID NO:10所示。

[0114] 术语“O-甲基转移酶(OMT)”可以指具有在甲基供体存在下催化物质的羟基的甲基化反应的活性的蛋白质(EC 2.1.1.68等)。此活性也可以称为“OMT活性”。编码OMT的基因也可以称为“OMT基因”。根据在如本文中所述的方法中生产目标物质的特定生物合成途径,OMT可以具有需要的底物特异性。例如,当经由原儿茶酸向香草酸的转化生产目标物质时,可以使用至少对原儿茶酸特异性的OMT。此外,例如,当经由原儿茶醛向香草醛的转化生产目标物质时,可以使用至少对原儿茶醛特异性的OMT。也就是说,具体地,术语“O-甲基转移酶(OMT)”可以指具有催化在存在甲基基团供体的情况下使原儿茶酸和/或原儿茶醛甲基化(即间位上的羟基基团的甲基化)以产生香草酸和/或香草醛的反应的活性的蛋白质。OMT可以对作为底物的原儿茶酸和原儿茶醛两者是特异性的,但不一定限于此。也就是说,可以采用具有所需底物特异性的OMT,这取决于在本发明的方法中生产目标物质的生物合成途径的类型。OMT通常可以使用原儿茶酸和原儿茶醛两者作为底物,但是不必然限于此。甲基基团供体的实例可以包括S-腺苷甲硫氨酸(SAM)。OMT的实例可以包括各种生物体天然的OMT,诸如智人(*Homo sapiens*) (Hs)天然的OMT(GenBank登录号NP_000745和NP_009294)、拟南芥(*Arabidopsis thaliana*)天然的OMT(GenBank登录号NP_200227和NP_009294)、*Fragaria x*

ananassa天然的OMT (GenBank登录号AAF28353) 以及W02013/022881A1中示例的哺乳动物、植物和微生物的其他各种OMT。对于智人天然的OMT基因, 已知四种转录物变体和两种OMT同种型。这四种转录物变体(转录物变体1-4, GenBank登录号NM_000754.3, NM_001135161.1, NM_001135162.1和NM_007310.2) 的核苷酸序列显示为SEQ ID NO:11至14, 较长OMT同种型(MB-COMT, GenBank登录号NP_000745.1) 的氨基酸序列如SEQ ID NO:15所示, 并且较短OMT同种型(S-COMT, GenBank登录号NP_009294.1) 的氨基酸序列如SEQ ID NO:16所示。SEQ ID NO:16对应于截短了N端的50个氨基酸残基的SEQ ID NO:15。

[0115] OMT还可以催化原儿茶酸和/或原儿茶醛的甲基化反应以产生异香草酸和/或异香草醛(即对位羟基的甲基化) 作为副反应。OMT可以选择性催化间位羟基的甲基化的。表述“选择性催化间位羟基的甲基化”可以意味着OMT选择性地从原儿茶酸生成香草酸的OMT和/或选择性地从原儿茶醛生成香草醛的OMT。表述“从原儿茶酸选择性产生香草酸”可以表示当允许OMT作用于原儿茶酸时按摩尔比计, OMT产生香草酸的量为异香草酸的量的例如3倍或更多, 5倍或更多, 10倍或更多, 15倍或更多, 20倍或更多, 25倍或更多, 或30倍或更多。此外, 表述“从原儿茶醛选择性产生香草酸”可以表示当允许OMT作用于原儿茶醛时按摩尔比计, OMT产生香草醛的量为异香草酸的量的例如3倍或更多, 5倍或更多, 10倍或更多, 15倍或更多, 20倍或更多, 25倍或更多, 或者30倍或更多。选择性催化间位羟基甲基化的OMT的实例可以包括本文所述的具有“特定突变”的OMT。

[0116] 具有“特定突变”的OMT也可以称为“突变体OMT”。编码突变体OMT的基因也可以称为“突变体OMT基因”。

[0117] 不具有“特定突变”的OMT也可以称为“野生型OMT”。编码野生型OMT的基因也可以称为“野生型OMT基因”。本文中提及的术语“野生型”是为了方便将“野生型”OMT与“突变体”OMT区分开而使用, 并且“野生型”OMT不限于作为天然物质获得的那些, 并且可以包括任何没有“特定突变”的OMT。野生型OMT的实例可以包括例如上面例示的OMT。另外, 上述例示的所有OMT保守变体应当认为包括在野生型OMT中, 条件是此类保守变体不具有“特定突变”。

[0118] “特定突变”的实例可以包括W02013/022881A1中描述的突变体OMT中包含的突变。也就是说, “特定突变”的实例可以包括野生型OMT的198位亮氨酸残基(L198) 被比亮氨酸残基的疏水性指数(hydrophobic index) 具有更低疏水性指数(亲水性指数(hydropathy index)) 的氨基酸残基取代的突变, 以及野生型OMT的199位谷氨酸残基(E199) 被在pH7.4具有中性或正侧链电荷的氨基酸残基取代的突变。突变体OMT可以具有这些突变中的一个或两者。

[0119] “比亮氨酸残基的疏水性指数具有更低疏水性指数(亲水指数) 的氨基酸残基”的实例可以包括Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Lys, Met, Phe, Pro, Ser, Thr, Trp和Tyr。作为“比亮氨酸残基的疏水性指数具有更低疏水性指数(亲水指数) 的氨基酸残基”, 特别地优选选自Ala, Arg, Asn, Asp, Glu, Gln, Gly, His, Lys, Met, Pro, Ser, Thr, Trp和Tyr的氨基酸残基, 并且Tyr是更具体的实例。

[0120] “在pH7.4具有中性或正侧链电荷的氨基酸残基”可以包括Ala, Arg, Asn, Cys, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr和Val。作为“在pH7.4具有中性或正侧链电荷的氨基酸残基”, Ala和Gln是具体的实例。

[0121] 任意的野生型OMT中的“L198”和“E199”可以指“与以SEQ ID NO:16所示的氨基酸

序列的第198位的亮氨酸残基对应的氨基酸残基”和“与SEQ ID NO:16所示的氨基酸序列的第199位谷氨酸残基对应的氨基酸残基”。这些氨基酸残基的位置表示相对位置,并且其绝对位置可以由于氨基酸残基的缺失,插入,添加等而改变。例如,若在SEQ ID NO:16所示的氨基酸序列中的位置X的N端侧的位置缺失或插入一个氨基酸残基,则原来在位置X处的氨基酸残基重新定位在位置X-1或X+1,但仍被认为是“与SEQ ID NO:16所示的氨基酸序列的X位的氨基酸残基对应的氨基酸残基”。此外,虽然“L198”和“E199”通常分别是亮氨酸残基和谷氨酸残基,但它们可以不是亮氨酸残基和谷氨酸残基。也就是说,当“L198”和“E199”分别不是亮氨酸残基和谷氨酸残基时,“特定突变”可以包括那些氨基酸残基各自被任何上述氨基酸残基替换的突变。

[0122] 在任意OMT的氨基酸序列中,哪个氨基酸残基为对应于“L198”或“E199”的氨基酸残基可以通过将任意OMT的氨基酸序列和SEQ ID NO:16的氨基酸序列比对确定。比对可以通过例如使用已知的基因分析软件来进行。此类软件的具体实例可以包括由Hitachi Solutions生产的DNASIS,Genetyx生产的GENETYX等(Elizabeth C.Tyler et al., Computers and Biomedical Research, 24(1) 72-96, 1991; Barton GJ et al., Journal of Molecular Biology, 198(2), 327-37, 1987)。

[0123] 例如,可以通过修饰野生型OMT基因使得由其编码的OMT具有“特定突变”来获得突变体OMT基因。待修饰的野生型OMT基因可以通过例如从具有野生型OMT基因的生物体的克隆或化学合成来获得。此外,也可以不使用野生型OMT基因而获得突变体OMT基因。例如,可以通过化学合成直接获得突变体OMT基因。获得的突变OMT基因可以原样使用,或者在使用前进一步修饰。

[0124] 基因可以使用已知的方法进行修饰。例如,可以通过位点特异性诱变方法将目标突变引入DNA的靶位点中。位点特异性诱变方法的实例可以包括使用PCR的方法(Higuchi, R., 61, 于PCR Technology, Erlich, H.A. 编, Stockton Press (1989); Carter P., Meth. In Enzymol., 154, 382 (1987)); 以及使用噬菌体的方法(Kramer, W. and Frits, H.J., Meth. in Enzymol., 154, 350 (1987); Kunkel, T.A. et al., Meth. in Enzymol., 154, 367 (1987))。

[0125] 可以通过例如在SAM存在下将酶与底物,诸如原儿茶酸或原儿茶醛温育并测量相应产物,诸如香草酸或香草醛的酶和底物依赖性产生来测量OMT活性(WO2013/022881A1)。此外,通过在相同条件下测量相应副产物,诸如异香草酸或异香草醛的产生,并将副产物的产生与产物的产生进行比较,可以确定OMT是否选择性地生成产物。

[0126] 术语“芳族醛氧化还原酶(芳族羧酸还原酶;ACAR)”可以指具有催化在电子供体和ATP存在下还原香草酸和/或原儿茶酸的反应以产生香草醛和/或原儿茶醛的活性的蛋白质(EC 1.2.99.6等)。此活性也可以称为“ACAR活性”。编码ACAR的基因也可称为“ACAR基因”。ACAR通常可以使用香草酸和原儿茶酸两者作为底物,但不必限于此。也就是说,ACAR可以具有需要的底物特异性,这取决于在如本文所述的方法中生产目标物质的特定生物合成途径。例如,当通过香草酸向香草醛的转化来生产目标物质时,可以使用至少对香草酸特异性的ACAR。另外,例如,当通过原儿茶酸向原儿茶醛的转化来生产目标物质时,可以使用至少对原儿茶酸特异性的ACAR。电子供体的实例可以包括NADH和NADPH。ACAR的实例可以包括各种生物体如诺卡菌属种(*Nocardia* sp.)菌株NRRL 5646、放线菌属种(*Actinomyces* sp.)、热乙酸梭菌(*Clostridium thermoaceticum*)、黑曲霉(*Aspergillus niger*)、甜瓜棒

孢(*Corynespora melonis*)、云芝属种(*Coriolus* sp.)和链孢霉属种(*Neurospora* sp.)天然的ACAR(J.Biol.Chem.,2007,Vol.282,No.1,pp.478-485)。诺卡尔菌属种菌株NRRL 5646已被分类为艾阿华诺卡氏菌(*Nocardia iowensis*)。ACAR的实例还可包括其他诺卡氏菌细菌,例如巴西诺卡氏菌(*Nocardia brasiliensis*)和伤口诺卡氏菌(*Nocardia vulneris*)的ACAR。巴西诺卡氏菌ATCC 700358天然的ACAR基因的核苷酸序列以SEQ ID NO:17显示,并且由此基因编码的ACAR的氨基酸序列以SEQ ID NO:18显示。巴西诺卡氏菌ATCC 700358天然的变体ACAR基因的实例的核苷酸序列以SEQ ID NO:19显示,并且由此基因编码的ACAR的氨基酸序列以SEQ ID NO:20显示。

[0127] ACAR活性可以通过例如在ATP和NADPH存在下将酶与底物,诸如香草酸或原儿茶酸一起温育,并且测量NADPH的酶和底物依赖性氧化来测量(J.Biol.Chem.,2007,Vol.282,No.1,pp.478-485中描述的方法的修改)。

[0128] 可以通过磷酸泛酰巯基乙胺化(phosphopantetheinylation)将ACAR变为活性酶(J.Biol.Chem.,2007,Vol.282,No.1,pp.478-485)。因此,ACAR活性也可以通过增加催化蛋白质的磷酸泛酰巯基乙胺化的酶(其也可以称为“磷酸泛酰巯基乙胺基化酶”)的活性来增加。也就是说,赋予或提高目标物质生产能力的方法的实例可以包括增加磷酸泛酰巯基乙胺基化酶的活性的方法。也就是说,微生物可以以使磷酸泛酰巯基乙胺基化酶的活性得以增加的方式进行了修饰。磷酸泛酰巯基乙胺基化酶的实例可以包括磷酸泛酰巯基乙胺基转移酶(PPT)。

[0129] 术语“磷酸泛酰巯基乙胺基转移酶(PPT)”可以指具有在磷酸泛酰巯基乙胺基供体存在下催化使ACAR磷酸泛酰巯基乙胺化的反应的活性的蛋白质。此活性也可以称为“PPT活性”。编码PPT的基因也可以称为“PPT基因”。磷酸泛酰巯基乙胺基供体的实例可以包括辅酶A(CoA)。PPT的实例可以包括EntD蛋白,其由entD基因编码。PPT例如EntD蛋白质的实例可以包括各种生物体天然的那些。PPT的具体实例可以包括大肠杆菌天然的EntD蛋白。大肠杆菌K-12MG1655菌株的entD基因的核苷酸序列如SEQ ID NO:21所示,并且由该基因编码的EntD蛋白的氨基酸序列如SEQ ID NO:22所示。PPT的具体实例还可以包括巴西诺卡菌天然的PPT,皮疽诺卡菌(*Nocardia farcinica*) IFM10152天然的PPT(J.Biol.Chem.,2007,Vol.282,No.1,pp.478-485)和谷氨酸棒杆菌天然的PPT(App.Env.Microbiol.2009,Vol.75,No.9,pp.2765-2774)。谷氨酸棒杆菌ATCC 13032菌株天然的PPT基因的核苷酸序列如SEQ ID NO:23所示,并且由该基因编码的PPT的氨基酸序列如SEQ ID NO:24所示。

[0130] PPT活性可以基于例如在CoA存在下与ACAR温育时观察到的ACAR活性的增强来测量(J.Biol.Chem.,2007,Vol.282,No.1,pp.478-485)。

[0131] 褪黑激素可以从L-色氨酸生产。也就是说,目标物质生物合成酶的实例还可以包括例如L-色氨酸生物合成酶和催化L-色氨酸转化成褪黑激素的酶。L-色氨酸生物合成酶的实例可以包括芳香族氨基酸的常见生物合成酶,例如3-脱氧-D-阿拉伯糖庚酮糖酸-7-磷酸合酶(aroF,aroG,aroH),3-脱氢喹啉合酶(aroB),3-脱氢喹啉脱水酶(aroD),莽草酸脱氢酶(aroE),莽草酸激酶(aroK,aroL),5-烯醇丙酮酰莽草酸-3-磷酸合酶(aroA)和分支酸合酶(aroC);以及邻氨基苯甲酸合酶(trpED)和色氨酸合酶(trpAB)。在酶的名称之后的括号中显示了编码酶的基因的名称的实例(这在下文也适用于相同的情况)。通过色氨酸5-羟化酶(EC 1.14.16.4),5-羟色氨酸脱羧酶(EC 4.1.1.28),芳烷基胺N-乙酰转移酶(AANAT;EC

2.3.1.87) 和乙酰血清素O-甲基转移酶 (EC 2.1.1.4) 的作用,L-色氨酸可以依次转化为羟色氨酸,5-羟色胺,N-乙酰血清素和褪黑激素。也就是说,催化L-色氨酸转化为褪黑激素的酶的实例可以包括这些酶。值得注意的是,乙酰血清素O-甲基转移酶是OMT的一个实例,其催化使用SAM作为甲基供体使N-乙酰血清素甲基化以产生褪黑激素的反应。

[0132] 麦角硫因可以从L-组氨酸生产。也就是说,目标物质生物合成酶的实例还可以包括例如L-组氨酸生物合成酶和催化L-组氨酸转化为麦角硫因的酶。L-组氨酸生物合成酶的实例可以包括ATP磷酸核糖基转移酶 (hisG),磷酸核糖基AMP环化水解酶 (hisI),磷酸核糖基-ATP焦磷酸水解酶 (hisI),磷酸核糖基亚胺甲基-5-氨基咪唑甲酰胺核苷酸异构酶 (phosphoribosylformimino-5-aminoimidazole carboxamide ribotide isomerase) (hisA),氨基转移酶 (hisH),磷酸组氨醇氨基转移酶 (hisC),组氨醇磷酸酶 (hisB) 和组氨醇脱氢酶 (hisD)。L-组氨酸可以通过分别由egtB、egtC、egtD和egtE基因编码的EgtB、EgtC、EgtD和EgtE蛋白的作用连续转化为组氨酸三甲基内盐 (hercynine),hercynyl- γ -L-谷氨酰-L-半胱氨酸亚砷,hercynyl-L-半胱氨酸亚砷和麦角硫因。组氨酸三甲基内盐还可以通过由egt1基因编码的Egt1蛋白的作用转化为hercynyl-L-半胱氨酸亚砷。也就是说,催化L-组氨酸转化为麦角硫因的酶的实例可以包括这些酶。值得注意的是,EgtD是S-腺苷-1-甲硫氨酸 (SAM) 依赖性组氨酸N,N,N-甲基转移酶,其催化使用SAM作为甲基供体使组氨酸甲基化以产生组氨酸三甲基内盐的反应。

[0133] 愈创木酚可以从香草酸生产。因此,上述关于香草酸的目标物质生物合成酶的描述可以在必要修饰后适用于愈创木酚的目标物质生物合成酶。香草酸可以通过香草酸脱羧酶 (VDC) 的作用转化为愈创木酚。也就是说,目标物质生物合成酶的实例还可以包括VDC。

[0134] 阿魏酸、4-乙烯基愈创木酚和4-乙基愈创木酚可以从L-苯丙氨酸或L-酪氨酸生产。也就是说,目标物质生物合成酶的实例还可以包括例如L-苯丙氨酸生物合成酶,L-酪氨酸生物合成酶,以及催化L-苯丙氨酸或L-酪氨酸转化为阿魏酸,4-乙烯基愈创木酚或4-乙基愈创木酚的酶。L-苯丙氨酸生物合成酶的实例可包括上文例举的芳香族氨基酸的常见生物合成酶,以及分支酸变位酶 (pheA),预苯酸脱水酶 (pheA) 和酪氨酸氨基转移酶 (tyrB)。分支酸变位酶和预苯酸脱水酶可以由pheA基因编码为双功能酶。L-酪氨酸生物合成酶的实例可以包括上文例举的芳香族氨基酸的常见生物合成酶,以及分支酸变位酶 (tyrA),预苯酸脱氢酶 (tyrA) 和酪氨酸氨基转移酶 (tyrB)。分支酸变位酶和预苯酸脱氢酶可以由tyrA基因编码为双功能酶。L-苯丙氨酸可以通过苯丙氨酸脱氢酶 (PAL;EC 4.3.1.24) 的作用转化为肉桂酸,然后通过肉桂酸4-羟化酶 (C4H;EC 1.14.13.11) 的作用转化为对香豆酸。此外,L-酪氨酸可通过酪氨酸脱氢酶 (TAL;EC 4.3.1.23) 的作用转化为对香豆酸。对香豆酸可以分别通过羟基肉桂酸3-羟化酶 (C3H),O-甲基转移酶 (OMT),阿魏酸脱羧酶 (FDC) 和乙烯基苯酚还原酶 (VPR) 的作用连续转化为咖啡酸,阿魏酸,4-乙烯基愈创木酚和4-乙基愈创木酚。也就是说,催化L-苯丙氨酸或L-酪氨酸转化为阿魏酸,4-乙烯基愈创木酚或4-乙基愈创木酚的酶的实例可以包括这些酶。为了生产阿魏酸,4-乙烯基愈创木酚或4-乙基愈创木酚,可以使用至少使用咖啡酸的OMT。

[0135] 多胺可以从L-精氨酸或L-鸟氨酸生产。也就是说,目标物质生物合成酶的实例还可包括,例如,L-精氨酸生物合成酶,L-鸟氨酸生物合成酶,以及催化L-精氨酸或L-鸟氨酸转化为多胺的酶。L-鸟氨酸生物合成酶的实例可包括N-乙酰谷氨酸合酶 (argA),N-乙酰谷

氨酸激酶 (argB), N-乙酰谷氨酰磷酸还原酶 (argC), 乙酰鸟氨酸转氨酶 (argD) 和乙酰鸟氨酸脱乙酰酶 (argE)。L-精氨酸生物合成酶的实例可包括上文例举的L-鸟氨酸生物合成酶, 以及氨甲酰磷酸合成酶 (carAB), 鸟氨酸氨甲酰转移酶 (argF, argI), 精氨酸琥珀酸合成酶 (argG), 精氨酸琥珀酸裂合酶 (argH)。L-精氨酸可以通过精氨酸脱羧酶 (speA; EC 4.1.1.19) 的作用转化为胍丁胺, 然后通过胍丁胺尿素水解酶 (speB; EC 3.5.3.11) 的作用转化为腐胺。此外, L-鸟氨酸可以通过鸟氨酸脱羧酶 (speC; EC 4.1.1.17) 的作用转化为腐胺。腐胺可以通过亚精胺合酶 (speE; EC 2.5.1.16) 的作用转化为亚精胺, 然后通过精胺合成酶 (EC 2.5.1.22) 的作用转化为精胺。胍丁胺也可以通过胍丁胺/三胺氨基丙基转移酶的作用转化为氨基丙基胍丁胺, 然后通过氨基丙基胍丁胺尿素水解酶的作用转化为亚精胺。也就是说, 催化L-精氨酸或L-鸟氨酸转化为多胺的酶的实例可以包括这些酶。值得注意的是, 亚精胺合酶, 精胺合成酶和胍丁胺/三胺氨基丙基转移酶各自催化将来自脱羧的S-腺苷甲硫氨酸 (dcSAM) (其可以通过脱羧作用从SAM产生) 的丙胺基团转移到相应的底物中的反应。

[0136] 肌酸可以从L-精氨酸和甘氨酸生产。也就是说, 目标物质生物合成酶的实例还可以包括例如L-精氨酸生物合成酶, 甘氨酸生物合成酶和催化L-精氨酸和甘氨酸转化成肌酸的酶。可以将L-精氨酸和甘氨酸组合以通过精氨酸:甘氨酸脒基转移酶 (AGAT, EC 2.1.4.1) 的作用产生胍基乙酸 (guanidinoacetate) 和鸟氨酸; 使用SAM作为甲基供体, 可以通过胍基乙酸N-甲基转移酶 (GAMT, EC 2.1.1.2) 的作用使胍基乙酸甲基化以生成肌酸。也就是说, 催化L-精氨酸和甘氨酸转化为肌酸的酶的实例可包括这些酶。

[0137] 麦根酸可以从SAM生产。也就是说, 目标物质生物合成酶的实例还可以包括例如催化SAM转化为麦根酸的酶。可以通过烟草胺 (nicotianamine) 合酶 (EC 2.5.1.43) 的作用从三个SAM分子合成一个烟草胺分子。烟草胺可以通过烟草胺转氨酶 (EC 2.6.1.80)、3''-脱氨基-3''-氧烟草胺 (oxonicotianamine) 还原酶 (EC 1.1.1.285) 和2'-脱氧麦根酸2'-双加氧酶 (EC 1.14.11.24) 的作用连续转化为3''-脱氨基-3''-氧烟草胺、2'-脱氧麦根酸和麦根酸。也就是说, 催化SAM转化为麦根酸的酶的实例可以包括这些酶。

[0138] L-甲硫氨酸可以从L-半胱氨酸生产。也就是说, 目标物质生物合成酶的实例还可包括, 例如, L-半胱氨酸生物合成酶和催化L-半胱氨酸转化为L-甲硫氨酸的酶。L-半胱氨酸生物合成酶的实例可包括CysIXHDNYZ蛋白、Fpr2蛋白和CysK蛋白, 它们分别由cysIXHDNYZ基因、fpr2基因和cysK基因编码。催化L-半胱氨酸转化为L-甲硫氨酸的酶的实例可包括胱硫醚- γ -合酶和胱硫醚- β -裂合酶。

[0139] 用于赋予或提高目标物质生产能力的方法的实例还可以包括提高目标物质以外物质, 例如在生产目标物质期间作为中间产物生成的物质和用作目标物质前体的物质的摄取系统的活性的方法。也就是说, 本发明的细菌可以以使此类摄取系统的活性得以增加的方式进行了修饰。术语“物质的摄取系统”可以指具有将来自细胞外部的物质掺入细胞的功能的蛋白质。此活性也可以称为“物质的摄取活性”。编码此类摄取系统的基因也可以称为“摄取系统基因”。此类摄取系统的实例可以包括香草酸摄取系统和原儿茶酸摄取系统。香草酸摄取系统的实例可以包括由vanK基因编码的VanK蛋白 (M.T.Chaudhry, et al., Microbiology, 2007, 153:857-865)。谷氨酸棒杆菌ATCC13869菌株天然的vanK基因 (NCgl2302) 的核苷酸序列如SEQ ID NO:25所示, 并且由该基因编码的VanK蛋白的氨基酸序

列如SEQ ID NO:26所示。原儿茶酸摄取系统基因可以包括由pcaK基因编码的PcaK蛋白(M.T.Chaudhry, et al., Microbiology, 2007, 153:857-865)。谷氨酸棒杆菌ATCC13869菌株天然的pcaK基因(NCgl1031)的核苷酸序列如SEQ ID NO:27所示,并且由该基因编码的PcaK蛋白的氨基酸序列如SEQ ID NO:28所示。

[0140] 可以根据例如已知方法(M.T.Chaudhry, et al., Microbiology, 2007, 153:857-865)测量物质的摄取活性。

[0141] 用于赋予或提高目标物质生产能力的方法的实例还可以包括降低参与目标物质以外的物质的副生成的酶的活性的方法。除了目标物质以外的此类物质也可以称为“副产物”。此类酶也可以称为“副产物生成酶”。副产物生成酶的实例可以包括例如参与目标物质利用的酶和催化远离目标物质的生物合成途径分支的反应以产生除目标物质之外的物质的酶。本文描述了降低蛋白质诸如酶等的活性的方法。通过例如破坏编码蛋白质的基因可以降低蛋白质诸如酶等的活性。例如,据报道,在棒状杆菌细菌中,香草醛按香草醛→香草酸→原儿茶酸的顺序代谢并利用(Current Microbiology, 2005, Vol. 51, pp. 59-65)。也就是说,副产物生成酶的具体实例可以包括催化香草醛转化为原儿茶酸的酶和催化原儿茶酸进一步代谢的酶。此类酶的实例可以包括香草酸脱甲基酶,原儿茶酸3,4-双加氧酶,以及将原儿茶酸3,4-双加氧酶的反应产物进一步分解为琥珀酰-CoA和乙酰-CoA的各种酶(Appl. Microbiol. Biotechnol., 2012, Vol. 95, p77-89)。此外,可以通过醇脱氢酶的作用将香草醛转化成香草醇(Kunjapur AM. et al., J. Am. Chem. Soc., 2014, Vol. 136, p11644-11654.; Hansen EH. et al., App. Env. Microbiol., 2009, Vol. 75, p2765-2774.)。也就是说,副产物生成酶的具体实例也可以包括醇脱氢酶(ADH)。此外,作为香草酸和香草醛生物合成途径中间体的3-脱氢莽草酸也可以通过莽草酸脱氢酶的作用转化为莽草酸。也就是说,副产物生成酶的具体实例还可以包括莽草酸脱氢酶。

[0142] 术语“香草酸脱甲基酶”可以指具有催化香草酸脱甲基化生成原儿茶酸的反应的活性的蛋白质。此活性也可以称为“香草酸脱甲基酶活性”。编码香草酸脱甲基酶的基因也可以称为“香草酸脱甲基酶基因”。香草酸脱甲基酶的实例可以包括由vanAB基因编码的VanAB蛋白(Current Microbiology, 2005, Vol. 51, pp. 59-65)。vanA基因和vanB基因分别编码香草酸脱甲基酶的亚基A和亚基B。为了降低香草酸脱甲基酶活性,这两种vanAB基因可以被破坏等,或者两者中仅一种可以被破坏等。谷氨酸棒杆菌ATCC13869天然菌株的vanAB基因的核苷酸序列分别如SEQ ID NO:29和31所示,并且由这些基因编码的VanAB蛋白的氨基酸序列分别如SEQ ID NO:30和32所示。vanAB基因通常与vanK基因一起构成vanABK操纵子。因此,为了降低香草酸脱甲基酶活性,vanABK操纵子可以被完全破坏等(例如缺失)。在此类情况下,可以将vanK基因再引入宿主。例如,当使用存在于细胞外的香草酸,并且vanABK操纵子完全被破坏等(例如缺失)时,优选再引入vanK基因。

[0143] 香草酸脱甲基酶活性可以通过例如将酶与底物诸如香草酸温育并测量原儿茶酸的酶和底物依赖性产生来测量(J Bacteriol, 2001, Vol. 183, p3276-3281)。

[0144] 术语“原儿茶酸3,4-双加氧酶”可以指具有催化氧化原儿茶酸反应生成 β -羧基-顺,顺-粘康酸的反应活性的蛋白质。此活性也可以称为“原儿茶酸3,4-双加氧酶活性”。编码原儿茶酸3,4-双加氧酶的基因也可以称为“原儿茶酸3,4-双加氧酶基因”。原儿茶酸3,4-双加氧酶的实例可以包括由pcaGH基因编码的PcaGH蛋白(Appl. Microbiol. Biotechnol.,

2012, Vol. 95, p77-89)。pcaG基因和pcaH基因分别编码原儿茶酸3,4-双加氧酶的 α 亚基和 β 亚基。为了降低原儿茶酸3,4-双加氧酶活性,这两种pcaGH基因可以被破坏等,或者两者中仅一种可以被破坏等。谷氨酸棒杆菌ATCC 13032菌株天然的pcaGH基因的核苷酸序列分别如SEQ ID NO:33和35所示,并且由这些基因编码的PcaGH蛋白的氨基酸序列分别如SEQ ID NO:34和36所示。

[0145] 原儿茶酸3,4-双加氧酶活性可以通过例如将酶与底物诸如原儿茶酸温育并测量酶和底物依赖性氧消耗量来测量(Meth.Enz., 1970, Vol. 17A, p526-529)。

[0146] 术语“醇脱氢酶(ADH)”可以指具有在电子供体存在下催化醛还原以产生醇的反应的活性的蛋白质(EC 1.1.1.1, EC 1.1.1.2, EC 1.1.1.71等)。此活性也可以称为“ADH活性”。编码ADH的基因也可以称为“ADH基因”。电子供体的实例可以包括NADH和NADPH。

[0147] 作为ADH,具有在电子供体存在下催化还原香草醛的反应以产生香草醇的活性的ADH是一个特别的例子。该活性也可以特别称为“香草醇脱氢酶活性”。此外,具有香草醇脱氢酶活性的ADH也可以特别称为“香草醇脱氢酶”。作为ADH,具有催化在电子供体存在下还原香草醛以产生香草醇的反应的活性的ADH是具体的实例。此活性也可以特别称为“香草醇脱氢酶活性”。此外,具有香草醇脱氢酶活性的ADH也可以特别称为“香草醇脱氢酶”。

[0148] ADH的实例可以包括分别由yqhD基因, NCg10324基因, NCg10313基因, NCg12709基因, NCg10219基因和NCg12382基因编码的YqhD蛋白, NCg10324蛋白质, NCg10313蛋白质, NCg12709蛋白质, NCg10219蛋白质和NCg12382蛋白质。yqhD基因和NCg10324基因编码香草醇脱氢酶。yqhD基因可以在例如属于肠杆菌科的细菌, 诸如大肠杆菌中找到。NCg10324基因, NCg10313基因, NCg12709基因, NCg10219基因和NCg12382基因可以例如在棒状杆菌细菌例如谷氨酸棒杆菌中找到。大肠杆菌K-12MG1655菌株天然的yqhD基因的核苷酸序列如SEQ ID NO:37所示, 并且由此基因编码的YqhD蛋白的氨基酸序列以SEQ ID NO:38所示。谷氨酸棒杆菌ATCC13869菌株天然的NCg10324基因, NCg10313基因和NCg12709基因的核苷酸序列分别如SEQ ID NO:39, 41和43所示, 并且由这些基因编码的蛋白质的氨基酸序列分别如SEQ ID NO:40, 42和44所示。谷氨酸棒杆菌ATCC 13032菌株天然的NCg10219基因和NCg12382基因的核苷酸序列分别如SEQ ID NO:45和47所示, 并且由这些基因编码的蛋白质的氨基酸序列如SEQ ID NO:46和48所示。可以降低一种ADH的活性, 或者可以降低两种或更多种ADH的活性。例如, 可以降低诸如NCg10324蛋白质, NCg12709蛋白质和NCg10313蛋白质中的一种或多种的活性。特别地, 可以至少降低NCg10324蛋白质的活性。

[0149] ADH活性可以通过例如在NADPH或NADH存在下将酶与底物(诸如醛, 如香草醛)一起温育, 并测量NADPH或NADH的酶和底物依赖性氧化来测量。

[0150] 术语“莽草酸脱氢酶”可以指具有催化在电子供体存在下还原3-脱氢莽草酸以产生莽草酸的反应的活性的蛋白质(EC 1.1.1.25)。该活性也可称为“莽草酸脱氢酶活性”。编码莽草酸脱氢酶的基因也可称为“莽草酸脱氢酶基因”。电子供体的实例可包括NADH和NADPH。莽草酸脱氢酶的实例可包括AroE蛋白, 其由aroE基因编码。大肠杆菌K-12MG1655菌株天然的aroE基因的核苷酸序列如SEQ ID NO:49所示, 并且由该基因编码的AroE蛋白的氨基酸序列如SEQ ID NO:50所示。

[0151] 莽草酸脱氢酶活性可以通过例如在NADPH或NADH存在下将酶与底物诸如3-脱氢莽草酸一起温育, 并测量NADPH或NADH的酶和底物依赖性氧化来测量。

[0152] 待修饰活性的蛋白质可以根据产生目标物质的生物合成途径的类型以及选定的微生物中固有存在的蛋白质的类型和活性来适当选择。例如,当通过原儿茶酸的生物转化生产香草醛时,可优选增强OMT,ACAR,PPT和原儿茶酸摄取系统中一种或多种的活性。此外,当通过原儿茶醛生物转化生产香草醛时,可以优选增强OMT的活性。

[0153] 用于育种出具有目标物质生产能力的微生物的基因和蛋白质可以分别具有例如上面列举的或其他已知的核苷酸序列和氨基酸序列。此外,用于育种出具有目标物质生产能力的微生物的基因和蛋白质可以是上面例示的基因和蛋白质的保守变体,例如分别具有上述例示或其他已知核苷酸序列和氨基酸序列的基因和蛋白质。具体地,例如,用于育种出具有目标物质生产能力的微生物的基因可以各自是编码蛋白质的基因,所述蛋白质具有上述例示的氨基酸序列或已知蛋白质的氨基酸序列但可以包含一个或几个位置处一个或几个氨基酸残基的取代、缺失、插入和/或添加,只要保持蛋白质的原始功能,即其酶活性,转运蛋白活性等即可。对于基因和蛋白质的保守变体,可以在必要修改后应用后面描述的关于NCg12048基因和NCg12048蛋白的保守变体的描述。

[0154] <1-2>NCg12048蛋白活性的降低

[0155] 微生物可以以NCg12048蛋白的活性得以降低的方式进行修饰。具体地,微生物可以以与非修饰的菌株相比NCg12048蛋白的活性得以降低的方式进行修饰。通过以NCg12048蛋白的活性得以降低的方式修饰微生物,微生物的目标物质生产能力可以得到改善,并且也就是说,可以提高通过使用微生物生产目标物质。此外,通过以NCg12048蛋白的活性得以降低的方式修饰微生物,微生物生成或再生SAM的能力可能得到改善。也就是说,具体地,微生物的目标物质生产能力的增加可以是由于微生物生成或再生SAM的能力增加所致。

[0156] 可以通过以NCg12048蛋白的活性得以降低的方式修饰具有目标物质生产能力的微生物来获得微生物。微生物也可以通过以NCg12048蛋白的活性得以降低的方式修饰微生物,然后赋予微生物以目标物质生产能力或增强微生物的目标物质生产能力来获得。另外,微生物可以由于降低NCg12048蛋白活性的修饰,或者由于降低NCg12048蛋白活性的修饰和其他用于赋予或增强目标物质生产能力的修饰的组合而获得目标物质生产能力。构建微生物的修饰可以以任意顺序进行。

[0157] 术语“NCg12048蛋白”可以指由NCg12048基因编码的蛋白质。NCg12048蛋白的实例可以包括各种生物体,例如肠杆菌科细菌和棒状杆菌细菌天然的那些蛋白。NCg12048蛋白的具体实例可以包括谷氨酸棒杆菌天然的NCg12048蛋白。谷氨酸棒杆菌ATCC 13869菌株天然的NCg12048基因的核苷酸序列以SEQ ID NO:92所示,并且由此基因编码的蛋白质的氨基酸序列以SEQ ID NO:93所示。

[0158] 也就是说,NCg12048基因可以是例如具有以SEQ ID NO:92所示的核苷酸序列的基因。此外,NCg12048蛋白可以是例如具有以SEQ ID NO:93所示的氨基酸序列的蛋白质。表述“基因或蛋白质具有核苷酸或氨基酸序列”涵盖当基因或蛋白质包含该核苷酸或氨基酸序列时,以及当基因或蛋白质仅包括核苷酸或氨基酸序列时。

[0159] NCg12048基因可以是上面例示的任何NCg12048基因的变体,即具有以SEQ ID NO:92所示的核苷酸序列的基因,只要其原始功能得以保持即可。类似地,NCg12048蛋白可以是上文例示的任何NCg12048蛋白的变体,即具有SEQ ID NO:93所示的氨基酸序列的蛋白质,只要其原始功能得以保持即可。保持其原始功能的变体也可以称为“保守变体”。术语

“NCg12048基因”不仅可以包括上文例示的NCg12048基因,诸如具有以SEQ ID NO:92所示的核苷酸序列的NCg12048基因,而且还可以包括其保守变体。类似地,术语“NCg12048蛋白”不仅可以包括上文例示的NCg12048蛋白,诸如具有以SEQ ID NO:93所示的氨基酸序列的NCg12048蛋白,而且还可以包括其保守变体。保守变体的实例可以包括例如上文例示的基因和蛋白质的同源物和人工修饰形式。

[0160] 表述“保持原始功能”指基因或蛋白质的变体具有对应于原始基因或蛋白质的功能(例如活性或性质)的功能(例如活性或性质)。当提到基因时,表述“保持原始功能”指该基因的变体编码保持原始功能的蛋白质。也就是说,当提及NCg12048基因时,表述“保持原始功能”指基因的变体编码具有NCg12048蛋白功能的蛋白质,诸如由以SEQ ID NO:93所示的氨基酸序列组成的蛋白质的功能。当提及NCg12048基因时,表述“保持原始功能”还可以指基因的变体具有微生物中的基因表达降低提供目标物质的生成增加的特性。当提及NCg12048蛋白时,表述“保持原始功能”指蛋白质的变体具有NCg12048蛋白的功能,诸如具有以SEQ ID NO:93所示的氨基酸序列的蛋白质的功能。当提及NCg12048蛋白时,表述“保持原始功能”还可以指蛋白质的变体具有微生物中的蛋白质活性降低提供目标物质的生成增加的特性。

[0161] 以下将解释保守变体的实例。

[0162] NCg12048基因或NCg12048蛋白的同源物可以使用上述例示的NCg12048基因的任何核苷酸序列或上文例示的NCg12048蛋白的任何氨基酸序列作为查询序列通过例如BLAST搜索或FASTA搜索从公共数据库容易地获得。此外,NCg12048基因的同源物可以通过例如使用棒状杆菌细菌等生物体的染色体作为模板和基于上述例示的NCg12048基因的任何核苷酸序列制备的寡核苷酸作为引物的PCR获得。

[0163] NCg12048基因可以编码蛋白质,所述蛋白质具有任何上述氨基酸序列如以SEQ ID NO:93所示的氨基酸序列,且包括在一个或几个位置处一个或几个氨基酸残基的取代,缺失,插入和/或添加,只要保持原始功能即可。例如,编码的蛋白质可以具有延伸的或缺失的N端和/或C端。尽管以上使用的术语“一个或几个”所表示的数目可以随蛋白质的三维结构中的氨基酸残基的位置或氨基酸残基的类型而不同,具体而言,它例如为1至50,1至40,1至30,1至20,1至10,1至5,或1至3。

[0164] 一个或几个氨基酸残基的上述取代,缺失,插入和/或添加可以各自是保持蛋白质原始功能的保守突变。保守突变的典型实例是保守取代。若取代位点是芳香氨基酸,则保守取代是其中取代在Phe,Trp和Tyr之间相互发生的突变;若它是疏水性氨基酸,则保守取代是取代在Leu,Ile和Val之间相互发生的突变;若它是极性氨基酸;则保守取代是取代在Gln和Asn之间相互发生的突变;若它是碱性氨基酸,则保守取代是取代在Lys,Arg和His之间相互发生的突变;若它是酸性氨基酸,则保守取代是取代在Asp和Glu之间相互发生的突变;以及若它是具有羟基的氨基酸,则保守取代是取代在Ser和Thr之间相互发生的突变。认为是保守取代的取代的实例可以具体包括用Ser或Thr取代Ala,用Gln,His,或Lys取代Arg,用Glu,Gln,Lys,His,或Asp取代Asn,用Asn,Glu,或Gln取代Asp,用Ser或Ala取代Cys,用Asn,Glu,Lys,His,Asp,或Arg取代Gln,用Gly,Asn,Gln,Lys,或Asp取代Glu,用Pro取代Gly,用Asn,Lys,Gln,Arg,或Tyr取代His,用Leu,Met,Val,或Phe取代Ile,用Ile,Met,Val,或用Phe取代Leu,用Asn,Glu,Gln,His,或Arg取代Lys,用Ile,Leu,Val,或Phe取代Met,用Trp,Tyr,

Met, Ile, 或Leu取代Phe, 用Thr或Ala取代Ser, 用Ser或Ala取代Thr, 用Phe或Tyr取代Trp, 用His, Phe, 或Trp取代Tyr, 以及用Met, Ile, 或Leu取代Val。此外, 如上所述的氨基酸残基的此类取代, 缺失, 插入, 添加等可以包括由于个体差异引起的天然存在的突变或者该基因来源的生物体的物种的差异(突变体或变体)。

[0165] 此外, NCg12048基因可以是编码与上述任何氨基酸序列的总氨基酸序列具有例如50%或更多, 65%或更多, 80%或更多, 90%或更多, 95%或更多, 97%或更多, 或99%或更多的同源性的氨基酸序列的蛋白质的基因, 只要保持原始功能。另外, 在本说明书中, “同源性”等同于“同一性”。

[0166] 此外, NCg12048基因可以是能够在严格条件下与可以从任何前述核苷酸序列, 如以SEQ ID NO:92所示的核苷酸序列制备的探针, 例如与任何上述核苷酸序列的全部序列或部分序列互补的序列杂交的基因(例如DNA), 只要保持原始功能。“严格条件”可以指形成所谓特异性杂合物, 并且不形成非特异性杂合物的条件。严格条件的实例可以包括高度同源的DNA彼此杂交(例如, 不小于50%, 65%, 或80%同源, 不小于90%同源, 不小于95%同源, 不小于97%同源, 或不小于99%同源的DNA彼此杂交, 并且同源性小于上述项的DNA不彼此杂交)的那些条件, 或典型的Southern杂交的清洗条件, 即在与1x SSC, 0.1%SDS于60℃; 0.1x SSC, 0.1%SDS于60℃; 或者0.1x SSC, 0.1%SDS于68℃对应的盐浓度和温度清洗1次, 或2或3次的条件。

[0167] 用于上述杂交的探针可以是与如上述的基因互补的序列的一部分。此类探针可以使用基于已知基因序列制备的寡核苷酸作为引物和含有任何上述基因的DNA片段作为模板通过PCR制备。作为探针, 例如可以使用长度约300bp的DNA片段。当使用长度为约300bp的DNA片段作为探针时, 具体地, 杂交的清洗条件可以是例如50℃, 2xSSC和0.1%SDS。

[0168] 此外, 由于关于密码子简并性的性质随宿主而变化, NCg12048基因可以包括相应等同密码子对任意密码子的替换。也就是说, 由于遗传密码的简并性, NCg12048基因可以是以上例示的任何NCg12048基因的变体。例如, NCg12048基因可以以使其根据选择的宿主中的密码子频率而具有最佳密码子的方式进行了修饰的基因。

[0169] 例如, 可以通过数学算法来确定两个序列之间的序列同一性的百分比。此类数学算法的非限制性实例可以包括Myers and Miller(1988)CABIOS 4:11-17的算法、Smith et al(1981)Adv. Appl. Math. 2:482的局部同源性算法、Needleman and Wunsch(1970)J. Mol. Biol. 48:443-453的同源性比对算法、Pearson and Lipman(1988)Proc. Natl. Acad. Sci. 85:2444-2448的同源性筛选方法和Karlin and Altschul(1990)Proc. Natl. Acad. Sci. USA 87:2264的算法的修改形式, 诸如记载于Karlin and Altschul(1993)Proc. Natl. Acad. Sci. USA 90:5873-5877的。

[0170] 通过使用基于此类数学算法的程序, 可以进行用于确定序列同一性的序列比较和比对。该程序可以由计算机适当地执行。此类程序的实例可以包括但不限于PC/Gene程序的CLUSTAL(可从Intelligenetics, Mountain View, Calif. 获得), ALIGN程序(版本2.0)以及Wisconsin Genetics Software Package, 版本8(可从Genetics Computer Group(GCG), 575Science Drive, Madison, Wis., USA获得)的GAP, BESTFIT, BLAST, FASTA和TFASTA。使用这些程序的比对可以通过使用例如初始参数来进行。CLUSTAL程序充分记载于Higgins et al. (1988)Gene 73:237-244(1988), Higgins et al. (1989)CABIOS 5:151-153, Corpet et

al.(1988)Nucleic Acids Res.16:10881-90,Huang et al.(1992)CABIOS 8:155-65,以及 Pearson et al.(1994)Meth.Mol.Biol.24:307-331。

[0171] 为了获得与靶核苷酸序列同源的核苷酸序列,具体而言,例如,可以通过使用BLASTN程序以100的得分和12的字长来进行BLAST核苷酸搜索。为了获得与靶蛋白质同源的氨基酸序列,特别地,例如可以通过以50的得分和3的字长使用BLASTX程序进行BLAST蛋白质搜索。关于BLAST核苷酸搜索和BLAST蛋白质搜索,参见ncbi.nlm.nih.gov。另外,为了比较目的,可以使用Gapped BLAST(BLAST 2.0)以获得包含缺口的比对。另外,可以使用PSI-BLAST来进行重复搜索以检测序列之间的远距离关系。关于Gapped BLAST和PSI-BLAST,参见Altschul et al.(1997)Nucleic Acids Res.25:3389。当使用BLAST、Gapped BLAST或PSI-BLAST时,可以使用每个程序的初始参数(例如核苷酸序列的BLASTN和氨基酸序列的BLASTX)。比对也可以手动进行。

[0172] 两个序列之间的序列同一性计算为在比对两个序列以便彼此最大限度配合时两个序列中匹配的残基的比率。

[0173] 前述关于基因和蛋白质的保守变体的描述可以在必要修改后应用于任意蛋白质的变体,例如目标物质生物合成酶和编码它们的基因。

[0174] <1-3>增加蛋白质活性的方法

[0175] 以下,将解释用于增加蛋白质活性的方法。

[0176] 表述“蛋白质的活性得以增加”指与非修饰菌株相比,蛋白质的活性增加。具体而言,表述“蛋白质的活性得以增加”可以意味着与非修饰菌株相比,每细胞的蛋白质的活性增加。术语“非修饰菌株”或“非修饰的微生物的菌株”可以指尚未以增加目标蛋白质活性的方式进行修饰的对照菌株。非修饰菌株的实例可以包括野生型菌株和亲本菌株。非修饰菌株的具体实例可以包括微生物物种的相应类型菌株。非修饰菌株的具体实例还可以包括以上关于微生物描述所例示的菌株。也就是说,在一个实施方案中,与类型菌株,即微生物所属的物种的类型菌株相比,蛋白质的活性可以得以增加。在另一个实施方案中,与谷氨酸棒杆菌ATCC13869菌株相比,蛋白质的活性也可以得以增加。在另一个实施方案中,与谷氨酸棒杆菌ATCC13032菌株相比,蛋白质的活性也可以得以增加。在另一个实施方案中,与大肠杆菌K-12 MG1655菌株相比,蛋白质的活性也可以得以增加。“蛋白质活性得以增加”的短语也可以表达为“蛋白质的活性得以增强”。更具体地说,表述“蛋白质的活性得以增加”可以意味着与非修饰菌株相比每细胞的蛋白质的分子数目增加,和/或蛋白质的每个分子的功能增加。也就是说,表述“蛋白质活性得以增加”中的术语“活性”不限于蛋白质的催化活性,而且还可以指编码蛋白质的基因的转录量(即mRNA的量),或蛋白质的翻译量(即蛋白质的量)。此外,“蛋白质的活性得以增加”的短语不仅包括当目标蛋白质的活性在固有具有目标蛋白质活性的菌株中增加时,而且还包括当将目标蛋白质的活性赋予不固有地具有目标蛋白质活性的菌株时。此外,只要蛋白质的活性最终得以增加,宿主中固有存在的目标蛋白质的活性可以弱化和/或消除,然后可以将适当类型的目标蛋白质赋予宿主。

[0177] 蛋白质活性增加的程度不受特别限制,只要蛋白质的活性与非修饰菌株相比增加即可。蛋白质的活性可以增加至例如非修饰菌株的蛋白质活性的1.2倍以上,1.5倍以上,2倍以上,或3倍以上。此外,当非修饰菌株不具有目标蛋白质的活性时,足够的是由于引入编码蛋白质的基因而产生蛋白质,并且例如,蛋白质可以以使其活性可以得到测量的程度生

产。

[0178] 提高蛋白质活性的修饰可以通过例如增加编码该蛋白质的基因的表达来实现。短语“基因的表达得以增加”意指与非修饰菌株如野生型菌株和亲本菌株相比,基因的表达得以增加。具体而言,短语“基因表达得以增加”可以意味着每个细胞的基因表达量与非修饰菌株相比得以增加。更具体地,短语“基因表达得以增加”可以意味着基因的转录量(即mRNA的量)得以增加,和/或基因的翻译量(即从基因表达的蛋白质量)得以增加。“基因表达得以增加”的短语也称为“基因表达得以增强”。基因的表达可以增加至例如非修饰菌株的表达的1.2倍以上,1.5倍以上,2倍以上,或3倍以上。此外,“基因表达得以增加”的短语不仅可以包括当目标基因的表达量在固有表达目标基因的菌株中得以增加时,而且还包括当将基因引入到不固有地表达目标基因并在其中表达的菌株中时。也就是说,短语“基因表达得以增加”也可以意指例如,将目标基因引入不拥有该基因的菌株中并在其中表达。

[0179] 例如,可以通过增加基因的拷贝数来增加基因的表达。

[0180] 可以通过将基因引入宿主的染色体来增加基因的拷贝数。例如可以通过使用同源重组(Miller, J.H., Experiments in Molecular Genetics, 1972, Cold Spring Harbor Laboratory)将基因引入染色体。使用同源重组的基因转移方法的实例可以包括使用线性DNA的方法,诸如Red驱动的综合(Datsenko, K.A., and Wanner, B.L., Proc. Natl. Acad. Sci. USA, 97:6640-6645 (2000)),使用含有温度敏感性复制起点的质粒的方法,使用能够接合转移的质粒的方法,使用不具有在宿主中起作用的复制起点的自杀载体的方法,使用噬菌体的转导方法。可以引入基因的仅一个拷贝,或可以引入基因的两个或更多个拷贝。例如,通过使用以染色体上的多个拷贝存在的序列作为靶标进行同源重组,可以将多个拷贝的基因引入染色体。以多个拷贝存在于染色体上的此类序列的实例可以包括重复DNA和位于转座子两端的反向重复序列。或者,可以通过使用染色体上的适当序列,如对于产生目标物质不必要的基因作为靶标进行同源重组。另外,也可以用转座子或Mini-Mu(日本专利公开(Kokai)No.2-109985,美国专利No.5,882,888,EP 805867 B1)将基因随机引入染色体。

[0181] 可以通过使用具有与全部基因或其一部分互补的序列的探针进行Southern杂交、使用基于该基因序列制备的引物的PCR等确认靶基因对染色体的引入。

[0182] 此外,通过将含有该基因的载体引入宿主中,也可以增加基因的拷贝数。例如,通过将含有靶基因的DNA片段与在宿主中起作用的载体连接以构建该基因的表达载体,并用该表达载体转化宿主,可以增加靶基因的拷贝数。可以通过例如使用具有靶基因的微生物的基因组DNA作为模板的PCR获得含有靶基因的DNA片段。作为载体,可以使用宿主细胞中可自主复制的载体。该载体可以是多拷贝载体。此外,载体可以具有用于选择转化体的标志物如抗生素抗性基因。此外,载体可以具有用于表达引入的基因的启动子和/或终止子。载体可以是例如源自细菌质粒的载体,源自酵母质粒的载体,源自噬菌体的载体,粘粒,噬菌粒等。可在肠杆菌科细菌诸如大肠杆菌中自主复制的载体的具体实例可以包括例如pUC19, pUC18, pHSG299, pHSG399, pHSG398, pBR322, pSTV29(所有这些可购自Takara Bio), pACYC184, pMW219(NIPPON GENE), pTrc99A(Pharmacia), pPROK系列载体(Clontech), pKK233-2(Clontech), pET系列载体(Novagen), pQE系列载体(QIAGEN), pCold TF DNA(TaKaRa), pACYC系列载体,和宽宿主范围的载体RSF1010。可在棒状杆菌细菌中自主复制的

载体的具体实例可以包括例如pHM1519(Agric.Biol.Chem., 48, 2901-2903(1984)); pAM330(Agric.Biol.Chem., 48, 2901-2903(1984)); 通过改善这些质粒获得并具有药物抗性基因的质粒; 日本专利公开(Kokai) No. 3-210184中描述的质粒pCRY30; 日本专利公开(Kokai) No. 2-72876和美国专利No. 5,185,262中描述的质粒pCRY21, pCRY2KE, pCRY2KX, pCRY31, pCRY3KE和pCRY3KX; 日本专利公开(Kokai) No. 1-191686中描述的质粒pCRY2和pCRY3; 日本专利公开(Kokai) No. 58-192900中描述的pAJ655, pAJ611和pAJ1844; 日本专利公开(Kokai) No. 57-134500中描述的pCG1; 日本专利公开(Kokai) No. 58-35197中描述的pCG2; 日本专利公开(Kokai) No. 57-183799中描述的pCG4和pCG11; 日本专利公开(Kokai) No. 10-215883中描述的pVK7; W02007/046389中描述的pVK9; W02013/069634中描述的pVS7; 和日本专利公开(Kokai) No. 9-070291中描述的pVC7。

[0183] 当引入基因时, 宿主可以表达基因是足够的。具体而言, 宿主中存在基因以便其在宿主中起作用的启动子的控制下表达是足够的。术语“在宿主中起作用的启动子”可以指在宿主中显示启动子活性的启动子。启动子可以是源自宿主的启动子或异源启动子。启动子可以是待引入基因的天然启动子或另一基因的启动子。作为启动子, 例如也可以使用如本文描述的更强的启动子。

[0184] 终止基因转录的终止子可以位于基因的下游。终止子没有特别限制, 只要其在选择的宿主中起作用。终止子可以是源自宿主的终止子或异源终止子。终止子可以是待引入基因的天然终止子或另一基因的终止子。终止子的具体实例可以包括例如T7终止子、T4终止子、fd噬菌体终止子、tet终止子和trpA终止子。

[0185] 在“Fundamental Microbiology Vol.8, Genetic Engineering, KYORITSU SHUPPAN CO., LTD, 1987”中详细公开了可用于各种微生物中的载体, 启动子和终止子, 并且可以使用那些。

[0186] 此外, 当引入两个或更多个基因时, 基因各自可以由宿主表达是足够的。例如, 所有基因可以由单个表达载体或染色体携带。此外, 基因可以分别由两个或更多个表达载体携带, 或者分开由单一或两个或更多个表达载体和染色体携带。也可以引入由两个或更多个基因构成的操纵子。“引入2个或更多个基因”的短语可以指例如引入编码2种或更多种蛋白质(酶等)的相应基因, 引入编码构成单一蛋白质复合物(例如酶复合物)的2种或更多种亚基的相应基因以及它们的组合。

[0187] 待引入的基因没有特别限制, 只要它编码在宿主中起作用的蛋白即可。待引入的基因可以是源自宿主的基因, 或者可以是异源基因。可以例如使用基于该基因的核苷酸序列设计的引物, 并使用具有该基因的生物体的基因组DNA, 携带该基因的质粒等作为模板通过PCR获得待引入的基因。待引入的基因也可以完全合成, 例如基于该基因的核苷酸序列(Gene, 60(1), 115-127(1987))。获得的基因可以原样使用, 或者根据需要在修饰后使用。也就是说, 可以通过修饰基因来获得基因的变体。基因可以通过已知技术进行修饰。例如, 可以通过位点特异性突变方法将目标突变引入DNA的目标位点。也就是说, 可以通过位点特异性突变方法修饰基因的编码区, 使得编码蛋白的特定位点包括氨基酸残基的取代, 缺失, 插入和/或添加。位点特异性突变方法的实例可以包括利用PCR的方法(Higuchi, R., 61, 于PCR Technology, Erlich, H.A. Eds., Stockton Press(1989); Carter, P., Meth. in Enzymol., 154, 382(1987)) 和利用噬菌体的方法(Kramer, W. and Frits, H.J., Meth. 于Enzymol., 154,

350(1987);Kunkel,T.A.et al.,Meth.in Enzymol.,154,367(1987))。或者,可以完全合成基因的变体。

[0188] 另外,当蛋白质作为由多个亚基构成的复合物起作用时,可以修饰多个亚基的部分或全部,只要蛋白质的活性最终得以增加。也就是说,例如,当通过增加基因的表达来增加蛋白质的活性时,可以增强编码亚基的多个基因的部分或全部的表达。通常优选增强编码亚基的多个基因全部的表达。此外,构成该复合物的亚基可以源自单一种类的生物体或两种或更多种生物体,只要该复合物具有目标蛋白质的功能即可。也就是说,例如,可以将编码多个亚基的相同生物体的基因引入宿主,或者可以将编码多个亚基的不同生物体的基因引入宿主中。

[0189] 此外,可以通过改善基因的转录效率增加基因的表达。另外,通过提高基因的翻译效率也可以增加基因的表达。基因的转录效率和基因的翻译效率可以通过例如修饰基因的表达控制序列来改善。术语“表达控制序列”可以共同指影响基因表达的位点。表达控制序列的实例可以包括例如启动子,Shine-Dalgarno (SD) 序列(其也可以称为核糖体结合位点(RBS))和RBS与起始密码子之间的间隔物区。可以通过使用启动子搜索载体或基因分析软件如GENETYX来鉴定表达控制序列。这些表达调控序列可以通过例如使用温度敏感性载体的方法或Red驱动整合方法(WO2005/010175)进行修饰。

[0190] 通过例如用更强的启动子替换染色体上的基因的启动子可以提高基因的转录效率。术语“更强的启动子”可以指与基因的固有的野生型启动子相比提供改善的基因转录的启动子。更强的启动子的实例可以包括例如已知的高表达启动子,诸如T7启动子、trp启动子、lac启动子、thr启动子、tac启动子、trc启动子、tet启动子、araBAD启动子、rpoH启动子、msrA启动子、Pml启动子(源自双歧杆菌属(Bifidobacterium))、PR启动子和PL启动子。在棒状杆菌细菌中可用的更强启动子的实例可以包括例如在棒状杆菌细菌中用乙酸、乙醇、丙酮酸等可诱导的人工修饰的P54-6启动子(Appl.Microbiol.Biotechnol.,53,674-679(2000)),pta,aceA,aceB,adh,和amyE启动子,cspB,SOD和tuf(EF-Tu)启动子,它们是能够在棒状杆菌细菌中提供大表达量的有力启动子(Journal of Biotechnology,104(2003)311-323;Appl.Environ.Microbiol.,2005Dec;71(12):8587-96),P2启动子(SEQ ID NO:81的位置942-1034)和P3启动子(SEQ ID NO:84)以及lac启动子,tac启动子和trc启动子。此外,作为更强的启动子,通过使用各种报道基因也可以获得高活性类型的现有启动子。例如,通过使启动子区中的-35和-10区更接近共有序列,可以增强启动子的活性(WO00/18935)。高活性型启动子的实例可以包括各种tac样启动子(Katashkina JI et al.,Russian Federation Patent Application No.2006134574)。在Goldstein et al.(Prokaryotic Promoters in Biotechnology,Biotechnol.Annu.Rev.,1,105-128(1995))的论文等中描述了评估启动子强度的方法和强启动子的实例。

[0191] 例如,通过用更强的SD序列替换染色体上的基因的Shine-Dalgarno (SD) 序列(其也可以称为核糖体结合位点(RBS)),可以改善基因的翻译效率。术语“更强的SD序列”可以指与基因固有的野生型SD序列相比提供改善的mRNA翻译的SD序列。更强的SD序列的实例可以包括例如源自噬菌体T7的基因10的RBS(Olins P.O.et al,Gene,1988,73,227-235)。此外,已知在RBS和起始密码子之间的间隔物区中,特别是在紧邻起始密码子上游的序列(5'-UTR)中,几个核苷酸的取代,插入或缺失显著影响mRNA的稳定性和翻译效率,因此,也可以

通过修饰来改善基因的翻译效率。

[0192] 基因的翻译效率也可以通过例如修饰密码子来改善。例如,通过用更频繁使用的同义密码子取代基因中存在的稀有密码子,可以提高基因的翻译效率。也就是说,例如,基因可以以使该基因含有根据在选择的主机中观察到的密码子频率的最佳密码子的方式进行了修饰。例如,可以通过用于将目标突变引入DNA的目标位点中的位点特异性突变方法替换密码子。或者,可以完全合成替换目标密码子的基因片段。“密码子选择数据库”(kazusa.or.jp/codon; Nakamura, Y. et al, Nucl. Acids Res., 28, 292 (2000)) 中公开了各种生物体中的密码子的频率。

[0193] 此外,基因的表达还可以通过扩增增加基因表达的调节子或缺失或弱化减少基因表达的调节子来增加。

[0194] 可以独立使用或以任意组合使用如上所述用于增加基因表达的此类方法。

[0195] 此外,也可以通过例如增强酶的比活性来获得增加蛋白质活性的修饰。比活性的增强也可以包括对反馈抑制的脱敏。也就是说,当蛋白质被代谢物反馈抑制时,可以通过使选择的主机中的基因或蛋白质突变以对反馈抑制脱敏来增加蛋白质的活性。短语“对反馈抑制脱敏”可以包括消除反馈抑制,和弱化反馈抑制,除非另有叙述。此外,“对反馈抑制脱敏”的短语,即当反馈抑制得到消除或弱化时,也可以称为“对反馈抑制耐受”。通过例如搜索各种生物体可以获得显示增强的比活性的蛋白质。此外,通过向现有蛋白质中引入突变也可以获得高活性类型的固有蛋白质。待引入的突变可以是例如在蛋白质的一个或几个位置上一个或几个氨基酸残基的取代,缺失,插入和/或添加。可以通过例如如上所述的此类位点特异性突变方法来引入突变。突变也可以通过例如诱变处理来引入。诱变处理的实例可以包括X射线照射,紫外线照射和用突变剂如N-甲基-N'-硝基-N-亚硝基胍(MNNG),甲磺酸乙酯(EMS)和甲磺酸甲酯(MMS)处理。此外,可以通过用羟胺在体外直接处理DNA来诱导随机突变。比活性的增强可以独立使用,或者可以与如上述用于增强基因表达的此类方法任意组合使用。

[0196] 用于转化的方法不受特别限制,并且可以使用常规已知的方法。可以使用例如用氯化钙处理接受细胞以增加其对DNA的通透性的方法,其已经对大肠杆菌K-12菌株报道(Mandel, M. and Higa, A., J. Mol. Biol., 1970, 53, 159-162),以及从生长阶段中的细胞制备感受态细胞,接着用DNA转化的方法,其已经对枯草芽孢杆菌(*Bacillus subtilis*)报告(Duncan, C. H., Wilson, G. A. and Young, F. E., Gene, 1977, 1: 153-167)。或者,可以使用使DNA-接受细胞变为原生质体或原生质球(其可以容易摄取重组DNA),接着将重组DNA引入DNA-接受细胞中的方法,已知该方法适用于枯草芽孢杆菌、放线菌和酵母(Chang, S. and Choen, S. N., 1979, Mol. Gen. Genet., 168: 111-115; Bibb, M. J., Ward, J. M. and Hopwood, O. A., 1978, Nature, 274: 398-400; Hinnen, A., Hicks, J. B. and Fink, G. R., 1978, Proc. Natl. Acad. Sci. USA, 75: 1929-1933)。此外,也可以使用对棒状杆菌细菌报告的电脉冲方法(日本专利公开(Kokai) No. 2-207791)。

[0197] 可以通过测量蛋白质的活性来确认蛋白质活性的增加。

[0198] 通过确认编码蛋白质的基因的表达增加,也可以确认蛋白质活性的增加。可以通过确认基因转录量的增加或通过确认从该基因表达的蛋白质量的增加来确认基因表达的增加。

[0199] 可以通过比较从基因转录的mRNA的量与非修饰菌株如野生型菌株或亲本菌株的mRNA量来确认基因转录量的增加。用于评估mRNA量的方法的实例可以包括Northern杂交, RT-PCR等(Sambrook, J., et al., Molecular Cloning A Laboratory Manual/第三版, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (USA), 2001)。mRNA的量可以增加例如非修饰菌株的mRNA量的1.2倍以上, 1.5倍以上, 2倍以上或3倍以上。

[0200] 可以通过使用抗体的Western印迹法 (Molecular Cloning, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (USA), 2001) 来确认蛋白质质量的增加。蛋白质的量(诸如每细胞的蛋白质分子的数目)可以增加例如非修饰菌株的蛋白质质量的1.2倍以上, 1.5倍以上, 2倍以上或3倍以上。

[0201] 上述提高蛋白质活性的方法可用于提高任意蛋白质如目标物质生物合成酶, 磷酸泛酰巯基乙胺基化酶和物质摄取系统的活性, 和增强任意基因如编码那些任意蛋白质的基因的表达。

[0202] <1-4>降低蛋白质活性的方法

[0203] 以下, 将解释用于降低蛋白质诸如NCg12048蛋白的活性的方法。

[0204] 表述“蛋白质的活性得以降低”指与非修饰菌株相比蛋白质的活性得以降低。具体而言, 表述“蛋白质的活性得以降低”可以意味着与非修饰菌株相比, 每个细胞的蛋白质的活性得以降低。术语“非修饰菌株”可以指尚未以使目的蛋白质的活性得以降低的方式进行修饰的对照菌株。非修饰菌株的实例可以包括野生型菌株和亲本菌株。非修饰菌株的具体实例可以包括微生物物种的相应类型菌株。非修饰菌株的具体实例还可以包括以上关于微生物描述所例示的菌株。也就是说, 在一个实施方案中, 与类型菌株, 即微生物所属的物种的类型菌株相比, 可以降低蛋白质的活性。在另一个实施方案中, 与谷氨酸棒杆菌ATCC 13869菌株相比, 蛋白质的活性也可以得以降低。在另一个实施方案中, 与谷氨酸棒杆菌ATCC 13032相比, 蛋白质的活性也可以得以降低。另一个实施方案中, 与大肠杆菌K-12 MG1655菌株相比, 蛋白质的活性也可以得以降低。“蛋白质活性得以降低”的短语还可以包括当蛋白质的活性已经完全消失时。更具体地说, 表述“蛋白质的活性得以降低”可以意味着与非修饰的菌株相比每个细胞的蛋白质的分子数目得以减少, 和/或蛋白质的每个分子的功能得以降低。也就是说, 表述“蛋白质的活性得以降低”中的术语“活性”不限于蛋白质的催化活性, 而且还可以指编码蛋白质的基因的转录量(即mRNA的量)或蛋白质的翻译量(即蛋白质的量)。“每个细胞的蛋白质的分子数目得以减少”的短语还可以包括当完全不存在蛋白质时。“蛋白质的每个分子的功能得以降低”的短语也可以包括当每个蛋白质分子的功能完全消失时。蛋白质活性降低的程度没有特别限制, 只要活性与非修饰菌株的活性相比降低即可。蛋白质的活性可以降低到例如非修饰菌株的蛋白质活性的50%或更少, 20%或更少, 10%或更少, 5%或更少, 或0%。

[0205] 可以通过例如减少编码蛋白质的基因的表达来实现用于降低蛋白质活性的修饰。短语“基因的表达得以降低”指与非修饰菌株如野生型菌株和亲本菌株相比, 该基因的表达得以降低。具体而言, 短语“基因表达得以降低”可以意味着与非修饰菌株相比, 每个细胞的基因表达得以降低。更具体地说, 短语“基因表达得以降低”可以意味着基因的转录量(即mRNA的量)得以减少, 和/或基因的翻译量(即从基因表达的蛋白质量)得以减少。“基因表达得以降低”的短语也可以包括当基因完全不表达时。“基因表达得以降低”的短语也可以称

为“基因表达得以弱化”。基因的表达可以降低到例如非修饰菌株的表达的50%或更少,20%或更少,10%或更少,5%或更少,或0%。

[0206] 基因表达的降低可以是例如由于例如转录效率降低,翻译效率降低或组合所致。可以通过修饰基因的表达控制序列诸如启动子,Shine-Dalgarno (SD) 序列(其也可以称为核糖体结合位点(RBS)),和基因的RBS和起始密码子之间的间隔物区来降低基因的表达。当修饰表达控制序列时,修饰表达控制序列的一个或多个核苷酸,两个或更多个核苷酸,或三个或更多个核苷酸。例如,通过用更弱的启动子替换染色体上基因的启动子,可以降低基因的转录效率。术语“更弱的启动子”可以指与基因固有的野生型启动子相比,提供基因的弱化转录的启动子。更弱启动子的实例可以包括例如诱导型启动子。也就是说,诱导型启动子可以在非诱导条件下,例如在不存在相应诱导物的情况下,起到更弱启动子的作用。更弱启动子的实例还可以包括例如P4和P8启动子(分别为SEQ ID NO:82的位置872-969和SEQ ID NO:83的位置901-1046)。此外,可以缺失表达控制序列的一部分或全部。基因的表达还可以通过例如操纵负责表达控制的因子来降低。负责表达控制的因子的实例可以包括负责转录或翻译控制的低分子(诸如诱导剂,抑制剂等),负责转录或翻译控制的蛋白质(转录因子等),负责转录或翻译控制的核酸(诸如siRNA等),等等。此外,通过例如对基因的编码区引入降低基因表达的突变,也可以降低基因的表达。例如,通过用在宿主中不太频繁使用的同义密码子替换基因编码区中的密码子,可以降低基因的表达。此外,例如,如本文所述,基因表达可以由于基因的破坏而降低。

[0207] 用于降低蛋白质活性的修饰也可以通过例如破坏编码该蛋白质的基因来获得。短语“基因受到破坏”可以指基因以使得不产生通常可以起作用的蛋白质的方式进行了修饰。“不产生通常可以起作用的蛋白质”的短语可以包括当完全不从基因产生蛋白质时和当从基因产生每个分子的功能(例如活性或性质)得以降低或消除的蛋白质时。

[0208] 可以通过例如缺失染色体上的基因来获得基因的破坏。术语“基因的缺失”可以指基因编码区的部分或整个区域的缺失。此外,可以缺失包括染色体上基因的编码区上游和下游序列的整个基因。只要可以降低蛋白质的活性,待缺失的区域可以是任何区域,例如N端区域(即编码蛋白质的N端区域的区域),内部区域或C端区域(即编码蛋白质的C端区域的区域)。缺失更长的区域通常可以更确定地使基因失活。待缺失的区域可以是具有例如基因的编码区的总长度的10%以上、20%以上、30%以上、40%以上、50%以上、60%以上、70%以上、80%以上、90%以上、或95%以上的长度的区域。此外,优选的是,待缺失的区域的上游和下游序列的阅读框不相同。读码框的不一致性可以引起待缺失的区域下游的移码。

[0209] 例如,也可以通过对染色体上基因的编码区引入用于氨基酸取代的突变(错义突变),终止密码子的突变(无义突变),添加或缺失一个或两个核苷酸残基的突变(移码突变)等获得基因的破坏(Journal of Biological Chemistry,272:8611-8617(1997); Proceedings of the National Academy of Sciences,USA,95 5511-5515(1998); Journal of Biological Chemistry,26 116,20833-20839(1991))。

[0210] 也可以通过例如将另一个核苷酸序列插入到染色体上的基因的编码区中获得基因的破坏。插入的位点可以位于基因的任何区域中,并且插入更长的核苷酸序列通常可以更确定地使基因失活。优选的是插入位点上游和下游序列的阅读框不相同。读码框的不一致性可以引起待缺失的区域下游的移码。其他核苷酸序列没有特别限制,只要选择降低或

消除所编码蛋白质活性的序列即可,并且其实例可以包括例如标志物基因如抗生素抗性基因,以及可用于产生目标物质的基因。

[0211] 特别地,可以以编码蛋白质的氨基酸序列得以缺失的方式进行基因的破坏。换言之,可以通过例如缺失蛋白质的氨基酸序列,特别地修饰基因以使该基因编码氨基酸序列得以的蛋白质来实现降低蛋白质活性的修饰。短语“蛋白质的氨基酸序列的缺失”可以指蛋白质的氨基酸序列的部分或整个区域的缺失。另外,短语“蛋白质的氨基酸序列的缺失”可以指原始氨基酸序列在蛋白质中消失,并且还可以包括当原始氨基酸序列变为另一个氨基酸序列时。也就是说,例如,通过移码改变为另一氨基酸序列的区域可以视为缺失区域。当蛋白质的氨基酸序列得以缺失时,通常缩短蛋白质的总长度,但也可以存在蛋白质的总长度没有改变或延长的情况。例如,通过缺失基因编码区的部分或整个区域,可以在编码的蛋白质中缺失由缺失区域编码的区域。另外,例如,通过将终止密码子引入基因的编码区中,可以在编码的蛋白质中缺失由引入位点的下游区域编码的区域。另外,例如,通过在基因的编码区中移码,可以在编码的蛋白质中缺失由移码区编码的区域。关于基因缺失中待缺失区域的位置和长度的上述描述可以在必要修改后应用于缺失蛋白质氨基酸序列中待缺失区域的位置和长度。

[0212] 如上所述的染色体上基因的此类修饰可以如下获得:通过例如制备以使基因不能产生通常起作用的蛋白质的方式进行了修饰的破坏类型基因,并用含有破坏类型基因的重组DNA转化宿主以引起破坏类型基因和染色体上的野生型基因之间的同源重组,从而用该破坏类型基因取代染色体上的野生型基因。在该程序中,若重组DNA中包括根据宿主特征如营养缺陷选择的标志物基因,则操作变得更容易。破坏类型基因的实例可以包括编码区的部分或整个区域得以缺失的基因,包含错义突变的基因,包括无义突变的基因,包含移码突变的基因,以及包含转座子或标志物基因插入的基因。由破坏类型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构象,因此其功能得以降低或消除。已经建立了基于使用同源重组的基因取代的此类基因破坏,并且存在使用线性DNA的方法,诸如称为“Red驱动整合”的方法(Datsenko, K.A, and Wanner, B.L., Proc. Natl. Acad. Sci. USA, 97:6640-6645 (2000)), 以及利用Red驱动整合与源自 λ 噬菌体的切除系统的组合的方法(Cho, E.H., Gumport, R.I., Gardner, J.F., J. Bacteriol., 184:5200-5203 (2002)) (参考W02005/010175), 使用具有温度敏感性复制起点的质粒的方法, 使用能够接合转移的质粒的方法, 利用没有在宿主中起作用的复制起点的自杀载体的方法(美国专利号6,303,383, 日本专利公开(Kokai)No.05-007491), 等等。

[0213] 用于降低蛋白质活性的修饰也可以通过例如诱变处理来获得。诱变处理的实例包括X射线或紫外线的照射和用突变剂如N-甲基-N'-硝基-N-亚硝基胍(MNNG), 甲磺酸乙酯(EMS)和甲磺酸甲酯(MMS)处理。

[0214] 如上文提及的用于降低蛋白质活性的此类方法可以独立使用或者以任意组合使用。

[0215] 当蛋白质作为由多个亚基构成的复合物起作用时,可以修饰多个亚基的部分或全部,只要蛋白质的活性最终得以降低。也就是说,例如,编码相应亚基的多个基因的部分或全部可以得以破坏等。此外,当蛋白质存在多种同工酶时,多种同工酶的部分或全部活性可以得以降低,只要蛋白质的活性最终得以降低。也就是说,例如,编码相应同工酶的多个基

因的部分或全部可以得以破坏等。

[0216] 可以通过测量蛋白质的活性确认蛋白质活性的降低。

[0217] 蛋白质活性的降低也可以通过确认编码该蛋白质的基因的表达的降低来确认。基因表达的降低可以通过确认基因转录量的减少或从基因表达的蛋白质量的降低来确认。

[0218] 通过比较从基因转录的mRNA的量与非修饰菌株中观察到的量,可以确认基因转录量的降低。用于评估mRNA量的方法的实例可以包括Northern杂交,RT-PCR等(Molecular Cloning,Cold Spring Harbor Laboratory Press,Cold Spring Harbor(USA),2001)。mRNA的量可以降低到例如在非修饰菌株中观察到的量的50%或更少,20%或更少,10%或更少,5%或更少,或0%。

[0219] 通过使用抗体的Western印迹法(Molecular Cloning,Cold Spring Harbor Laboratory Press,Cold Spring Harbor(USA) 2001)可以确认蛋白质量的减少。蛋白质的量(诸如每个细胞的蛋白质的分子数目)可以降低到例如在非修饰菌株中观察到的蛋白质量的50%或更少,20%或更少,10%或更少,5%或更少,或0%。

[0220] 可以根据用于破坏的手段通过确定基因的部分或全部的核苷酸序列,基因的限制酶图谱,全长等来确认基因的破坏。

[0221] 除NCg12048蛋白的活性弱化以外,上述的降低蛋白质活性的方法还可以适用于任意蛋白质诸如副产物生成酶的活性降低,以及任意基因诸如编码那些任意蛋白的基因的表达降低。

[0222] <2>用于生产目标物质的方法

[0223] 如本文中描述的方法是通过使用如本文中描述的微生物生产目标物质的方法。

[0224] <2-1>发酵方法

[0225] 可以通过例如如本文中描述的微生物的发酵来生产目标物质。也就是说,如本文中描述的方法的实施方案可以是通过发酵微生物来生产目标物质的方法。此实施方案也可以称为“发酵方法”。此外,通过发酵如本文中所述的微生物生产目标物质的步骤也可以称为“发酵步骤”。

[0226] 可以通过培养如本文中描述的微生物来进行发酵步骤。具体而言,在发酵方法中,可以从碳源生产目标物质。也就是说,发酵步骤可以是例如在培养基,诸如含有碳源的培养基中培养微生物以在培养基中生产和积累目标物质的步骤。也就是说,发酵方法可以是用于生产目标物质的方法,其包括以下步骤:例如在培养基,诸如含有碳源的培养基中培养微生物以在培养基中生产和积累目标物质的方法。此外,换言之,发酵步骤可以是例如通过使用微生物从碳源生产目标物质的步骤。

[0227] 待使用的培养基没有特别限制,只要微生物可以在其中增殖并产生目标物质即可。作为培养基,例如可以使用用于培养微生物诸如细菌和酵母的典型培养基。培养基可以根据需要含有碳源,氮源,磷酸盐源,硫源,以及其他培养基组分,如各种有机组分和无机组分。培养基成分的类型和浓度可以根据诸如选择的微生物的类型等各种条件适当确定。

[0228] 碳源没有特别限制,只要微生物可以利用它并产生目标物质即可。碳源的具体实例可以包括例如糖类,诸如葡萄糖,果糖,蔗糖,乳糖,半乳糖,木糖,阿拉伯糖,赤糖糊(blackstrap molasses),淀粉的水解物和生物质的水解物;有机酸如乙酸,柠檬酸,琥珀酸和葡萄糖酸;醇如乙醇,甘油和粗甘油;和脂肪酸。作为碳源,特别地,可以使用植物来源的材

料。植物的实例可以包括例如玉米, 稻, 小麦, 大豆, 甘蔗, 甜菜和棉。植物来源材料的实例可以包括例如器官, 诸如根, 茎, 干, 枝, 叶, 花和种子, 包括它们的植物体和这些植物器官的分解产物。植物来源材料的在其使用时的形式没有特别限定, 并且它们可以以诸如未加工产物, 汁, 研磨产物, 纯化产物等任何形式使用。戊糖如木糖, 己糖如葡萄糖或它们的混合物可从例如植物生物质获得并且得到使用。具体而言, 可以通过将植物生物质进行处理, 诸如蒸汽处理, 浓酸水解, 稀酸水解, 纤维素酶等酶水解, 碱处理而得到这些糖类。由于半纤维素与纤维素相比通常更容易水解, 因此可预先水解植物生物质中的半纤维素以释放戊糖, 然后可以水解纤维素以产生己糖。此外, 木糖可以通过例如通过对微生物赋予用于将己糖如葡萄糖转化为木糖的途径从己糖转化供应。作为碳源, 可以使用一种碳源, 或者可以组合使用两种或更多种碳源。

[0229] 培养基中碳源的浓度没有特别限制, 只要微生物可以增殖并产生目标物质即可。培养基中碳源的浓度可以在使得目标物质的生产不受抑制的范围内尽可能高。碳源在培养基中的初始浓度通常可以是例如通常5至30% (w/v), 或10至20% (w/v)。此外, 可以根据需要对培养基添加碳源。例如, 可以与伴随发酵进展的碳源的减少或消减成比例地对培养基添加碳源。虽然可以暂时消减碳源, 但是只要最终可以生产目标物质, 可以优选以不消减碳源或不继续消减碳源的方式进行培养。

[0230] 氮源的具体实例可以包括例如铵盐如硫酸铵, 氯化铵和磷酸铵, 有机氮源如蛋白胨, 酵母提取物, 肉膏和大豆蛋白分解产物, 氨和尿素。用于pH调节的氨气和氨水也可用作氮源。作为氮源, 可以使用一种氮源, 也可以组合使用两种或更多种的氮源。

[0231] 磷酸盐源的具体实例可以包括例如磷酸盐如磷酸二氢钾和磷酸氢二钾, 以及磷酸聚合物如焦磷酸。作为磷酸盐源, 可以使用一种磷酸盐源, 或者可以组合使用两种或更多种磷酸盐源。

[0232] 硫源的具体实例可以包括例如无机硫化物如硫酸盐, 硫代硫酸盐和亚硫酸盐, 以及含硫氨基酸如半胱氨酸, 胱氨酸和谷胱甘肽。作为硫源, 可以使用一种硫源, 或者可以组合使用两种或更多种硫源。

[0233] 其他各种有机和无机组分的具体实例可以包括例如无机盐如氯化钠和氯化钾; 微量金属如铁, 锰, 镁和钙; 维生素如维生素B1, 维生素B2, 维生素B6, 烟酸, 烟酰胺和维生素B12; 氨基酸; 核酸; 和含有这些的有机成分如蛋白胨, 酪蛋白氨基酸, 酵母提取物和大豆蛋白分解产物。作为其他各种有机和无机组分, 可以使用一种组分, 或者可以组合使用两种或更多种组分。

[0234] 此外, 当使用需要营养物如氨基酸进行生长的营养缺陷型突变株时, 优选培养基含有此类需要的营养物。此外, 培养基可以含有用于生产目标物质的组分。此类组分的具体实例可以包括例如甲基供体例如SAM及其前体例如甲硫氨酸。

[0235] 培养条件没有特别限制, 只要微生物可以增殖并且产生目标物质即可。培养可以用例如用于培养微生物诸如细菌和酵母的典型条件进行。培养条件可以根据诸如选择的微生物的种类等各种条件适当确定。

[0236] 培养可以通过使用液体培养基来进行。在培养时, 例如, 可以将诸如琼脂培养基等固体培养基上培养的微生物直接接种到液体培养基中, 或者将在液体培养基中作为种子培养物培养的微生物接种到液体培养基中进行主要培养。也就是说, 培养可以作为种子培

养和主要培养分开进行。在此类情况下,种子培养和主要培养的培养条件可以是相同的或者可以是不相同。至少在主要培养过程中产生目标物质是足够的。培养开始时培养基中存在的微生物的量没有特别限制。例如,显示OD660为4-100的种子培养液可以在开始培养时以0.1-100质量%,或1-50质量%的量接种到用于主要培养的培养基中。

[0237] 培养可以以分批培养,补料分批培养,连续培养或这些的组合进行。在培养开始时使用的培养基也可以称为“起始培养基”。在补料分批培养或连续培养中添加到培养系统(例如发酵罐)的培养基也可以称为“补料培养基”。为了在补料分批培养或者连续培养中向培养系统添加补料培养基也可以称为“补料”。此外,当作为种子培养和主要培养分开进行培养时,种子培养和主要培养的培养方案可以相同或可以不同。例如,种子培养和主要培养两者都可以作为分批培养进行。或者,例如,种子培养可以作为分批培养进行,并且主要培养可以以补料分批培养或连续培养进行。

[0238] 各种组分诸如碳源可以存在于起始培养基,补料培养基或两者中。也就是说,各种组分诸如碳源可以在培养期间独立地或以任意组合的方式添加到培养基。这些组分可以添加一次或多次,或者可以连续添加。起始培养基中存在的组分的类型可以与补料培养基中存在的组分的类型相同或不同。此外,起始培养基中存在的组分的浓度可以与补料培养基中存在的组分的浓度相同或不同。此外,可以使用两种或更多种含有不同类型和/或不同浓度组分的补料培养基。例如,当间歇进行两次或更多次补料时,补料培养基中存在的组分的类型和/或浓度对于每次补料可以是相同的或者可以是不同的。

[0239] 培养可以例如在需氧条件下进行。术语“需氧条件”可以指其中培养基中的溶解氧浓度为0.33ppm或更高,或1.5ppm或更高的条件。可以将氧浓度控制为例如饱和氧浓度的1-50%,或约5%。培养可以例如在通气或摇动下进行。培养基的pH值可以是例如3至10,或4.0至9.5。培养基的pH可以根据需要在培养过程中进行调整。培养基的pH可以通过使用各种碱性和酸性物质如氨气,氨水,碳酸钠,碳酸氢钠,碳酸钾,碳酸氢钾,碳酸镁,氢氧化钠,氢氧化钙和氢氧化镁来调节。培养温度可以是例如20至45℃,或25至37℃。培养时间可以是例如10至120小时。例如,可以继续培养直到培养基中存在的碳源被消耗,或者直到微生物的活性丧失。

[0240] 通过在诸如上文所述的条件下培养微生物,在培养基中积累目标物质。

[0241] 可以通过用于检测或鉴定化合物的已知方法来确认目标物质的产生。此类方法的实例可以包括例如HPLC,UPLC,LC/MS,GC/MS和NMR。这些方法可以独立使用,或者可以以适当的组合使用。这些方法也可以用于确定培养基中存在的各种组分的浓度。

[0242] 生产的目标物质可以适当收集。也就是说,发酵方法可以进一步包括收集目标物质的步骤。此步骤也可以称为“收集步骤”。收集步骤可以是培养肉汤,特别是从培养基中收集目标物质的步骤。目标物质可以通过用于分离和纯化化合物的已知方法收集。此类方法的实例可以包括例如离子交换树脂法,膜处理,沉淀,提取,蒸馏和结晶。具体地,目标物质也可以通过用有机溶剂如乙酸乙酯萃取或通过蒸汽蒸馏来收集。这些方法可以独立使用,或者可以以适当的组合使用。

[0243] 此外,当目标物质沉淀在培养基中时,它可以通过例如离心或过滤收集。在培养基中溶解的目标物质结晶后,可以将沉淀在培养基中的目标物质和溶解在培养基中的目标物质一起分离。

[0244] 除了目标物质之外,收集的目标物质可以包含例如微生物细胞,培养基组分,水分和微生物的副产物代谢物。收集的目标物质的纯度可以是例如30% (w/w)或更高,50% (w/w)或更高,70% (w/w)或更高,80% (w/w)或更高,90% (w/w)或更高,或95% (w/w)或更高。

[0245] <2-2>生物转化方法

[0246] 目标物质也可以通过例如使用如本文中所述的微生物的生物转化来产生。也就是说,如本文中所述的方法的另一个实施方案可以是使用微生物通过生物转化来生产目标物质的方法。该实施方案也可以称为“生物转化方法”。此外,使用微生物通过生物转化来生产目标物质的步骤也可以称为“生物转化步骤”。

[0247] 具体而言,在生物转化方法中,目标物质可以从目标物质的前体产生。更具体而言,在生物转化方法中,通过使用微生物,可以通过将目标物质的前体转化为目标物质来生产目标物质。也就是说,生物转化步骤可以通过使用微生物将目标物质的前体转化为目标物质的步骤。

[0248] 目标物质的前体也可以简称为“前体”。前体的实例可以包括下述物质,该物质向目标物质的转化包括需要SAM。前体的具体实例可以包括目标物质的生物合成途径的中间体,如就目标物质生物合成酶的描述而言所述的那些,条件是中间体向目标物质的转化需要SAM。前体的更具体实例可以包括例如原儿茶酸,原儿茶醛,L-色氨酸,L-组氨酸,香草酸,苯甲酸,L-苯丙氨酸,L-酪氨酸,L-精氨酸,L-鸟氨酸和甘氨酸。原儿茶酸可以用作生产例如香草醛,香草酸或愈创木酚的前体。原儿茶醛可以用作生产例如香草醛的前体。L-色氨酸可以用作生产例如褪黑激素的前体。L-组氨酸可以用作生产例如麦角硫因的前体。L-苯丙氨酸和L-酪氨酸各自可以用作生产例如阿魏酸,4-乙烯基愈创木酚或4-乙基愈创木酚的前体。L-精氨酸和L-鸟氨酸各自可以用作生产例如多胺的前体。L-精氨酸和甘氨酸各自可以用作生产例如肌酸的前体。作为前体,可以使用一种前体,或者可以组合使用两种或更多种前体。在前体是可以形成盐的化合物的情况下,前体可以作为游离化合物,其盐或其混合物使用。也就是说,术语“前体”可以指游离形式的前体,其盐或其混合物,除非另有说明。盐的实例可以包括例如硫酸盐,盐酸盐,碳酸盐,铵盐,钠盐和钾盐。作为前体的盐,可以使用一种盐,或者可以组合使用两种或更多种盐。

[0249] 作为前体,可以使用商业产品,或者可以使用适当制备和获得的产品。也就是说,生物转化方法可以进一步包括生产前体的步骤。生产前体的方法没有特别限定,例如可以使用已知的方法。前体可以通过例如化学合成法,酶法,生物转化法,发酵法,提取方法或它们的组合来生产。也就是说,例如,目标物质的前体可以使用酶(该酶也可以称为“前体生物合成酶”)从其另外的前体产生,所述酶催化此类另外的前体转化成目标物质的前体。此外,例如,通过使用具有前体生产能力的微生物,可以从碳源或此类另外的前体生产目标物质的前体。短语“具有前体生产能力的微生物”可以指能够从碳源或其另外的前体生产目标物质的前体的微生物。例如,根据酶法或生物转化法生产原儿茶酸的方法的实例可以包括使用恶臭假单胞菌(*Pseudomonas putida*)KS-0180(日本专利公开(Kokai)No.7-75589)将对甲酚转化为原儿茶酸的方法,使用NADH依赖性对羟基苯甲酸羟化酶将对羟基苯甲酸转化为原儿茶酸的方法(日本专利公开(Kokai)No.5-244941),通过在含有对苯二甲酸的培养基中培养含有参与从对苯二甲酸生成原儿茶酸的反应的基因的转化体来生产原儿茶酸的方法(日本专利公开(Kokai)No.2007-104942),以及通过使用具有原儿茶酸生产能力并且具有

原儿茶酸5-氧化酶活性降低或该活性缺乏的微生物从其前体生产原儿茶酸的方法(日本专利公开(Kokai)No.2010-207094)。此外,通过发酵生产原儿茶酸的方法的实例可以包括通过使用短杆菌属细菌和乙酸作为碳源生产原儿茶酸的方法(日本专利公开(Kokai)No.50-89592),通过使用引入有编码3-二氢莽草酸脱氢酶的基因的埃希氏菌属或克雷伯氏菌属细菌和葡萄糖作为碳源生产原儿茶酸的方法(美国专利号5,272,073)。此外,原儿茶醛可以通过根据使用ACAR的酶法或使用具有ACAR的微生物的生物转化法使用原儿茶酸作为前体生产。生产的前体可以原样或者根据需要进行适当的处理如浓缩,稀释,干燥,溶解,分馏,提取和纯化后用于生物转化方法。也就是说,作为前体,例如可以使用纯化至所需程度的纯化产物,或者可以使用含有前体的材料。含有前体的材料没有特别限制,只要微生物可以使用前体即可。含有前体的材料的具体实例可以包括通过培养具有前体生产能力的微生物获得的培养肉汤,与培养肉汤分离的培养上清液及其加工产物如其浓缩产物(如浓缩液体)和其干燥产物。

[0250] 在一个实施方案中,生物转化步骤可以通过例如培养如本文中所述的微生物来进行。该实施方案也可以称为“生物转化方法的第一实施方案”。也就是说,生物转化步骤可以是例如在含有目标物质的前体的培养基中培养微生物以将该前体转化为目标物质的步骤。生物转化步骤具体可以是在含有目标物质的前体的培养基中培养微生物以在培养基中产生并积累目标物质的步骤。

[0251] 要使用的培养基没有特别限制,只要培养基含有目标物质的前体,并且微生物可以在其中增殖并产生目标物质。培养条件没有特别限制,只要微生物可以增殖并且产生目标物质即可。关于发酵方法所提及的文化的描述,诸如关于培养基和培养条件的描述可以在必要修改后适用于生物转化方法的第一实施方案中的培养,只是培养基在第一实施方案中含有前体。

[0252] 前体可以在整个培养期内存在在培养基中,或者可以在培养的部分时期期间存在在培养基中。也就是说,短语“在含有前体的培养基中培养微生物”并不一定意味着培养基在整个培养期内都存在前体。例如,从培养开始起培养基中可以存在或可以不存在前体。当在培养开始时培养基中不存在前体时,在培养开始后将前体添加到培养基。添加的时机可以根据诸如培养期的长度等各种条件适当确定。例如,在微生物充分生长后,可以将前体添加到培养基中。此外,在任何情况下,可以根据需要将前体添加到培养基。例如,可以与伴随目标物质生成的前体的减少或消减成比例地对培养基添加前体。用于将前体添加到培养基的方法没有特别限制。例如,可以通过将含有前体的补料培养基补料到培养基来将前体供应给培养基。此外,例如,如本文中所述的微生物和具有前体生产能力的微生物可以共培养以允许具有前体生产能力的微生物在培养基中产生前体,从而将前体添加到培养基。这些添加方法可以独立使用,或者可以以适当的组合使用。前体在培养基中的浓度没有特别限制,只要微生物可以使用该前体作为目标物质的原料即可。前体在培养基中的浓度例如可以为0.1g/L或更高,1g/L或更高,2g/L或更高,5g/L或更高,10g/L或更高,或者15g/L或更高,或者可以是200g/L或更低,100g/L或更低,50g/L或更低,或者20g/L或更低,或者可以在由其组合限定的范围内(按游离化合物的重量计)。在整个培养期内,前体可以以或不以上文例示的范围内的浓度存在于培养基中。例如,在培养开始时前体可以以上文例示的范围内的浓度存在于培养基中,或者可以将其添加到培养基,从而在培养开始后获得上文例示

的范围。在培养分别作为种子培养和主要培养进行的情况下,至少在主要培养期间产生目标物质是足够的。因此,足够的是至少在主要培养期间,即在主要培养的整个时期内或在主要培养的部分时期期间,培养基中存在前体,并且也就是说,在种子培养期间培养基中可以存在或可以不存在前体。在此类情况下,关于培养的术语,如“培养期(培养时期)”和“培养开始”可以理解为与主要培养有关的术语。

[0253] 在另一个实施方案中,生物转化步骤也可以通过例如使用如本文中所述的微生物的细胞来进行。该实施方案也可以称为“生物转化方法的第二实施方案”。也就是说,生物转化步骤可以是例如通过使用微生物的细胞将反应混合物中的目标物质的前体转化成目标物质的步骤。生物转化步骤具体可以是使微生物的细胞作用于反应混合物中的目标物质的前体以在反应混合物中生成并积累目标物质的步骤。通过使用此类细胞进行的生物转化步骤也可以称为“转化反应”。

[0254] 微生物的细胞可以通过培养微生物来获得。用于获得细胞的培养方法没有特别限制,只要微生物可以增殖即可。在用于获得细胞的培养时,前体可以存在或可以不存在于培养基中。此外,在用于获得细胞的培养时,培养基中可以产生或可以不产生目标物质。关于发酵方法中所述的培养的描述,例如关于培养基和培养条件的描述可以在必要修改后应用于获得用于生物转化方法的第二实施方案的细胞的培养。

[0255] 细胞可以在存在于培养肉汤(具体是培养基)中时,或者在从培养肉汤(具体是培养基)中收集后用于转化反应。也可以根据需要在进行处理后将细胞用于转化反应。也就是说,细胞的实例可以包括含有细胞的培养肉汤,从培养肉汤中收集的细胞和其加工产物。换言之,细胞的实例可以包括存在于微生物的培养肉汤中的细胞,从培养肉汤中收集的细胞或存在于其加工产物中的细胞。加工产物的实例可以包括通过对细胞进行处理,特别地通过对含有细胞的培养肉汤或从培养肉汤中收集的细胞进行处理而获得的产物。这些形式的细胞可以独立使用,或者可以以适当的组合使用。

[0256] 从培养基中收集细胞的方法没有特别限制,例如可以使用已知的方法。此类方法的实例可以包括例如自发沉淀,离心和过滤。也可以使用絮凝剂。这些方法可以独立使用,或者可以以适当的组合使用。收集的细胞可以根据需要通过使用合适的培养基进行清洗。收集的细胞可以根据需要通过使用适当的培养基重悬浮。可用于清洗或悬浮细胞的培养基的实例可以包括例如水性介质(水性溶剂),如水和缓冲水溶液。

[0257] 细胞处理的实例可以包括例如稀释,冷凝,在载体如丙烯酰胺和角叉菜胶上固定化,冷冻和融化处理,以及增加细胞膜通透性的处理。例如,可以通过使用表面活性剂或有机溶剂来增加细胞膜的通透性。这些处理可以独立使用,或者可以以适当的组合使用。

[0258] 用于转化反应的细胞没有特别限制,只要细胞具有目标物质生产能力。优选细胞保持其代谢活性。短语“细胞保持其代谢活性”可以表示细胞具有利用碳源产生或再生产生目标物质所需物质的能力。此类物质的实例可以包括例如ATP,电子供体如NADH和NADPH,以及甲基供体如SAM。细胞可以具有或可以不具有增殖能力。

[0259] 转化反应可以在适当的反应混合物中进行。具体而言,转化反应可以通过使细胞和前体共存于适当的反应混合物中进行。转化反应可以通过分批方法进行或可以通过柱方法进行。在分批法的情况下,转化反应可以通过例如将微生物的细胞和前体混合在反应容器中所含的反应混合物中进行。转化反应可以静态进行,或者可以在搅拌或摇动反应

混合物的情况下进行。在柱方法的情况下,转化反应可以通过例如使含有前体的反应混合物通过填充有固定化细胞的柱来进行。反应混合物的实例可以包括基于水性介质(水性溶剂)如水和水性缓冲液的那些。

[0260] 除了前体之外,反应混合物可以根据需要含有除前体之外的组分。除前体以外的成分的实例可以包括ATP,电子供体诸如NADH,NADPH,甲基供体诸如SAM,金属离子,缓冲剂,表面活性剂,有机溶剂,碳源,磷酸盐源,和其它各种培养基成分。也就是说,例如,含有前体的培养基也可以用作反应混合物。也就是说,关于生物转化方法的第一实施方案中提及的培养基的描述也可以在必要修改后适用于生物转化方法的第二实施方案中的反应混合物。存在于反应混合物中的组分的类型和浓度可以根据各种条件来确定,例如待使用的前体的类型和待使用的细胞的形式。

[0261] 转化反应的条件,诸如溶解氧浓度,反应混合物的pH,反应温度,反应时间,各种成分的浓度等没有特别限制,只要产生目标物质。转化反应可以用例如用于利用微生物细胞诸如静止细胞进行物质转换的典型条件来进行。转化反应的条件可以根据各种条件来确定,例如选择的微生物的类型。转化反应可以例如在需氧条件下进行。术语“需氧条件”可以指其中反应混合物中的溶解氧浓度为0.33ppm或更高,或1.5ppm或更高的条件。氧浓度可以控制为例如饱和氧浓度的1-50%,或约5%。反应混合物的pH值通常可以是例如6.0至10.0,或6.5至9.0。反应温度可以是例如15至50℃,或15至45℃,或20至40℃。反应时间可以是例如5分钟至200小时。在柱方法的情况下,反应混合物的加载速率可以是例如使得反应时间落入上述例示的反应时间范围内的速率。此外,转化反应也可以在例如培养条件下进行,例如用于培养微生物诸如细菌和真菌的典型条件。在转化反应期间,细胞可以增殖或可以不增殖。也就是说,关于生物转化方法的第一实施方案的培养条件的描述也可以在必要修改后适用于生物转化方法的第二实施方案中转化反应的条件,只是细胞可以或不可以在第二实施方案中增殖。在此类情况下,用于获得细胞的培养条件和转化反应的条件可以相同或不同。前体在反应混合物中的浓度例如可以为0.1g/L或更高,1g/L或更高,2g/L或更高,5g/L或更高,10g/L或更高,或者15g/L或更高,或者可以是200g/L或更低,100g/L或更低,50g/L或更低,或者20g/L或更低,或者可以在由其组合限定的范围内(按游离化合物的重量计)。例如,反应混合物中细胞的密度可以是1或更高,或者可以是300或更低,或者可以在其组合限定的范围内(按600nm的光密度(OD)计)。

[0262] 在转化反应期间,可以将细胞,前体和其他组分独立地或以其任何任意组合添加到反应混合物。例如,可以与伴随目标物质产生的前体的减少或消耗成比例将前体添加到反应混合物。这些组分可以添加一次或多次,或者可以连续添加。

[0263] 用于向反应混合物添加诸如前体的各种组分的方法没有特别限制。这些组分各自可以通过例如直接将它们加入到反应混合物中而添加到反应混合物中。此外,例如,可以将如本文中所述的微生物和具有前体生产能力的微生物共培养,以使具有前体生产能力的微生物在反应混合物中产生前体,从而将前体提供给反应混合物。此外,例如,组分如ATP,电子给体和甲基供体可以各自在反应混合物中产生或再生,可以在微生物的细胞中产生或再生,或者可以通过不同细胞之间的偶联反应产生或再生。例如,当微生物的细胞保持其代谢活性时,它们可以通过使用碳源在它们内部产生或再生诸如ATP,电子供体和甲基供体等组分。例如,具体地,微生物可以具有增强的SAM生成或再生能力,并且其生成或再生的SAM可

以用于转化反应。SAM的生成或再生可以与用于生成或再生SAM的任何其他方法组合而得到进一步增强。另外,ATP生成或再生方法的实例可以包括例如通过使用棒状杆菌属细菌从碳源提供ATP的方法(Hori,H.et al.,Appl.Microbiol.Biotechnol.,48(6):693-698(1997)),通过使用酵母细胞和葡萄糖再生ATP的方法(Yamamoto,S et al.,Biosci.Biotechnol.Biochem.,69(4):784-789(2005)),使用磷酸烯醇丙酮酸和丙酮酸激酶再生ATP的方法(C.Aug'e and Ch.Gautheron,Tetrahedron Lett.,29:789-790(1988)),以及通过使用多磷酸和多磷酸激酶再生ATP的方法(Murata,K.et al.,Agric.Biol.Chem.,52(6):1471-1477(1988))。

[0264] 此外,反应条件可以从转化反应的开始至结束是恒定的,或者它们可以在转化反应期间改变。表述“反应条件在转化反应期间改变”不仅可以包括当反应条件在时间上改变时,而且还包括当反应条件在空间上改变时。表述“反应条件在空间上改变”表示例如当通过柱法进行转化反应时,反应温度和孔密度等反应条件随流动中的位置而不同。

[0265] 通过进行如上所述的生物转化步骤,获得培养肉汤(具体地培养基)或含有目标物质的反应混合物。目标物质的产生和目标物质的收集的确认可以以与上述发酵方法相同的方式进行。也就是说,生物转化方法可以进一步包括收集步骤,例如从培养肉汤(具体地培养基)或反应混合物中收集目标物质的步骤。除了目标物质之外,收集的目标物质可以包含例如微生物细胞,培养基组分,反应混合物组分,水分和微生物的副产物代谢物。收集的目标物质的纯度可以是,例如,30%(w/w)或更高,50%(w/w)或更高,70%(w/w)或更高,80%(w/w)或更高,90%(w/w)或更高,或95%(w/w)或更高。

[0266] <2-3>生产香草醛和其他目标物质的方法

[0267] 当通过使用如本文所述的微生物,即通过发酵方法或生物转化方法生产目标物质时,由此产生的目标物质可以进一步转化为另一种目标物质。因此,本发明提供了制备第二目标物质,即目标物质B的方法,所述方法包括以下步骤:通过使用微生物,即通过发酵方法或生物转化方法产生第一目标物质,即目标物质A,并且将由此产生的第一目标物质A转化为第二目标物质B。

[0268] 例如,当通过使用如本文所述的微生物,即通过发酵方法或生物转化方法生产香草酸时,由此产生的香草酸可以进一步转化为香草醛。因此,本发明提供了生产香草醛的方法,其包括通过使用微生物,即通过发酵方法或生物转化方法生产香草酸,并将由此产生的香草酸转化为香草醛的步骤。该方法也可以称为“香草醛生产方法”。

[0269] 通过使用微生物产生的香草酸可以原样用于转化成香草醛,或者在根据需要进行适当的处理如浓缩,稀释,干燥,溶解,分馏,提取和纯化之后用于转化成香草醛。也就是说,作为香草酸,例如,可以使用纯化至期望程度的纯化产物,或者可以使用含有香草酸的物质。含有香草酸的物质没有特别限制,只要催化转化的组分(例如微生物和酶)可以使用香草酸。含有香草酸的材料的实例可以包括含有香草酸的培养肉汤或反应混合物,从培养肉汤或反应混合物中分离的上清液,及其加工产物,如浓缩产物,如其浓缩液及其干燥产物。

[0270] 将香草酸转化为香草醛的方法没有特别限制。

[0271] 可以通过例如使用具有ACAR的微生物的生物转化方法将香草酸转化为香草醛。具有ACAR的微生物可以以NCg12048蛋白的活性得以降低的方式或者可以不以NCg12048蛋白

的活性得以降低的方式进行修饰。关于如本文所述的微生物的描述可以在必要的修改后适用于具有ACAR的微生物,只是具有ACAR的微生物可以以使NCg12048蛋白的活性得以降低的方式进行修饰或可以不以使NCg12048蛋白的活性得以降低的方式进行修饰。具有ACAR的微生物可以以ACAR、PPT和香草酸摄取系统中的一种或多种的活性得以增强的方式进行修饰。另外,关于使用微生物生产目标物质的生物转化方法的描述可以在必要的修改后适用于使用具有ACAR的微生物将香草酸转化为香草醛的生物转化方法。

[0272] 也可以例如通过使用ACAR的酶方法将香草酸转化成香草醛。

[0273] ACAR可以通过允许具有ACAR基因的宿主表达ACAR基因来生成。ACAR也可以用无细胞蛋白质表达系统来生成。

[0274] 具有ACAR基因的宿主也可以称为“具有ACAR的宿主”。具有ACAR基因的宿主可以是固有地具有ACAR基因的宿主,或者可以是经修饰而具有ACAR基因的宿主。固有地具有ACAR基因的宿主的实例可以包括衍生出上述示例的ACAR的生物体。经修饰而具有ACAR基因的宿主的实例可以包括已经引入有ACAR基因的宿主。此外,固有地具有ACAR基因的宿主可以以ACAR得以增加的方式进行修饰。用于表达ACAR的宿主不受特别限制,只要宿主可以表达可以起作用的ACAR即可。宿主的实例可以包括例如,微生物,如细菌和酵母(真菌),植物细胞,昆虫细胞和动物细胞。

[0275] 可以通过培养具有ACAR基因的宿主来表达ACAR基因。培养方法没有特别限制,只要具有ACAR基因的宿主可以增殖并表达ACAR即可。关于发酵方法培养的描述可以在必要的修改后适用于具有ACAR基因的宿主的培养。必要时,可以诱导ACAR基因的表达。作为培养的结果,可以获得含有ACAR的培养肉汤。ACAR可以在宿主细胞和/或培养基中积累。

[0276] 在宿主细胞,培养基等中包含的ACAR可以原样用于酶促反应,或者自其纯化的ACAR可以用于酶促反应。可以进行纯化到期望的程度。也就是说,作为ACAR,可以使用纯化的ACAR,或者可以使用含有ACAR的级分。此类级分没有特别限制,只要其中含有的ACAR可以作用于香草酸即可。此类级分的实例可以包括具有ACAR基因的宿主的培养肉汤,即具有ACAR的宿主;从培养肉汤中收集细胞;细胞的加工产物,如细胞破碎物,细胞裂解物,细胞提取物和固定化细胞,如用丙烯酰胺,角叉菜胶等固定化的细胞;从培养肉汤中收集的培养物上清液;其部分纯化产物,例如粗产物;及其组合。这些级分可以独立使用或与纯化的ACAR组合使用。

[0277] 酶促反应可以通过允许ACAR作用于香草酸来进行。酶促反应的条件没有特别限制,只要产生香草醛即可。酶促反应可以用例如用于使用酶或微生物细胞如静息细胞进行物质转化的典型条件进行。例如,关于生物转化方法的第二实施方案中的转化反应的描述也可以在必要的修改后适用于香草醛生产方法中的酶促反应。

[0278] 通过进行如上所述的转化获得含有香草醛的反应混合物。香草醛的产生和香草醛的收集的确认可以以与上述发酵方法相同的方式进行。也就是说,香草醛生产方法可以进一步包括从反应混合物中收集香草醛的步骤。除香草醛外,收集的香草醛可以含有例如微生物细胞,培养基组分,反应混合物组分,ACAR,水分和微生物的副产物代谢物。收集的香草醛的纯度可以是例如30% (w/w)或更高,50% (w/w)或更高,70% (w/w)或更高,80% (w/w)或更高,90% (w/w)或更高,或95% (w/w)更高。

[0279] 香草酸也可以通过例如使用具有VDC的微生物的生物转化方法或使用VDC的酶促

方法转化为愈创木酚。阿魏酸可以通过例如使用具有FDC的微生物的生物转化方法或使用FDC的酶促方法转化为4-乙基愈创木酚。4-乙基愈创木酚可以通过例如使用具有VPR的微生物的生物转化方法或使用VPR的酶促方法转化为4-乙基愈创木酚。通过这些方法的组合,阿魏酸也可以转化为4-乙基愈创木酚。具体地,阿魏酸可以通过例如与VPR或具有VPR的微生物组合同时或依次使用FDC或具有FDC的微生物,或使用具有FDC和VPR两者的微生物转化为4-乙基愈创木酚。关于香草醛生产方法的上述描述可以在必要的修改后适用于生产其他目标物质的方法。

实施例

[0280] 下文,将参考以下非限制性实施例更具体地解释本发明。

[0281] 在本实施例中,从作为亲本菌株的谷氨酸棒杆菌2256菌株(ATCC 13869)构建具有NCgl2048基因的弱化表达的菌株,并且用构建的菌株进行香草酸生产。

[0282] <1>香草酸脱甲基酶基因缺陷的菌株(FKS0165菌株)的构建

[0283] 已经报道了在棒状杆菌细菌中,香草醛按香草醛→香草酸→原儿茶酸的顺序代谢,并利用(Current Microbiology,2005,Vol.51,pp.59-65)。从香草酸到原儿茶酸的转化反应由香草酸脱甲基酶催化。vanA基因和vanB基因分别编码香草酸脱甲基酶的亚基A和亚基B。vanK基因编码香草酸摄取系统,并与vanAB基因一起构成vanABK操纵子(M.T.Chaudhry,et al.,Microbiology,2007,153:857-865)。因此,首先通过缺失vanABK操纵子从谷氨酸棒杆菌2256菌株构建目标物质诸如香草醛和香草醛利用能力缺陷的菌株(FKS0165菌株)。程序如下所示。

[0284] <1-1>用于缺失vanABK基因的质粒pBS4S Δ vanABK56的构建

[0285] 通过使用谷氨酸棒杆菌2256菌株的基因组DNA作为模板,和SEQ ID NO:51和52的合成DNA作为引物进行PCR,以获得含有vanA基因的N端侧编码区的PCR产物。分别通过使用谷氨酸棒杆菌2256菌株的基因组DNA作为模板和SEQ ID NO:53和54的合成DNA作为引物进行PCR,以获得含有vanK基因的C端侧编码区的PCR产物。SEQ ID NO:52和53的序列彼此部分互补。然后,以近似等摩尔量混合含有vanA基因的N端侧编码区的PCR产物和含有vanK基因的C端侧编码区的PCR产物,并通过使用In Fusion HD克隆试剂盒(Clontech)将其插入用BamHI和PstI处理的pBS4S载体(W02007/046389)中。利用该DNA,转化大肠杆菌JM109的感受态细胞(Takara Bio),并且将细胞应用于含有100μM IPTG,40μg/mL X-Gal和40μg/mL卡那霉素的LB培养基,并培养过夜。然后,挑取出现的白色菌落,并且分离成单菌落以获得转化体。从所获得的转化体中提取质粒,并将插入了靶PCR产物的质粒命名为pBS4S Δ vanABK56。

[0286] <1-2>FKS0165菌株的构建

[0287] 上面获得的pBS4S Δ vanABK56不含能使质粒在棒状杆菌细菌细胞中自主复制的区域。因此,若用该质粒转化棒状杆菌细菌,则通过同源重组将该质粒掺入基因组中的菌株作为转化体出现,尽管它以非常低的频率发生。因此,通过电脉冲法将pBS4S Δ vanABK56导入谷氨酸棒杆菌2256菌株。将细胞应用于含有25μg/mL卡那霉素的CM-Dex琼脂培养基(5g/L葡萄糖、10g/L聚脲、10g/L酵母提取物、1g/L KH_2PO_4 、0.4g/L $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 、0.01g/L $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ 、0.01g/L $\text{MnSO}_4 \cdot 7\text{H}_2\text{O}$ 、3g/L尿素、1.2g/L大豆水解物、10μg/L生物素、15g/L琼脂,用NaOH调节到pH 7.5),并在31.5℃培养。通过PCR确认生长的菌株是一次重组菌株,其中通过

同源重组将pBS4S Δ vanABK56掺入基因组中。此一次重组的菌株具有野生型vanABK基因和缺陷型vanABK基因两者。

[0288] 将一次重组菌株在CM-Dex液体培养基(与CM-Dex琼脂培养基具有相同的组成,只是它不含琼脂)中培养过夜,并将该培养肉汤应用于S10琼脂培养基(100g/L蔗糖,10g/L聚脲,10g/L酵母提取物,1g/L KH_2PO_4 ,0.4g/L $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$,0.01g/L $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$,0.01g/L $\text{MnSO}_4 \cdot 4\text{H}_2\text{O}$,3g/L尿素,1.2g/L大豆蛋白水解物溶液,10 μg /L生物素,20 μg /L琼脂,用NaOH调节至pH7.5,并在120 $^\circ\text{C}$ 高压灭菌20分钟),并在31.5 $^\circ\text{C}$ 培养。在出现的菌落中,在CM-Dex琼脂培养基上纯化显示卡那霉素易感性的菌株。通过从纯化的菌株中制备基因组DNA,并使用它用SEQ ID NO:55和56的合成DNA作为引物进行PCR,确认了vanABK基因的缺失,并将该菌株命名为FKS0165菌株。

[0289] <2>醇脱氢酶同源基因缺陷菌株(FKFC14菌株)的构建

[0290] 随后,通过使用谷氨酸棒杆菌菌株FKS0165作为亲本株,通过以下程序构建菌株FKFC14,其是醇脱氢酶同源基因(即NCg10324基因(adhC)、NCg10313基因(adhE)、和NCg12709 gene(adhA))缺陷的。

[0291] <2-1>FKFC5菌株(FKS0165 Δ NCg10324菌株)的构建

[0292] <2-1-1>用于缺失NCg10324基因的质粒pBS4S Δ 2256adhC的构建

[0293] 通过使用谷氨酸棒杆菌2256菌株的基因组DNA作为模板,和SEQ ID NO:57和58的合成DNA作为引物进行PCR,以获得含有NCg10324基因的N端侧编码区的PCR产物。分别通过使用谷氨酸棒杆菌2256菌株的基因组DNA作为模板,和SEQ ID NO:59和60的合成DNA作为引物进行PCR,以获得含有NCg10324基因的C端侧编码区的PCR产物。SEQ ID NO:58和59的序列彼此部分互补。然后,将约等摩尔量的含有NCg10324基因的N端侧编码区的PCR产物和含有NCg10324基因的C端侧编码区的PCR产物混合,并通过使用In Fusion HD克隆试剂盒(Clontech)插入已经用BamHI和PstI处理的pBS4S载体(W02007/046389)中。利用该DNA,转化大肠杆菌JM109的感受态细胞(Takara Bio),并且将细胞应用于含有100 μM IPTG,40 μg /mL X-Gal和40 μg /mL卡那霉素的LB培养基,并培养过夜。然后,挑取出现的白色菌落,分离成单菌落以获得转化体。从所获得的转化体中提取质粒,并将其中插入了靶PCR产物的质粒命名为pBS4S Δ 2256adhC。

[0294] <2-1-2>FKFC5菌株(FKS0165 Δ NCg10324菌株)的构建

[0295] 由于上述获得的pBS4S Δ 2256adhC不含使质粒能够在棒状杆菌细菌细胞中自主复制的区域,若用该质粒转化棒状杆菌细菌,则通过同源重组将该质粒掺入基因组中的菌株作为转化体出现,尽管它以非常低的频率发生。因此,通过电脉冲法将pBS4S Δ 2256adhC导入谷氨酸棒杆菌FKS0165菌株。将细胞应用于含有25 μg /mL卡那霉素的CM-Dex琼脂培养基,并在31.5 $^\circ\text{C}$ 培养。通过PCR确认,生长的菌株是通过同源重组将pBS4S Δ 2256adhC掺入基因组中的一次重组菌株。此一次重组菌株具有野生型NCg10324基因和缺陷型NCg10324基因两者。

[0296] 将一次重组的菌株在CM-Dex液体培养基中培养过夜,将培养肉汤应用到S10琼脂培养基,并在31.5 $^\circ\text{C}$ 进行培养。在出现的菌落中,在CM-Dex琼脂培养基上纯化显示卡那霉素易感性的菌株。从纯化的菌株制备基因组DNA,并用SEQ ID NO:61和62的合成DNA作为引物进行PCR以确认NCg10324基因的缺失,并将该菌株命名为FKFC5菌株。

[0297] <2-2>FKFC11菌株(2256 Δ vanABK Δ NCg10324 Δ NCg10313菌株)的构建

[0298] <2-2-1>用于缺失NCg10313基因的质粒pBS4S Δ 2256adhE的构建

[0299] 通过使用谷氨酸棒杆菌2256菌株的基因组DNA作为模板和SEQ ID NO:63和64的合成DNA作为引物进行PCR,以获得含有NCg10313基因的N端侧编码区的PCR产物。分别通过使用谷氨酸棒杆菌2256菌株的基因组DNA为模板和SEQ ID NO:65和66的合成DNA为引物进行PCR,以获得含有NCg10313基因的C端侧编码区的PCR产物。SEQ ID NO:64和65的序列彼此部分互补。然后,将约等摩尔量的含有NCg10313基因的N端侧编码区的PCR产物和含有NCg10313基因的C端侧编码区的PCR产物混合,并通过使用In Fusion HD克隆试剂盒(Clontech)将其插入用BamHI和PstI处理的pBS4S载体(W02007/046389)中。利用该DNA,转化大肠杆菌JM109的感受态细胞(Takara Bio),并且将细胞应用于含有100 μ M IPTG,40 μ g/mL X-Gal和40 μ g/mL卡那霉素的LB培养基,并培养过夜。然后,挑取出现的白色菌落,并且分离成单菌落以获得转化体。从获得的转化体中提取质粒,并将其中插入了靶PCR产物的质粒命名为pBS4S Δ 2256adhE。

[0300] <2-2-2>FKFC11菌株(2256 Δ vanABK Δ NCg10324 Δ NCg10313株)的构建

[0301] 由于上述获得的pBS4S Δ 2256adhE不含有使质粒能够在棒状杆菌细菌细胞中的自主复制的区域,若用该质粒转化棒状杆菌,则通过同源重组将该质粒掺入基因组中的菌株作为转化体出现,尽管它在非常低的频率发生。因此,通过电脉冲方法将pBS4S Δ 2256adhE导入谷氨酸棒杆菌FKFC5菌株。将细胞应用于含有25 μ g/mL卡那霉素的CM-Dex琼脂培养基,并在31.5 $^{\circ}$ C下培养。通过PCR证实,生长的菌株是一次重组菌株,其中通过同源重组将pBS4S Δ 2256adhE掺入基因组中。此一次重组菌株具有野生型NCg10313基因和缺陷型NCg10313基因两者。

[0302] 将一次重组的菌株在CM-Dex液体培养基中培养过夜,将培养肉汤应用到S10琼脂培养基上,并在31.5 $^{\circ}$ C进行培养。在出现的菌落中,显示卡那霉素易感性的菌株在CM-Dex琼脂培养基上纯化。从纯化的菌株制备基因组DNA,并使用SEQ ID NO:67和68的合成DNA作为引物进行PCR以确认NCg10313基因的缺失,并将该菌株命名为FKFC11菌株。

[0303] <2-3>FKFC14菌株(2256 Δ vanABK Δ NCg10324 Δ NCg10313 Δ NCg12709株)的构建

[0304] <2-3-1>用于缺失NCg12709基因的质粒pBS4S Δ 2256adhA的构建

[0305] 通过使用谷氨酸棒杆菌2256菌株的基因组DNA作为模板和SEQ ID NO:69和70的合成DNA作为引物进行PCR以获得含有NCg12709基因的N端侧编码区的PCR产物。分别通过使用谷氨酸棒杆菌2256菌株的基因组DNA为模板和SEQ ID NO:71和72的合成DNA作为引物进行PCR,以得到含有NCg12709基因的C端侧编码区的PCR产物。SEQ ID NO:70和71的序列彼此部分互补。然后,将约等摩尔量的含有NCg12709基因的N端侧编码区的PCR产物和含有NCg12709基因的C端侧编码区的PCR产物混合,并通过使用In Fusion HD克隆试剂盒(Clontech)插入到用BamHI和PstI处理的pBS4S载体中。利用该DNA,转化大肠杆菌JM109的感受态细胞(Takara Bio),并且将细胞应用于含有100 μ M IPTG,40 μ g/mL X-Gal和40 μ g/mL卡那霉素的LB培养基,并培养过夜。然后,挑取出现的白色菌落,并且分离成单菌落以获得转化体。从获得的转化体中提取质粒,并将其中插入了靶PCR产物的质粒命名为pBS4S Δ 2256adhA。

[0306] <2-3-2>FKFC14菌株(2256 Δ vanABK Δ NCg10324 Δ NCg10313 Δ NCg12709菌株)的构

建

[0307] 由于上文获得的pBS4S Δ 2256adhA不含能够使质粒在棒状杆菌细菌细胞中自主复制的区域,若用该质粒转化棒状杆菌,则通过同源重组将该质粒掺入基因组的菌株作为转化体出现,尽管它在非常低的频率发生。因此,通过电脉冲法将pBS4S Δ 2256adhA导入谷氨酸棒杆菌FKFC11菌株。将细胞应用于含有25 μ g/mL卡那霉素的CM-Dex琼脂培养基,并在31.5 $^{\circ}$ C培养。通过PCR确认,生长的菌株是通过同源重组将pBS4S Δ 2256adhA掺入基因组中的一次重组菌株。该一次重组菌株具有野生型NCg12709基因和缺陷型NCg12709基因两者。

[0308] 将一次重组的菌株在CM-Dex液体培养基中培养过夜,将培养肉汤应用到S10琼脂培养基上,并在31.5 $^{\circ}$ C进行培养。在出现的菌落中,显示卡那霉素易感性的菌株在CM-Dex琼脂培养基上纯化。从纯化的菌株制备基因组DNA,并用SEQ ID NO:73和74的合成DNA作为引物进行PCR以确认NCg12709基因的缺失,并将该菌株命名为FKFC14菌株。

[0309] <3>原儿茶酸双加氧酶基因缺陷菌株(FKFC14 Δ pcaGH菌株)的构建

[0310] 随后,通过使用谷氨酸棒杆菌FKFC14菌株作为亲本菌株,通过外包构建菌株FKFC14 Δ pcaGH,其是编码原儿茶酸3,4-双加氧酶的 α 亚基和 β 亚基的NCg12314基因(pcaG)和NCg12315基因(pcaH)缺陷的。也可以经由以下程序构建FKFC14 Δ pcaGH菌株。

[0311] <3-1>用于缺失NCg12314和NCg12315基因的质粒pBS4S Δ 2256pcaGH的构建

[0312] NCg12314和NCg12315基因彼此相邻,因此可以一起缺失这些基因。通过使用谷氨酸棒杆菌2256菌株的基因组DNA作为模板,和SEQ ID NO:75和76的合成DNA作为引物进行PCR以获得含有NCg12315基因的上游区的PCR产物。分别通过使用谷氨酸棒杆菌2256菌株的基因组DNA作为模板和SEQ ID NO:77和78的合成DNA作为引物进行PCR,以获得含有NCg12314基因的下游区的PCR产物。SEQ ID NO:76和77的序列彼此部分互补。然后,混合近似等摩尔量的含有NCg12315基因的上游区的PCR产物和含有NCg12314基因的下游区的PCR产物,并通过使用In Fusion HD克隆试剂盒(Clontech)将其插入用BamHI和PstI处理的pBS4S载体(W02007/046389)中。利用该DNA,转化大肠杆菌JM109的感受态细胞(Takara Bio),并且将细胞应用于含有100 μ M IPTG,40 μ g/mL X-Gal和40 μ g/mL卡那霉素的LB培养基,并培养过夜。然后,挑取出现的白色菌落,并且分离成单菌落以获得转化体。从所获得的转化体中提取质粒,并将插入了靶PCR产物的质粒命名为pBS4S Δ 2256pcaGH。

[0313] <3-2>FKFC14 Δ pcaGH菌株的构建

[0314] 由于上面获得的pBS4S Δ 2256pcaGH不含能使质粒在棒状杆菌细菌细胞中自主复制的区域,若用该质粒转化棒状杆菌细菌,则通过同源重组将该质粒掺入基因组中的菌株作为转化体出现,尽管它以非常低的频率发生。因此,通过电脉冲法将pBS4S Δ 2256pcaGH导入谷氨酸棒杆菌FKFC14菌株。将细胞应用于含有25 μ g/mL卡那霉素的CM-Dex琼脂培养基,并在31.5 $^{\circ}$ C培养。通过PCR确认生长的菌株是一次重组菌株,其中通过同源重组将pBS4S Δ 2256pcaGH掺入基因组中。此一次重组的菌株具有野生型NCg12314和NCg12315基因,和缺陷型NCg12314和NCg12315基因两者。

[0315] 将一次重组菌株在CM-Dex液体培养基中培养过夜,并将该培养基应用于S10琼脂培养基,并在31.5 $^{\circ}$ C培养。在出现的菌落中,在CM-Dex琼脂培养基上纯化显示卡那霉素易感性的菌株。通过从纯化的菌株中制备基因组DNA,并使用它用SEQ ID NO:79和80的合成DNA作为引物进行PCR,确认了NCg12314和NCg12315基因的缺失,并将该菌株命名为FKFC14 Δ

pcaGH菌株。

[0316] <4>Ap1_0112菌株(FKFC14 Δ pcaGH P8::NCg12048菌株)的构建

[0317] 随后,通过使用谷氨酸棒杆菌FKFC14 Δ pcaGH菌株作为亲本菌株,通过外包构建菌株Ap1_0112,其中NCg12048基因的启动子区已经用P8启动子替换。此菌株中含有P8启动子的基因组区域的核苷酸序列以SEQ ID NO:83显示,其中位置901-1046对应于P8启动子。也可以经由以下程序构建Ap1_0112菌株。

[0318] <4-1>用于替换NCg12048基因启动子的质粒pBS4SP8::NCg12048的构建

[0319] 通过使用谷氨酸棒杆菌2256菌株的基因组DNA作为模板,和SEQ ID NO:85和86的合成DNA作为引物进行PCR以获得含有NCg12048基因的上游区的PCR产物。分别通过使用谷氨酸棒杆菌2256菌株的基因组DNA作为模板和SEQ ID NO:87和88的合成DNA作为引物进行PCR,以获得含有NCg12048基因的N端侧编码区的PCR产物。另外,通过人工基因合成获得含有P8启动子区的SEQ ID NO:89的DNA片段。然后,通过使用SEQ ID NO:89的DNA片段作为模板,和SEQ ID NO:90和91的合成DNA作为引物进行PCR,以获得含有P8启动子的PCR产物。SEQ ID NO:86和90的序列彼此部分互补,并且SEQ ID NO:87和91的序列彼此部分互补。然后,混合近似等摩尔量的含有NCg12048基因的上游区的PCR产物、含有NCg12048基因的N端侧编码区的PCR产物和含有P8启动子的PCR产物,并通过使用In Fusion HD克隆试剂盒(Clontech)将其插入用BamHI和PstI处理的pBS4S载体(W02007/046389)中。利用该DNA,转化大肠杆菌JM109的感受态细胞(Takara Bio),并且将细胞应用于含有100 μ M IPTG,40 μ g/mL X-Gal和40 μ g/mL卡那霉素的LB培养基,并培养过夜。然后,挑取出现的白色菌落,并且分离成单菌落以获得转化体。从所获得的转化体中提取质粒,并将插入了靶PCR产物的质粒命名为pBS4SP8::NCg12048。

[0320] <4-2>Ap1_0112菌株的构建

[0321] 由于上面获得的pBS4SP8::NCg12048不含能使质粒在棒状杆菌细菌细胞中自主复制的区域,若用该质粒转化棒状杆菌细菌,则通过同源重组将该质粒掺入基因组中的菌株作为转化体出现,尽管它以非常低的频率发生。因此,通过电脉冲法将pBS4SP8::NCg12048导入谷氨酸棒杆菌FKFC14 Δ pcaGH菌株。将细胞应用于含有25 μ g/mL卡那霉素的CM-Dex琼脂培养基,并在31.5 $^{\circ}$ C培养。通过PCR确认生长的菌株是一次重组菌株,其中通过同源重组将pBS4SP8::NCg12048掺入基因组中。

[0322] 将一次重组菌株在CM-Dex液体培养基中培养过夜,并将该培养基应用于S10琼脂培养基,并在31.5 $^{\circ}$ C培养。在出现的菌落中,在CM-Dex琼脂培养基上纯化显示卡那霉素易感性的菌株。通过从纯化的菌株中制备基因组DNA,并用于进行核苷酸序列分析以确认P8启动子位于NCg12048基因上游,并且该菌株命名为Ap1_0112菌株。

[0323] <5>用于表达智人OMT基因的质粒pVK9::PcspB-hsomt的构建

[0324] <5-1>质粒pEPlac-COMT2的构建

[0325] 对于智人OMT基因已知两种OMT同种型,即较短的OMT同种型(S-COMT)和较长的OMT同种型(MB-COMT)。S-COMT的氨基酸序列以SEQ ID NO:16显示,并且编码S-COMT的野生型cDNA的核苷酸序列以SEQ ID NO:94显示。对S-COMT的野生型cDNA进行了密码子优化以在大肠杆菌中表达并且使用由ATG Service Gen(Russian Federation,Saint-Petersburg)提供的服务进行化学合成。为了便于进一步克隆,合成在3'和5'端分别具有限制酶NdeI和

SacI的位点的基因的DNA片段。经密码子优化的S-COMT cDNA也可以称为COMT2基因。含有COMT2基因的合成DNA片段的核苷酸序列以SEQ ID NO:95显示。在pUC57载体(GenScript)中获得包含COMT2基因的合成DNA片段。

[0326] 在T7系统中确认COMT2基因的表达。将pUC57载体中插入的COMT2基因再克隆到pET22(+)载体(Novagen)的NdeI和SacI限制性位点中。将获得的质粒导入大肠杆菌BL21(DE3)细胞(Novagen)中。在含有200mg/l氨苄青霉素的LB培养基(胰蛋白胨,10g/l;酵母提取物,5g/l;NaCl,10g/l)中培养含有质粒的细胞,并且在指数生长期中在2小时内通过IPTG,1mM诱导。通过超声处理破坏细胞。使用12%SDS-PAGE中的电泳分析粗制蛋白质提取物。鉴定出对应于S-OMT的条带(约24kDa),并且从凝胶中切出。从凝胶中分离目标蛋白,并且用胰蛋白酶处理。使用质谱法分析获得的胰蛋白酶水解物以确认COMT2基因表达。

[0327] 将pUC57载体中插入的COMT2基因再克隆到pELAC载体的NdeI和SacI限制性位点中(SEQ ID NO:96,Smirnov S.V.et al.,Appl.Microbiol.Biotechnol.,2010,88(3):719-726)。通过用含有P_{lacUV5}启动子的合成BglIII-XbaI片段替换pET22b(+) (Novagen)的BglIII-XbaI片段构建pELAC载体。为了将COMT2基因插入pELAC载体中,使用T4DNA连接酶(Fermentas,Lithuania)如供应商推荐的那样进行连接反应。用乙醇处理连接混合物,并且将获得的沉淀物溶于水中,并且使用电穿孔(Micro Pulser,BioRad)在由供应商推荐的条件下导入大肠杆菌TG1细胞中。将细胞应用到补充有氨苄青霉素(200mg/L)的LA板(Sambrook J.and Russell D.W.,Molecular Cloning:A Laboratory Manual(第3版),Cold Spring Harbor Laboratory Press,2001)上,并且于37℃培养过夜。使用PCR分析测试获得的菌落以选择需要的克隆。使用引物P1和P2(SEQ ID NO:97和98)来选择含有COMT2基因的菌落。当使用载体特异性引物P1和用于终止COMT2基因的反向引物P2时获得DNA片段(713bp)。因此,构建载体pEPlac-COMT2。使用引物P1和P3(SEQ ID NO:97和99)测定克隆的COMT2基因的序列。

[0328] <5-2>质粒pVK9::PcspB-hsomt的构建

[0329] 通过使用谷氨酸棒杆菌2256菌株的基因组DNA作为模板,和SEQ ID NO:100和101的合成DNA作为引物进行PCR以获得含有cspB基因的启动子区和SD序列的PCR产物。分开地,也通过使用质粒pEPlac-COMT2作为模板和SEQ ID NO:102和103的合成DNA作为引物进行PCR以获得含有COMT2基因的PCR产物。然后,通过使用In Fusion HD克隆试剂盒(Clontech)将这些PCR产物插入用BamHI和PstI处理的pVK9载体(W02007/046389)中。pVK9载体是棒状杆菌细菌和大肠杆菌的穿梭载体。用此DNA转化大肠杆菌JM109的感受态细胞(Takara Bio),并且将细胞应用于含有100μM IPTG、40μg/mL X-Gal和25μg/mL卡那霉素的LB培养基,并且培养过夜。然后,挑出出现的白色菌落,并且分成单菌落以获得转化体。从获得的转化体提取质粒,并且插入有靶PCR产物的质粒命名为pVK9::PcspB-hsomt。

[0330] <6>香草酸生产菌株的构建

[0331] 通过外包构建谷氨酸棒杆菌FKFC14 ΔpcaGH/pVK9::PcspB-hsomt和Ap1_0112/pVK9::PcspB-hsomt菌株(其含有质粒pVK9::PcspB-hsomt)。也可以经由以下程序构建这些菌株。

[0332] 通过电脉冲方法将质粒pVK9::PcspB-hsomt导入谷氨酸棒杆菌FKFC14 ΔpcaGH和Ap1_0112菌株中。将细胞应用于含有25μg/mL卡那霉素的CM-Dex琼脂培养基中,并且于31.5

℃培养。在相同的琼脂培养基上纯化生长的菌株,并且分别命名为FKFC14 Δ pcaGH/pVK9::PcspB-hsomt和Ap1_0112/pVK9::PcspB-hsomt。

[0333] 将这些菌株各自接种到试管中存在的4mL含有25 μ g/mL卡那霉素的不含mameno的CM-Dex培养基(5g/L葡萄糖、10g/L聚脲、10g/L酵母提取物、1g/L KH_2PO_4 、0.4g/L $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 、0.01g/L $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ 、0.01g/L $\text{MnSO}_4 \cdot 7\text{H}_2\text{O}$ 、3g/L尿素、10 μ g/L生物素,用KOH调节到pH 7.5)中,并且于31.5℃在摇动的情况下培养约16小时。将获得的培养肉汤的0.9mL等分试样与0.6mL 50%甘油水溶液混合以获得甘油储液,并且于-80℃贮存。

[0334] <7>谷氨酸棒杆菌FKFC14 Δ pcaGH/pVK9::PcspB-hsomt和Ap1_0112/pVK9::PcspB-hsomt菌株的香草酸生产

[0335] 将FKFC14 Δ pcaGH/pVK9::PcspB-hsomt和Ap1_0112/pVK9::PcspB-hsomt菌株的甘油储液各5 μ L等分试样接种到试管中存在的4mL含有25 μ g/mL卡那霉素的不含mameno的CM-Dex培养基,并且于31.5℃在摇动的情况下培养20小时作为预培养物。将获得的预培养物肉汤的0.5mL等分试样接种到具有挡板的锥形瓶中存在的50mL含有25 μ g/mL卡那霉素的不含mameno的CM-Dex培养基,并且于31.5℃在摇动的情况下培养20小时。以8000rpm将获得的培养物肉汤离心5分钟,除去上清液,并且在经灭菌的生理盐水中悬浮细胞。测量细胞悬浮液的光密度(OD),并且用生理盐水稀释细胞悬浮液以获得600nm处的OD为50。将稀释的细胞悬浮液的5mL等分试样接种到具有挡板的锥形瓶中存在的20mL含有25 μ g/mL卡那霉素的香草酸生产培养基(75g/L葡萄糖、0.6g/L $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 、6.3g/L $(\text{NH}_4)_2\text{SO}_4$ 、2.5g/L KH_2PO_4 、12.5mg/L $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ 、12.5mg/L $\text{MnSO}_4 \cdot 4\text{H}_2\text{O}$ 、2.5g/L酵母提取物、150 μ g/L维生素B1、150 μ g/L生物素、6.9g/L原儿茶酸,用KOH调节到pH 7,然后与37.5g/L CaCO_3 (用180℃的热空气灭菌3小时)混合)中,并且于31.5℃在摇动的情况下培养24小时。

[0336] 在培养开始和完成时,用Biotech分析仪AS-310(Sakura SI)分析培养基中的葡萄糖浓度。也通过使用超高效液相层析NEXERA X2系统(SHIMADZU)用以下条件分析培养基中的原儿茶酸和香草酸浓度。

[0337] UPLC分析条件:

[0338] 柱:KINETEX 2.6 μ m XB-C18,150x 30mm(Phenomenex)

[0339] 炉温度:40℃

[0340] 流动相(A):0.1%三氟乙酸

[0341] 流动相(B):0.1%三氟乙酸/80%乙腈

[0342] 梯度程度(时间,A(%),B(%)):(0,90,10)→(3,80,20)

[0343] 流速:1.5ml/min

[0344] 表1中显示了结果。对Ap1_0112/pVK9::PcspB-hsomt菌株观察到的培养基中的香草酸浓度是对FKFC14 Δ pcaGH/pVK9::PcspB-hsomt菌株观察到的香草酸浓度的约1.2倍。

[0345] [表1]

[0346] 表1:谷氨酸棒杆菌香草酸生产菌株的香草酸生产

[0347]

		在培养开始时	
菌株	葡萄糖浓度 (g/L)	原儿茶酸浓度 (g/L)	
FKFC14ΔpcaGH/ pVK9::PcspB-hsomt	57.8 ± 0.5	5.70 ± 0.2	
Ap1_0112/ pVK9::PcspB-hsomt	61.3 ± 0.3	5.67 ± 0.1	

[0348]

		在培养完成时	
菌株	残留葡萄糖的 浓度(g/L)	残留原儿茶酸浓 度 (g/L)	生成的香草酸 浓度 (mg/L)
FKFC14ΔpcaGH/ pVK9::PcspB-hsomt	13.0 ± 0.2	5.64 ± 0.1	73.0 ± 1.8
Ap1_0112/ pVK9::PcspB-hsomt	19.3 ± 1.2	5.66 ± 0.2	88.1 ± 5.0

[0349] <8>通过定量PCR分析NCg12048基因的表达量

[0350] 随后,通过定量PCR分析FKFC14 Δ pcaGH/pVK9::PcspB-hsomt和Ap1_0112/pVK9::PcspB-hsomt菌株中的NCg12048基因的表达量。

[0351] <8-1>RNA的制备

[0352] 将含有细胞的培养肉汤的250 μ L等分试样与500 μ L RNA Protect Bacteria Reagent (QIAGEN) 混合,并且于-80 $^{\circ}$ C贮存,所述培养肉汤对于FKFC14 Δ pcaGH/pVK9::PcspB-hsomt和Ap1_0112/pVK9::PcspB-hsomt菌株的每种在实施例<7>中在培养开始后5小时获得。于室温融化冷冻的混合物,与200 μ L含有溶菌酶的TE缓冲液(10mM Tris,1mM EDTA, pH 8.0) 以及10 μ L蛋白酶K (20mg/mL) 一起添加,混合,然后于室温温育40分钟。使用RNeasy Mini试剂盒(QIAGEN) 进行以下程序。将经处理的产物与700 μ L含有1%2-巯基乙醇的RLT缓冲液一起添加,混合,并且离心以获得上清液。将上清液与500 μ L乙醇一起添加,混合,并且应用到试剂盒中包含的柱,并且将柱离心。用350 μ L RW1缓冲液清洗柱,然后将80 μ L DNA酶I溶液添加到柱以于室温进行DNA酶处理15分钟。此外,用350 μ L RW1缓冲液清洗柱,并且用500 μ L RPE缓冲液清洗两次,并且用无RNA酶的灭菌水洗脱以获得RNA。使用NanoDrop (Thermo Fisher Scientific) 量化获得的RNA,并且使用BioAnalyser (Agilent Technologies) 用RNA 6000Nano试剂盒(Agilent Technologies) 通过进行电泳分析以确认获得的RNA具有足够的纯度。

[0353] <8-2>通过反向转录合成cDNA

[0354] 使用具有gDNA Eraser的PrimeScript RT试剂试剂盒(TAKARA BIO) 进行逆转录。将1 μ g RNA等分试样与1 μ L gDNA Eraser和2 μ L 5 $\times$ DNA Eraser缓冲液一起添加,用直至总体积10 μ L的灭菌水稀释,并且于42 $^{\circ}$ C温育2min以降解染色体DNA。将所得的混合物与4 μ L 5 $\times$ PrimeScript Buffer2、1 μ L PrimeScript RT Enzyme MixI、1 μ L RT Primer Mix和4 μ L灭菌水进一步一起添加,于37 $^{\circ}$ C温育15分钟和85 $^{\circ}$ C温育5秒以获得cDNA。

[0355] <8-3>定量PCR

[0356] 用以下程序从cDNA扩增NCg12048基因作为靶基因:2μL cDNA、10μL Power SYBR Green PCR主混合物(Life Technologies)、SEQ ID NO:104和105的引物(各500nM,作为终浓度),并且混合灭菌水以获得总体积20μL;使用7000实时PCR系统(Applied Bio Systems)以95℃变性10分钟,接着95℃达15秒和60℃达1min的40个循环进行PCR。另外,使用与用于靶基因扩增的程序相同的程序从cDNA扩增16S rRNA基因作为持家基因,只是使用2μL 32倍稀释的cDNA作为模板并且使用SEQ ID NO:106和107的引物。在扩增反应后,PCR产物进行熔解曲线分析以确认PCR产物的一致性。此外,通过琼脂糖凝胶电泳分析PCT产物以确认PCR产物具有用使用的引物可获得的长度。

[0357] <8-4>表达量的分析

[0358] 使用 $\Delta\Delta Ct$ 方法(METHODS, 25, 402 (2001))来分析NCg12048基因的表达量。提供通过从NCg12048基因的Ct值扣除持家基因的Ct值获得的数值作为 ΔCt 值。然而,由于使用32倍稀释(即 2^5 倍稀释)的cDNA作为扩增持家基因的模板,作为持家基因的Ct值,使用通过将5加到实际测量的持家基因的 ΔCt 值获得的值。提供通过从Ap1_0112/pVK9::PcspB-hsomt菌株的 ΔCt 值扣除FKFC14 Δ pcaGH/pVK9::PcspB-hsomt的 ΔCt 值获得的数值作为 $\Delta\Delta Ct$ 值。基于FKFC14 Δ pcaGH/pVK9::PcspB-hsomt菌株,Ap1_0112/pVK9::PcspB-hsomt菌株中的NCg12048基因的相对表达量作为 $2^{-\Delta\Delta Ct}$ 计算。

[0359] 表2中显示了结果。Ap1_0112/pVK9::PcspB-hsomt菌株中的NCg12048基因的相对表达量是FKFC14 Δ pcaGH/pVK9::PcspB-hsomt菌株中的NCg12048基因的相对表达量的约二十五分之一(1/25)。

[0360] [表2]

[0361] 表2:NCg12048基因的相对表达量

菌株	$2^{-\Delta\Delta Ct}$
FKFC14 Δ pcaGH/pVK9::PcspB-hsomt	1.00
Ap1_0112/pVK9::PcspB-hsomt	0.04

[0363] [工业适用性]

[0364] 根据本发明,微生物用于生产目标物质,诸如香草醛和香草酸的能力可以得到改善,并且可以有效生产目标物质。

[0365] <序列表的解释>

[0366] SEQ ID NOS:

[0367] 1:大肠杆菌MG1655的aroG基因的核苷酸序列

[0368] 2:大肠杆菌MG1655的AroG蛋白的氨基酸序列

[0369] 3:大肠杆菌MG1655的aroB基因的核苷酸序列

[0370] 4:大肠杆菌MG1655的AroB蛋白的氨基酸序列

[0371] 5:大肠杆菌MG1655的aroD基因的核苷酸序列

[0372] 6:大肠杆菌MG1655的AroD蛋白的氨基酸序列

[0373] 7:苏云金芽孢杆菌BMB171的asbF基因的核苷酸序列

[0374] 8:苏云金芽孢杆菌BMB171的AsbF蛋白的氨基酸序列

[0375] 9:大肠杆菌MG1655的tyrR基因的核苷酸序列

- [0376] 10:大肠杆菌MG1655的TyrR蛋白的氨基酸序列
- [0377] 11-14:智人OMT基因的转录物变体1至4的核苷酸序列
- [0378] 15:智人OMT同种型(MB-COMT)的氨基酸序列
- [0379] 16:智人OMT同种型(S-COMT)的氨基酸序列
- [0380] 17:巴西诺卡菌的ACAR基因的核苷酸序列
- [0381] 18:巴西诺卡菌的ACAR蛋白的氨基酸序列
- [0382] 19:巴西诺卡菌的ACAR基因的核苷酸序列
- [0383] 20:巴西诺卡菌的ACAR蛋白的氨基酸序列
- [0384] 21:大肠杆菌MG1655的entD基因的核苷酸序列
- [0385] 22:大肠杆菌MG1655的EntD蛋白的氨基酸序列
- [0386] 23:谷氨酸棒杆菌ATCC 13032的PPT基因的核苷酸序列
- [0387] 24:谷氨酸棒杆菌ATCC 13032的PPT蛋白的氨基酸序列
- [0388] 25:谷氨酸棒杆菌2256(ATCC 13869)的vanK基因的核苷酸序列
- [0389] 26:谷氨酸棒杆菌2256(ATCC 13869)的VanK蛋白的氨基酸序列
- [0390] 27:谷氨酸棒杆菌2256(ATCC 13869)的pcaK基因的核苷酸序列
- [0391] 28:谷氨酸棒杆菌2256(ATCC 13869)的PcaK蛋白的氨基酸序列
- [0392] 29:谷氨酸棒杆菌2256(ATCC 13869)的vanA基因的核苷酸序列
- [0393] 30:谷氨酸棒杆菌2256(ATCC 13869)的VanA蛋白的氨基酸序列
- [0394] 31:谷氨酸棒杆菌2256(ATCC 13869)的vanB基因的核苷酸序列
- [0395] 32:谷氨酸棒杆菌2256(ATCC 13869)的VanB蛋白的氨基酸序列
- [0396] 33:谷氨酸棒杆菌ATCC 13032的pcaG基因的核苷酸序列
- [0397] 34:谷氨酸棒杆菌ATCC 13032的PcaG蛋白的氨基酸序列
- [0398] 35:谷氨酸棒杆菌ATCC 13032的pcaH基因的核苷酸序列
- [0399] 36:谷氨酸棒杆菌ATCC 13032的PcaH蛋白的氨基酸序列
- [0400] 37:大肠杆菌MG1655的yqhD基因的核苷酸序列
- [0401] 38:大肠杆菌MG1655的YqhD蛋白的氨基酸序列
- [0402] 39:谷氨酸棒杆菌2256(ATCC 13869)的NCg10324基因的核苷酸序列
- [0403] 40:谷氨酸棒杆菌2256(ATCC 13869)的NCg10324蛋白的氨基酸序列
- [0404] 41:谷氨酸棒杆菌2256(ATCC 13869)的NCg10313基因的核苷酸序列
- [0405] 42:谷氨酸棒杆菌2256(ATCC 13869)的NCg10313蛋白的氨基酸序列
- [0406] 43:谷氨酸棒杆菌2256(ATCC 13869)的NCg12709基因的核苷酸序列
- [0407] 44:谷氨酸棒杆菌2256(ATCC 13869)的NCg12709蛋白的氨基酸序列
- [0408] 45:谷氨酸棒杆菌ATCC 13032的NCg10219基因的核苷酸序列
- [0409] 46:谷氨酸棒杆菌ATCC 13032的NCg10219蛋白的氨基酸序列
- [0410] 47:谷氨酸棒杆菌ATCC 13032的NCg12382基因的核苷酸序列
- [0411] 48:谷氨酸棒杆菌ATCC 13032的NCg12382蛋白的氨基酸序列
- [0412] 49:大肠杆菌MG1655的aroE基因的核苷酸序列
- [0413] 50:大肠杆菌MG1655的AroE蛋白的氨基酸序列
- [0414] 51-80:引物

- [0415] 81:含有P2启动子的核苷酸序列
- [0416] 82:含有P4启动子的核苷酸序列
- [0417] 83:含有P8启动子的核苷酸序列
- [0418] 84:含有P3启动子的核苷酸序列
- [0419] 85-88:引物
- [0420] 89:含有P8启动子区的DNA片段的核苷酸序列
- [0421] 90-91:引物
- [0422] 92:谷氨酸棒杆菌2256(ATCC 13869)的NCg12048基因的核苷酸序列
- [0423] 93:谷氨酸棒杆菌2256(ATCC 13869)的NCg12048蛋白的氨基酸序列
- [0424] 94:编码智人S-COMT的cDNA的核苷酸序列
- [0425] 95:含有COMT2基因的合成DNA片段的核苷酸序列
- [0426] 96:pELAC载体
- [0427] 97-107:引物

- [0001] 序列表
- [0002] <110> 味之素株式会社
- [0003] <120> 用于生产L-甲硫氨酸或需要S-腺苷甲硫氨酸进行合成的代谢物的方法
- [0004] <130> D886-16115
- [0005] <150> US62/413044
- [0006] <151> 2016-10-26
- [0007] <150> US62/417609
- [0008] <151> 2016-11-04
- [0009] <160> 107
- [0010] <170> PatentIn version 3.5
- [0011] <210> 1
- [0012] <211> 1053
- [0013] <212> DNA
- [0014] <213> 大肠杆菌
- [0015] <400> 1
- [0016] atgaattatc agaacgacga tttacgcatac aaagaaatca aagagttact tcctcctgtc 60
- [0017] gcattgctgg aaaaattccc cgctactgaa aatgccgcga atacggttgc ccatgcccga 120
- [0018] aaagcgatcc ataagatcct gaaagtaat gatgatcgcc tgttggttgt gattggccca 180
- [0019] tgctcaattc atgatcctgt cgcggaacaa gagtatgcca ctgccttgct ggcgctgcgt 240
- [0020] gaagagctga aagatgagct ggaaatcgta atgcgcgtct attttgaaaa gccgcgtacc 300
- [0021] acggtgggct ggaaagggt gattaacgat ccgcatatgg ataatagctt ccagatcaac 360
- [0022] gacggtctgc gtatagcccg taaattgctg cttgatatta acgacagcgg tctgccagcg 420
- [0023] gcaggtgagt ttctcgatat gatcaccca caatatctcg ctgacctgat gagctggggc 480
- [0024] gcaattggcg cacgtaccac cgaatcgtag gtgcaccgcg aactggcatc agggctttct 540
- [0025] tgtccggtcg gcttcaaaaa tggcaccgac ggtacgatta aagtggctat cgatgccatt 600
- [0026] aatgccgcg gtgcgccgca ctgcttcctg tccgtaacga aatgggggca ttcggcgatt 660
- [0027] gtgaatacca gcgtaaacgg cgattgccat atcattctgc gcggcggtaa agagcctaac 720
- [0028] tacagcgcga agcagcttgc tgaagtgaac gaagggtga acaaagcagg cctgccagca 780
- [0029] caggtgatga tcgatttcag ccatgctaac tcgtccaaac aattcaaaaa gcagatggat 840
- [0030] gtttgtgctg acgtttgcca gcagattgcc ggtggcgaaa aggccattat tggcgtgatg 900
- [0031] gtggaaagcc atctggtgga aggcaatcag agcctcgaga gcggggagcc gctggcctac 960
- [0032] ggtaagagca tcaccgatgc ctgcatcggc tgggaagata ccgatgctct gttacgtcaa 1020
- [0033] ctggcgaatg cagtaaaagc gcgtcgcggg taa 1053
- [0034] <210> 2
- [0035] <211> 350
- [0036] <212> PRT
- [0037] <213> 大肠杆菌
- [0038] <400> 2
- [0039] Met Asn Tyr Gln Asn Asp Asp Leu Arg Ile Lys Glu Ile Lys Glu Leu
- [0040] 1 5 10 15
- [0041] Leu Pro Pro Val Ala Leu Leu Glu Lys Phe Pro Ala Thr Glu Asn Ala

[0042]	20	25	30
[0043]	Ala Asn Thr Val Ala His Ala Arg Lys Ala Ile His Lys Ile Leu Lys		
[0044]	35	40	45
[0045]	Gly Asn Asp Asp Arg Leu Leu Val Val Ile Gly Pro Cys Ser Ile His		
[0046]	50	55	60
[0047]	Asp Pro Val Ala Ala Lys Glu Tyr Ala Thr Arg Leu Leu Ala Leu Arg		
[0048]	65	70	75
[0049]	Glu Glu Leu Lys Asp Glu Leu Glu Ile Val Met Arg Val Tyr Phe Glu		
[0050]	85	90	95
[0051]	Lys Pro Arg Thr Thr Val Gly Trp Lys Gly Leu Ile Asn Asp Pro His		
[0052]	100	105	110
[0053]	Met Asp Asn Ser Phe Gln Ile Asn Asp Gly Leu Arg Ile Ala Arg Lys		
[0054]	115	120	125
[0055]	Leu Leu Leu Asp Ile Asn Asp Ser Gly Leu Pro Ala Ala Gly Glu Phe		
[0056]	130	135	140
[0057]	Leu Asp Met Ile Thr Pro Gln Tyr Leu Ala Asp Leu Met Ser Trp Gly		
[0058]	145	150	155
[0059]	Ala Ile Gly Ala Arg Thr Thr Glu Ser Gln Val His Arg Glu Leu Ala		
[0060]	165	170	175
[0061]	Ser Gly Leu Ser Cys Pro Val Gly Phe Lys Asn Gly Thr Asp Gly Thr		
[0062]	180	185	190
[0063]	Ile Lys Val Ala Ile Asp Ala Ile Asn Ala Ala Gly Ala Pro His Cys		
[0064]	195	200	205
[0065]	Phe Leu Ser Val Thr Lys Trp Gly His Ser Ala Ile Val Asn Thr Ser		
[0066]	210	215	220
[0067]	Gly Asn Gly Asp Cys His Ile Ile Leu Arg Gly Gly Lys Glu Pro Asn		
[0068]	225	230	235
[0069]	Tyr Ser Ala Lys His Val Ala Glu Val Lys Glu Gly Leu Asn Lys Ala		
[0070]	245	250	255
[0071]	Gly Leu Pro Ala Gln Val Met Ile Asp Phe Ser His Ala Asn Ser Ser		
[0072]	260	265	270
[0073]	Lys Gln Phe Lys Lys Gln Met Asp Val Cys Ala Asp Val Cys Gln Gln		
[0074]	275	280	285
[0075]	Ile Ala Gly Gly Glu Lys Ala Ile Ile Gly Val Met Val Glu Ser His		
[0076]	290	295	300
[0077]	Leu Val Glu Gly Asn Gln Ser Leu Glu Ser Gly Glu Pro Leu Ala Tyr		
[0078]	305	310	315
[0079]	Gly Lys Ser Ile Thr Asp Ala Cys Ile Gly Trp Glu Asp Thr Asp Ala		
[0080]	325	330	335
[0081]	Leu Leu Arg Gln Leu Ala Asn Ala Val Lys Ala Arg Arg Gly		
[0082]	340	345	350
[0083]	<210> 3		

[0084]	<211> 1089
[0085]	<212> DNA
[0086]	<213> 大肠杆菌
[0087]	<400> 3
[0088]	atggagagga ttgtcgttac tctcggggaa cgtagttacc caattacat cgcactctggt 60
[0089]	ttgtttaatg aaccagcttc attcttaccg ctgaaatcgg gcgagcaggt catgttggtc 120
[0090]	accaacgaaa ccctggctcc tctgtatctc gataaggtcc gcggcgtaact tgaacaggcg 180
[0091]	ggtgttaacg tcgatagcgt tatcctccct gacggcgagc agtataaaaag cctggctgta 240
[0092]	ctcgataccg tctttacggc gttgtttacaa aaaccgcatg gtcgcgatac tacgctgggtg 300
[0093]	gcgcttgccg gcggcgtagt gggcgatctg accggcttcg cggcgcgag ttatcagcgc 360
[0094]	ggtgtccgtt tcattcaagt cccgacgacg ttactgtcgc aggtcgattc ctccgttggtc 420
[0095]	ggcaaaactg cgggtcaacca tcccctcggt aaaaacatga ttggcgcggt ctaccaacct 480
[0096]	gcttcagtgg tggatgatct cgactgtctg aaaacgctc ccccgcgta gttagcgtcg 540
[0097]	gggctggcag aagtcatcaa atacggcatt attcttgacg gtgcgttttt taactggctg 600
[0098]	gaagagaatc tggatgcgtt gttgcgtctg gacgggtccg caatggcgta ctgtattcgc 660
[0099]	cgttgtgtg aactgaagc agaagttgc gccgccgacg agcgcgaaac cgggttacgt 720
[0100]	gctttactga atctgggaca cacctttggt catgccattg aagctgaaat ggggtatggc 780
[0101]	aattggttac atggtgaagc ggtcgctcgc ggtatggtga tggcggcgcg gacgtcgga 840
[0102]	cgtctcgggc agtttagttc tgccgaaacg cagcgtatta taaccctgct caagcgggct 900
[0103]	gggttaccgg tcaatgggcc gcgcgaaatg tccgcgcagg cgtatttacc gcatatgctg 960
[0104]	cgtgacaaga aagtccttgc gggagagatg cgcttaattc ttccgttggtc aattggtgaag 1020
[0105]	agtgaagttc gcagcggcgt ttcgcacgag cttgttctta acgccattgc cgattgtcaa 1080
[0106]	tcagcgtaa 1089
[0107]	<210> 4
[0108]	<211> 362
[0109]	<212> PRT
[0110]	<213> 大肠杆菌
[0111]	<400> 4
[0112]	Met Glu Arg Ile Val Val Thr Leu Gly Glu Arg Ser Tyr Pro Ile Thr
[0113]	1 5 10 15
[0114]	Ile Ala Ser Gly Leu Phe Asn Glu Pro Ala Ser Phe Leu Pro Leu Lys
[0115]	20 25 30
[0116]	Ser Gly Glu Gln Val Met Leu Val Thr Asn Glu Thr Leu Ala Pro Leu
[0117]	35 40 45
[0118]	Tyr Leu Asp Lys Val Arg Gly Val Leu Glu Gln Ala Gly Val Asn Val
[0119]	50 55 60
[0120]	Asp Ser Val Ile Leu Pro Asp Gly Glu Gln Tyr Lys Ser Leu Ala Val
[0121]	65 70 75 80
[0122]	Leu Asp Thr Val Phe Thr Ala Leu Leu Gln Lys Pro His Gly Arg Asp
[0123]	85 90 95
[0124]	Thr Thr Leu Val Ala Leu Gly Gly Gly Val Val Gly Asp Leu Thr Gly
[0125]	100 105 110

[0126]	Phe Ala Ala Ala Ser Tyr Gln Arg Gly Val Arg Phe Ile Gln Val Pro	
[0127]	115	120 125
[0128]	Thr Thr Leu Leu Ser Gln Val Asp Ser Ser Val Gly Gly Lys Thr Ala	
[0129]	130	135 140
[0130]	Val Asn His Pro Leu Gly Lys Asn Met Ile Gly Ala Phe Tyr Gln Pro	
[0131]	145	150 155 160
[0132]	Ala Ser Val Val Val Asp Leu Asp Cys Leu Lys Thr Leu Pro Pro Arg	
[0133]	165	170 175
[0134]	Glu Leu Ala Ser Gly Leu Ala Glu Val Ile Lys Tyr Gly Ile Ile Leu	
[0135]	180	185 190
[0136]	Asp Gly Ala Phe Phe Asn Trp Leu Glu Glu Asn Leu Asp Ala Leu Leu	
[0137]	195	200 205
[0138]	Arg Leu Asp Gly Pro Ala Met Ala Tyr Cys Ile Arg Arg Cys Cys Glu	
[0139]	210	215 220
[0140]	Leu Lys Ala Glu Val Val Ala Ala Asp Glu Arg Glu Thr Gly Leu Arg	
[0141]	225	230 235 240
[0142]	Ala Leu Leu Asn Leu Gly His Thr Phe Gly His Ala Ile Glu Ala Glu	
[0143]	245	250 255
[0144]	Met Gly Tyr Gly Asn Trp Leu His Gly Glu Ala Val Ala Ala Gly Met	
[0145]	260	265 270
[0146]	Val Met Ala Ala Arg Thr Ser Glu Arg Leu Gly Gln Phe Ser Ser Ala	
[0147]	275	280 285
[0148]	Glu Thr Gln Arg Ile Ile Thr Leu Leu Lys Arg Ala Gly Leu Pro Val	
[0149]	290	295 300
[0150]	Asn Gly Pro Arg Glu Met Ser Ala Gln Ala Tyr Leu Pro His Met Leu	
[0151]	305	310 315 320
[0152]	Arg Asp Lys Lys Val Leu Ala Gly Glu Met Arg Leu Ile Leu Pro Leu	
[0153]	325	330 335
[0154]	Ala Ile Gly Lys Ser Glu Val Arg Ser Gly Val Ser His Glu Leu Val	
[0155]	340	345 350
[0156]	Leu Asn Ala Ile Ala Asp Cys Gln Ser Ala	
[0157]	355	360
[0158]	<210> 5	
[0159]	<211> 759	
[0160]	<212> DNA	
[0161]	<213> 大肠杆菌	
[0162]	<400> 5	
[0163]	atgaaaaccg taactgtaaa agatctcgtc attggtacgg ggcacacctaa aatcatcgtc	60
[0164]	tcgctgatgg cgaaagatat cgccagcgtg aaatccgaag ctctcgcccta tcgtgaagcg	120
[0165]	gactttgata ttctggaatg gcgtgtggac cactatgccg acctctccaa tgtggagtct	180
[0166]	gtcatggcgg cagcaaaaat tctccgtgag accatgccag aaaaaccgct gctgtttacc	240
[0167]	ttccgcagtg ccaaagaagg cggcgagcag gcgatttcca ccgaggetta tattgcactc	300

[0168]	aatcgtgcag ccatcgacag cggcctggtt gatatgatcg atctggagtt atttaccggt	360
[0169]	gatgatcagg ttaaagaaac cgtcgctac gccacgcgc atgatgtgaa agtagtcatg	420
[0170]	tccaaccatg acttccataa aacgccgaa gccgaagaaa tcattgcccg tctgcgcaa	480
[0171]	atgcaatcct tcgacgccga tattcctaag attgcgctga tgccgcaaag taccagcgat	540
[0172]	gtgctgacgt tgcttgccgc gaccctggag atgcaggagc agtatgccga tcgtccaatt	600
[0173]	atcacgatgt cgatggcaaa aactggcgta atttctcgtc tggctgggtga agtatttggc	660
[0174]	tcggcggcaa cttttgggtgc ggtaaaaaa gcgtctgcgc cagggcaaat ctcggtaaat	720
[0175]	gatttgcgca cggtattaac tattttacac caggcataa	759
[0176]	<210>	6
[0177]	<211>	252
[0178]	<212>	PRT
[0179]	<213>	大肠杆菌
[0180]	<400>	6
[0181]	Met Lys Thr Val Thr Val Lys Asp Leu Val Ile Gly Thr Gly Ala Pro	
[0182]	1 5 10 15	
[0183]	Lys Ile Ile Val Ser Leu Met Ala Lys Asp Ile Ala Ser Val Lys Ser	
[0184]	20 25 30	
[0185]	Glu Ala Leu Ala Tyr Arg Glu Ala Asp Phe Asp Ile Leu Glu Trp Arg	
[0186]	35 40 45	
[0187]	Val Asp His Tyr Ala Asp Leu Ser Asn Val Glu Ser Val Met Ala Ala	
[0188]	50 55 60	
[0189]	Ala Lys Ile Leu Arg Glu Thr Met Pro Glu Lys Pro Leu Leu Phe Thr	
[0190]	65 70 75 80	
[0191]	Phe Arg Ser Ala Lys Glu Gly Gly Glu Gln Ala Ile Ser Thr Glu Ala	
[0192]	85 90 95	
[0193]	Tyr Ile Ala Leu Asn Arg Ala Ala Ile Asp Ser Gly Leu Val Asp Met	
[0194]	100 105 110	
[0195]	Ile Asp Leu Glu Leu Phe Thr Gly Asp Asp Gln Val Lys Glu Thr Val	
[0196]	115 120 125	
[0197]	Ala Tyr Ala His Ala His Asp Val Lys Val Val Met Ser Asn His Asp	
[0198]	130 135 140	
[0199]	Phe His Lys Thr Pro Glu Ala Glu Glu Ile Ile Ala Arg Leu Arg Lys	
[0200]	145 150 155 160	
[0201]	Met Gln Ser Phe Asp Ala Asp Ile Pro Lys Ile Ala Leu Met Pro Gln	
[0202]	165 170 175	
[0203]	Ser Thr Ser Asp Val Leu Thr Leu Leu Ala Ala Thr Leu Glu Met Gln	
[0204]	180 185 190	
[0205]	Glu Gln Tyr Ala Asp Arg Pro Ile Ile Thr Met Ser Met Ala Lys Thr	
[0206]	195 200 205	
[0207]	Gly Val Ile Ser Arg Leu Ala Gly Glu Val Phe Gly Ser Ala Ala Thr	
[0208]	210 215 220	
[0209]	Phe Gly Ala Val Lys Lys Ala Ser Ala Pro Gly Gln Ile Ser Val Asn	

[0210]	225	230	235	240
[0211]	Asp Leu Arg Thr Val Leu Thr Ile Leu His Gln Ala			
[0212]	245	250		
[0213]	<210> 7			
[0214]	<211> 843			
[0215]	<212> DNA			
[0216]	<213> 苏云金芽孢杆菌			
[0217]	<400> 7			
[0218]	atgaaatatt cgctatgtac catttcattt cgtcatcaat taatttcatt tactgatatt 60			
[0219]	gttcaatttg catatgaaaa cggttttgaa ggaattgaat tatgggggac gcatgcacaa 120			
[0220]	aatttgtaca tgcaagaacg tgaacgaca gaacgagaat tgaattttct aaaggataaa 180			
[0221]	aacttagaaa ttacgatgat aagtgattac ttagatatat cattatcagc agattttgaa 240			
[0222]	aaaacgatag agaaaagtga acaacttgta gtactagcta attggtttaa tacgaataaa 300			
[0223]	attcgcacgt ttgctgggca aaaaggagc aaggacttct cggaacaaga gagaaaagag 360			
[0224]	tatgtgaagc gaatacgtaa gatttgtgat gtgtttgctc agaacaatat gtatgtgctg 420			
[0225]	ttagaaacac atcccaatac actaacggac acattgcctt ctactataga gttattagaa 480			
[0226]	gaagtaaacc atccgaattt aaaaataaat cttgattttc ttcatatatg ggagtctggc 540			
[0227]	gcagatccaa tagacagttt ccatcgatta aagccgtgga cactacatta ccattttaag 600			
[0228]	aatatatctt cagcggatta tttgcatgtg tttgaaccta ataatgtata tgctgcagca 660			
[0229]	ggaagtcgta taggtatggt tccgttatit gaaggtattg taaattatga tgagattatt 720			
[0230]	caggaagtga gaaatacga tctttttgct tccttagaat ggtttgaca taattcaaaa 780			
[0231]	gagatattaa aagaagaaat gaaagtatta ataaatagaa aattagaagt agtaacttcg 840			
[0232]	taa 843			
[0233]	<210> 8			
[0234]	<211> 280			
[0235]	<212> PRT			
[0236]	<213> 苏云金芽孢杆菌			
[0237]	<400> 8			
[0238]	Met Lys Tyr Ser Leu Cys Thr Ile Ser Phe Arg His Gln Leu Ile Ser			
[0239]	1	5	10	15
[0240]	Phe Thr Asp Ile Val Gln Phe Ala Tyr Glu Asn Gly Phe Glu Gly Ile			
[0241]	20	25	30	
[0242]	Glu Leu Trp Gly Thr His Ala Gln Asn Leu Tyr Met Gln Glu Arg Glu			
[0243]	35	40	45	
[0244]	Thr Thr Glu Arg Glu Leu Asn Phe Leu Lys Asp Lys Asn Leu Glu Ile			
[0245]	50	55	60	
[0246]	Thr Met Ile Ser Asp Tyr Leu Asp Ile Ser Leu Ser Ala Asp Phe Glu			
[0247]	65	70	75	80
[0248]	Lys Thr Ile Glu Lys Ser Glu Gln Leu Val Val Leu Ala Asn Trp Phe			
[0249]	85	90	95	
[0250]	Asn Thr Asn Lys Ile Arg Thr Phe Ala Gly Gln Lys Gly Ser Lys Asp			
[0251]	100	105	110	

[0252]	Phe Ser Glu Gln Glu Arg Lys Glu Tyr Val Lys Arg Ile Arg Lys Ile		
[0253]	115	120	125
[0254]	Cys Asp Val Phe Ala Gln Asn Asn Met Tyr Val Leu Leu Glu Thr His		
[0255]	130	135	140
[0256]	Pro Asn Thr Leu Thr Asp Thr Leu Pro Ser Thr Ile Glu Leu Leu Glu		
[0257]	145	150	155
[0258]	Glu Val Asn His Pro Asn Leu Lys Ile Asn Leu Asp Phe Leu His Ile		
[0259]	165	170	175
[0260]	Trp Glu Ser Gly Ala Asp Pro Ile Asp Ser Phe His Arg Leu Lys Pro		
[0261]	180	185	190
[0262]	Trp Thr Leu His Tyr His Phe Lys Asn Ile Ser Ser Ala Asp Tyr Leu		
[0263]	195	200	205
[0264]	His Val Phe Glu Pro Asn Asn Val Tyr Ala Ala Ala Gly Ser Arg Ile		
[0265]	210	215	220
[0266]	Gly Met Val Pro Leu Phe Glu Gly Ile Val Asn Tyr Asp Glu Ile Ile		
[0267]	225	230	235
[0268]	Gln Glu Val Arg Asn Thr Asp Leu Phe Ala Ser Leu Glu Trp Phe Gly		
[0269]	245	250	255
[0270]	His Asn Ser Lys Glu Ile Leu Lys Glu Glu Met Lys Val Leu Ile Asn		
[0271]	260	265	270
[0272]	Arg Lys Leu Glu Val Val Thr Ser		
[0273]	275	280	
[0274]	<210> 9		
[0275]	<211> 1542		
[0276]	<212> DNA		
[0277]	<213> 大肠杆菌		
[0278]	<400> 9		
[0279]	atgcgtctgg aagtcttttg tgaagaccga ctcggtctga cccgcgaatt actcgatcta	60	
[0280]	ctcgtgctaa gaggcattga tttacgcggt attgagattg atccattgg gccaatctac	120	
[0281]	ctcaattttg ctgaactgga gtttgagagt ttcagcagtc tgatggccga aatacgccgt	180	
[0282]	attgcgggtg ttaccgatgt gcgtactgtc ccgtggatgc cttccgaacg tgagcatctg	240	
[0283]	gcgttgagcg cgttactgga ggcgttgccct gaacctgtgc tctctgtcga tatgaaaagc	300	
[0284]	aaagtggata tggcgaaacc ggcgagctgt cagctttttg ggcaaaaatt ggatcgctg	360	
[0285]	cgcaaccata ccgccgcaca attgattaac ggctttaatt ttttacgttg gctggaaagc	420	
[0286]	gaaccgcaag attcgcataa cgagcatgtc gttattaatg ggcagaattt cctgatggag	480	
[0287]	attacgcctg tttatcttca ggatgaaaat gatcaacacg tcctgaccgg tgcggtggtg	540	
[0288]	atgttgcgat caacgattcg tatgggcccgc cagttgcaaa atgtcgcccgc ccaggacgtc	600	
[0289]	agcgcttca gtcaaattgt cgccgtcagc ccgaaaatga agcatgttgt cgaacaggcg	660	
[0290]	cagaaactgg cgatgctaag cgcgccgtg ctgattacgg gtgacacagg tacaggtaaa	720	
[0291]	gatctctttg cctacgcctg ccatcaggca agccccagag cgggcaaacc ttacctggcg	780	
[0292]	ctgaactgtg cgtctatacc ggaagatgcg gtcgagagtg aactgtttgg tcatgctccg	840	
[0293]	gaagggaaga aaggattctt tgagcaggcg aacggtggtt cggtgctggt ggatgaaata	900	

[0294] ggggaaatgt caccacggat gcaggcgaat ttactgcgtt tccttaatga tggcactttc 960
 [0295] cgtcgggttg gcgaagacca tgaggtgcat gtcgatgtgc ggggtgatttg cgctacgcag 1020
 [0296] aagaatctgg tcgaactggt gcaaaaaggc atgttccgtg aagatctcta ttatcgtctg 1080
 [0297] aacgtgttga cgctcaatct gccgccgcta cgtgactgtc cgcaggacat catgccgtta 1140
 [0298] actgagctgt tcgtcgcccg ctttgccgac gagcagggcg tgccgcgtcc gaaactggcc 1200
 [0299] gctgacctga atactgtact tacgcgttat gcgtggcccg gaaatgtgcg gcagttaaag 1260
 [0300] aacgetatct atcgcgcact gacacaactg gacggttatg agctgcgtcc acaggatatt 1320
 [0301] ttgttgccgg attatgacgc cgcaacggtg gccgtggcg aagatgcgat ggaaggttcg 1380
 [0302] ctggacgaaa tcaccagccg ttttgaacgc tcggtattaa cccagcttta tcgcaattat 1440
 [0303] cccagcacgc gcaaacggc aaaacgtctc ggcgtttcac ataccgcgat tgccaataag 1500
 [0304] ttgcgggaat atggtctgag tcagaagaag aacgaagagt aa 1542
 [0305] <210> 10
 [0306] <211> 513
 [0307] <212> PRT
 [0308] <213> 大肠杆菌
 [0309] <400> 10
 [0310] Met Arg Leu Glu Val Phe Cys Glu Asp Arg Leu Gly Leu Thr Arg Glu
 [0311] 1 5 10 15
 [0312] Leu Leu Asp Leu Leu Val Leu Arg Gly Ile Asp Leu Arg Gly Ile Glu
 [0313] 20 25 30
 [0314] Ile Asp Pro Ile Gly Arg Ile Tyr Leu Asn Phe Ala Glu Leu Glu Phe
 [0315] 35 40 45
 [0316] Glu Ser Phe Ser Ser Leu Met Ala Glu Ile Arg Arg Ile Ala Gly Val
 [0317] 50 55 60
 [0318] Thr Asp Val Arg Thr Val Pro Trp Met Pro Ser Glu Arg Glu His Leu
 [0319] 65 70 75 80
 [0320] Ala Leu Ser Ala Leu Leu Glu Ala Leu Pro Glu Pro Val Leu Ser Val
 [0321] 85 90 95
 [0322] Asp Met Lys Ser Lys Val Asp Met Ala Asn Pro Ala Ser Cys Gln Leu
 [0323] 100 105 110
 [0324] Phe Gly Gln Lys Leu Asp Arg Leu Arg Asn His Thr Ala Ala Gln Leu
 [0325] 115 120 125
 [0326] Ile Asn Gly Phe Asn Phe Leu Arg Trp Leu Glu Ser Glu Pro Gln Asp
 [0327] 130 135 140
 [0328] Ser His Asn Glu His Val Val Ile Asn Gly Gln Asn Phe Leu Met Glu
 [0329] 145 150 155 160
 [0330] Ile Thr Pro Val Tyr Leu Gln Asp Glu Asn Asp Gln His Val Leu Thr
 [0331] 165 170 175
 [0332] Gly Ala Val Val Met Leu Arg Ser Thr Ile Arg Met Gly Arg Gln Leu
 [0333] 180 185 190
 [0334] Gln Asn Val Ala Ala Gln Asp Val Ser Ala Phe Ser Gln Ile Val Ala
 [0335] 195 200 205

[0336]	Val Ser Pro Lys Met Lys His Val Val Glu Gln Ala Gln Lys Leu Ala
[0337]	210 215 220
[0338]	Met Leu Ser Ala Pro Leu Leu Ile Thr Gly Asp Thr Gly Thr Gly Lys
[0339]	225 230 235 240
[0340]	Asp Leu Phe Ala Tyr Ala Cys His Gln Ala Ser Pro Arg Ala Gly Lys
[0341]	245 250 255
[0342]	Pro Tyr Leu Ala Leu Asn Cys Ala Ser Ile Pro Glu Asp Ala Val Glu
[0343]	260 265 270
[0344]	Ser Glu Leu Phe Gly His Ala Pro Glu Gly Lys Lys Gly Phe Phe Glu
[0345]	275 280 285
[0346]	Gln Ala Asn Gly Gly Ser Val Leu Leu Asp Glu Ile Gly Glu Met Ser
[0347]	290 295 300
[0348]	Pro Arg Met Gln Ala Lys Leu Leu Arg Phe Leu Asn Asp Gly Thr Phe
[0349]	305 310 315 320
[0350]	Arg Arg Val Gly Glu Asp His Glu Val His Val Asp Val Arg Val Ile
[0351]	325 330 335
[0352]	Cys Ala Thr Gln Lys Asn Leu Val Glu Leu Val Gln Lys Gly Met Phe
[0353]	340 345 350
[0354]	Arg Glu Asp Leu Tyr Tyr Arg Leu Asn Val Leu Thr Leu Asn Leu Pro
[0355]	355 360 365
[0356]	Pro Leu Arg Asp Cys Pro Gln Asp Ile Met Pro Leu Thr Glu Leu Phe
[0357]	370 375 380
[0358]	Val Ala Arg Phe Ala Asp Glu Gln Gly Val Pro Arg Pro Lys Leu Ala
[0359]	385 390 395 400
[0360]	Ala Asp Leu Asn Thr Val Leu Thr Arg Tyr Ala Trp Pro Gly Asn Val
[0361]	405 410 415
[0362]	Arg Gln Leu Lys Asn Ala Ile Tyr Arg Ala Leu Thr Gln Leu Asp Gly
[0363]	420 425 430
[0364]	Tyr Glu Leu Arg Pro Gln Asp Ile Leu Leu Pro Asp Tyr Asp Ala Ala
[0365]	435 440 445
[0366]	Thr Val Ala Val Gly Glu Asp Ala Met Glu Gly Ser Leu Asp Glu Ile
[0367]	450 455 460
[0368]	Thr Ser Arg Phe Glu Arg Ser Val Leu Thr Gln Leu Tyr Arg Asn Tyr
[0369]	465 470 475 480
[0370]	Pro Ser Thr Arg Lys Leu Ala Lys Arg Leu Gly Val Ser His Thr Ala
[0371]	485 490 495
[0372]	Ile Ala Asn Lys Leu Arg Glu Tyr Gly Leu Ser Gln Lys Lys Asn Glu
[0373]	500 505 510
[0374]	Glu
[0375]	<210> 11
[0376]	<211> 2304
[0377]	<212> DNA

[0378]	<213> 智人
[0379]	<400> 11
[0380]	eggccctgctg cgcgccaccg aagcgccctc ctaatccccg cagcgccacc gccattgccg 60
[0381]	ccatcgtcgt ggggcttctg gggcagctag ggctgcccgc cgcgctgcct gcgccgacc 120
[0382]	ggggcgggtc cagtccccgg cgggcccgtc cgggagagaa ataacatctg ctttgctgcc 180
[0383]	gagctcagag gagaccccag acccctcccc cagccagagg gctggagcct gctcagaggt 240
[0384]	gctttgaaga tgccggagge cccgcctctg ctgttggcag ctgtgttgct gggcctgggtg 300
[0385]	ctgctgggtg tgctgctgct gcttctgagg cactggggct ggggcctgtg ccttatcggc 360
[0386]	tggaacgagt tcacctgca gcccatccac aacctgctca tgggtgacac caaggagcag 420
[0387]	cgcacctga accacgtgct gcagcatgcg gagcccggga acgcacagag cgtgctggag 480
[0388]	gccattgaca cctactgca gcagaaggag tgggccatga acgtgggcga caagaaaggc 540
[0389]	aagatcgtgg acgccgtgat tcaggagcac cagccctccg tgctgctgga gctgggggcc 600
[0390]	tactgtggct actcagctgt gcgcatggcc cgcctgctgt caccaggggc gaggtcatc 660
[0391]	accatcgaga tcaaccccga ctgtgccgcc atcaccagc ggatggtgga tttcgtggc 720
[0392]	gtgaaggaca aggtcacctc tgtggttga gcgtcccagg acatcatccc ccagctgaag 780
[0393]	aagaagtatg atgtggacac actggacatg gtcttctcgc accactggaa ggaccggtac 840
[0394]	ctgccggaca cgcttctctt ggaggaatgt ggctgctgc ggaaggggac agtgctactg 900
[0395]	gctgacaacg tgatctgccc aggtgcgcca gacttcctag cacacgtgcg cgggagcagc 960
[0396]	tgctttgagt gcacacacta ccaatcgttc ctggaataga gggaggtggt ggacggcctg 1020
[0397]	gagaaggcca tctacaaggg cccaggcagc gaagcagggc cctgactgcc ccccgggccc 1080
[0398]	ccctctggg ctctctcacc cagcctggta ctgaaggtgc cagacgtgct cctgctgacc 1140
[0399]	ttctgcggct ccgggctgtg tcctaaatgc aaagcacacc tcggccgagg cctgcgcct 1200
[0400]	gacatgctaa cctctctgaa ctgcaacact ggattgttct tttttaagac tcaatcatga 1260
[0401]	cttctttact aacctggct agctatatta tcttatatac taatatcatg ttttaaaaat 1320
[0402]	ataaaaataga aattaagaat ctaaataatt agatataact cgacttagta catecttctc 1380
[0403]	aactgccatt ccctgctgc ccttgacttg ggcaccaaac attcaaagct ccccttgacg 1440
[0404]	gacgctaacg ctaagggcgg ggcccctagc tggctgggtt ctgggtggca cgcctggccc 1500
[0405]	actggcctcc cagccacagt ggtgcagagg tcagccctcc tgcagctagg ccaggggcac 1560
[0406]	ctgttagccc catggggagc actgccggcc tgggaaacga agaggagtca gccagcattc 1620
[0407]	acacctttct gaccaagcag gcgctgggga caggtggacc ccgcagcagc accagccct 1680
[0408]	ctgggccccca tgtggcacag agtgggaagca tctccttccc tactccccac tgggccttgc 1740
[0409]	ttacagaaga ggcaatggct cagaccagct cccgcatccc tgtagttgcc tccctggccc 1800
[0410]	atgagtgagg atgcagtgt ggtttctgcc cacctacacc tagagctgtc cccatctcct 1860
[0411]	ccaaggggtc agactgctag ccacctcaga ggctccaagg gccagttcc caggcccagg 1920
[0412]	acaggaatca accctgtgct agctgagttc acctgcaccg agaccagccc ctagccaaga 1980
[0413]	ttctactcct gggctcaagg cctggctagc cccagccag cccactccta tggatagaca 2040
[0414]	gaccagttag cccaagtgga caagtgtgg gccaccagg gaccagaaac agagcctctg 2100
[0415]	caggacacag cagatgggca cctgggacca cctccacca gggcctgcc ccagacgcgc 2160
[0416]	agaggcccga cacaaggag aagccagcca cttgtgccag acctgagtgg cagaaagcaa 2220
[0417]	aaagtccctt tgctgcttta atttttaaat tttcttaca aaatttaggt gtttaccat 2280
[0418]	agtcttattt tggcttattt ttaa 2304
[0419]	<210> 12

[0420]	<211> 2262	
[0421]	<212> DNA	
[0422]	<213> 智人	
[0423]	<400> 12	
[0424]	ctccccacggg aggagcaaga acacagaaca gagggggcaa gacagctcca ccaggagtca	60
[0425]	ggagtgaatc ccctctggga acgaggcact aggaagaaga acttcagcc caggagaaat	120
[0426]	aacatctgct ttgctgccga gctcagagga gacccagac ccctcccga gccagagggc	180
[0427]	tggagcctgc tcagaggtgc tttgaagatg ccggaggccc cgcctctgct gttggcagct	240
[0428]	gtgttgctgg gcctgggtgct gctgggtgtg ctgctgctgc ttctgaggca ctggggctgg	300
[0429]	ggcctgtgcc ttatcggtg gaacgagttc atcctgcagc ccatccaaa cctgctcatg	360
[0430]	ggtgacacca aggagcagcg catcctgaac cacgtgctgc agcatgcgga gcccggaac	420
[0431]	gcacagagcg tgctggagge cattgacacc tactgcgagc agaaggagtg ggccatgaac	480
[0432]	gtgggcgaca agaaaggcaa gatcgtggac gccgtgattc aggagcacca gccctccgtg	540
[0433]	ctgctggagc tgggggccta ctgtggctac tcagctgtgc gcatggccc cctgctgtca	600
[0434]	ccagggcgga ggctcatcac catcgagatc aacccgact gtgccgcat caccagcgg	660
[0435]	atggtggatt tcgctggcgt gaaggacaag gtcaccctt tggttgagc gtcccaggac	720
[0436]	atcatcccc agctgaagaa gaagtatgat gtggacacac tggacatgtt ctctctgac	780
[0437]	cactggaagg accggtacct gccggacag cttctcttg aggaatgtgg cctgctgcgg	840
[0438]	aaggggacag tgctactggc tgacaacgtg atctgcccag gtgcgccaga ctctctagca	900
[0439]	cacgtgcgcg ggagcagctg ctttgagtgc acacactacc aatcgttcct ggaatacagg	960
[0440]	gaggtgggtg acggcctgga gaaggccatc tacaagggcc caggcagcga agcaggggcc	1020
[0441]	tgactgcccc cccggcccc ctctcgggct ctctcaccca gcctggtact gaaggtgcca	1080
[0442]	gacgtgctcc tgctgacctt ctgcggctcc gggctgtgtc ctaaagcaa agcacacctc	1140
[0443]	ggccgaggcc tgcgcctga catgctaacc tctctgaact gcaacactgg attgttcttt	1200
[0444]	tttaagactc aatcatgact tctttactaa cactggctag ctatattatc ttatatacta	1260
[0445]	atatcatgtt ttaaaaatat aaaatagaaa ttaagaatct aaatatattag atataactcg	1320
[0446]	acttagtaca tccttctcaa ctgccattcc cctgctgccc ttgacttggg caccaaacat	1380
[0447]	tcaaagctcc cttgacgga cgctaacgct aagggcgggg ccctagctg gctgggttct	1440
[0448]	gggtggcagc cctggccac tggcctccca gccacagtgg tgcagagtc agcctctctg	1500
[0449]	cagctaggcc aggggcacct gttagcccca tggggacgac tgccggcctg ggaaacgaag	1560
[0450]	aggagtcagc cagcattcac acctttctga ccaagcaggc gctggggaca ggtggacccc	1620
[0451]	gcagcagcac cagcccctct gggcccatg tggcacagag tggaagcatc tccttcccta	1680
[0452]	ctccccactg ggcttgctt acagaagagg caatggctca gaccagctcc cgcctccctg	1740
[0453]	tagttgcctc cctggcccat gagtgaggat gcagtgtggt tttctgccc cctacaccta	1800
[0454]	gagctgtccc catctcctcc aaggggtcag actgctagcc acctcagagg ctccaagggc	1860
[0455]	ccagttccca ggcccaggac aggaatcaac cctgtgctag ctgagttcac ctgcaccgag	1920
[0456]	accagcccct agccaagatt ctactcctgg gctcaaggcc tggctagccc ccagccagcc	1980
[0457]	cactcctatg gatagacaga ccagtggacc caagtggaca agtttggggc caccagggga	2040
[0458]	ccagaaacag agcctctgca ggacacagca gatgggcacc tgggaccacc tccaccagc	2100
[0459]	gccctgcccc agacgcgcag agggccgaca caaggagaa gccagccact tgtgccagac	2160
[0460]	ctgagtggca gaaagcaaaa agttcctttg ctgctttaat ttttaattt tcttacaaaa	2220
[0461]	atttaggtgt ttaccaatag tcttattttg gcttattttt aa	2262

[0462]	<210> 13
[0463]	<211> 2279
[0464]	<212> DNA
[0465]	<213> 智人
[0466]	<400> 13
[0467]	tggagataac acggatcgct gtgtacactg tgtgctccgg ttgttgcatc cgagggttga 60
[0468]	tcggatgggtg gttcccatcc agatccaagt cctggcccct gatcacagag aaacacagct 120
[0469]	ggacattaaa gtgaaataac atctgctttg ctgccgagct cagaggagac cccagacccc 180
[0470]	tcccgcagcc agagggtctg agcctgctca gaggtgcttt gaagatgccg gaggccccgc 240
[0471]	ctctgctgtt ggcagctgtg ttgctgggcc tgggtgctgt ggtggtgctg ctgctgcttc 300
[0472]	tgaggcactg gggctggggc ctgtgcctta tcggctggaa cgagttcatc ctgcagccca 360
[0473]	tccacaacct gctcatgggt gacaccaagg agcagcgcac cctgaaccac gtgctgcagc 420
[0474]	atgcggagcc cgggaacgca cagagcgtgc tggaggccat tgacacctac tgcgagcaga 480
[0475]	aggagtgggc catgaacgtg ggcgacaaga aaggcaagat cgtggacgcc gtgattcagg 540
[0476]	agcaccagcc ctccgtgctg ctggagctgg gggcctactg tggctactca gctgtgcgca 600
[0477]	tggccgcct gctgtcacca ggggcgaggc tcatcaccat cgagatcaac cccgactgtg 660
[0478]	ccgccatcac ccagcggatg gtggatttcg ctggcgtgaa ggacaaggtc acccttgttg 720
[0479]	ttggagcgtc ccaggacatc atccccagc tgaagaagaa gtatgatgtg gacacactgg 780
[0480]	acatggtctt cctcgaccac tggaaggacc ggtacctgcc ggacacgctt ctcttgagg 840
[0481]	aatgtggcct gctgcggaag gggacagtgc tactggctga caacgtgatc tgcccagggtg 900
[0482]	cgccagactt cctagcacac gtgcgcggga gcagctgctt tgagtgcaca cactaccaat 960
[0483]	cgttcctgga atacaggag gtggtggacg gcctggagaa ggccatctac aaggggccag 1020
[0484]	gcagcgaagc agggccctga ctgccccccc ggccccctc tcgggctctc tcaccagcc 1080
[0485]	tgggtactgaa ggtgccagac gtgctcctgc tgacctctg cggtccggg ctgtgtccta 1140
[0486]	aatgcaaagc acacctcggc cgaggcctgc gccctgacat gctaacctct ctgaactgca 1200
[0487]	acactggatt gttctttttt aagactcaat catgacttct ttactaacac tggctagcta 1260
[0488]	tattatctta tatactaata tcatgtttta aaaatataaa atagaaatta agaattctaa 1320
[0489]	tatttagata taactcgact tagtacatcc ttctcaactg ccattcccct gctgcccttg 1380
[0490]	acttgggcac caaacattca aagctcccct tgacggacgc taacgctaag ggcgggggcc 1440
[0491]	ctagctggct gggttctggg tggcacgcct ggcccactgg cctcccagcc acagtgtgct 1500
[0492]	agaggtcagc cctcctgcag ctaggccagg ggcacctgtt agcccatgg ggacgactgc 1560
[0493]	cggcctggga aacgaagagg agtcagccag cattcacacc tttctgacca agcaggcgct 1620
[0494]	ggggacaggt ggaccccgca gcagcaccag cccctctggg ccccatgtgg cacagagtgg 1680
[0495]	aagcatctcc ttccctactc cccactgggc cttgcttaca gaagaggcaa tggctcagac 1740
[0496]	cagctccgc atccctgtag ttgcctccct ggcccatgag tgaggatgca gtgctggttt 1800
[0497]	ctgcccacct acacctagag ctgtcccat ctcctccaag gggtcagact gctagccacc 1860
[0498]	tcagaggctc caaggccca gttcccaggc ccaggacagg aatcaacctt gtgctagctg 1920
[0499]	agttcacctg caccgagacc agcccctagc caagattcta ctctgggt caaggcctgg 1980
[0500]	ctagcccca gccagccac tcctatggat agacagacca gtgagccaa gtggacaagt 2040
[0501]	ttggggccac ccaggacca gaaacagagc ctctgcagga cacagcagat gggcacctgg 2100
[0502]	gaccacctcc acccagggcc ctgcccaga cgcgagagg cccgacaaa gggagaagcc 2160
[0503]	agccacttgt gccagacctg agtggcagaa agcaaaaagt tcctttgctg ctttaatttt 2220

[0504]	taaattttct tacaaaaatt taggtgttta ccaatagtct tatttttggt tatttttaa	2279
[0505]	<210>	14
[0506]	<211>	2035
[0507]	<212>	DNA
[0508]	<213>	智人
[0509]	<400>	14
[0510]	gctgttggca gctgtgttgc tgggcctggt gctgctggtg gtgctgctgc tgcttctgag	60
[0511]	gcactggggc tggggcctgt gccttatcgg ctggaacgag ttcactcctgc agcccatcca	120
[0512]	caacctgctc atgggtgaca ccaaggagca gcgcacctg aaccacgtgc tgcagcatgc	180
[0513]	ggagcccggg aacgcacaga gcgtgctgga ggccattgac acctactgcg agcagaagga	240
[0514]	gtgggccatg aacgtgggcg acaagaaagg caagatcgtg gacgccgtga ttcaggagca	300
[0515]	ccagccctcc gtgctgctgg agctgggggc ctactgtggc tactcagctg tgcgcatggc	360
[0516]	ccgcctgctg tcaccagggg cgaggctcat caccatcgag atcaacccc actgtgccgc	420
[0517]	catcaccag cggtggtgg atttcgctgg cgtgaaggac aaggtcaccc ttgtggttgg	480
[0518]	agcgtcccag gacatcatcc cccagctgaa gaagaagtat gatgtggaca cactggacat	540
[0519]	ggtcttctc gacctgga aggaccgta cctgccggac acgcttctct tggaggaatg	600
[0520]	tggcctgctg cggaagggga cagtgtact ggctgacaac gtgatctgcc cagggtgcc	660
[0521]	agacttcta gcacacgtgc gcgggagcag ctgctttgag tgcacacact accaatcgtt	720
[0522]	cctggaatac agggaggtgg tggacggcct ggagaaggcc atctacaagg gccaggcag	780
[0523]	cgaagcagg ccctgactgc cccccggcc cccctctcgg gctctctcac ccagcctggt	840
[0524]	actgaagggt ccagacgtgc tctgtgtgac cttctgcggc tccgggctgt gtctaaatg	900
[0525]	caaagcacac ctcgccgag gcctgcgcc tgacatgcta acctctctga actgcaacac	960
[0526]	tggattgttc tttttaaga ctcaatcatg acttctttac taacactggc tagctatatt	1020
[0527]	atcttatata ctaatatcat gttttaaaaa tataaaatag aaattaagaa tctaaatatt	1080
[0528]	tagatataac tcgacttagt acatccttct caactgccat tcccctgtg cccctgactt	1140
[0529]	gggcacaaa cattcaaagc tccccttgac ggacgctaac gctaaggcg gggcccctag	1200
[0530]	ctggctgggt tctgggtggc acgcctggcc cactggcctc ccagccacag tgggtgcagag	1260
[0531]	gtcagccctc ctgcagctag gccaggggca cctgttagcc ccatggggac gactgccggc	1320
[0532]	ctgggaaacg aagaggagtc agccagcatt cacaccttct tgaccaagca ggcgctgggg	1380
[0533]	acaggtggac cccgcagcag caccagcccc tctgggcccc atgtggcaca gagtggaagc	1440
[0534]	atctccttcc ctactccca ctgggccttg cttacagaag aggcaatggc tcagaccagc	1500
[0535]	tcccgcatcc ctgtagttgc ctccctggcc catgagttag gatgcagtgc tggtttctgc	1560
[0536]	ccacctacac ctagagctgt ccccatctcc tccaaggggt cagactgcta gccacctcag	1620
[0537]	aggtccaag ggcccagttc ccaggcccag gacaggaatc aaccctgtgc tagctgagtt	1680
[0538]	cacctgcacc gagaccagcc cctagccaag attctactcc tgggctcaag gcctggctag	1740
[0539]	ccccagcca gccactcct atggatagac agaccagtga gcccaagtgg acaagtttgg	1800
[0540]	ggccaccag ggaccagaaa cagagcctct gcaggacaca gcagatgggc acctgggacc	1860
[0541]	acctccacc agggccctgc ccagacgcg cagaggcccg acacaaggga gaagccagcc	1920
[0542]	acttggtcca gacctgagt gcagaaagca aaaagttcct ttgctgcttt aatttttaa	1980
[0543]	ttttcttaca aaaatttagg tgtttacaa tagtcttatt ttgcttatt tttaa	2035
[0544]	<210>	15
[0545]	<211>	271

[0546] <212> PRT
 [0547] <213> 智人
 [0548] <400> 15
 [0549] Met Pro Glu Ala Pro Pro Leu Leu Leu Ala Ala Val Leu Leu Gly Leu
 [0550] 1 5 10 15
 [0551] Val Leu Leu Val Val Leu Leu Leu Leu Leu Arg His Trp Gly Trp Gly
 [0552] 20 25 30
 [0553] Leu Cys Leu Ile Gly Trp Asn Glu Phe Ile Leu Gln Pro Ile His Asn
 [0554] 35 40 45
 [0555] Leu Leu Met Gly Asp Thr Lys Glu Gln Arg Ile Leu Asn His Val Leu
 [0556] 50 55 60
 [0557] Gln His Ala Glu Pro Gly Asn Ala Gln Ser Val Leu Glu Ala Ile Asp
 [0558] 65 70 75 80
 [0559] Thr Tyr Cys Glu Gln Lys Glu Trp Ala Met Asn Val Gly Asp Lys Lys
 [0560] 85 90 95
 [0561] Gly Lys Ile Val Asp Ala Val Ile Gln Glu His Gln Pro Ser Val Leu
 [0562] 100 105 110
 [0563] Leu Glu Leu Gly Ala Tyr Cys Gly Tyr Ser Ala Val Arg Met Ala Arg
 [0564] 115 120 125
 [0565] Leu Leu Ser Pro Gly Ala Arg Leu Ile Thr Ile Glu Ile Asn Pro Asp
 [0566] 130 135 140
 [0567] Cys Ala Ala Ile Thr Gln Arg Met Val Asp Phe Ala Gly Val Lys Asp
 [0568] 145 150 155 160
 [0569] Lys Val Thr Leu Val Val Gly Ala Ser Gln Asp Ile Ile Pro Gln Leu
 [0570] 165 170 175
 [0571] Lys Lys Lys Tyr Asp Val Asp Thr Leu Asp Met Val Phe Leu Asp His
 [0572] 180 185 190
 [0573] Trp Lys Asp Arg Tyr Leu Pro Asp Thr Leu Leu Leu Glu Glu Cys Gly
 [0574] 195 200 205
 [0575] Leu Leu Arg Lys Gly Thr Val Leu Leu Ala Asp Asn Val Ile Cys Pro
 [0576] 210 215 220
 [0577] Gly Ala Pro Asp Phe Leu Ala His Val Arg Gly Ser Ser Cys Phe Glu
 [0578] 225 230 235 240
 [0579] Cys Thr His Tyr Gln Ser Phe Leu Glu Tyr Arg Glu Val Val Asp Gly
 [0580] 245 250 255
 [0581] Leu Glu Lys Ala Ile Tyr Lys Gly Pro Gly Ser Glu Ala Gly Pro
 [0582] 260 265 270
 [0583] <210> 16
 [0584] <211> 221
 [0585] <212> PRT
 [0586] <213> 智人
 [0587] <400> 16

[0588]	Met Gly Asp Thr Lys Glu Gln Arg Ile Leu Asn His Val Leu Gln His
[0589]	1 5 10 15
[0590]	Ala Glu Pro Gly Asn Ala Gln Ser Val Leu Glu Ala Ile Asp Thr Tyr
[0591]	20 25 30
[0592]	Cys Glu Gln Lys Glu Trp Ala Met Asn Val Gly Asp Lys Lys Gly Lys
[0593]	35 40 45
[0594]	Ile Val Asp Ala Val Ile Gln Glu His Gln Pro Ser Val Leu Leu Glu
[0595]	50 55 60
[0596]	Leu Gly Ala Tyr Cys Gly Tyr Ser Ala Val Arg Met Ala Arg Leu Leu
[0597]	65 70 75 80
[0598]	Ser Pro Gly Ala Arg Leu Ile Thr Ile Glu Ile Asn Pro Asp Cys Ala
[0599]	85 90 95
[0600]	Ala Ile Thr Gln Arg Met Val Asp Phe Ala Gly Val Lys Asp Lys Val
[0601]	100 105 110
[0602]	Thr Leu Val Val Gly Ala Ser Gln Asp Ile Ile Pro Gln Leu Lys Lys
[0603]	115 120 125
[0604]	Lys Tyr Asp Val Asp Thr Leu Asp Met Val Phe Leu Asp His Trp Lys
[0605]	130 135 140
[0606]	Asp Arg Tyr Leu Pro Asp Thr Leu Leu Leu Glu Glu Cys Gly Leu Leu
[0607]	145 150 155 160
[0608]	Arg Lys Gly Thr Val Leu Leu Ala Asp Asn Val Ile Cys Pro Gly Ala
[0609]	165 170 175
[0610]	Pro Asp Phe Leu Ala His Val Arg Gly Ser Ser Cys Phe Glu Cys Thr
[0611]	180 185 190
[0612]	His Tyr Gln Ser Phe Leu Glu Tyr Arg Glu Val Val Asp Gly Leu Glu
[0613]	195 200 205
[0614]	Lys Ala Ile Tyr Lys Gly Pro Gly Ser Glu Ala Gly Pro
[0615]	210 215 220
[0616]	<210> 17
[0617]	<211> 3453
[0618]	<212> DNA
[0619]	<213> 巴西诺卡菌
[0620]	<400> 17
[0621]	ttgttcgcgc aggacgagca ggtgaaagcc gcggtgccgg accaggaggt ggtcgaggcg 60
[0622]	atccggggcgc ccggcctgcg cctggcacag atcatggcca ccgtgatgga gcgctatgcg 120
[0623]	gaccgccccg cgggtgggaca gcgggcgagc gagccgggtca ccgagagcgg tcgcaccacc 180
[0624]	ttccggctgc tcccggaatt cgagaccctg acctaccgcg agctgtgggc gcgcgtccgc 240
[0625]	gcggtggccg ccgcgtggca cggagatgcc gaaaggcctt tgcgggcccgg ggatttcgtt 300
[0626]	gctctgctgg gtttcgccgg catcgattac ggcaccctcg atctcgcgaa catccatctc 360
[0627]	ggcctcgtca cggtgccgct gcaatccggc gccacggccc cgcaactcgc cgcgatcctg 420
[0628]	gccgagacca cgccccgggt gctggccgcg acaccgacc atctcgatat cgccgtcgaa 480
[0629]	ttgctgaccg ggggagcctc gccggaacgg ctggtggtat tcgactaccg ccccgcgga 540

[0630]	gacgatcacc	gggcggcgct	cgagtccgcg	cgcagacggt	tgagcgacgc	gggcagtgcg	600
[0631]	gtggtggtcg	agacgctcga	cgcggtccgc	gcccgcggca	gcgaattgcc	ggccgcgccg	660
[0632]	ctgttcgttc	ccgccgcgga	cgaggaccgc	ctggctctgc	tcatctacac	ctccggcagc	720
[0633]	accggcacgc	ctaagggcgc	catgtacacc	gaaagactga	accgcacgac	gtggctgagc	780
[0634]	ggggcgaaag	gcgtcggcct	cacgctcggc	tacatgccga	tgagtcatat	tgccggggcg	840
[0635]	gcctcgttcg	ccggtgtgct	ggcccgcggc	ggcacgggtct	acttcaccgc	ccgcagcgat	900
[0636]	atgtcgacgc	tgttcgaaga	tctggccctg	gtgcggccga	ccgagatgtt	cttcgtcccg	960
[0637]	cgcgtgtgcg	acatgatctt	ccagcgtat	caggccgaac	tgtcgcggcg	cgcgcccgc	1020
[0638]	gcggccgcga	gcccgggaact	cgagcaggaa	ctgaagaccg	aactgcgctt	gtccgcggtc	1080
[0639]	ggggaccgct	tactcggggc	gatcgcgggc	agcgcgccgc	tgtcggccga	gatgcgggag	1140
[0640]	ttcatggagt	cgtctgtgga	tctggaactg	cacgacggct	acggctcgac	cgaggcgggt	1200
[0641]	atcggcgta	tgcaagacaa	tatcgtccag	cgtccgccgg	tcatcgatta	caagctcgtc	1260
[0642]	gacgtgccgg	aattgggcta	cttcgggacg	gaccagccgc	atccccgcgg	tgagttgctg	1320
[0643]	ttgaaaaccg	aagggatgat	tccgggttac	ttccggcggc	ccgaggtgac	cgcggagatc	1380
[0644]	ttcgacgagg	acggtttcta	caggaccggg	gacatcgtcg	ccgaactcga	accggatcgg	1440
[0645]	ctgatctacc	tggaccgccg	caacaatgtg	ctgaaactgg	cccagggcga	gttcgtcacg	1500
[0646]	gtcgcctatc	tggaagcggg	gttcgcgacc	agtccgctga	tccggcagat	ctacatctac	1560
[0647]	ggcaacagcg	agcgtctggt	cctgctggcg	gtgatcgtgc	ccaccgcgga	cgcgtggcc	1620
[0648]	gacggtgtca	ccgacgcgct	gaacacggcg	ctgaccgaat	ccttgcgaca	gctcgcgaaa	1680
[0649]	gaagccgggc	tgcaatccta	tgagctgccg	cgcgagttcc	tggtcgaaac	cgaaccgttc	1740
[0650]	accgtcgaga	acggtctgct	ctccgggtatc	gcgaaactgt	tgcgggccaa	gctcaaggag	1800
[0651]	cactacggcg	agcgactcga	gcagctgtac	cgcgatatcg	aggcgaaccg	caacgacgag	1860
[0652]	ctgatcgagc	tgccggcgac	cgcggccgag	ctgccgggtg	tcgaaaccgt	cacgcgggct	1920
[0653]	gcacgttcga	tgctcggact	ggccgcgtcg	gagttgcggc	cggacgcgca	tttcaccgat	1980
[0654]	ctcggcggtg	attcactgtc	cgcgtgtcg	ttttcgacct	tgctgcagga	catgctcgag	2040
[0655]	gtcgaggtcc	cggctcgggt	catcgtgagc	cccgccaaact	cgtctgccga	tctggcgaaa	2100
[0656]	tacatcgagg	ccgaacggca	ttcgggggtg	cggcggccga	gcctgatctc	ggtgcacggt	2160
[0657]	cccggcaccg	agatccgtgc	cgcggtatc	accctggaca	agttcatcga	cgagcgcacc	2220
[0658]	ctcgtgcgcg	cgaagcggg	tccggccgcg	cggcccagg	cgcagaccgt	cctgctcacc	2280
[0659]	ggggcgaaac	gctatctcgg	ccgcttcctg	tgcttggaat	ggctgcagcg	actggaccag	2340
[0660]	accggcggca	cgtcgtctg	catcgtgcgc	ggtaccgacg	cggccgcgcg	gcggaagcgc	2400
[0661]	ctggatgcgg	tgttcgacag	cggatgatccg	gagctgctcg	accactaccg	gaagctggcc	2460
[0662]	gccgagcacc	tcgaggtgct	cgcgggcgat	atcggcgacc	cgaatctcgg	cctggacgaa	2520
[0663]	gcgacttggc	agcggctcgc	cgcgaccgtc	gacctgatcg	tgcacccgcg	cgcctcgtc	2580
[0664]	aaccatgtgc	tgccgtacag	ccagctgttc	gggccgaatg	tggtcggcac	cgccgagatc	2640
[0665]	atccggtcgg	ccatcaccga	gcgccgtaag	cccgtgacgt	acctgtcgac	ggtcgcgggt	2700
[0666]	gccgcacagg	tcgatccgcg	cggcttcgac	gaggagcgcg	atatccggga	gatgagcgcg	2760
[0667]	gtgcgtcca	tcgaccccg	gtacgcgaac	ggttacggca	acagcaagtg	ggccggcgag	2820
[0668]	gtgctgctgc	gcgaggccca	tgatctgtgc	gggctgccgg	tcgccgtgtt	ccgctcggac	2880
[0669]	atgatcctgg	cgcacagcaa	atacgtcggg	cagctcaacg	tccccgatgt	gttcaccggg	2940
[0670]	ctcatcctga	gcctggcgct	caccggcatc	gcaccgtatt	cgttctacgg	gacggacagc	3000
[0671]	gccgggcagc	gcaggcgggc	ccactacgac	ggtctgcccc	ccgatttcgt	cgcgaggcg	3060

[0672] atcaccaccc tcggcgcgcg agccgagtcg gggttccata cctacgacgt gtggaacccg 3120
 [0673] tacgacgacg gcatctcgct ggacgaattc gtcgactggc tcggcgattt cggcgtgccg 3180
 [0674] atccagcgga tcgacgacta cgacgaatgg ttccggcggtt tcgagaccgc gatccgcgcg 3240
 [0675] ctgcccgaaa agcagcgcgga tgcttcgctg ctaccgctgc tggacgcaca ccggcggcca 3300
 [0676] ctgcgcgcggtg tcgcgcggttc gctgttggcc gccaaagaact tccaggcggc ggtgcagtcc 3360
 [0677] gcgcggatcg gccccgatca ggacatcccg catctttccc cgcagttgat cgacaagtac 3420
 [0678] gtcaccgacc tgcgccacct cggcctgctc tga 3453
 [0679] <210> 18
 [0680] <211> 1150
 [0681] <212> PRT
 [0682] <213> 巴西诺卡菌
 [0683] <400> 18
 [0684] Met Phe Ala Glu Asp Glu Gln Val Lys Ala Ala Val Pro Asp Gln Glu
 [0685] 1 5 10 15
 [0686] Val Val Glu Ala Ile Arg Ala Pro Gly Leu Arg Leu Ala Gln Ile Met
 [0687] 20 25 30
 [0688] Ala Thr Val Met Glu Arg Tyr Ala Asp Arg Pro Ala Val Gly Gln Arg
 [0689] 35 40 45
 [0690] Ala Ser Glu Pro Val Thr Glu Ser Gly Arg Thr Thr Phe Arg Leu Leu
 [0691] 50 55 60
 [0692] Pro Glu Phe Glu Thr Leu Thr Tyr Arg Glu Leu Trp Ala Arg Val Arg
 [0693] 65 70 75 80
 [0694] Ala Val Ala Ala Ala Trp His Gly Asp Ala Glu Arg Pro Leu Arg Ala
 [0695] 85 90 95
 [0696] Gly Asp Phe Val Ala Leu Leu Gly Phe Ala Gly Ile Asp Tyr Gly Thr
 [0697] 100 105 110
 [0698] Leu Asp Leu Ala Asn Ile His Leu Gly Leu Val Thr Val Pro Leu Gln
 [0699] 115 120 125
 [0700] Ser Gly Ala Thr Ala Pro Gln Leu Ala Ala Ile Leu Ala Glu Thr Thr
 [0701] 130 135 140
 [0702] Pro Arg Val Leu Ala Ala Thr Pro Asp His Leu Asp Ile Ala Val Glu
 [0703] 145 150 155 160
 [0704] Leu Leu Thr Gly Gly Ala Ser Pro Glu Arg Leu Val Val Phe Asp Tyr
 [0705] 165 170 175
 [0706] Arg Pro Ala Asp Asp Asp His Arg Ala Ala Leu Glu Ser Ala Arg Arg
 [0707] 180 185 190
 [0708] Arg Leu Ser Asp Ala Gly Ser Ala Val Val Val Glu Thr Leu Asp Ala
 [0709] 195 200 205
 [0710] Val Arg Ala Arg Gly Ser Glu Leu Pro Ala Ala Pro Leu Phe Val Pro
 [0711] 210 215 220
 [0712] Ala Ala Asp Glu Asp Pro Leu Ala Leu Leu Ile Tyr Thr Ser Gly Ser
 [0713] 225 230 235 240

[0714]	Thr Gly Thr Pro Lys Gly Ala Met Tyr Thr Glu Arg Leu Asn Arg Thr		
[0715]		245	250 255
[0716]	Thr Trp Leu Ser Gly Ala Lys Gly Val Gly Leu Thr Leu Gly Tyr Met		
[0717]		260	265 270
[0718]	Pro Met Ser His Ile Ala Gly Arg Ala Ser Phe Ala Gly Val Leu Ala		
[0719]		275	280 285
[0720]	Arg Gly Gly Thr Val Tyr Phe Thr Ala Arg Ser Asp Met Ser Thr Leu		
[0721]		290	295 300
[0722]	Phe Glu Asp Leu Ala Leu Val Arg Pro Thr Glu Met Phe Phe Val Pro		
[0723]	305	310	315 320
[0724]	Arg Val Cys Asp Met Ile Phe Gln Arg Tyr Gln Ala Glu Leu Ser Arg		
[0725]		325	330 335
[0726]	Arg Ala Pro Ala Ala Ala Ala Ser Pro Glu Leu Glu Gln Glu Leu Lys		
[0727]		340	345 350
[0728]	Thr Glu Leu Arg Leu Ser Ala Val Gly Asp Arg Leu Leu Gly Ala Ile		
[0729]		355	360 365
[0730]	Ala Gly Ser Ala Pro Leu Ser Ala Glu Met Arg Glu Phe Met Glu Ser		
[0731]		370	375 380
[0732]	Leu Leu Asp Leu Glu Leu His Asp Gly Tyr Gly Ser Thr Glu Ala Gly		
[0733]	385	390	395 400
[0734]	Ile Gly Val Leu Gln Asp Asn Ile Val Gln Arg Pro Pro Val Ile Asp		
[0735]		405	410 415
[0736]	Tyr Lys Leu Val Asp Val Pro Glu Leu Gly Tyr Phe Arg Thr Asp Gln		
[0737]		420	425 430
[0738]	Pro His Pro Arg Gly Glu Leu Leu Leu Lys Thr Glu Gly Met Ile Pro		
[0739]		435	440 445
[0740]	Gly Tyr Phe Arg Arg Pro Glu Val Thr Ala Glu Ile Phe Asp Glu Asp		
[0741]		450	455 460
[0742]	Gly Phe Tyr Arg Thr Gly Asp Ile Val Ala Glu Leu Glu Pro Asp Arg		
[0743]	465	470	475 480
[0744]	Leu Ile Tyr Leu Asp Arg Arg Asn Asn Val Leu Lys Leu Ala Gln Gly		
[0745]		485	490 495
[0746]	Glu Phe Val Thr Val Ala His Leu Glu Ala Val Phe Ala Thr Ser Pro		
[0747]		500	505 510
[0748]	Leu Ile Arg Gln Ile Tyr Ile Tyr Gly Asn Ser Glu Arg Ser Phe Leu		
[0749]		515	520 525
[0750]	Leu Ala Val Ile Val Pro Thr Ala Asp Ala Leu Ala Asp Gly Val Thr		
[0751]		530	535 540
[0752]	Asp Ala Leu Asn Thr Ala Leu Thr Glu Ser Leu Arg Gln Leu Ala Lys		
[0753]	545	550	555 560
[0754]	Glu Ala Gly Leu Gln Ser Tyr Glu Leu Pro Arg Glu Phe Leu Val Glu		
[0755]		565	570 575

[0756]	Thr	Glu	Pro	Phe	Thr	Val	Glu	Asn	Gly	Leu	Leu	Ser	Gly	Ile	Ala	Lys
[0757]							580									590
[0758]	Leu	Leu	Arg	Pro	Lys	Leu	Lys	Glu	His	Tyr	Gly	Glu	Arg	Leu	Glu	Gln
[0759]							595									605
[0760]	Leu	Tyr	Arg	Asp	Ile	Glu	Ala	Asn	Arg	Asn	Asp	Glu	Leu	Ile	Glu	Leu
[0761]							610									620
[0762]	Arg	Arg	Thr	Ala	Ala	Glu	Leu	Pro	Val	Leu	Glu	Thr	Val	Thr	Arg	Ala
[0763]							625									640
[0764]	Ala	Arg	Ser	Met	Leu	Gly	Leu	Ala	Ala	Ser	Glu	Leu	Arg	Pro	Asp	Ala
[0765]							645									655
[0766]	His	Phe	Thr	Asp	Leu	Gly	Gly	Asp	Ser	Leu	Ser	Ala	Leu	Ser	Phe	Ser
[0767]							660									670
[0768]	Thr	Leu	Leu	Gln	Asp	Met	Leu	Glu	Val	Glu	Val	Pro	Val	Gly	Val	Ile
[0769]							675									685
[0770]	Val	Ser	Pro	Ala	Asn	Ser	Leu	Ala	Asp	Leu	Ala	Lys	Tyr	Ile	Glu	Ala
[0771]							690									700
[0772]	Glu	Arg	His	Ser	Gly	Val	Arg	Arg	Pro	Ser	Leu	Ile	Ser	Val	His	Gly
[0773]							705									720
[0774]	Pro	Gly	Thr	Glu	Ile	Arg	Ala	Ala	Asp	Leu	Thr	Leu	Asp	Lys	Phe	Ile
[0775]							725									735
[0776]	Asp	Glu	Arg	Thr	Leu	Ala	Ala	Ala	Lys	Ala	Val	Pro	Ala	Ala	Pro	Ala
[0777]							740									750
[0778]	Gln	Ala	Gln	Thr	Val	Leu	Leu	Thr	Gly	Ala	Asn	Gly	Tyr	Leu	Gly	Arg
[0779]							755									765
[0780]	Phe	Leu	Cys	Leu	Glu	Trp	Leu	Gln	Arg	Leu	Asp	Gln	Thr	Gly	Gly	Thr
[0781]							770									780
[0782]	Leu	Val	Cys	Ile	Val	Arg	Gly	Thr	Asp	Ala	Ala	Ala	Ala	Arg	Lys	Arg
[0783]							785									800
[0784]	Leu	Asp	Ala	Val	Phe	Asp	Ser	Gly	Asp	Pro	Glu	Leu	Leu	Asp	His	Tyr
[0785]							805									815
[0786]	Arg	Lys	Leu	Ala	Ala	Glu	His	Leu	Glu	Val	Leu	Ala	Gly	Asp	Ile	Gly
[0787]							820									830
[0788]	Asp	Pro	Asn	Leu	Gly	Leu	Asp	Glu	Ala	Thr	Trp	Gln	Arg	Leu	Ala	Ala
[0789]							835									845
[0790]	Thr	Val	Asp	Leu	Ile	Val	His	Pro	Ala	Ala	Leu	Val	Asn	His	Val	Leu
[0791]							850									860
[0792]	Pro	Tyr	Ser	Gln	Leu	Phe	Gly	Pro	Asn	Val	Val	Gly	Thr	Ala	Glu	Ile
[0793]							865									880
[0794]	Ile	Arg	Leu	Ala	Ile	Thr	Glu	Arg	Arg	Lys	Pro	Val	Thr	Tyr	Leu	Ser
[0795]							885									895
[0796]	Thr	Val	Ala	Val	Ala	Ala	Gln	Val	Asp	Pro	Ala	Gly	Phe	Asp	Glu	Glu
[0797]							900									910

[0798]	Arg Asp Ile Arg Glu Met Ser Ala Val Arg Ser Ile Asp Ala Gly Tyr
[0799]	915 920 925
[0800]	Ala Asn Gly Tyr Gly Asn Ser Lys Trp Ala Gly Glu Val Leu Leu Arg
[0801]	930 935 940
[0802]	Glu Ala His Asp Leu Cys Gly Leu Pro Val Ala Val Phe Arg Ser Asp
[0803]	945 950 955 960
[0804]	Met Ile Leu Ala His Ser Lys Tyr Val Gly Gln Leu Asn Val Pro Asp
[0805]	965 970 975
[0806]	Val Phe Thr Arg Leu Ile Leu Ser Leu Ala Leu Thr Gly Ile Ala Pro
[0807]	980 985 990
[0808]	Tyr Ser Phe Tyr Gly Thr Asp Ser Ala Gly Gln Arg Arg Arg Ala His
[0809]	995 1000 1005
[0810]	Tyr Asp Gly Leu Pro Ala Asp Phe Val Ala Glu Ala Ile Thr Thr
[0811]	1010 1015 1020
[0812]	Leu Gly Ala Arg Ala Glu Ser Gly Phe His Thr Tyr Asp Val Trp
[0813]	1025 1030 1035
[0814]	Asn Pro Tyr Asp Asp Gly Ile Ser Leu Asp Glu Phe Val Asp Trp
[0815]	1040 1045 1050
[0816]	Leu Gly Asp Phe Gly Val Pro Ile Gln Arg Ile Asp Asp Tyr Asp
[0817]	1055 1060 1065
[0818]	Glu Trp Phe Arg Arg Phe Glu Thr Ala Ile Arg Ala Leu Pro Glu
[0819]	1070 1075 1080
[0820]	Lys Gln Arg Asp Ala Ser Leu Leu Pro Leu Leu Asp Ala His Arg
[0821]	1085 1090 1095
[0822]	Arg Pro Leu Arg Ala Val Arg Gly Ser Leu Leu Pro Ala Lys Asn
[0823]	1100 1105 1110
[0824]	Phe Gln Ala Ala Val Gln Ser Ala Arg Ile Gly Pro Asp Gln Asp
[0825]	1115 1120 1125
[0826]	Ile Pro His Leu Ser Pro Gln Leu Ile Asp Lys Tyr Val Thr Asp
[0827]	1130 1135 1140
[0828]	Leu Arg His Leu Gly Leu Leu
[0829]	1145 1150
[0830]	<210> 19
[0831]	<211> 3501
[0832]	<212> DNA
[0833]	<213> 巴西诺卡菌
[0834]	<400> 19
[0835]	atggcgactg attcgcaag cgatcggeta cggcgctcgaa ttgcacagtt gttegccgag 60
[0836]	gacgagcagg tgaagccgc ggtgccggac caggaggtgg tcgaggcgat ccgggcgccc 120
[0837]	ggcctgcgcc tggcacagat catggccacc gtgatggagc gctatgcgga ccgccccgcg 180
[0838]	gtgggacagc gggcgagcga gccggtcacc gagagcggtc gcaccacatt ccggtgctc 240
[0839]	ccggaattcg agacctgac ctaccgcgag ctgtgggcgc gcgtccgcgc ggtggccgcc 300

[0840]	gcgtggcacg gagatgccga aaggcctttg cgggccgggg atttcgttgc tctgctgggt	360
[0841]	ttcgccggca tcgattacgg caccctcgat ctgcgaaca tccatctcgg cctcgtcacg	420
[0842]	gtgccgtgc aatccggcgc caccggcccg caactcgccg cgatcctggc cgagaccacg	480
[0843]	ccccgggtgc tggccgcgac acccgaccat ctcgatatcg ccgtcgaatt gctgaccggg	540
[0844]	ggagcctcgc cggaacggct ggtggtattc gactaccgcc ccgcggacga cgatcaccgg	600
[0845]	gcggcgctcg agtccgcgcg cagacggttg agcgacgcgg gcagtgcggt ggtggtcgag	660
[0846]	acgtcgcacg cggtcgcgcg ccgcggcagc gaattgccgg ccgcgccgct gttcgttccc	720
[0847]	gccgcggacg aggaccgct ggctctgctc atctacacct ccggcagcac cggcacgcct	780
[0848]	aaggcgcca tgtacaccga aagactgaac cgcacgacgt ggctgagcgg ggcgaaaggc	840
[0849]	gtcggcctca cgctcgcta catgccgatg agtcatattg ccgggcgggc ctcgttcgcc	900
[0850]	ggtgtgctgg ccgcggcgg caccggtctac ttaccgccc gcagcgatat gtcgacgctg	960
[0851]	ttcgaagatc tggccctggt gcggccgacc gagatgttct tcgtcccgcg cgtgtgcgac	1020
[0852]	atgatcttcc agcgtatca ggccgaactg tcgcggcgcg cgcgccgcgc ggccgcgagc	1080
[0853]	ccggaactcg agcaggaact gaagaccgaa ctgcgcttgt ccgcggtcgg ggaccgctta	1140
[0854]	ctcggggcga tcgcgggcag cgcgccgctg tcggccgaga tgcgggagtt catggagtgc	1200
[0855]	ctgctggatc tggaactgca cgacggctac ggctcgaccg aggccgggtat cggcgactg	1260
[0856]	caagacaata tcgtccagcg tccgccggtc atcgattaca agctcgtcga cgtgccggaa	1320
[0857]	ttgggctact tccggacgga ccagccgcat ccccgcggtg agttgctgtt gaaaaccgaa	1380
[0858]	gggatgattc cgggctactt ccggcggccc gaggtgaccg cggagatctt cgacgaggac	1440
[0859]	ggtttctaca ggaccggtga catcgtcgcc gaactcgaac cggatcggct gatctacctg	1500
[0860]	gaccgccgca acaatgtgct gaaactggcc cagggcgagt tcgtcacggt cgcctatctg	1560
[0861]	gaagcgggtg tcgcgaccag tccgctgac cggcagatct acatctacgg caacagcgag	1620
[0862]	cgctcgttcc tgctggcggg gatcgtgccc accgcggacg cgctggccga cgggtgcacc	1680
[0863]	gacgcgctga acacggcgct gaccgaatcc ttgcgacagc tcgcgaaaga agccgggctg	1740
[0864]	caatcctatg agctgccgcg cgagttcctg gtcgaaaccg aaccgttcac cgtcgagaac	1800
[0865]	ggtctgctct ccggtatcgc gaaactgttg cggcccaagc tcaaggagca ctacggcgag	1860
[0866]	cgactcgagc agctgtaccg cgatatcgag gcgaaccgca acgacgagct gatcgagctg	1920
[0867]	cggcgcaccg cggccgagct gccggtgctc gaaaccgtca cgcgggctgc acgttcgatg	1980
[0868]	ctcggactgg ccgcgtcgga gttgcggccg gacgcgcatt tcaccgatct cggcgggtgat	2040
[0869]	tcaactgtcg cgctgtcgtt ttcgaccctg ctgcaggaca tgctcgaggt cgaggtccc	2100
[0870]	gtcgggtgca tcgtgagccc cgccaactcg ctgcgccgac tggcgaaata catcgaggcc	2160
[0871]	gaacggcatt cgggggtgcg gcggccgagc ctgatctcgg tgcacggtcc cggcaccgag	2220
[0872]	atccgtgccg ccgatctcac cctggacaag ttcatcgacg agcgcaccct cgctgccgcg	2280
[0873]	aaagcggttc cggccgcgcc ggcccaggcg cagaccgtcc tgctcaccgg ggcgaacggc	2340
[0874]	tatctcggcc gcttcctgtg cctggaatgg ctgcagcgac tggaccagac cggcggcacg	2400
[0875]	ctggtctgca tcgtgcgcgg taccgacgcg gccgccgcgc ggaagcgctt ggatgcgggtg	2460
[0876]	ttcgacagcg gtgatccgga gctgctcgac cactaccgga agctggccgc cgagcacctc	2520
[0877]	gaggtgctcg cgggcgatat cggcgacccg aatctcggcc tggacgaagc gacttggcag	2580
[0878]	cggctcgccg cgaccgtcga cctgatcgtg caccgccgcg ccctcgtcaa ccatgtgctg	2640
[0879]	ccgtacagcc agctgttcgg gccgaatgtg gtcggcaccg ccgagatcat ccggctggcc	2700
[0880]	atcaccgagc gccgtaagcc cgtgacgtac ctgtcgacgg tcgcgggtggc cgcacaggtc	2760
[0881]	gatcccgcgc gcttcgacga ggagcgcgat atccgggaga tgagcgcggt gcgtccatc	2820

[0882] gacgccgggt acgcgaacgg ttacggcaac agcaagtggg ccggcgaggt gctgctgcgc 2880
 [0883] gaggcccatg atctgtgcgg gctgccggtc gccgtgttcc gctcggacat gatcctggcg 2940
 [0884] cacagcaaat acgtcggtea gctcaacgtc cccgatgtgt tcaccggct catectgagc 3000
 [0885] ctggcgctca ccggcatcgc accgtattcg ttctacggga cggacagcgc cgggcagcgc 3060
 [0886] aggcggggccc actacgacgg tctgcccgcc gatttcgtcg ccgaggcgat caccaccctc 3120
 [0887] ggcgcgcgag ccgagtcggg gttccatacc tacgacgtgt ggaaccgta cgacgacggc 3180
 [0888] atctcgctgg acgaattcgt cgactggctc ggcgatttcg gcgtgccgat ccagcggatc 3240
 [0889] gacgactacg acgaatggtt ccggcgtttc gagaccgca tccgcgcgct gcccgaaaag 3300
 [0890] cagcgcgatg cttcgctgct accgctgctg gacgcacacc ggcggccact gcgcgcggtg 3360
 [0891] cgcggttcgc tgttccccgc caagaacttc caggcggcgg tgcagtccgc gcggatcggc 3420
 [0892] cccgatcagg acatcccga tctttcccg cagttgatcg acaagtacgt caccgacctg 3480
 [0893] cgccacctcg gctgctctg a 3501
 [0894] <210> 20
 [0895] <211> 1166
 [0896] <212> PRT
 [0897] <213> 巴西诺卡菌
 [0898] <400> 20
 [0899] Met Ala Thr Asp Ser Arg Ser Asp Arg Leu Arg Arg Arg Ile Ala Gln
 [0900] 1 5 10 15
 [0901] Leu Phe Ala Glu Asp Glu Gln Val Lys Ala Ala Val Pro Asp Gln Glu
 [0902] 20 25 30
 [0903] Val Val Glu Ala Ile Arg Ala Pro Gly Leu Arg Leu Ala Gln Ile Met
 [0904] 35 40 45
 [0905] Ala Thr Val Met Glu Arg Tyr Ala Asp Arg Pro Ala Val Gly Gln Arg
 [0906] 50 55 60
 [0907] Ala Ser Glu Pro Val Thr Glu Ser Gly Arg Thr Thr Phe Arg Leu Leu
 [0908] 65 70 75 80
 [0909] Pro Glu Phe Glu Thr Leu Thr Tyr Arg Glu Leu Trp Ala Arg Val Arg
 [0910] 85 90 95
 [0911] Ala Val Ala Ala Ala Trp His Gly Asp Ala Glu Arg Pro Leu Arg Ala
 [0912] 100 105 110
 [0913] Gly Asp Phe Val Ala Leu Leu Gly Phe Ala Gly Ile Asp Tyr Gly Thr
 [0914] 115 120 125
 [0915] Leu Asp Leu Ala Asn Ile His Leu Gly Leu Val Thr Val Pro Leu Gln
 [0916] 130 135 140
 [0917] Ser Gly Ala Thr Ala Pro Gln Leu Ala Ala Ile Leu Ala Glu Thr Thr
 [0918] 145 150 155 160
 [0919] Pro Arg Val Leu Ala Ala Thr Pro Asp His Leu Asp Ile Ala Val Glu
 [0920] 165 170 175
 [0921] Leu Leu Thr Gly Gly Ala Ser Pro Glu Arg Leu Val Val Phe Asp Tyr
 [0922] 180 185 190
 [0923] Arg Pro Ala Asp Asp Asp His Arg Ala Ala Leu Glu Ser Ala Arg Arg

[0924]	195	200	205
[0925]	Arg Leu Ser Asp Ala Gly Ser Ala Val Val Val Glu Thr Leu Asp Ala		
[0926]	210	215	220
[0927]	Val Arg Ala Arg Gly Ser Glu Leu Pro Ala Ala Pro Leu Phe Val Pro		
[0928]	225	230	235
[0929]	Ala Ala Asp Glu Asp Pro Leu Ala Leu Leu Ile Tyr Thr Ser Gly Ser		
[0930]	245	250	255
[0931]	Thr Gly Thr Pro Lys Gly Ala Met Tyr Thr Glu Arg Leu Asn Arg Thr		
[0932]	260	265	270
[0933]	Thr Trp Leu Ser Gly Ala Lys Gly Val Gly Leu Thr Leu Gly Tyr Met		
[0934]	275	280	285
[0935]	Pro Met Ser His Ile Ala Gly Arg Ala Ser Phe Ala Gly Val Leu Ala		
[0936]	290	295	300
[0937]	Arg Gly Gly Thr Val Tyr Phe Thr Ala Arg Ser Asp Met Ser Thr Leu		
[0938]	305	310	315
[0939]	Phe Glu Asp Leu Ala Leu Val Arg Pro Thr Glu Met Phe Phe Val Pro		
[0940]	325	330	335
[0941]	Arg Val Cys Asp Met Ile Phe Gln Arg Tyr Gln Ala Glu Leu Ser Arg		
[0942]	340	345	350
[0943]	Arg Ala Pro Ala Ala Ala Ala Ser Pro Glu Leu Glu Gln Glu Leu Lys		
[0944]	355	360	365
[0945]	Thr Glu Leu Arg Leu Ser Ala Val Gly Asp Arg Leu Leu Gly Ala Ile		
[0946]	370	375	380
[0947]	Ala Gly Ser Ala Pro Leu Ser Ala Glu Met Arg Glu Phe Met Glu Ser		
[0948]	385	390	395
[0949]	Leu Leu Asp Leu Glu Leu His Asp Gly Tyr Gly Ser Thr Glu Ala Gly		
[0950]	405	410	415
[0951]	Ile Gly Val Leu Gln Asp Asn Ile Val Gln Arg Pro Pro Val Ile Asp		
[0952]	420	425	430
[0953]	Tyr Lys Leu Val Asp Val Pro Glu Leu Gly Tyr Phe Arg Thr Asp Gln		
[0954]	435	440	445
[0955]	Pro His Pro Arg Gly Glu Leu Leu Leu Lys Thr Glu Gly Met Ile Pro		
[0956]	450	455	460
[0957]	Gly Tyr Phe Arg Arg Pro Glu Val Thr Ala Glu Ile Phe Asp Glu Asp		
[0958]	465	470	475
[0959]	Gly Phe Tyr Arg Thr Gly Asp Ile Val Ala Glu Leu Glu Pro Asp Arg		
[0960]	485	490	495
[0961]	Leu Ile Tyr Leu Asp Arg Arg Asn Asn Val Leu Lys Leu Ala Gln Gly		
[0962]	500	505	510
[0963]	Glu Phe Val Thr Val Ala His Leu Glu Ala Val Phe Ala Thr Ser Pro		
[0964]	515	520	525
[0965]	Leu Ile Arg Gln Ile Tyr Ile Tyr Gly Asn Ser Glu Arg Ser Phe Leu		

[0966]	530	535	540
[0967]	Leu Ala Val Ile Val Pro Thr Ala Asp Ala Leu Ala Asp Gly Val Thr		
[0968]	545	550	555 560
[0969]	Asp Ala Leu Asn Thr Ala Leu Thr Glu Ser Leu Arg Gln Leu Ala Lys		
[0970]	565	570	575
[0971]	Glu Ala Gly Leu Gln Ser Tyr Glu Leu Pro Arg Glu Phe Leu Val Glu		
[0972]	580	585	590
[0973]	Thr Glu Pro Phe Thr Val Glu Asn Gly Leu Leu Ser Gly Ile Ala Lys		
[0974]	595	600	605
[0975]	Leu Leu Arg Pro Lys Leu Lys Glu His Tyr Gly Glu Arg Leu Glu Gln		
[0976]	610	615	620
[0977]	Leu Tyr Arg Asp Ile Glu Ala Asn Arg Asn Asp Glu Leu Ile Glu Leu		
[0978]	625	630	635 640
[0979]	Arg Arg Thr Ala Ala Glu Leu Pro Val Leu Glu Thr Val Thr Arg Ala		
[0980]	645	650	655
[0981]	Ala Arg Ser Met Leu Gly Leu Ala Ala Ser Glu Leu Arg Pro Asp Ala		
[0982]	660	665	670
[0983]	His Phe Thr Asp Leu Gly Gly Asp Ser Leu Ser Ala Leu Ser Phe Ser		
[0984]	675	680	685
[0985]	Thr Leu Leu Gln Asp Met Leu Glu Val Glu Val Pro Val Gly Val Ile		
[0986]	690	695	700
[0987]	Val Ser Pro Ala Asn Ser Leu Ala Asp Leu Ala Lys Tyr Ile Glu Ala		
[0988]	705	710	715 720
[0989]	Glu Arg His Ser Gly Val Arg Arg Pro Ser Leu Ile Ser Val His Gly		
[0990]	725	730	735
[0991]	Pro Gly Thr Glu Ile Arg Ala Ala Asp Leu Thr Leu Asp Lys Phe Ile		
[0992]	740	745	750
[0993]	Asp Glu Arg Thr Leu Ala Ala Ala Lys Ala Val Pro Ala Ala Pro Ala		
[0994]	755	760	765
[0995]	Gln Ala Gln Thr Val Leu Leu Thr Gly Ala Asn Gly Tyr Leu Gly Arg		
[0996]	770	775	780
[0997]	Phe Leu Cys Leu Glu Trp Leu Gln Arg Leu Asp Gln Thr Gly Gly Thr		
[0998]	785	790	795 800
[0999]	Leu Val Cys Ile Val Arg Gly Thr Asp Ala Ala Ala Ala Arg Lys Arg		
[1000]	805	810	815
[1001]	Leu Asp Ala Val Phe Asp Ser Gly Asp Pro Glu Leu Leu Asp His Tyr		
[1002]	820	825	830
[1003]	Arg Lys Leu Ala Ala Glu His Leu Glu Val Leu Ala Gly Asp Ile Gly		
[1004]	835	840	845
[1005]	Asp Pro Asn Leu Gly Leu Asp Glu Ala Thr Trp Gln Arg Leu Ala Ala		
[1006]	850	855	860
[1007]	Thr Val Asp Leu Ile Val His Pro Ala Ala Leu Val Asn His Val Leu		

[1008]	865	870	875	880
[1009]	Pro Tyr Ser Gln Leu Phe Gly Pro Asn Val Val Gly Thr Ala Glu Ile			
[1010]		885	890	895
[1011]	Ile Arg Leu Ala Ile Thr Glu Arg Arg Lys Pro Val Thr Tyr Leu Ser			
[1012]		900	905	910
[1013]	Thr Val Ala Val Ala Ala Gln Val Asp Pro Ala Gly Phe Asp Glu Glu			
[1014]		915	920	925
[1015]	Arg Asp Ile Arg Glu Met Ser Ala Val Arg Ser Ile Asp Ala Gly Tyr			
[1016]		930	935	940
[1017]	Ala Asn Gly Tyr Gly Asn Ser Lys Trp Ala Gly Glu Val Leu Leu Arg			
[1018]	945	950	955	960
[1019]	Glu Ala His Asp Leu Cys Gly Leu Pro Val Ala Val Phe Arg Ser Asp			
[1020]		965	970	975
[1021]	Met Ile Leu Ala His Ser Lys Tyr Val Gly Gln Leu Asn Val Pro Asp			
[1022]		980	985	990
[1023]	Val Phe Thr Arg Leu Ile Leu Ser Leu Ala Leu Thr Gly Ile Ala Pro			
[1024]		995	1000	1005
[1025]	Tyr Ser Phe Tyr Gly Thr Asp Ser Ala Gly Gln Arg Arg Arg Ala			
[1026]		1010	1015	1020
[1027]	His Tyr Asp Gly Leu Pro Ala Asp Phe Val Ala Glu Ala Ile Thr			
[1028]		1025	1030	1035
[1029]	Thr Leu Gly Ala Arg Ala Glu Ser Gly Phe His Thr Tyr Asp Val			
[1030]		1040	1045	1050
[1031]	Trp Asn Pro Tyr Asp Asp Gly Ile Ser Leu Asp Glu Phe Val Asp			
[1032]		1055	1060	1065
[1033]	Trp Leu Gly Asp Phe Gly Val Pro Ile Gln Arg Ile Asp Asp Tyr			
[1034]		1070	1075	1080
[1035]	Asp Glu Trp Phe Arg Arg Phe Glu Thr Ala Ile Arg Ala Leu Pro			
[1036]		1085	1090	1095
[1037]	Glu Lys Gln Arg Asp Ala Ser Leu Leu Pro Leu Leu Asp Ala His			
[1038]		1100	1105	1110
[1039]	Arg Arg Pro Leu Arg Ala Val Arg Gly Ser Leu Leu Pro Ala Lys			
[1040]		1115	1120	1125
[1041]	Asn Phe Gln Ala Ala Val Gln Ser Ala Arg Ile Gly Pro Asp Gln			
[1042]		1130	1135	1140
[1043]	Asp Ile Pro His Leu Ser Pro Gln Leu Ile Asp Lys Tyr Val Thr			
[1044]		1145	1150	1155
[1045]	Asp Leu Arg His Leu Gly Leu Leu			
[1046]		1160	1165	
[1047]	<210> 21			
[1048]	<211> 621			
[1049]	<212> DNA			

[1050] <213> 大肠杆菌
 [1051] <400> 21
 [1052] atgaaaacta cgcatacctc cctccccctt gccggacata cgtgcattt tgttgagttc 60
 [1053] gatccggcga atttttgtga gcaggattta ctctggctgc cgcactacgc acaactgcaa 120
 [1054] cacgctggac gtaaacgtaa aacagagcat ttagccggac ggatcgctgc tgtttatgct 180
 [1055] ttgcgggaat atggctataa atgtgtgccc gcaatcggcg agctacgcca acctgtctgg 240
 [1056] cctgcggagg tatacggcag tattagccac tgtgggacta cggcattagc cgtggtatct 300
 [1057] cgtcaaccga ttggcattga tatagaagaa attttttctg taaaaccgc aagagaattg 360
 [1058] acagacaaca ttattacacc agcggaacac gagcgactcg cagactgcgg tttagccttt 420
 [1059] tctctggcgc tgacactggc attttccgcc aaagagagcg catttaaggc aagtgcagtc 480
 [1060] caaactgatg caggttttct ggactatcag ataattagct ggaataaaca gcaggtcatc 540
 [1061] attcatcgctg agaatgagat gtttgcctgtg cactggcaga taaaagaaaa gatagtcata 600
 [1062] acgctgtgcc aacacgatta a 621
 [1063] <210> 22
 [1064] <211> 206
 [1065] <212> PRT
 [1066] <213> 大肠杆菌
 [1067] <400> 22
 [1068] Met Lys Thr Thr His Thr Ser Leu Pro Phe Ala Gly His Thr Leu His
 [1069] 1 5 10 15
 [1070] Phe Val Glu Phe Asp Pro Ala Asn Phe Cys Glu Gln Asp Leu Leu Trp
 [1071] 20 25 30
 [1072] Leu Pro His Tyr Ala Gln Leu Gln His Ala Gly Arg Lys Arg Lys Thr
 [1073] 35 40 45
 [1074] Glu His Leu Ala Gly Arg Ile Ala Ala Val Tyr Ala Leu Arg Glu Tyr
 [1075] 50 55 60
 [1076] Gly Tyr Lys Cys Val Pro Ala Ile Gly Glu Leu Arg Gln Pro Val Trp
 [1077] 65 70 75 80
 [1078] Pro Ala Glu Val Tyr Gly Ser Ile Ser His Cys Gly Thr Thr Ala Leu
 [1079] 85 90 95
 [1080] Ala Val Val Ser Arg Gln Pro Ile Gly Ile Asp Ile Glu Glu Ile Phe
 [1081] 100 105 110
 [1082] Ser Val Gln Thr Ala Arg Glu Leu Thr Asp Asn Ile Ile Thr Pro Ala
 [1083] 115 120 125
 [1084] Glu His Glu Arg Leu Ala Asp Cys Gly Leu Ala Phe Ser Leu Ala Leu
 [1085] 130 135 140
 [1086] Thr Leu Ala Phe Ser Ala Lys Glu Ser Ala Phe Lys Ala Ser Glu Ile
 [1087] 145 150 155 160
 [1088] Gln Thr Asp Ala Gly Phe Leu Asp Tyr Gln Ile Ile Ser Trp Asn Lys
 [1089] 165 170 175
 [1090] Gln Gln Val Ile Ile His Arg Glu Asn Glu Met Phe Ala Val His Trp
 [1091] 180 185 190

[1092]	Gln Ile Lys Glu Lys Ile Val Ile Thr Leu Cys Gln His Asp	
[1093]	195	200 205
[1094]	<210> 23	
[1095]	<211> 654	
[1096]	<212> DNA	
[1097]	<213> 谷氨酸棒杆菌	
[1098]	<400> 23	
[1099]	atgctggatg agtctttgtt tccaaattcg gcaaagtttt ctttcattaa aactggcgat	60
[1100]	gctgttaatt tagaccattt ccatcagttg catccgttgg aaaaggcact ggtagcgcac	120
[1101]	tcggttgata ttagaaaagc agagtttggg gatgccaggt ggtgtgcaca tcaggcactc	180
[1102]	caagcttttg gacgagatag cggatgatccc attttgcgtg gggaacgagg aatgccattg	240
[1103]	tggccttctt cgggtgtctg ttcatgacc cacactgacg gattccgagc tgctgttgtg	300
[1104]	gcgccacgat tgttggtgcg ttctatggga ttggatgccg aacctgcgga gccgttgccc	360
[1105]	aaggatgttt tgggttcaat cgctcgggtg ggggagattc ctcaactta gcgcttgag	420
[1106]	gaacaaggtg tgcactgcgc ggatcgctg ctgttttgtg ccaaggaagc aacatacaaa	480
[1107]	gcgtggttcc cgctgacgca taggtgctt ggttttgaac aagctgagat cgacttgcgt	540
[1108]	gatgatggca cttttgtgtc ctatttctg gttcgaccaa ctccagtgc gtttatttca	600
[1109]	ggtaaatggg tactgcgtga tggttatgtc atagctgcga ctgcagtgc ttga	654
[1110]	<210> 24	
[1111]	<211> 217	
[1112]	<212> PRT	
[1113]	<213> 谷氨酸棒杆菌	
[1114]	<400> 24	
[1115]	Met Leu Asp Glu Ser Leu Phe Pro Asn Ser Ala Lys Phe Ser Phe Ile	
[1116]	1 5 10 15	
[1117]	Lys Thr Gly Asp Ala Val Asn Leu Asp His Phe His Gln Leu His Pro	
[1118]	20 25 30	
[1119]	Leu Glu Lys Ala Leu Val Ala His Ser Val Asp Ile Arg Lys Ala Glu	
[1120]	35 40 45	
[1121]	Phe Gly Asp Ala Arg Trp Cys Ala His Gln Ala Leu Gln Ala Leu Gly	
[1122]	50 55 60	
[1123]	Arg Asp Ser Gly Asp Pro Ile Leu Arg Gly Glu Arg Gly Met Pro Leu	
[1124]	65 70 75 80	
[1125]	Trp Pro Ser Ser Val Ser Gly Ser Leu Thr His Thr Asp Gly Phe Arg	
[1126]	85 90 95	
[1127]	Ala Ala Val Val Ala Pro Arg Leu Leu Val Arg Ser Met Gly Leu Asp	
[1128]	100 105 110	
[1129]	Ala Glu Pro Ala Glu Pro Leu Pro Lys Asp Val Leu Gly Ser Ile Ala	
[1130]	115 120 125	
[1131]	Arg Val Gly Glu Ile Pro Gln Leu Lys Arg Leu Glu Glu Gln Gly Val	
[1132]	130 135 140	
[1133]	His Cys Ala Asp Arg Leu Leu Phe Cys Ala Lys Glu Ala Thr Tyr Lys	

[1134]	145	150	155	160
[1135]	Ala Trp Phe Pro Leu Thr His Arg Trp Leu Gly Phe Glu Gln Ala Glu			
[1136]		165	170	175
[1137]	Ile Asp Leu Arg Asp Asp Gly Thr Phe Val Ser Tyr Leu Leu Val Arg			
[1138]		180	185	190
[1139]	Pro Thr Pro Val Pro Phe Ile Ser Gly Lys Trp Val Leu Arg Asp Gly			
[1140]		195	200	205
[1141]	Tyr Val Ile Ala Ala Thr Ala Val Thr			
[1142]		210	215	
[1143]	<210> 25			
[1144]	<211> 1428			
[1145]	<212> DNA			
[1146]	<213> 谷氨酸棒杆菌			
[1147]	<400> 25			
[1148]	atgcgcctgc gtgtctcgag tagtctcctc cccttcctcg tccccaacct cgaccattac	60		
[1149]	ggtcgccctc tcctaaagga gcctggcatg gatatccgcc aaacaattaa cgacacagca	120		
[1150]	atgtcgagat atcagtgggt cattgtatct atcgcagtgc tgctcaacgc actggacggc	180		
[1151]	tttgatgtcc tcgccatgct ttttactgct aatgcagtga ccgaagaatt tggactgagt	240		
[1152]	ggcagccagc ttgggtgtgt gctgagttcc gcgctgttcg gcatgaccgc tggatctttg	300		
[1153]	ctgttcggtc cgatcgggtga ccgtttcggc cgtaagaatg ccctgatgat cgcgctgctg	360		
[1154]	ttcaacgtgg tgggattgggt attgtccgcc accgcgcagt ccgcaggcca gttgggcgtg	420		
[1155]	tggcgtttga tcaactggtat cggcatcggc ggaatcctcg cctgcatcac agtgggtgatc	480		
[1156]	agtgagttct ccaacaacaa aaaccgcggc atggccatgt ccatctacgc tgctggttac	540		
[1157]	ggcatcggcg cgtccttggg cggattcggc gcagcgcagc tcatcccaac atttgatgg	600		
[1158]	cgtcccggtg tcgcagccgg tgcgatcgca actgggtatcg ccaccatcgc tactttcttc	660		
[1159]	ttcctgccag aatccgttga ttggctgagc actcgcgcc ctgcgggcgc tcgcgacaag	720		
[1160]	atcaattaca ttgcgcgccg cctgggcaaa gtcggtacct ttgagcttcc aggccaacaa	780		
[1161]	agcttgtcga cgaaaaaagc cgggtctcaa tcgtatgcag tgctcgttaa caaagagaac	840		
[1162]	cgtggaacca gcatcaagct gtgggttgcg ttcggcacgc tgatgttcgg cttctacttc	900		
[1163]	gccaacactt ggaccccgaa gctgctcgtg gaaaccggaa tgtcagaaca gcagggcac	960		
[1164]	atcggtgggt tgatgtgtgc catgggtgga gcattcggtt ccctgctcta cggtttcctc	1020		
[1165]	accaccaagt tcagctcccg aaacacactg atgaccttca tgggtgctgc cggcctgacg	1080		
[1166]	ctgacccgtg tcatttcctc cacctctgtt ccatccatcg cgtttgccag cggcgttgc	1140		
[1167]	gtgggcatgc tgatcaatgg ttgtgtggct ggtctgtaca ccctgtcccc acagctgtac	1200		
[1168]	tccgctgaag tacgcaccac tgggtgtggc gctgcgattg gtatgggtcg tgctgggtgcg	1260		
[1169]	atttccgcgc cactgtgggt gggtagcctg ctggattctg gctgggtccc aacgcagctg	1320		
[1170]	tatgttggtg tggcagtgat tgttattgcc ggtgcaaccg cattgattgg gatgcgcact	1380		
[1171]	caggcagtag ccgtcgaaaa gcagcctgaa gccctagcga ccaaatag	1428		
[1172]	<210> 26			
[1173]	<211> 475			
[1174]	<212> PRT			
[1175]	<213> 谷氨酸棒杆菌			

[1176]	<400> 26
[1177]	Met Arg Leu Arg Val Ser Ser Ser Leu Leu Pro Phe Leu Val Pro Asn
[1178]	1 5 10 15
[1179]	Leu Asp His Tyr Gly Arg Pro Leu Leu Lys Glu Pro Gly Met Asp Ile
[1180]	20 25 30
[1181]	Arg Gln Thr Ile Asn Asp Thr Ala Met Ser Arg Tyr Gln Trp Phe Ile
[1182]	35 40 45
[1183]	Val Phe Ile Ala Val Leu Leu Asn Ala Leu Asp Gly Phe Asp Val Leu
[1184]	50 55 60
[1185]	Ala Met Ser Phe Thr Ala Asn Ala Val Thr Glu Glu Phe Gly Leu Ser
[1186]	65 70 75 80
[1187]	Gly Ser Gln Leu Gly Val Leu Leu Ser Ser Ala Leu Phe Gly Met Thr
[1188]	85 90 95
[1189]	Ala Gly Ser Leu Leu Phe Gly Pro Ile Gly Asp Arg Phe Gly Arg Lys
[1190]	100 105 110
[1191]	Asn Ala Leu Met Ile Ala Leu Leu Phe Asn Val Val Gly Leu Val Leu
[1192]	115 120 125
[1193]	Ser Ala Thr Ala Gln Ser Ala Gly Gln Leu Gly Val Trp Arg Leu Ile
[1194]	130 135 140
[1195]	Thr Gly Ile Gly Ile Gly Gly Ile Leu Ala Cys Ile Thr Val Val Ile
[1196]	145 150 155 160
[1197]	Ser Glu Phe Ser Asn Asn Lys Asn Arg Gly Met Ala Met Ser Ile Tyr
[1198]	165 170 175
[1199]	Ala Ala Gly Tyr Gly Ile Gly Ala Ser Leu Gly Gly Phe Gly Ala Ala
[1200]	180 185 190
[1201]	Gln Leu Ile Pro Thr Phe Gly Trp Arg Ser Val Phe Ala Ala Gly Ala
[1202]	195 200 205
[1203]	Ile Ala Thr Gly Ile Ala Thr Ile Ala Thr Phe Phe Phe Leu Pro Glu
[1204]	210 215 220
[1205]	Ser Val Asp Trp Leu Ser Thr Arg Arg Pro Ala Gly Ala Arg Asp Lys
[1206]	225 230 235 240
[1207]	Ile Asn Tyr Ile Ala Arg Arg Leu Gly Lys Val Gly Thr Phe Glu Leu
[1208]	245 250 255
[1209]	Pro Gly Glu Gln Ser Leu Ser Thr Lys Lys Ala Gly Leu Gln Ser Tyr
[1210]	260 265 270
[1211]	Ala Val Leu Val Asn Lys Glu Asn Arg Gly Thr Ser Ile Lys Leu Trp
[1212]	275 280 285
[1213]	Val Ala Phe Gly Ile Val Met Phe Gly Phe Tyr Phe Ala Asn Thr Trp
[1214]	290 295 300
[1215]	Thr Pro Lys Leu Leu Val Glu Thr Gly Met Ser Glu Gln Gln Gly Ile
[1216]	305 310 315 320
[1217]	Ile Gly Gly Leu Met Leu Ser Met Gly Gly Ala Phe Gly Ser Leu Leu

[1218]	325	330	335
[1219]	Tyr Gly Phe Leu Thr Thr Lys Phe Ser Ser Arg Asn Thr Leu Met Thr		
[1220]	340	345	350
[1221]	Phe Met Val Leu Ser Gly Leu Thr Leu Ile Leu Phe Ile Ser Ser Thr		
[1222]	355	360	365
[1223]	Ser Val Pro Ser Ile Ala Phe Ala Ser Gly Val Val Val Gly Met Leu		
[1224]	370	375	380
[1225]	Ile Asn Gly Cys Val Ala Gly Leu Tyr Thr Leu Ser Pro Gln Leu Tyr		
[1226]	385	390	395
[1227]	Ser Ala Glu Val Arg Thr Thr Gly Val Gly Ala Ala Ile Gly Met Gly		
[1228]	405	410	415
[1229]	Arg Val Gly Ala Ile Ser Ala Pro Leu Leu Val Gly Ser Leu Leu Asp		
[1230]	420	425	430
[1231]	Ser Gly Trp Ser Pro Thr Gln Leu Tyr Val Gly Val Ala Val Ile Val		
[1232]	435	440	445
[1233]	Ile Ala Gly Ala Thr Ala Leu Ile Gly Met Arg Thr Gln Ala Val Ala		
[1234]	450	455	460
[1235]	Val Glu Lys Gln Pro Glu Ala Leu Ala Thr Lys		
[1236]	465	470	475
[1237]	<210> 27		
[1238]	<211> 1296		
[1239]	<212> DNA		
[1240]	<213> 谷氨酸棒杆菌		
[1241]	<400> 27		
[1242]	gtgtcaacga ccacccaac ccgcgcaacc aaaagtgtcg gaacagttct cgcactcctg	60	
[1243]	tggttcgcaa ttgtcctcga cggttttgac ctagtcgtcc tgggcgcaac aatcccgctc	120	
[1244]	atgctggagg atcccgctg ggatctcact gctggacagg ccacacagat ttccaccatc	180	
[1245]	ggcctcgtcg gcatgaccat cggcgccactg accattgggt tcttaactga ccgtctgggt	240	
[1246]	cgaccccgcg tcatgtgtt ctctgtggca gtgttttctg tattcaccct cctgctggca	300	
[1247]	ttcaccacca acgtccagct cttcagcctg tggcgtttcc tcgcaggtgt tggccttggt	360	
[1248]	ggagcactcc ccaccgcaat tgccatggtg accgagtttc gccccggcac caaagcgggc	420	
[1249]	tctgcatcaa ctacctgat gaccggatac cacgtcgggg cagtagcaac cgctttcctt	480	
[1250]	ggctctcttc ttatcgacgg ctttgggttg cactccatgt tcatcgcagg cgctgtgcca	540	
[1251]	ggactactcc tgctgccact gctgtatttc ttcttccag aatccccgca gtacctcaaa	600	
[1252]	atctccggca agttggatga ggcgcaggca gttgcagcat cttatggact ttccctggat	660	
[1253]	gatgatcttg atcgcaaca cgaagaagaa cttggcgagt cctcctcact ttctcctg	720	
[1254]	ttcaagccct cgttccgccg caacaccctg gcgatttggg gcacctcatt catgggactc	780	
[1255]	ctctgggtct acggcctgaa cacatggctg ccacaaatca tgcgccaagc agactacgac	840	
[1256]	atgggtaact ccctgggctt cctcatggtt cttaacatcg gcgcagtgat cggcctttat	900	
[1257]	attgcagggc gaattgccga taagaactcc cctcgcaaaa cagcactcgt atggttcgtg	960	
[1258]	ttctctgcat ttttctcgc actacttgcg gtccggatgc cactgatcgg tctgtatggc	1020	
[1259]	atcgtgctgc tcaccggcat ctttgtgttc agtcccagg tactcateta cgccttcgtt	1080	

[1260] ggtgagaatc accctgccaa gatgcgtgca actgccatgg gattctccgc aggaattggt 1140
 [1261] cgccctcgcg cgatctcggg tccgttgctg ggcggcctgc ttgtcagtgc caaccttgct 1200
 [1262] tacccatggg gcttcttcgc cttegtgtgc gttggactgc tgggcgcgct gattttctcc 1260
 [1263] gcatcgaaga ctctgaggca tcgcgagaac gcttag 1296
 [1264] <210> 28
 [1265] <211> 431
 [1266] <212> PRT
 [1267] <213> 谷氨酸棒杆菌
 [1268] <400> 28
 [1269] Met Ser Thr Thr Thr Pro Thr Arg Ala Thr Lys Ser Val Gly Thr Val
 [1270] 1 5 10 15
 [1271] Leu Ala Leu Leu Trp Phe Ala Ile Val Leu Asp Gly Phe Asp Leu Val
 [1272] 20 25 30
 [1273] Val Leu Gly Ala Thr Ile Pro Ser Met Leu Glu Asp Pro Ala Trp Asp
 [1274] 35 40 45
 [1275] Leu Thr Ala Gly Gln Ala Thr Gln Ile Ser Thr Ile Gly Leu Val Gly
 [1276] 50 55 60
 [1277] Met Thr Ile Gly Ala Leu Thr Ile Gly Phe Leu Thr Asp Arg Leu Gly
 [1278] 65 70 75 80
 [1279] Arg Arg Arg Val Met Leu Phe Ser Val Ala Val Phe Ser Val Phe Thr
 [1280] 85 90 95
 [1281] Leu Leu Leu Ala Phe Thr Thr Asn Val Gln Leu Phe Ser Leu Trp Arg
 [1282] 100 105 110
 [1283] Phe Leu Ala Gly Val Gly Leu Gly Gly Ala Leu Pro Thr Ala Ile Ala
 [1284] 115 120 125
 [1285] Met Val Thr Glu Phe Arg Pro Gly Thr Lys Ala Gly Ser Ala Ser Thr
 [1286] 130 135 140
 [1287] Thr Leu Met Thr Gly Tyr His Val Gly Ala Val Ala Thr Ala Phe Leu
 [1288] 145 150 155 160
 [1289] Gly Leu Phe Leu Ile Asp Gly Phe Gly Trp His Ser Met Phe Ile Ala
 [1290] 165 170 175
 [1291] Gly Ala Val Pro Gly Leu Leu Leu Leu Pro Leu Leu Tyr Phe Phe Leu
 [1292] 180 185 190
 [1293] Pro Glu Ser Pro Gln Tyr Leu Lys Ile Ser Gly Lys Leu Asp Glu Ala
 [1294] 195 200 205
 [1295] Gln Ala Val Ala Ala Ser Tyr Gly Leu Ser Leu Asp Asp Asp Leu Asp
 [1296] 210 215 220
 [1297] Arg Glu His Glu Glu Glu Leu Gly Glu Ser Ser Ser Leu Ser Ser Leu
 [1298] 225 230 235 240
 [1299] Phe Lys Pro Ser Phe Arg Arg Asn Thr Leu Ala Ile Trp Gly Thr Ser
 [1300] 245 250 255
 [1301] Phe Met Gly Leu Leu Leu Val Tyr Gly Leu Asn Thr Trp Leu Pro Gln

[1302]	260	265	270
[1303]	Ile Met Arg Gln Ala Asp Tyr Asp Met Gly Asn Ser Leu Gly Phe Leu		
[1304]	275	280	285
[1305]	Met Val Leu Asn Ile Gly Ala Val Ile Gly Leu Tyr Ile Ala Gly Arg		
[1306]	290	295	300
[1307]	Ile Ala Asp Lys Asn Ser Pro Arg Lys Thr Ala Leu Val Trp Phe Val		
[1308]	305	310	315
[1309]	Phe Ser Ala Phe Phe Leu Ala Leu Leu Ala Val Arg Met Pro Leu Ile		
[1310]	325	330	335
[1311]	Gly Leu Tyr Gly Ile Val Leu Leu Thr Gly Ile Phe Val Phe Ser Ser		
[1312]	340	345	350
[1313]	Gln Val Leu Ile Tyr Ala Phe Val Gly Glu Asn His Pro Ala Lys Met		
[1314]	355	360	365
[1315]	Arg Ala Thr Ala Met Gly Phe Ser Ala Gly Ile Gly Arg Leu Gly Ala		
[1316]	370	375	380
[1317]	Ile Ser Gly Pro Leu Leu Gly Gly Leu Leu Val Ser Ala Asn Leu Ala		
[1318]	385	390	395
[1319]	Tyr Pro Trp Gly Phe Phe Ala Phe Ala Gly Val Gly Leu Leu Gly Ala		
[1320]	405	410	415
[1321]	Leu Ile Phe Ser Ala Ser Lys Thr Leu Arg His Arg Glu Asn Ala		
[1322]	420	425	430
[1323]	<210> 29		
[1324]	<211> 1131		
[1325]	<212> DNA		
[1326]	<213> 谷氨酸棒杆菌		
[1327]	<400> 29		
[1328]	atgacactgt ccgaacgcaa gtcaccacc accgccaaga ttcttcccca cccactcaac	60	
[1329]	gcctggtacg tcgccgttg ggattatgaa gtcacatcta aaaagcccat ggccaggaca	120	
[1330]	atcgccaaca aaccactcgc tttgtaccgc accaaagatg gccgagccgt tgcccttgca	180	
[1331]	gacgcctgct ggcaccgcct cgcaccgcta tccaaggga aactcgtggg cacagacgga	240	
[1332]	atccaatgcc cttatcacgg cttggagtac aactccgcgg gccgctgcat gaaaatgccc	300	
[1333]	gcgcaggaaa ccctcaaccc gtcagcagcc gtcaactcct acccgtgggt ggaagccac	360	
[1334]	cgttttgtgt ggggtgtggct gggcgatccc acattggcag atcccacca agtaccgat	420	
[1335]	atgcaccaga tgagccaccc cgaatgggca ggcgatggac gcaccatctc cgctgactgc	480	
[1336]	aactaccaat tagtgctgga caacttgatg gacctcacc acgaagaatt cgtgcactcc	540	
[1337]	tccagcatcg gccaaacga acttagtgaa tcagagttcg tggtcaccca cactgaagat	600	
[1338]	tccgtgacgg tcaccgcgtg gatgatgac atagatgcac caccgttttg gcaaaagaac	660	
[1339]	atgaatgata agttcccagg atttgaagc aaggtggatc gttggcagat catccactac	720	
[1340]	tactaccctt ccaccatctg cattgatgtt ggtgtagcaa aggctggaac cggcgcgag	780	
[1341]	gaaggcgacc gcagccaggg cgtaaatggg tatgtaatga acaccattac cccagattca	840	
[1342]	gatcgttcct ctactactt ctgggcattc atgcgcaact accgcctgga aagccaaacc	900	
[1343]	atcaccaccc agctgcgca cggtgtatcc ggtgtattca aagaagacga agacatgctg	960	

[1344] accgctcagc aagatgccat cgacgccaac accgactatg agttttacag cctcaacatt 1020
 [1345] gatgccggtg gcatgtgggt gcgccgaatc ctcgaggaag cactctccaa ggaaggccga 1080
 [1346] ctggatatcc ccaccacatt cccccgcga acaccgaagc cggaggcata a 1131
 [1347] <210> 30
 [1348] <211> 376
 [1349] <212> PRT
 [1350] <213> 谷氨酸棒杆菌
 [1351] <400> 30
 [1352] Met Thr Leu Ser Glu Arg Lys Leu Thr Thr Thr Ala Lys Ile Leu Pro
 [1353] 1 5 10 15
 [1354] His Pro Leu Asn Ala Trp Tyr Val Ala Ala Trp Asp Tyr Glu Val Thr
 [1355] 20 25 30
 [1356] Ser Lys Lys Pro Met Ala Arg Thr Ile Ala Asn Lys Pro Leu Ala Leu
 [1357] 35 40 45
 [1358] Tyr Arg Thr Lys Asp Gly Arg Ala Val Ala Leu Ala Asp Ala Cys Trp
 [1359] 50 55 60
 [1360] His Arg Leu Ala Pro Leu Ser Lys Gly Lys Leu Val Gly Thr Asp Gly
 [1361] 65 70 75 80
 [1362] Ile Gln Cys Pro Tyr His Gly Leu Glu Tyr Asn Ser Ala Gly Arg Cys
 [1363] 85 90 95
 [1364] Met Lys Met Pro Ala Gln Glu Thr Leu Asn Pro Ser Ala Ala Val Asn
 [1365] 100 105 110
 [1366] Ser Tyr Pro Val Val Glu Ala His Arg Phe Val Trp Val Trp Leu Gly
 [1367] 115 120 125
 [1368] Asp Pro Thr Leu Ala Asp Pro Thr Gln Val Pro Asp Met His Gln Met
 [1369] 130 135 140
 [1370] Ser His Pro Glu Trp Ala Gly Asp Gly Arg Thr Ile Ser Ala Asp Cys
 [1371] 145 150 155 160
 [1372] Asn Tyr Gln Leu Val Leu Asp Asn Leu Met Asp Leu Thr His Glu Glu
 [1373] 165 170 175
 [1374] Phe Val His Ser Ser Ser Ile Gly Gln Asp Glu Leu Ser Glu Ser Glu
 [1375] 180 185 190
 [1376] Phe Val Val Thr His Thr Glu Asp Ser Val Thr Val Thr Arg Trp Met
 [1377] 195 200 205
 [1378] His Asp Ile Asp Ala Pro Pro Phe Trp Gln Lys Asn Met Asn Asp Lys
 [1379] 210 215 220
 [1380] Phe Pro Gly Phe Glu Gly Lys Val Asp Arg Trp Gln Ile Ile His Tyr
 [1381] 225 230 235 240
 [1382] Tyr Tyr Pro Ser Thr Ile Cys Ile Asp Val Gly Val Ala Lys Ala Gly
 [1383] 245 250 255
 [1384] Thr Gly Ala Gln Glu Gly Asp Arg Ser Gln Gly Val Asn Gly Tyr Val
 [1385] 260 265 270

[1386]	Met Asn Thr Ile Thr Pro Asp Ser Asp Arg Ser Ser His Tyr Phe Trp
[1387]	275 280 285
[1388]	Ala Phe Met Arg Asn Tyr Arg Leu Glu Ser Gln Thr Ile Thr Thr Gln
[1389]	290 295 300
[1390]	Leu Arg Asp Gly Val Ser Gly Val Phe Lys Glu Asp Glu Asp Met Leu
[1391]	305 310 315 320
[1392]	Thr Ala Gln Gln Asp Ala Ile Asp Ala Asn Thr Asp Tyr Glu Phe Tyr
[1393]	325 330 335
[1394]	Ser Leu Asn Ile Asp Ala Gly Gly Met Trp Val Arg Arg Ile Leu Glu
[1395]	340 345 350
[1396]	Glu Ala Leu Ser Lys Glu Gly Arg Leu Asp Ile Pro Thr Thr Phe Pro
[1397]	355 360 365
[1398]	Arg Ala Thr Pro Lys Pro Glu Ala
[1399]	370 375
[1400]	<210> 31
[1401]	<211> 978
[1402]	<212> DNA
[1403]	<213> 谷氨酸棒杆菌
[1404]	<400> 31
[1405]	atgaactcgc aatggcaaga tgcacatgtt gtttccagcg aaatcatcgc tgcagacatt 60
[1406]	cgacgaatag aactatcccc gaaatttgcg attccagtaa aaccggcgca acatctcaag 120
[1407]	atcatggtgc ccctaaaaac tggacaggaa aagagatcgt actccatcgt tgacgctcgt 180
[1408]	cacgacgggt cgactctcgc cctgagcgta ctcaaaacca gaaactcccg tggaggatct 240
[1409]	gagttcatgc atacgttcg agctggagac acagttactg tctccaggcc gtctcaggat 300
[1410]	tttctctccc gcgtgggtgc gcctgagtat gtacttgttg ccggcggaat tggaatcaca 360
[1411]	gcgatccgtt caatggcatc tttattaaag aaattgggag caaactaccg cattcatttc 420
[1412]	gcagcacgca gccttgatgc catggcttac aaagatgagc tcgtggcaga acacggcgac 480
[1413]	aagctgcacc tgcatctaga ttctgaaggc accaccatcg atgtcccagc attgatcgaa 540
[1414]	accttaaac cccacactga gctttatatg tgcggcccca tccgcttgat ggatgccatc 600
[1415]	cggcgcgcat ggaacacccg cggacttgac cccaccaatc tgcgtttcga aacgtttgga 660
[1416]	aacagtggat ggttctcccc agaggtttcc cacatccaag taccagagct ggggcttcac 720
[1417]	gccacagtca acaaggatga aagcatgctg gaggctttgc aaaaggctgg ggcgaaatag 780
[1418]	atgtttgatt gtcgaaaagg cgaatgtggt ttgtgccagg ttcgcgttct agaagtcgat 840
[1419]	ggccaggttg atcaccgcga tgtgttcttc tctgatcgtc aaaaagaatc cgacgcaaag 900
[1420]	gcatgcgcct gcgtgtctcg agtagtctcc tccccttcct cgtccccaac ctcgaccatt 960
[1421]	acggtcgccc tctcctaa 978
[1422]	<210> 32
[1423]	<211> 325
[1424]	<212> PRT
[1425]	<213> 谷氨酸棒杆菌
[1426]	<400> 32
[1427]	Met Asn Ser Gln Trp Gln Asp Ala His Val Val Ser Ser Glu Ile Ile

[1428]	1	5	10	15
[1429]	Ala Ala Asp Ile Arg Arg Ile Glu Leu Ser Pro Lys Phe Ala Ile Pro			
[1430]		20	25	30
[1431]	Val Lys Pro Gly Glu His Leu Lys Ile Met Val Pro Leu Lys Thr Gly			
[1432]		35	40	45
[1433]	Gln Glu Lys Arg Ser Tyr Ser Ile Val Asp Ala Arg His Asp Gly Ser			
[1434]		50	55	60
[1435]	Thr Leu Ala Leu Ser Val Leu Lys Thr Arg Asn Ser Arg Gly Gly Ser			
[1436]	65	70	75	80
[1437]	Glu Phe Met His Thr Leu Arg Ala Gly Asp Thr Val Thr Val Ser Arg			
[1438]		85	90	95
[1439]	Pro Ser Gln Asp Phe Pro Leu Arg Val Gly Ala Pro Glu Tyr Val Leu			
[1440]		100	105	110
[1441]	Val Ala Gly Gly Ile Gly Ile Thr Ala Ile Arg Ser Met Ala Ser Leu			
[1442]		115	120	125
[1443]	Leu Lys Lys Leu Gly Ala Asn Tyr Arg Ile His Phe Ala Ala Arg Ser			
[1444]		130	135	140
[1445]	Leu Asp Ala Met Ala Tyr Lys Asp Glu Leu Val Ala Glu His Gly Asp			
[1446]	145	150	155	160
[1447]	Lys Leu His Leu His Leu Asp Ser Glu Gly Thr Thr Ile Asp Val Pro			
[1448]		165	170	175
[1449]	Ala Leu Ile Glu Thr Leu Asn Pro His Thr Glu Leu Tyr Met Cys Gly			
[1450]		180	185	190
[1451]	Pro Ile Arg Leu Met Asp Ala Ile Arg Arg Ala Trp Asn Thr Arg Gly			
[1452]		195	200	205
[1453]	Leu Asp Pro Thr Asn Leu Arg Phe Glu Thr Phe Gly Asn Ser Gly Trp			
[1454]		210	215	220
[1455]	Phe Ser Pro Glu Val Phe His Ile Gln Val Pro Glu Leu Gly Leu His			
[1456]	225	230	235	240
[1457]	Ala Thr Val Asn Lys Asp Glu Ser Met Leu Glu Ala Leu Gln Lys Ala			
[1458]		245	250	255
[1459]	Gly Ala Asn Met Met Phe Asp Cys Arg Lys Gly Glu Cys Gly Leu Cys			
[1460]		260	265	270
[1461]	Gln Val Arg Val Leu Glu Val Asp Gly Gln Val Asp His Arg Asp Val			
[1462]		275	280	285
[1463]	Phe Phe Ser Asp Arg Gln Lys Glu Ser Asp Ala Lys Ala Cys Ala Cys			
[1464]		290	295	300
[1465]	Val Ser Arg Val Val Ser Ser Pro Ser Ser Ser Pro Thr Ser Thr Ile			
[1466]	305	310	315	320
[1467]	Thr Val Ala Leu Ser			
[1468]		325		
[1469]	<210> 33			

[1470]	<211> 615
[1471]	<212> DNA
[1472]	<213> 谷氨酸棒杆菌
[1473]	<400> 33
[1474]	atgattgata caggaagaa cggcgagttc cgctacgagc agtcgaatat catcgatcag 60
[1475]	aacgaagccg agttcggcat cactccttca cagaccgtgg gcccttacgt ccacatcggt 120
[1476]	ttgacccttg aaggtgcgga gcatctcgtg gagccaggtt cggaaggcgc ggtgtccttt 180
[1477]	actgtttccg caactgatgg caacggcgac cccatcgcg atgccatgtt tgaactgtgg 240
[1478]	caggccgatc cagaggcat ccacaactct gatttggatc caaacgcac agcaccagca 300
[1479]	accgcagatg gcttcgcgg gcttggtcgc gcatggcaa acgcgcaggg tgaggcaacg 360
[1480]	ttcaccactt tggttccggg agcattcgca gatgaggcac cacacttcaa ggttggtgtg 420
[1481]	ttcgcccggt gcatgctgga gcgtctgtac actcgcgcat acctgccaga cgccgatttg 480
[1482]	agcaccgacc cagttttggc tgttggtcca gctgatcgac gtgacctcct ggtgggtcaa 540
[1483]	aagaccgatg atggattccg cttcgacatc actgtccagg ctgaagacaa tgaacccca 600
[1484]	ttttttggac tctaa 615
[1485]	<210> 34
[1486]	<211> 204
[1487]	<212> PRT
[1488]	<213> 谷氨酸棒杆菌
[1489]	<400> 34
[1490]	Met Ile Asp Thr Gly Lys Asn Gly Glu Phe Arg Tyr Glu Gln Ser Asn
[1491]	1 5 10 15
[1492]	Ile Ile Asp Gln Asn Glu Ala Glu Phe Gly Ile Thr Pro Ser Gln Thr
[1493]	20 25 30
[1494]	Val Gly Pro Tyr Val His Ile Gly Leu Thr Leu Glu Gly Ala Glu His
[1495]	35 40 45
[1496]	Leu Val Glu Pro Gly Ser Glu Gly Ala Val Ser Phe Thr Val Ser Ala
[1497]	50 55 60
[1498]	Thr Asp Gly Asn Gly Asp Pro Ile Ala Asp Ala Met Phe Glu Leu Trp
[1499]	65 70 75 80
[1500]	Gln Ala Asp Pro Glu Gly Ile His Asn Ser Asp Leu Asp Pro Asn Arg
[1501]	85 90 95
[1502]	Thr Ala Pro Ala Thr Ala Asp Gly Phe Arg Gly Leu Gly Arg Ala Met
[1503]	100 105 110
[1504]	Ala Asn Ala Gln Gly Glu Ala Thr Phe Thr Thr Leu Val Pro Gly Ala
[1505]	115 120 125
[1506]	Phe Ala Asp Glu Ala Pro His Phe Lys Val Gly Val Phe Ala Arg Gly
[1507]	130 135 140
[1508]	Met Leu Glu Arg Leu Tyr Thr Arg Ala Tyr Leu Pro Asp Ala Asp Leu
[1509]	145 150 155 160
[1510]	Ser Thr Asp Pro Val Leu Ala Val Val Pro Ala Asp Arg Arg Asp Leu
[1511]	165 170 175

[1512]	Leu Val Ala Gln Lys Thr Asp Asp Gly Phe Arg Phe Asp Ile Thr Val	
[1513]	180	185 190
[1514]	Gln Ala Glu Asp Asn Glu Thr Pro Phe Phe Gly Leu	
[1515]	195	200
[1516]	<210> 35	
[1517]	<211> 693	
[1518]	<212> DNA	
[1519]	<213> 谷氨酸棒杆菌	
[1520]	<400> 35	
[1521]	atggacatcc cacacttcgc cccgacggga ggccaataact cccactgca cttcccggag	60
[1522]	taccggacca ccatcaagcg caacccaagc aacgatctca tcatggttcc tagtcgcctc	120
[1523]	ggcgagtcca cgggacctgt cttcggcgac cgcgacttgg gagacatcga caacgacatg	180
[1524]	accaaggtga acggtggcga ggctatcggc cagcgcatct tcgttcacgg ccgtgtcctc	240
[1525]	ggtttcgatg gcaagccagt tccgcacacc ttggtcgagg cgtggcaggc aaacgccgca	300
[1526]	ggccgttacc gccacaagaa tgactcctgg ccagcgccac tggatccaca cttcaacggt	360
[1527]	gttgacgta ctctcaccga caaggacggc cagtaccact tctggaccgt tatgccaggt	420
[1528]	aattaccctt ggggtaacca ccacaacgca tggcgcccgg cgcacattca cttctcgctc	480
[1529]	tatggtcgtc agtttacgga gcgtctggtc acccagatgt acttcccga cgatccattg	540
[1530]	ttcttccagg atccgatcta caacgcggtg ccaaagggtg cacgtgagcg catgatcgca	600
[1531]	acgttcgact atgacgagac ccgtgaaaac ttgcgccttg gttacaagtt cgacatcgtc	660
[1532]	cttcgtggcc gcaacgccac ccattttgag taa	693
[1533]	<210> 36	
[1534]	<211> 230	
[1535]	<212> PRT	
[1536]	<213> 谷氨酸棒杆菌	
[1537]	<400> 36	
[1538]	Met Asp Ile Pro His Phe Ala Pro Thr Gly Gly Glu Tyr Ser Pro Leu	
[1539]	1 5 10 15	
[1540]	His Phe Pro Glu Tyr Arg Thr Thr Ile Lys Arg Asn Pro Ser Asn Asp	
[1541]	20 25 30	
[1542]	Leu Ile Met Val Pro Ser Arg Leu Gly Glu Ser Thr Gly Pro Val Phe	
[1543]	35 40 45	
[1544]	Gly Asp Arg Asp Leu Gly Asp Ile Asp Asn Asp Met Thr Lys Val Asn	
[1545]	50 55 60	
[1546]	Gly Gly Glu Ala Ile Gly Gln Arg Ile Phe Val His Gly Arg Val Leu	
[1547]	65 70 75 80	
[1548]	Gly Phe Asp Gly Lys Pro Val Pro His Thr Leu Val Glu Ala Trp Gln	
[1549]	85 90 95	
[1550]	Ala Asn Ala Ala Gly Arg Tyr Arg His Lys Asn Asp Ser Trp Pro Ala	
[1551]	100 105 110	
[1552]	Pro Leu Asp Pro His Phe Asn Gly Val Ala Arg Thr Leu Thr Asp Lys	
[1553]	115 120 125	

[1554]	Asp Gly Gln Tyr His Phe Trp Thr Val Met Pro Gly Asn Tyr Pro Trp	
[1555]	130	140
[1556]	Gly Asn His His Asn Ala Trp Arg Pro Ala His Ile His Phe Ser Leu	
[1557]	145	155
[1558]	Tyr Gly Arg Gln Phe Thr Glu Arg Leu Val Thr Gln Met Tyr Phe Pro	
[1559]	165	175
[1560]	Asn Asp Pro Leu Phe Phe Gln Asp Pro Ile Tyr Asn Ala Val Pro Lys	
[1561]	180	190
[1562]	Gly Ala Arg Glu Arg Met Ile Ala Thr Phe Asp Tyr Asp Glu Thr Arg	
[1563]	195	205
[1564]	Glu Asn Phe Ala Leu Gly Tyr Lys Phe Asp Ile Val Leu Arg Gly Arg	
[1565]	210	220
[1566]	Asn Ala Thr Pro Phe Glu	
[1567]	225	230
[1568]	<210> 37	
[1569]	<211> 1164	
[1570]	<212> DNA	
[1571]	<213> 大肠杆菌	
[1572]	<400> 37	
[1573]	atgaacaact ttaatctgca caccccaacc cgcattctgt ttggtaaagg cgcaatcgct	60
[1574]	ggtttacgcg aacaaattcc tcacgatgct cgcgtattga ttacctacgg cggcggcagc	120
[1575]	gtgaaaaaaa ccggcgttct cgatcaagtt ctggatgccc tgaaaggcat ggacgtgctg	180
[1576]	gaatttggcg gtattgagcc aaaccgcgct tatgaaacgc tgatgaacgc cgtgaaactg	240
[1577]	gttcgcgaac agaaagtac tttcctgctg gcggttggcg gcggttctgt actggacggc	300
[1578]	accaaattta tcgccgcagc ggctaactat ccggaataaata tcgatccgtg gcacattctg	360
[1579]	caaacgggcg gtaaagagat taaaagcgcc atcccgatgg gctgtgtgct gacgctgcc	420
[1580]	gcaaccggtt cagaatccaa cgcaggcgcg gtgatctccc gtaaaaccac aggcgacaag	480
[1581]	caggcgttcc attctgccc	540
[1582]	tacacctgc cgccgcgtca ggtggttaac ggcgtagtgg acgcctttgt acacaccgtg	600
[1583]	gaacagtatg ttacaaacc ggttgatgcc aaaattcagg accgtttcgc agaaggcatt	660
[1584]	ttgctgacgc taatcgaaga tggctcgaaa gccctgaaag agccagaaaa ctacgatgtg	720
[1585]	cgcgccaacg tcatgtgggc ggcgactcag gcgctgaacg gtttgattgg cgctggcgta	780
[1586]	ccgcaggact gggcaacgca tatgctgggc cacgaactga ctgcgatgca cggctctggat	840
[1587]	cacgcgcaaa cactggctat cgtcctgcct gcaactgtga atgaaaaacg cgataccaag	900
[1588]	cgcgctaagc tgctgcaata tgctgaacgc gtctggaaca tcaactgaagg ttccgatgat	960
[1589]	gagcgtattg acgccgcgat tgccgcaacc cgcaatttct ttgagcaatt aggcgtgccg	1020
[1590]	accacctct ccgactacgg tctggacggc agctccatcc cggctttgct gaaaaaactg	1080
[1591]	gaagagcacg gcatgacca actggcgcaa aatcatgaca ttacgttgga tgtcagccgc	1140
[1592]	cgtatatacg aagccgcccg ctaa	1164
[1593]	<210> 38	
[1594]	<211> 387	
[1595]	<212> PRT	

[1596]	<213> 大肠杆菌															
[1597]	<400> 38															
[1598]	Met	Asn	Asn	Phe	Asn	Leu	His	Thr	Pro	Thr	Arg	Ile	Leu	Phe	Gly	Lys
[1599]	1				5					10					15	
[1600]	Gly	Ala	Ile	Ala	Gly	Leu	Arg	Glu	Gln	Ile	Pro	His	Asp	Ala	Arg	Val
[1601]					20				25					30		
[1602]	Leu	Ile	Thr	Tyr	Gly	Gly	Gly	Ser	Val	Lys	Lys	Thr	Gly	Val	Leu	Asp
[1603]					35				40					45		
[1604]	Gln	Val	Leu	Asp	Ala	Leu	Lys	Gly	Met	Asp	Val	Leu	Glu	Phe	Gly	Gly
[1605]		50					55					60				
[1606]	Ile	Glu	Pro	Asn	Pro	Ala	Tyr	Glu	Thr	Leu	Met	Asn	Ala	Val	Lys	Leu
[1607]	65					70				75						80
[1608]	Val	Arg	Glu	Gln	Lys	Val	Thr	Phe	Leu	Leu	Ala	Val	Gly	Gly	Gly	Ser
[1609]					85				90						95	
[1610]	Val	Leu	Asp	Gly	Thr	Lys	Phe	Ile	Ala	Ala	Ala	Ala	Asn	Tyr	Pro	Glu
[1611]					100				105					110		
[1612]	Asn	Ile	Asp	Pro	Trp	His	Ile	Leu	Gln	Thr	Gly	Gly	Lys	Glu	Ile	Lys
[1613]					115				120					125		
[1614]	Ser	Ala	Ile	Pro	Met	Gly	Cys	Val	Leu	Thr	Leu	Pro	Ala	Thr	Gly	Ser
[1615]		130					135					140				
[1616]	Glu	Ser	Asn	Ala	Gly	Ala	Val	Ile	Ser	Arg	Lys	Thr	Thr	Gly	Asp	Lys
[1617]	145					150				155						160
[1618]	Gln	Ala	Phe	His	Ser	Ala	His	Val	Gln	Pro	Val	Phe	Ala	Val	Leu	Asp
[1619]					165				170						175	
[1620]	Pro	Val	Tyr	Thr	Tyr	Thr	Leu	Pro	Pro	Arg	Gln	Val	Ala	Asn	Gly	Val
[1621]					180				185					190		
[1622]	Val	Asp	Ala	Phe	Val	His	Thr	Val	Glu	Gln	Tyr	Val	Thr	Lys	Pro	Val
[1623]					195				200					205		
[1624]	Asp	Ala	Lys	Ile	Gln	Asp	Arg	Phe	Ala	Glu	Gly	Ile	Leu	Leu	Thr	Leu
[1625]		210					215					220				
[1626]	Ile	Glu	Asp	Gly	Pro	Lys	Ala	Leu	Lys	Glu	Pro	Glu	Asn	Tyr	Asp	Val
[1627]	225					230				235						240
[1628]	Arg	Ala	Asn	Val	Met	Trp	Ala	Ala	Thr	Gln	Ala	Leu	Asn	Gly	Leu	Ile
[1629]					245				250					255		
[1630]	Gly	Ala	Gly	Val	Pro	Gln	Asp	Trp	Ala	Thr	His	Met	Leu	Gly	His	Glu
[1631]					260				265					270		
[1632]	Leu	Thr	Ala	Met	His	Gly	Leu	Asp	His	Ala	Gln	Thr	Leu	Ala	Ile	Val
[1633]					275				280					285		
[1634]	Leu	Pro	Ala	Leu	Trp	Asn	Glu	Lys	Arg	Asp	Thr	Lys	Arg	Ala	Lys	Leu
[1635]		290					295					300				
[1636]	Leu	Gln	Tyr	Ala	Glu	Arg	Val	Trp	Asn	Ile	Thr	Glu	Gly	Ser	Asp	Asp
[1637]	305					310				315						320

[1638]	Glu Arg Ile Asp Ala Ala Ile Ala Ala Thr Arg Asn Phe Phe Glu Gln		
[1639]		325	330 335
[1640]	Leu Gly Val Pro Thr His Leu Ser Asp Tyr Gly Leu Asp Gly Ser Ser		
[1641]		340	345 350
[1642]	Ile Pro Ala Leu Leu Lys Lys Leu Glu Glu His Gly Met Thr Gln Leu		
[1643]		355	360 365
[1644]	Gly Glu Asn His Asp Ile Thr Leu Asp Val Ser Arg Arg Ile Tyr Glu		
[1645]		370	375 380
[1646]	Ala Ala Arg		
[1647]	385		
[1648]	<210> 39		
[1649]	<211> 1062		
[1650]	<212> DNA		
[1651]	<213> 谷氨酸棒杆菌		
[1652]	<400> 39		
[1653]	atgagcatcc aagtaaaagc actccagaaa accggccccc aagcaccttt cgaggtcaaa	60	
[1654]	atcattgagc gtcgtgagcc tcgcgtgac gacgtagtta tcgacatcaa agctgccggc	120	
[1655]	atctgccaca gcgatatcca caccatccgc aacgaatggg gcgaggcaca cttcccgtc	180	
[1656]	accgtcggcc acgaaatcgc aggcgttgtc tctgcggttg gctccgatgt aaccaagtgg	240	
[1657]	aaagtccggc accgcgttg cgtcggtgc ctagttaact cctgcggcga atgtgaacag	300	
[1658]	tgtgtcggg gatttgaaaa caactgcctt cgcggaaacg tcggaaccta caactccgac	360	
[1659]	gacgtcgacg gcaccatcac gcaaggtggc tacgccgaaa aggtagtgtt caacgaacgt	420	
[1660]	ttcctctgca gcatcccaga ggaactcgac ttcgatgtcg cagcaccact gctgtgcgca	480	
[1661]	ggcatcacca cctactcccc gatcgctcgc tggaacgtta aagaaggcga caaagtagca	540	
[1662]	gtcatgggcc tcggcgggct cggccacatg ggtgtccaaa tcgccgcagc caagggcgt	600	
[1663]	gacgttaccg ttctgtcccg ttccctgcgc aaggtgaac ttgccaagga actcggcgca	660	
[1664]	gctcgcacgc ttgcgacttc tgatgaggat ttcttcaccg aacacgccgg tgaattcgac	720	
[1665]	ttcatcctca acaccattag cgcattcatc ccagtcgaca agtacctgag ctttctcaag	780	
[1666]	ccacacggtg tcatggtgt tgctcggtctg ccaccagaga agcagccact gagcttcggt	840	
[1667]	gcgtcatcg gcggcggaag agtcctcacc ggatccaaca ttggcggcat ccttgaaacc	900	
[1668]	caggaaatgc tcgacttctg tgcaaaacac ggcctcggcg cgatgatcga aactgtcggc	960	
[1669]	gtcaacgatg ttgatgcagc ctacgaccgc gttgttgccg gcgacgttca gttccgcgtt	1020	
[1670]	gtcattgata ctgcttcgtt tgcagaggta gaggcggttt ag	1062	
[1671]	<210> 40		
[1672]	<211> 353		
[1673]	<212> PRT		
[1674]	<213> 谷氨酸棒杆菌		
[1675]	<400> 40		
[1676]	Met Ser Ile Gln Val Lys Ala Leu Gln Lys Thr Gly Pro Glu Ala Pro		
[1677]	1	5	10 15
[1678]	Phe Glu Val Lys Ile Ile Glu Arg Arg Glu Pro Arg Ala Asp Asp Val		
[1679]		20	25 30

[1680]	Val	Ile	Asp	Ile	Lys	Ala	Ala	Gly	Ile	Cys	His	Ser	Asp	Ile	His	Thr
[1681]			35					40					45			
[1682]	Ile	Arg	Asn	Glu	Trp	Gly	Glu	Ala	His	Phe	Pro	Leu	Thr	Val	Gly	His
[1683]			50					55					60			
[1684]	Glu	Ile	Ala	Gly	Val	Val	Ser	Ala	Val	Gly	Ser	Asp	Val	Thr	Lys	Trp
[1685]	65					70					75					80
[1686]	Lys	Val	Gly	Asp	Arg	Val	Gly	Val	Gly	Cys	Leu	Val	Asn	Ser	Cys	Gly
[1687]					85					90					95	
[1688]	Glu	Cys	Glu	Gln	Cys	Val	Ala	Gly	Phe	Glu	Asn	Asn	Cys	Leu	Arg	Gly
[1689]					100					105					110	
[1690]	Asn	Val	Gly	Thr	Tyr	Asn	Ser	Asp	Asp	Val	Asp	Gly	Thr	Ile	Thr	Gln
[1691]					115					120					125	
[1692]	Gly	Gly	Tyr	Ala	Glu	Lys	Val	Val	Val	Asn	Glu	Arg	Phe	Leu	Cys	Ser
[1693]					130					135					140	
[1694]	Ile	Pro	Glu	Glu	Leu	Asp	Phe	Asp	Val	Ala	Ala	Pro	Leu	Leu	Cys	Ala
[1695]	145					150					155					160
[1696]	Gly	Ile	Thr	Thr	Tyr	Ser	Pro	Ile	Ala	Arg	Trp	Asn	Val	Lys	Glu	Gly
[1697]					165					170					175	
[1698]	Asp	Lys	Val	Ala	Val	Met	Gly	Leu	Gly	Gly	Leu	Gly	His	Met	Gly	Val
[1699]					180					185					190	
[1700]	Gln	Ile	Ala	Ala	Ala	Lys	Gly	Ala	Asp	Val	Thr	Val	Leu	Ser	Arg	Ser
[1701]					195					200					205	
[1702]	Leu	Arg	Lys	Ala	Glu	Leu	Ala	Lys	Glu	Leu	Gly	Ala	Ala	Arg	Thr	Leu
[1703]					210					215					220	
[1704]	Ala	Thr	Ser	Asp	Glu	Asp	Phe	Phe	Thr	Glu	His	Ala	Gly	Glu	Phe	Asp
[1705]	225					230					235					240
[1706]	Phe	Ile	Leu	Asn	Thr	Ile	Ser	Ala	Ser	Ile	Pro	Val	Asp	Lys	Tyr	Leu
[1707]					245					250					255	
[1708]	Ser	Leu	Leu	Lys	Pro	His	Gly	Val	Met	Ala	Val	Val	Gly	Leu	Pro	Pro
[1709]					260					265					270	
[1710]	Glu	Lys	Gln	Pro	Leu	Ser	Phe	Gly	Ala	Leu	Ile	Gly	Gly	Gly	Lys	Val
[1711]					275					280					285	
[1712]	Leu	Thr	Gly	Ser	Asn	Ile	Gly	Gly	Ile	Pro	Glu	Thr	Gln	Glu	Met	Leu
[1713]					290					295					300	
[1714]	Asp	Phe	Cys	Ala	Lys	His	Gly	Leu	Gly	Ala	Met	Ile	Glu	Thr	Val	Gly
[1715]	305					310					315					320
[1716]	Val	Asn	Asp	Val	Asp	Ala	Ala	Tyr	Asp	Arg	Val	Val	Ala	Gly	Asp	Val
[1717]					325					330					335	
[1718]	Gln	Phe	Arg	Val	Val	Ile	Asp	Thr	Ala	Ser	Phe	Ala	Glu	Val	Glu	Ala
[1719]					340					345					350	
[1720]	Val															
[1721]	<210>															41

[1722]	<211> 1113
[1723]	<212> DNA
[1724]	<213> 谷氨酸棒杆菌
[1725]	<400> 41
[1726]	gtgtccatga gcactgtcgt gcctggaatt gtcgccctgt ccaagggggc accggtagaa 60
[1727]	aaagtaaacg ttgtttgtccc tgatccaggt gctaacgatg tcatcgtcaa gattcaggcc 120
[1728]	tgcggtgtgt gccacaccga cttggcctac cgcgatggcg atatttcaga tgagtccct 180
[1729]	tacctcctcg gccacgaggc agcaggtatt gttgaggagg taggcgagtc cgtcaccac 240
[1730]	gttgaggctg gcgatttcgt catcttgaac tggcgtgcag tgtgcggcga gtgccgtgca 300
[1731]	tgtaagaagg gcgagccaaa gtactgcttt aacaccaca acgcatctaa gaagatgacc 360
[1732]	ctggaagacg gcaccgagct gtccccagca ctgggtattg gcgcgttctt ggaaaagacc 420
[1733]	ctggtccacg aaggccagtg caccaaggtt aacctgagg aagatccagc agcagctggc 480
[1734]	cttctgggtt gcggcatcat ggcaggtctt ggtgctgcgg taaacaccgg tgatattaag 540
[1735]	cgcggcgagt ccgtggcagt cttcggcctt ggtggcgtgg gcatggcagc tattgctggc 600
[1736]	gccaagattg ctggtgcatc gaagattatt gctgttgata tcgatgagaa gaagttggag 660
[1737]	tgggcgaagg aattcggcgc aaccacacc attaatcct ctggtcttgg tggcgagggt 720
[1738]	gatgcctctg aggtcgtggc aaaggttcgt gagctcactg atggtttcgg tactgacgtc 780
[1739]	tccatcgatg cggtaggcat catgccgacc tggcagcagg cgttttactc ccgtgatcat 840
[1740]	gcaggccgca tggatgatgt gggcgttcca aacctgacgt ctgcgtaga tgttcctgcg 900
[1741]	attgattttt acggtcgcgg tggctctgtg cgccctgcat ggtacggcga ctgcctgcct 960
[1742]	gagcgtgatt tcccaactta tgtggatctg cacctgcagg gtcgtttccc gctggataag 1020
[1743]	tttgtttctg agcgtattgg tcttgatgat gttgaagagg ctttcaacac catgaaggct 1080
[1744]	ggcgacgtgc tgcgttctgt ggtggagatc taa 1113
[1745]	<210> 42
[1746]	<211> 370
[1747]	<212> PRT
[1748]	<213> 谷氨酸棒杆菌
[1749]	<400> 42
[1750]	Met Ser Met Ser Thr Val Val Pro Gly Ile Val Ala Leu Ser Lys Gly
[1751]	1 5 10 15
[1752]	Ala Pro Val Glu Lys Val Asn Val Val Val Pro Asp Pro Gly Ala Asn
[1753]	20 25 30
[1754]	Asp Val Ile Val Lys Ile Gln Ala Cys Gly Val Cys His Thr Asp Leu
[1755]	35 40 45
[1756]	Ala Tyr Arg Asp Gly Asp Ile Ser Asp Glu Phe Pro Tyr Leu Leu Gly
[1757]	50 55 60
[1758]	His Glu Ala Ala Gly Ile Val Glu Glu Val Gly Glu Ser Val Thr His
[1759]	65 70 75 80
[1760]	Val Glu Val Gly Asp Phe Val Ile Leu Asn Trp Arg Ala Val Cys Gly
[1761]	85 90 95
[1762]	Glu Cys Arg Ala Cys Lys Lys Gly Glu Pro Lys Tyr Cys Phe Asn Thr
[1763]	100 105 110

[1764]	His Asn Ala Ser Lys Lys Met Thr Leu Glu Asp Gly Thr Glu Leu Ser
[1765]	115 120 125
[1766]	Pro Ala Leu Gly Ile Gly Ala Phe Leu Glu Lys Thr Leu Val His Glu
[1767]	130 135 140
[1768]	Gly Gln Cys Thr Lys Val Asn Pro Glu Glu Asp Pro Ala Ala Ala Gly
[1769]	145 150 155 160
[1770]	Leu Leu Gly Cys Gly Ile Met Ala Gly Leu Gly Ala Ala Val Asn Thr
[1771]	165 170 175
[1772]	Gly Asp Ile Lys Arg Gly Glu Ser Val Ala Val Phe Gly Leu Gly Gly
[1773]	180 185 190
[1774]	Val Gly Met Ala Ala Ile Ala Gly Ala Lys Ile Ala Gly Ala Ser Lys
[1775]	195 200 205
[1776]	Ile Ile Ala Val Asp Ile Asp Glu Lys Lys Leu Glu Trp Ala Lys Glu
[1777]	210 215 220
[1778]	Phe Gly Ala Thr His Thr Ile Asn Ser Ser Gly Leu Gly Gly Glu Gly
[1779]	225 230 235 240
[1780]	Asp Ala Ser Glu Val Val Ala Lys Val Arg Glu Leu Thr Asp Gly Phe
[1781]	245 250 255
[1782]	Gly Thr Asp Val Ser Ile Asp Ala Val Gly Ile Met Pro Thr Trp Gln
[1783]	260 265 270
[1784]	Gln Ala Phe Tyr Ser Arg Asp His Ala Gly Arg Met Val Met Val Gly
[1785]	275 280 285
[1786]	Val Pro Asn Leu Thr Ser Arg Val Asp Val Pro Ala Ile Asp Phe Tyr
[1787]	290 295 300
[1788]	Gly Arg Gly Gly Ser Val Arg Pro Ala Trp Tyr Gly Asp Cys Leu Pro
[1789]	305 310 315 320
[1790]	Glu Arg Asp Phe Pro Thr Tyr Val Asp Leu His Leu Gln Gly Arg Phe
[1791]	325 330 335
[1792]	Pro Leu Asp Lys Phe Val Ser Glu Arg Ile Gly Leu Asp Asp Val Glu
[1793]	340 345 350
[1794]	Glu Ala Phe Asn Thr Met Lys Ala Gly Asp Val Leu Arg Ser Val Val
[1795]	355 360 365
[1796]	Glu Ile
[1797]	370
[1798]	<210> 43
[1799]	<211> 1047
[1800]	<212> DNA
[1801]	<213> 谷氨酸棒杆菌
[1802]	<400> 43
[1803]	gtgagtttta tgaccactgc tgcaccccaa gaatttaccg ctgctgttgt tgaaaaattc 60
[1804]	ggtcatgacg tgaccgtgaa ggatattgac cttccaaagc cagggccaca ccaggcattg 120
[1805]	gtgaaggtac tcacctccgg catctgccac accgacctcc acgccttgga gggcgattgg 180

[1806]	ccagtaaagc cggaaccacc attcgtacca ggacacgaag gtgtaggtga agttgttgag	240
[1807]	ctcggaccag gtgaacacga tgtgaaggtc ggcgatattg tcggaatgc gtggctctgg	300
[1808]	tcagcgtgcg gcacctgcga atactgcatc acaggcaggg aaactcagt taacgaagct	360
[1809]	gagtacggtg gctacaccca aaatggatcc ttcggccagt acatgctggt ggataccga	420
[1810]	tacgccgtc gcatcccaga cggcgtggac tacctcgaag cagcgccaat tctgtgtgca	480
[1811]	ggcgtgactg tctacaaggc actcaaagtc tctgaaaccc gcccgggcca attcatggtg	540
[1812]	atctccggtg tcggcggact tggccacatc gcagtccaat acgcagcggc gatgggcatg	600
[1813]	cgtgtcattg cggtagatat tgccgaggac aagctggaac ttgcccgtaa gcacggtgcg	660
[1814]	gaatttaccg tgaatgcgcg taatgaagat ccaggcgaag ctgtacagaa gtacaccaac	720
[1815]	ggtggcgcac acggcgtgct tgtgactgca gttcacgagg cagcattcgg ccaggcactg	780
[1816]	gatatggctc gacgtgcagg aacaattgtg ttcaacggtc tgccaccggg agagttccca	840
[1817]	gcatccgtgt tcaacatcgt attcaagggc ctgaccatcc gtggatccct cgtgggaacc	900
[1818]	cgccaagact tggccgaagc gctcgatttc tttgcacgcg gactaatcaa gccaacctg	960
[1819]	agtgagtgt ccctcgatga ggtcaatgga gttcttgacc gcatgcgaaa cggcaagatc	1020
[1820]	gatggtcgtg tggcgattcg tttctaa	1047
[1821]	<210>	44
[1822]	<211>	348
[1823]	<212>	PRT
[1824]	<213>	谷氨酸棒杆菌
[1825]	<400>	44
[1826]	Met Ser Phe Met Thr Thr Ala Ala Pro Gln Glu Phe Thr Ala Ala Val	
[1827]	1 5 10 15	
[1828]	Val Glu Lys Phe Gly His Asp Val Thr Val Lys Asp Ile Asp Leu Pro	
[1829]	20 25 30	
[1830]	Lys Pro Gly Pro His Gln Ala Leu Val Lys Val Leu Thr Ser Gly Ile	
[1831]	35 40 45	
[1832]	Cys His Thr Asp Leu His Ala Leu Glu Gly Asp Trp Pro Val Lys Pro	
[1833]	50 55 60	
[1834]	Glu Pro Pro Phe Val Pro Gly His Glu Gly Val Gly Glu Val Val Glu	
[1835]	65 70 75 80	
[1836]	Leu Gly Pro Gly Glu His Asp Val Lys Val Gly Asp Ile Val Gly Asn	
[1837]	85 90 95	
[1838]	Ala Trp Leu Trp Ser Ala Cys Gly Thr Cys Glu Tyr Cys Ile Thr Gly	
[1839]	100 105 110	
[1840]	Arg Glu Thr Gln Cys Asn Glu Ala Glu Tyr Gly Gly Tyr Thr Gln Asn	
[1841]	115 120 125	
[1842]	Gly Ser Phe Gly Gln Tyr Met Leu Val Asp Thr Arg Tyr Ala Ala Arg	
[1843]	130 135 140	
[1844]	Ile Pro Asp Gly Val Asp Tyr Leu Glu Ala Ala Pro Ile Leu Cys Ala	
[1845]	145 150 155 160	
[1846]	Gly Val Thr Val Tyr Lys Ala Leu Lys Val Ser Glu Thr Arg Pro Gly	
[1847]	165 170 175	

[1848]	Gln Phe Met Val Ile Ser Gly Val Gly Gly Leu Gly His Ile Ala Val	
[1849]	180	185 190
[1850]	Gln Tyr Ala Ala Ala Met Gly Met Arg Val Ile Ala Val Asp Ile Ala	
[1851]	195	200 205
[1852]	Glu Asp Lys Leu Glu Leu Ala Arg Lys His Gly Ala Glu Phe Thr Val	
[1853]	210	215 220
[1854]	Asn Ala Arg Asn Glu Asp Pro Gly Glu Ala Val Gln Lys Tyr Thr Asn	
[1855]	225	230 235 240
[1856]	Gly Gly Ala His Gly Val Leu Val Thr Ala Val His Glu Ala Ala Phe	
[1857]	245	250 255
[1858]	Gly Gln Ala Leu Asp Met Ala Arg Arg Ala Gly Thr Ile Val Phe Asn	
[1859]	260	265 270
[1860]	Gly Leu Pro Pro Gly Glu Phe Pro Ala Ser Val Phe Asn Ile Val Phe	
[1861]	275	280 285
[1862]	Lys Gly Leu Thr Ile Arg Gly Ser Leu Val Gly Thr Arg Gln Asp Leu	
[1863]	290	295 300
[1864]	Ala Glu Ala Leu Asp Phe Phe Ala Arg Gly Leu Ile Lys Pro Thr Val	
[1865]	305	310 315 320
[1866]	Ser Glu Cys Ser Leu Asp Glu Val Asn Gly Val Leu Asp Arg Met Arg	
[1867]	325	330 335
[1868]	Asn Gly Lys Ile Asp Gly Arg Val Ala Ile Arg Phe	
[1869]	340	345
[1870]	<210> 45	
[1871]	<211> 1020	
[1872]	<212> DNA	
[1873]	<213> 谷氨酸棒杆菌	
[1874]	<400> 45	
[1875]	atgccccaaat acattgccat gcaggtatcc gaatccggtg caccgttagc cgcgaaatctc	60
[1876]	gtgcaacctg ctccgttgaa atcgaggga gtcgcgtgg aaatcgctgc tagtggtgtg	120
[1877]	tgccatgcag atattggcac ggcagcagca tcggggaagc aactgtttt tctgttacc	180
[1878]	cctggtcatg agattgcagg aaccatcgcg gaaattggtg aaaacgtatc tcggtggacg	240
[1879]	gttggtgata gcgttgcaat cggttggttt ggtggcaatt gcggtgactg cgctttttgt	300
[1880]	cgtgcaggtg atcctgtgca ttgcagagag cggaagattc ctggcgtttc ttatgcgggt	360
[1881]	ggttgggcac agaattattgt tgttcagcg gaggtcttg ctgcgattcc agatggcatg	420
[1882]	gacttttacg aggccgcccc gatgggctgc gcaggtgtga caacattcaa tgcgttgcca	480
[1883]	aacctgaagc tggatcccg tgcggctgtc gcggtcttg gaatcggcgg ttagtgcg	540
[1884]	ctagctattc agtttgctgc gaaatgggt tatcgaacca tcaccatcg ccgcggttta	600
[1885]	gagcgtgagg agctagctag gcaacttggc gccaaacct acatcgatag caatgatctg	660
[1886]	caccctggcc aggcgttatt tgaacttggc ggggctgact tgatcttgc tactgcgtcc	720
[1887]	accacggagc ctcttcgga gttgtctacc ggtctttcta ttggcgggca gctaaccatt	780
[1888]	atcggagttg atgggggaga tatcaccgtt tcggcagccc aattgatgat gaaccgtcag	840
[1889]	atcatcacag gtcacctcac tggaagtgcg aatgacacgg aacagactat gaaatttgc	900

[1890] catctccatg gcgtgaaacc gcttattgaa cggatgcctc tcgatcaagc caacgaggct 960
 [1891] attgcacgta tttcagctgg taaaccacgt ttccgtattg tcttgagacc gaattcataa 1020
 [1892] <210> 46
 [1893] <211> 339
 [1894] <212> PRT
 [1895] <213> 谷氨酸棒杆菌
 [1896] <400> 46
 [1897] Met Pro Lys Tyr Ile Ala Met Gln Val Ser Glu Ser Gly Ala Pro Leu
 [1898] 1 5 10 15
 [1899] Ala Ala Asn Leu Val Gln Pro Ala Pro Leu Lys Ser Arg Glu Val Arg
 [1900] 20 25 30
 [1901] Val Glu Ile Ala Ala Ser Gly Val Cys His Ala Asp Ile Gly Thr Ala
 [1902] 35 40 45
 [1903] Ala Ala Ser Gly Lys His Thr Val Phe Pro Val Thr Pro Gly His Glu
 [1904] 50 55 60
 [1905] Ile Ala Gly Thr Ile Ala Glu Ile Gly Glu Asn Val Ser Arg Trp Thr
 [1906] 65 70 75 80
 [1907] Val Gly Asp Arg Val Ala Ile Gly Trp Phe Gly Gly Asn Cys Gly Asp
 [1908] 85 90 95
 [1909] Cys Ala Phe Cys Arg Ala Gly Asp Pro Val His Cys Arg Glu Arg Lys
 [1910] 100 105 110
 [1911] Ile Pro Gly Val Ser Tyr Ala Gly Gly Trp Ala Gln Asn Ile Val Val
 [1912] 115 120 125
 [1913] Pro Ala Glu Ala Leu Ala Ala Ile Pro Asp Gly Met Asp Phe Tyr Glu
 [1914] 130 135 140
 [1915] Ala Ala Pro Met Gly Cys Ala Gly Val Thr Thr Phe Asn Ala Leu Arg
 [1916] 145 150 155 160
 [1917] Asn Leu Lys Leu Asp Pro Gly Ala Ala Val Ala Val Phe Gly Ile Gly
 [1918] 165 170 175
 [1919] Gly Leu Val Arg Leu Ala Ile Gln Phe Ala Ala Lys Met Gly Tyr Arg
 [1920] 180 185 190
 [1921] Thr Ile Thr Ile Ala Arg Gly Leu Glu Arg Glu Glu Leu Ala Arg Gln
 [1922] 195 200 205
 [1923] Leu Gly Ala Asn His Tyr Ile Asp Ser Asn Asp Leu His Pro Gly Gln
 [1924] 210 215 220
 [1925] Ala Leu Phe Glu Leu Gly Gly Ala Asp Leu Ile Leu Ser Thr Ala Ser
 [1926] 225 230 235 240
 [1927] Thr Thr Glu Pro Leu Ser Glu Leu Ser Thr Gly Leu Ser Ile Gly Gly
 [1928] 245 250 255
 [1929] Gln Leu Thr Ile Ile Gly Val Asp Gly Gly Asp Ile Thr Val Ser Ala
 [1930] 260 265 270
 [1931] Ala Gln Leu Met Met Asn Arg Gln Ile Ile Thr Gly His Leu Thr Gly

[1932]	275	280	285
[1933]	Ser Ala Asn Asp Thr Glu Gln Thr Met Lys Phe Ala His Leu His Gly		
[1934]	290	295	300
[1935]	Val Lys Pro Leu Ile Glu Arg Met Pro Leu Asp Gln Ala Asn Glu Ala		
[1936]	305	310	315 320
[1937]	Ile Ala Arg Ile Ser Ala Gly Lys Pro Arg Phe Arg Ile Val Leu Glu		
[1938]	325	330	335
[1939]	Pro Asn Ser		
[1940]	<210> 47		
[1941]	<211> 879		
[1942]	<212> DNA		
[1943]	<213> 谷氨酸棒杆菌		
[1944]	<400> 47		
[1945]	atgcaaacc ttgctgctat tgttcgtgcc acgaagcaac cttttgagat caccaccatt	60	
[1946]	gatctggatg caccacgacc agatgaagtt caaatccgtg ttattgctgc cggagtgcgc	120	
[1947]	cacactgacg caattgttcg tgatcagatt tacccaactt ttcttcccgc agttttcggc	180	
[1948]	cacgaaggcg ccggagtagt tgtcgccgtg ggttctgcag tcacctcgtt gaaaccagat	240	
[1949]	gacaaggtag tgctgggatt caactcttgt ggccagtgtc tgaagtgttt gggcggttaag	300	
[1950]	cctgcgtact gtgagaaatt ctatgaccgc aacttcgcat gcacccgcga tgccgggcac	360	
[1951]	actactttgt ttaccctgac aacaaaagag caggcagagg ccatcatcga cacccttgat	420	
[1952]	gatgttttct acgatgcgga tgcgggttct ctggcatacc cagcaactcc cccagaggct	480	
[1953]	tcgggagtaa gcgtgttggt tgtcgcggct ggtacctctg atctcccca agcaaaggaa	540	
[1954]	gcactacaca ctgcctccta cttggggcgc tccacctcac tgattgttga ttttgagtg	600	
[1955]	gctggcatcc accgcctgct ttcatacgaa gaagaactcc gcgctgcggg cgtgctcatc	660	
[1956]	gttcccgtg gaatggatgg tgcgtaccc ggagttgtc caggcttagt gtccgcacct	720	
[1957]	gtcgtgcac tgccaacctc cgtgggatac ggcgaggtg ctggaggaat cgcaccactt	780	
[1958]	ctgaccatgc ttaacgcctg cgcgccggga gttggagtgg tcaacattga taacggctat	840	
[1959]	ggagcaggac acctggctgc gcagattgcg gcgaggtaa	879	
[1960]	<210> 48		
[1961]	<211> 292		
[1962]	<212> PRT		
[1963]	<213> 谷氨酸棒杆菌		
[1964]	<400> 48		
[1965]	Met Gln Thr Leu Ala Ala Ile Val Arg Ala Thr Lys Gln Pro Phe Glu		
[1966]	1 5 10 15		
[1967]	Ile Thr Thr Ile Asp Leu Asp Ala Pro Arg Pro Asp Glu Val Gln Ile		
[1968]	20 25 30		
[1969]	Arg Val Ile Ala Ala Gly Val Arg His Thr Asp Ala Ile Val Arg Asp		
[1970]	35 40 45		
[1971]	Gln Ile Tyr Pro Thr Phe Leu Pro Ala Val Phe Gly His Glu Gly Ala		
[1972]	50 55 60		
[1973]	Gly Val Val Val Ala Val Gly Ser Ala Val Thr Ser Val Lys Pro Asp		

[1974]	65	70	75	80
[1975]	Asp Lys Val Val Leu Gly Phe Asn Ser Cys Gly Gln Cys Leu Lys Cys			
[1976]		85	90	95
[1977]	Leu Gly Gly Lys Pro Ala Tyr Cys Glu Lys Phe Tyr Asp Arg Asn Phe			
[1978]		100	105	110
[1979]	Ala Cys Thr Arg Asp Ala Gly His Thr Thr Leu Phe Thr Arg Ala Thr			
[1980]		115	120	125
[1981]	Lys Glu Gln Ala Glu Ala Ile Ile Asp Thr Leu Asp Asp Val Phe Tyr			
[1982]		130	135	140
[1983]	Asp Ala Asp Ala Gly Phe Leu Ala Tyr Pro Ala Thr Pro Pro Glu Ala			
[1984]	145	150	155	160
[1985]	Ser Gly Val Ser Val Leu Val Val Ala Ala Gly Thr Ser Asp Leu Pro			
[1986]		165	170	175
[1987]	Gln Ala Lys Glu Ala Leu His Thr Ala Ser Tyr Leu Gly Arg Ser Thr			
[1988]		180	185	190
[1989]	Ser Leu Ile Val Asp Phe Gly Val Ala Gly Ile His Arg Leu Leu Ser			
[1990]		195	200	205
[1991]	Tyr Glu Glu Glu Leu Arg Ala Ala Gly Val Leu Ile Val Ala Ala Gly			
[1992]		210	215	220
[1993]	Met Asp Gly Ala Leu Pro Gly Val Val Ala Gly Leu Val Ser Ala Pro			
[1994]	225	230	235	240
[1995]	Val Val Ala Leu Pro Thr Ser Val Gly Tyr Gly Ala Gly Ala Gly Gly			
[1996]		245	250	255
[1997]	Ile Ala Pro Leu Leu Thr Met Leu Asn Ala Cys Ala Pro Gly Val Gly			
[1998]		260	265	270
[1999]	Val Val Asn Ile Asp Asn Gly Tyr Gly Ala Gly His Leu Ala Ala Gln			
[2000]		275	280	285
[2001]	Ile Ala Ala Arg			
[2002]	290			
[2003]	<210> 49			
[2004]	<211> 819			
[2005]	<212> DNA			
[2006]	<213> 大肠杆菌			
[2007]	<400> 49			
[2008]	atggaaacct atgctgtttt tggtaatccg atagcccaca gcaaatcgcc attcattcat	60		
[2009]	cagcaatttg ctcagcaact gaatatgaa catccctatg ggcgcggtgt ggcacccatc	120		
[2010]	aatgatttca tcaacacact gaacgctttc tttagtgtg gtggtaaagg tgcgaatgtg	180		
[2011]	acggtgccct ttaaagaaga ggcttttgcc agagcggatg agcttactga acgggcagcg	240		
[2012]	ttggctggtg ctgttaatac cctcatgcgg ttagaagatg gacgcctgct gggtgacaat	300		
[2013]	accgatggtg taggcttggt aagcgatctg gaacgtctgt cttttatccg ccctgggtta	360		
[2014]	cgtattctgc ttatcggcgc tgggtggagca tctcggcg tactactgcc actcctttcc	420		
[2015]	ctggactgtg cggtgacaat aactaatcgg acggtatccc gcgcggaaga gttggctaaa	480		

[2016] ttgtttgcgc acactggcag tattcaggcg ttgagtatgg acgaactgga aggtcatgag 540
 [2017] tttgatctca ttattaatgc aacatccagt ggcacatcagtg gtgatattcc ggcatcccg 600
 [2018] tcatcgctca ttcatccagg cattttattgc tatgacatgt tctatcagaa aggaaaaact 660
 [2019] ccttttctgg catggtgtga gcagcgaggc tcaaagcgta atgctgatgg tttaggaatg 720
 [2020] ctggtggcac aggcggctca tgcctttctt ctctggcacg gtgttctgcc tgacgtagaa 780
 [2021] ccagttataa agcaattgca ggaggaattg tccgcgtga 819
 [2022] <210> 50
 [2023] <211> 272
 [2024] <212> PRT
 [2025] <213> 大肠杆菌
 [2026] <400> 50
 [2027] Met Glu Thr Tyr Ala Val Phe Gly Asn Pro Ile Ala His Ser Lys Ser
 [2028] 1 5 10 15
 [2029] Pro Phe Ile His Gln Gln Phe Ala Gln Gln Leu Asn Ile Glu His Pro
 [2030] 20 25 30
 [2031] Tyr Gly Arg Val Leu Ala Pro Ile Asn Asp Phe Ile Asn Thr Leu Asn
 [2032] 35 40 45
 [2033] Ala Phe Phe Ser Ala Gly Gly Lys Gly Ala Asn Val Thr Val Pro Phe
 [2034] 50 55 60
 [2035] Lys Glu Glu Ala Phe Ala Arg Ala Asp Glu Leu Thr Glu Arg Ala Ala
 [2036] 65 70 75 80
 [2037] Leu Ala Gly Ala Val Asn Thr Leu Met Arg Leu Glu Asp Gly Arg Leu
 [2038] 85 90 95
 [2039] Leu Gly Asp Asn Thr Asp Gly Val Gly Leu Leu Ser Asp Leu Glu Arg
 [2040] 100 105 110
 [2041] Leu Ser Phe Ile Arg Pro Gly Leu Arg Ile Leu Leu Ile Gly Ala Gly
 [2042] 115 120 125
 [2043] Gly Ala Ser Arg Gly Val Leu Leu Pro Leu Leu Ser Leu Asp Cys Ala
 [2044] 130 135 140
 [2045] Val Thr Ile Thr Asn Arg Thr Val Ser Arg Ala Glu Glu Leu Ala Lys
 [2046] 145 150 155 160
 [2047] Leu Phe Ala His Thr Gly Ser Ile Gln Ala Leu Ser Met Asp Glu Leu
 [2048] 165 170 175
 [2049] Glu Gly His Glu Phe Asp Leu Ile Ile Asn Ala Thr Ser Ser Gly Ile
 [2050] 180 185 190
 [2051] Ser Gly Asp Ile Pro Ala Ile Pro Ser Ser Leu Ile His Pro Gly Ile
 [2052] 195 200 205
 [2053] Tyr Cys Tyr Asp Met Phe Tyr Gln Lys Gly Lys Thr Pro Phe Leu Ala
 [2054] 210 215 220
 [2055] Trp Cys Glu Gln Arg Gly Ser Lys Arg Asn Ala Asp Gly Leu Gly Met
 [2056] 225 230 235 240
 [2057] Leu Val Ala Gln Ala Ala His Ala Phe Leu Leu Trp His Gly Val Leu

[2058]	245	250	255
[2059]	Pro Asp Val Glu Pro Val Ile Lys Gln Leu Gln Glu Glu Leu Ser Ala		
[2060]	260	265	270
[2061]	<210> 51		
[2062]	<211> 35		
[2063]	<212> DNA		
[2064]	<213> 人工序列		
[2065]	<220>		
[2066]	<223> 引物		
[2067]	<400> 51		
[2068]	cggtacccgg ggatccttac ttccgcgtat ccaac 35		
[2069]	<210> 52		
[2070]	<211> 41		
[2071]	<212> DNA		
[2072]	<213> 人工序列		
[2073]	<220>		
[2074]	<223> 引物		
[2075]	<400> 52		
[2076]	ctaggaatcg cggccggtga actcctaaag aactatataa c 41		
[2077]	<210> 53		
[2078]	<211> 20		
[2079]	<212> DNA		
[2080]	<213> 人工序列		
[2081]	<220>		
[2082]	<223> 引物		
[2083]	<400> 53		
[2084]	ggccgcgatt cctagcatgc 20		
[2085]	<210> 54		
[2086]	<211> 35		
[2087]	<212> DNA		
[2088]	<213> 人工序列		
[2089]	<220>		
[2090]	<223> 引物		
[2091]	<400> 54		
[2092]	ccaagcttgc atgccagtca tcatcaacgg tgccg 35		
[2093]	<210> 55		
[2094]	<211> 21		
[2095]	<212> DNA		
[2096]	<213> 人工序列		
[2097]	<220>		
[2098]	<223> 引物		
[2099]	<400> 55		

[2100]	atctccgcag aagacgtact g 21
[2101]	<210> 56
[2102]	<211> 20
[2103]	<212> DNA
[2104]	<213> 人工序列
[2105]	<220>
[2106]	<223> 引物
[2107]	<400> 56
[2108]	tccgatcatg tatgacctcc 20
[2109]	<210> 57
[2110]	<211> 36
[2111]	<212> DNA
[2112]	<213> 人工序列
[2113]	<220>
[2114]	<223> 引物
[2115]	<400> 57
[2116]	cggtacccgg ggatcggeat agtgcttcca acgctc 36
[2117]	<210> 58
[2118]	<211> 36
[2119]	<212> DNA
[2120]	<213> 人工序列
[2121]	<220>
[2122]	<223> 引物
[2123]	<400> 58
[2124]	tagtccact caagattcct cgatattacc tacagg 36
[2125]	<210> 59
[2126]	<211> 20
[2127]	<212> DNA
[2128]	<213> 人工序列
[2129]	<220>
[2130]	<223> 引物
[2131]	<400> 59
[2132]	tcttgagtgg agctagggcc 20
[2133]	<210> 60
[2134]	<211> 37
[2135]	<212> DNA
[2136]	<213> 人工序列
[2137]	<220>
[2138]	<223> 引物
[2139]	<400> 60
[2140]	ccaagcttgc atgcccataat agagcccagg agctctc 37
[2141]	<210> 61

[2142]	<211> 23
[2143]	<212> DNA
[2144]	<213> 人工序列
[2145]	<220>
[2146]	<223> 引物
[2147]	<400> 61
[2148]	cgccgcaaag tccaaataga aag 23
[2149]	<210> 62
[2150]	<211> 20
[2151]	<212> DNA
[2152]	<213> 人工序列
[2153]	<220>
[2154]	<223> 引物
[2155]	<400> 62
[2156]	ggattcttcc tgaactcagc 20
[2157]	<210> 63
[2158]	<211> 36
[2159]	<212> DNA
[2160]	<213> 人工序列
[2161]	<220>
[2162]	<223> 引物
[2163]	<400> 63
[2164]	cggtacccgg ggatcgggct cgtcctgaaa ttgcac 36
[2165]	<210> 64
[2166]	<211> 35
[2167]	<212> DNA
[2168]	<213> 人工序列
[2169]	<220>
[2170]	<223> 引物
[2171]	<400> 64
[2172]	tccgtcgtga gccatgttgt gcccacgaga ctacc 35
[2173]	<210> 65
[2174]	<211> 20
[2175]	<212> DNA
[2176]	<213> 人工序列
[2177]	<220>
[2178]	<223> 引物
[2179]	<400> 65
[2180]	atggctcacg acggattgcg 20
[2181]	<210> 66
[2182]	<211> 36
[2183]	<212> DNA

[2184]	<213> 人工序列
[2185]	<220>
[2186]	<223> 引物
[2187]	<400> 66
[2188]	ccaagcttgc atgcccggtt gcagccttca taaacg 36
[2189]	<210> 67
[2190]	<211> 21
[2191]	<212> DNA
[2192]	<213> 人工序列
[2193]	<220>
[2194]	<223> 引物
[2195]	<400> 67
[2196]	agaccaatga gtaccaacc g 21
[2197]	<210> 68
[2198]	<211> 20
[2199]	<212> DNA
[2200]	<213> 人工序列
[2201]	<220>
[2202]	<223> 引物
[2203]	<400> 68
[2204]	tcagcgtctg gctcagctac 20
[2205]	<210> 69
[2206]	<211> 36
[2207]	<212> DNA
[2208]	<213> 人工序列
[2209]	<220>
[2210]	<223> 引物
[2211]	<400> 69
[2212]	cggtaccgg ggatcaaccc cagctcaaatt aacacc 36
[2213]	<210> 70
[2214]	<211> 37
[2215]	<212> DNA
[2216]	<213> 人工序列
[2217]	<220>
[2218]	<223> 引物
[2219]	<400> 70
[2220]	tttcaacaca atccgtcctt ctcgcttgga ttacttg 37
[2221]	<210> 71
[2222]	<211> 23
[2223]	<212> DNA
[2224]	<213> 人工序列
[2225]	<220>

[2226]	<223> 引物
[2227]	<400> 71
[2228]	cggattgtgt tgaaattgct ctg 23
[2229]	<210> 72
[2230]	<211> 35
[2231]	<212> DNA
[2232]	<213> 人工序列
[2233]	<220>
[2234]	<223> 引物
[2235]	<400> 72
[2236]	ccaagcttgc atgcctcacc acgggaatct tcagg 35
[2237]	<210> 73
[2238]	<211> 20
[2239]	<212> DNA
[2240]	<213> 人工序列
[2241]	<220>
[2242]	<223> 引物
[2243]	<400> 73
[2244]	ccggactggg gtgtgttttg 20
[2245]	<210> 74
[2246]	<211> 20
[2247]	<212> DNA
[2248]	<213> 人工序列
[2249]	<220>
[2250]	<223> 引物
[2251]	<400> 74
[2252]	cccggaaaat acggtatagc 20
[2253]	<210> 75
[2254]	<211> 35
[2255]	<212> DNA
[2256]	<213> 人工序列
[2257]	<220>
[2258]	<223> 引物
[2259]	<400> 75
[2260]	ccaagcttgc atgccccatc gcattgccga aaagc 35
[2261]	<210> 76
[2262]	<211> 35
[2263]	<212> DNA
[2264]	<213> 人工序列
[2265]	<220>
[2266]	<223> 引物
[2267]	<400> 76

[2268]	aaagatcggg tcaatgcagt tcgcggggcg aacat 35
[2269]	<210> 77
[2270]	<211> 35
[2271]	<212> DNA
[2272]	<213> 人工序列
[2273]	<220>
[2274]	<223> 引物
[2275]	<400> 77
[2276]	cgccccgcga actgcattga cccgatcttt atacc 35
[2277]	<210> 78
[2278]	<211> 35
[2279]	<212> DNA
[2280]	<213> 人工序列
[2281]	<220>
[2282]	<223> 引物
[2283]	<400> 78
[2284]	cggtacccgg ggatcaacgt tgacggtgat gccat 35
[2285]	<210> 79
[2286]	<211> 25
[2287]	<212> DNA
[2288]	<213> 人工序列
[2289]	<220>
[2290]	<223> 引物
[2291]	<400> 79
[2292]	gaaatgtcat acttcagcca tcagg 25
[2293]	<210> 80
[2294]	<211> 25
[2295]	<212> DNA
[2296]	<213> 人工序列
[2297]	<220>
[2298]	<223> 引物
[2299]	<400> 80
[2300]	tgcgagtgat gaaatcctga aactt 25
[2301]	<210> 81
[2302]	<211> 2180
[2303]	<212> DNA
[2304]	<213> 人工序列
[2305]	<220>
[2306]	<223> 启动子P2
[2307]	<400> 81
[2308]	tttcgcggtg aatcaaccca cccgaacggc gatccttcga agtattttgc cgattgatca 60
[2309]	attagcgteg gcaacatcac tgaatatgct gctcatgcag accggcgcaa tcgttggccc 120

[2310]	gctgatcgca ggtgcgttga ttccgctgat cggtttcggg tggctgtatt tccttgatgt	180
[2311]	tgtctccatc atccccacac tgtgggctgt atggtcactg cttcaatca agccatccgg	240
[2312]	caaggtcatg aaggccggtt tcgccagtgt ggtggatggc ctgaagtatt tggctggcca	300
[2313]	accctgtgtg ttgatgggtga tgggtctgga tcttatcgcc atgattttcg gcatgccacg	360
[2314]	tgcgctttac cccgagatcg cggaagtga cttcggtggt ggtgacgccg gtgcaacgat	420
[2315]	gctggcggtt atgtactcat ccatggctgt tggcgcagtt cttggcggcg tgctgtctgg	480
[2316]	ttgggtttcc cggattagcc gccagggtgt tgcagtttat tggatgcatca tcgcctgggg	540
[2317]	cgcagccgtt gctttgggtg gcgtagcaat tgttgtcagc cccggcgtg tgaccgcgtg	600
[2318]	ggcgtggatg ttcacatca tgatggatcat tggatggcatg gctgacatgt ttagctcggc	660
[2319]	tggtcgaaat gctattttgc agcagctcgc agcggaaat gtgcagggcc gaatccaagg	720
[2320]	tgtgtggatc atcgtcgtgg tgggtggacc tcgttttagct gacgtccttc acggttgggc	780
[2321]	cgtgagccc ttgggtgcag gttggacggt attatggggc ggagtagcgg tggttgtact	840
[2322]	cactgcaatt tgtatgggtg cgggtgcctaa attctgaaa tacgagaaac caaaaattac	900
[2323]	cggcatctaa atacttatcc atgcccattt acagacaatg ccttagcttt gacctgcaca	960
[2324]	aatagttgca aattgtccca catacacata aagtagcttg cgtattttaa attatgaacc	1020
[2325]	taaggggttt agcaatgccc aatcaggccc acttctctgc gtcttttgcc cgccttctta	1080
[2326]	ccccggtgc aaagtgcacg caccatatcc gcctcggcca gcaactcatt agaaatgagc	1140
[2327]	tggctcaggc cacaggtctg tcccaaccga ctgtcaccg cgcagtcacc gctttaatgc	1200
[2328]	aggcaggttt ggttcgtgaa cgccctgac tcacactctc atcgggccct ggtcgtccca	1260
[2329]	atattcctct agaactcgt ccaagtccat ggattcatgc aggcgtggca atcggcacca	1320
[2330]	agtcttcta cgtcgtttg tttgatacca agggctgcac cttcgtgat gccatactgg	1380
[2331]	aaatctcagc agctgattta gatccagaca cttcatcga acacctcgc gctggtgtca	1440
[2332]	accgctcac cactggtctt gatctaccac tggtaggtat tgggtgtggc acctcaggaa	1500
[2333]	aagtcaccaa cgcgggcgtt gtcaccgcaa gcaacttggg ctgggatggc gttgatatcg	1560
[2334]	ccggccgtct gaactaccaa ttcagcgttc cagcaaccgt ggcatcagca attcctgcca	1620
[2335]	tcgcagcttc tgaactgcag gcttccccac tccccaccc tgagcagcca actcccatca	1680
[2336]	ccttgacctt ctacccgat gactctgtgg gcgcggccta cagcaatgat ttgggagtac	1740
[2337]	atgtcattgg accactggct acaactcgtg gatcagggtt ggatactttg ggcattgctg	1800
[2338]	ctgaagatgc gctgagcacc caaggtttct taagcagggt ttctgatcag ggtatctttg	1860
[2339]	ccaacagcct tggtagcta gtcaccattg ctaaagacaa tgaaaccgca cgggaattcc	1920
[2340]	tcaacgatcg cgcgacctg ctggtcaca ctgccgcaga agctgctgaa acagttaagc	1980
[2341]	catccacctt ggttctctcg ggatcggcgt tttccgaaga tccacaaggt cggttggtgt	2040
[2342]	tcgttccca attgaagaag gaatacgacg cagacattga gctccgcttg atccccacc	2100
[2343]	accgggaaaa tgtccgcga gcagctcgcg cagtcgcaact tgatcgacta ctcaacgagc	2160
[2344]	cattaccct cgtaccctaa	2180
[2345]	<210>	82
[2346]	<211>	2247
[2347]	<212>	DNA
[2348]	<213>	人工序列
[2349]	<220>	
[2350]	<223>	启动子P4
[2351]	<400>	82

[2352]	ccctcatgag ttcaggggtt agaaaagcaa tgggatttgg atgcggttcg gttttggccg	60
[2353]	tcatcatggt tatctcatTT gttggatggg cgcttagctt catggatgga acggcaccta	120
[2354]	ttcgccaact ccagcaaate cctgaagatg ttccgccggc gcgtggtgta gaagttccgc	180
[2355]	aaattgatac aggggcagat ggacgcacgt cagatcattt gcgttttttg gcggaaccaa	240
[2356]	ttgtcaaga tgctggtgtg tccgctcaag cgattgcggc ttatggaaac gcagagctca	300
[2357]	ttgcgagtac tgctggcct ggctgcaatc tggggtggaa taccttggca ggtatcgcc	360
[2358]	aggtggaaac ccgtcacgtt acctacaacg gcaaatgtt cgggggcagt tccctggatg	420
[2359]	aaaatggagt tgcaaccctt ccaatcatcg gcgttccact tgatggttca ccgggggttg	480
[2360]	cggaaattcc cgacactgat ggtggggaat tagatggcga tactgaatat gatcgcgcg	540
[2361]	taggtcccat gcagttcatt ccgaaacgt ggcgacttat gggattggat gcaaacggtg	600
[2362]	atggggtagc ggaccccaac caaattgatg acgcagcatt gaggccgca aacctgttgt	660
[2363]	gttccaacga tcgtgacttg tccactcctg aaggatggac cgcagctgtt cattcttaca	720
[2364]	acatgtctaa tcagtatttg atggacgttc gagatgctgc cgcgtcctac gctttacgac	780
[2365]	agccggcgat ctaaaactta acaagcgaac cccccgaaa tgtgagatta tgtccggtcg	840
[2366]	gacacgtgcg ggctggggat atgggtagtt taataaattt ataccacaca gtctattgca	900
[2367]	atagaccaag ctgttcagta ggggtcatgg gagaagaatt tcctaataaa aactcttaag	960
[2368]	gacctccaag tggctgaaat catgcacgta ttgctcgcg aaattctcga ctcccgcgtt	1020
[2369]	aaccaaccg tcgaggcaga ggttttctt gatgacggtt cccacggtg cgcaggtgtt	1080
[2370]	ccatccggcg catccaccgg cgtccacgag gctcatgagc tgcgtgacgg tggcgatcgc	1140
[2371]	tacctgggca agggcgtttt gaaggcagtt gaaaacgtca acgaagaaat cggcgacgag	1200
[2372]	ctcgttgcc tagagctga cgatcagcgc ctcatcgacg aagcaatgat caagcttgat	1260
[2373]	ggcaccgcca acaagtcccg cctgggtgca aacgcaatcc ttggtgtttc catggctgtt	1320
[2374]	gcaaaggctg ctgctgattc cgcaggctc ccactgttc gctacatcgg tggaccaaac	1380
[2375]	gcacacgttc ttccagttcc aatgatgaac atcatcaacg gtggcgctca cgctgactcc	1440
[2376]	ggtgttgacg ttccaggaatt catgatcgtt ccaatcggtg cagagacctt ctctgaggt	1500
[2377]	ctccgaacg gcgcagaggt ctaccacgca ctgaagtccg tcatcaagga aaagggcctg	1560
[2378]	tccaccggac ttggcgatga gggcggcttc gctccttcg tcggctccac ccgtgaggt	1620
[2379]	cttgacctta tcgttgaggc aatcgagaag gctggcttca cccaggcaa ggacatcgt	1680
[2380]	cttgetctgg acgttgcttc ctctgagttc ttcaaggacg gcacctacca ctctgaaggt	1740
[2381]	ggccagcact ccgcagctga gatggcaaac gtttacgtg agctcgttga cgcgtacca	1800
[2382]	atcgtctcca tcgaggacc actgcaggaa gatgactggg agggttacac caacctcacc	1860
[2383]	gcaaccatcg gcgacaaggt tcagatcgtt ggcgacgact tcttcgtcac caacctgag	1920
[2384]	cgctgaagg agggcatcgc taagaaggct gccaaactca tcttggttaa ggtgaaccag	1980
[2385]	atcggtagcc tcaccgagac cttcgacgt gtcgacatgg ctcaccgcgc aggtacacc	2040
[2386]	tccatgatgt cccaccgttc cggtagacc gaggacacca ccattgctga cctcgagtt	2100
[2387]	gcactcaact gtggccagat caagactggt gctccagcac gttccgaccg tgctgcaaag	2160
[2388]	tacaaccagc ttctccgcat cgagcagttg cttggcgacg ccggcgtcta cgcaggtcgc	2220
[2389]	agcgcattcc cacgtttca gggctaa	2247
[2390]	<210>	83
[2391]	<211>	2192
[2392]	<212>	DNA
[2393]	<213>	人工序列

[2394]	<220>	
[2395]	<223> 启动子P8	
[2396]	<400> 83	
[2397]	tcaccgagtc tttgatcaag ggtggcgctt ttgactccct tggacacgca cgaaaaggcc	60
[2398]	tcatgctggt cttcgaagat gccgttgatt ccgtcatcgc taccaaaaaa gctgctgaca	120
[2399]	agggacaatt tgatctcttt gcagctttcg actcggataa caacgacgat gtggcaagtt	180
[2400]	tcttccagat caccgttcct gatgacgaat gggaccgtaa gcatgagctc gcactcgagc	240
[2401]	gagaaatgct gggctctgtat gtttctggac acccactcga tggctatgaa gatgccattg	300
[2402]	ctgcccaggt tgatacagca ctgaccacca ttgttgccgg tgaactcaag cacggcgcag	360
[2403]	aagtgaccgt ggggtggcatt atctctggtg tggatcgacg gttctccaag aaggacggtt	420
[2404]	cccttgggc gattgtcacc attgaagatc acaacggcgc gtccgttgaa ttgttggtct	480
[2405]	tcaacaaggt gtattccatc gttggatcca tgattgtgga agacaacatc attttgcca	540
[2406]	aggcacacat ctccattcga gatgatcgtg tgagcctttt ctgtgatgat ctccgcgttc	600
[2407]	cagagcttgg gccaggaaac gggcaaggac ttccgcttcg tttgtccatg cgtactgatc	660
[2408]	agtgacccat gtccaacatt gccaaagctc agcagggtct ggtggacaac aagggtgaat	720
[2409]	ctgatgtgta cctcaatttg atcgatgggg ataactccac ggtcatgatt ttgggtgatc	780
[2410]	acttaagagt caaccgatcc gcaagtttga tgggcgacct caaggcaacg atggggccag	840
[2411]	gcatcctcgg ttaatcacat cactctggga ttaccccgctg taggggtgaa aaccgaatg	900
[2412]	tggctaaaac ttttggaac ttaagttacc tttaatcgga aacttattga attcgggtga	960
[2413]	ggcaactgca actctggact taaagcatga gccagaaccg catcaggacc actcacgttg	1020
[2414]	gttccttgcc ccgtaccca gagctacttg atgcaaacat caagcgtctt aacggtgaga	1080
[2415]	ttggggagga ggaattcttc cagatcctgc agtcttctgt agatgacgtg atcaagcgc	1140
[2416]	aggtgacct gggatcgcac atcctcaacg agggcgaata cggccacgtc acctccggtg	1200
[2417]	cagttgactt cggatcatgg tggaaactact cttcaccgg cctgggcgga ctgaccatga	1260
[2418]	ccgataccga ccgttgggca agccaggaaag cagtgcgttc caccctggc aacatcaagc	1320
[2419]	tgaccagctt ctctgatcgt cgcgaccgcg cattgttcag cgaagcatac gaggatccag	1380
[2420]	tatctggcat cttcaccggc cgcgcttctg tgggcaacc agagttcacc ggacctatta	1440
[2421]	cctacattgg ccaggaaagaa actcagacgg atgttgatct gctgaagaag ggcataaacg	1500
[2422]	cagcgggagc taccgacggc ttctgttcag cactatcccc aggatctgca gctcgattga	1560
[2423]	ccaacaagtt ctacgacact gatgaagaag tcgtcgcagc atgtgccgat gcgctttccc	1620
[2424]	aggaatacaa gatcatcacc gatgcaggtc tgaccgttca gctcgacgca ccggacttgg	1680
[2425]	cagaagcatg ggatcagatc aaccagagc caagcgtgaa ggattactta gactggatcg	1740
[2426]	gtacacgcat cgatgccatc aacagtgcag tgaagggcct tccaaaggaa cagaccgcc	1800
[2427]	tgcacatctg ctggggctct tggcacggac cacacgtcac tgacatccca ttcggtgaca	1860
[2428]	tcattggtga gatcctgcgc gcagaggtcg gtggcttctc cttcgaaggc gcatctctc	1920
[2429]	gtcacgcaca cgagtggcgt gtatgggaag aaaacaagct tcctgaaggc tctgttatct	1980
[2430]	acctgggtgt tgtgtctcac tccatcaacg ctgtggagca cccacgctg gttgctgatc	2040
[2431]	gtatcgttca gttcgccaag cttgttggcc ctgagaacgt cattgcgtcc actgactgtg	2100
[2432]	gtctgggagg acgtctgcat tcccagatcg catgggcaaa gctggagtcc ctagtagagg	2160
[2433]	gcgctcgcac tgcatcaaag gaactgttct aa	2192
[2434]	<210> 84	
[2435]	<211> 97	

- [2436] <212> DNA
[2437] <213> 人工序列
[2438] <220>
[2439] <223> 启动子P3
[2440] <400> 84
[2441] tgccgtttct cgcgttgtgt gtggtactac gtggggacct aagcgtgtaa gatggaaacg 60
[2442] tctgtatcgg ataagtagcg aggagtgttc gttaaaa 97
[2443] <210> 85
[2444] <211> 35
[2445] <212> DNA
[2446] <213> 人工序列
[2447] <220>
[2448] <223> 引物
[2449] <400> 85
[2450] ccaagcttgc atgcctcacc gagtctttga tcaag 35
[2451] <210> 86
[2452] <211> 35
[2453] <212> DNA
[2454] <213> 人工序列
[2455] <220>
[2456] <223> 引物
[2457] <400> 86
[2458] caaaagtttt agccacattc gggttttcac cccta 35
[2459] <210> 87
[2460] <211> 35
[2461] <212> DNA
[2462] <213> 人工序列
[2463] <220>
[2464] <223> 引物
[2465] <400> 87
[2466] ctctggactt aaagcatgag ccagaaccgc atcag 35
[2467] <210> 88
[2468] <211> 35
[2469] <212> DNA
[2470] <213> 人工序列
[2471] <220>
[2472] <223> 引物
[2473] <400> 88
[2474] cggtaccgg ggatcttaga acagttcctt tgatg 35
[2475] <210> 89
[2476] <211> 87
[2477] <212> DNA

[2478]	<213> 人工序列
[2479]	<220>
[2480]	<223> P8启动子的DNA片段
[2481]	<400> 89
[2482]	gtggctaaaa cttttggaaa cttaagttac ctttaatcgg aaacttattg aattcgggtg 60
[2483]	aggcaactgc aactctggac ttaaagc 87
[2484]	<210> 90
[2485]	<211> 35
[2486]	<212> DNA
[2487]	<213> 人工序列
[2488]	<220>
[2489]	<223> 引物
[2490]	<400> 90
[2491]	gtgaaaaccc gaatgtggct aaaacttttg gaaac 35
[2492]	<210> 91
[2493]	<211> 35
[2494]	<212> DNA
[2495]	<213> 人工序列
[2496]	<220>
[2497]	<223> 引物
[2498]	<400> 91
[2499]	gcggttctgg ctcatgcttt aagtccagag ttgca 35
[2500]	<210> 92
[2501]	<211> 1206
[2502]	<212> DNA
[2503]	<213> 谷氨酸棒杆菌
[2504]	<400> 92
[2505]	atgagccaga accgcatcag gaccactcac gttggttcct tgccccgtac cccagagcta 60
[2506]	cttgatgcaa acatcaagcg ttctaacggg gagattgggg aggaggaatt cttccagatt 120
[2507]	ctgcagtctt ctgtagatga cgtgatcaag cgccaggttg acctgggtat cgacatcctt 180
[2508]	aacgagggcg aatacggcca cgtcacctcc ggtgcagttg acttcgggtgc atggtggaac 240
[2509]	tactccttca cccgcctggg cggactgacc atgaccgata ccgaccgttg ggcaagccag 300
[2510]	gaagcagtgc gttccacccc tggcaacatc aagctgacca gcttctctga tcgtcgcgac 360
[2511]	cgcgcattgt tcagcgaagc atacgaggat ccagtatctg gcattctcac cggtcgcgct 420
[2512]	tctgtgggca acccagagtt caccggacct attacctaca ttggccagga agaaactcag 480
[2513]	acggatgttg atctgctgaa gaagggcatg aacgcagcgg gagctaccga cggcttcggt 540
[2514]	gcagcactat ccccaggatc tgcagctcga ttgaccaaca agttctacga cactgatgaa 600
[2515]	gaagtcgtcg cagcatgtgc tgatgcgctt tcccaggaat acaagatcat caccgatgca 660
[2516]	ggtctgaccg ttcagctcga cgcaccggac ttggcagaag catgggatca gatcaacca 720
[2517]	gagccaagcg tgaaggatta cttggactgg atcggtagac gcacgatgc catcaacagt 780
[2518]	gcagtgaagg gccttccaaa ggaacagacc cgcctgcaca tctgctgggg ctcttggcac 840
[2519]	ggaccacacg tcactgacat cccattcggg gacatcattg gtgagatcct gcgcgcagag 900

[2520] gtcggtggct tctccttcga aggcgcacgt cctcgtcacg cacacgagtg gcgtgtatgg 960
 [2521] gaagaaaaca agcttcctga aggcctctgtt atctaccctg gtgttgtgtc tcaactccatc 1020
 [2522] aacgctgtgg agcaccacag cctgggttgc gatcgatcgc ttcagttcgc caagcttggtt 1080
 [2523] ggccctgaga acgtcattgc gtccactgac tgtggctctgg gcggacgtct gcattcccag 1140
 [2524] atcgcatggg caaagctgga gtccctagta gagggcgctc gcattgcac aaaggaactg 1200
 [2525] ttctaa 1206
 [2526] <210> 93
 [2527] <211> 401
 [2528] <212> PRT
 [2529] <213> 谷氨酸棒杆菌
 [2530] <400> 93
 [2531] Met Ser Gln Asn Arg Ile Arg Thr Thr His Val Gly Ser Leu Pro Arg
 [2532] 1 5 10 15
 [2533] Thr Pro Glu Leu Leu Asp Ala Asn Ile Lys Arg Ser Asn Gly Glu Ile
 [2534] 20 25 30
 [2535] Gly Glu Glu Glu Phe Phe Gln Ile Leu Gln Ser Ser Val Asp Asp Val
 [2536] 35 40 45
 [2537] Ile Lys Arg Gln Val Asp Leu Gly Ile Asp Ile Leu Asn Glu Gly Glu
 [2538] 50 55 60
 [2539] Tyr Gly His Val Thr Ser Gly Ala Val Asp Phe Gly Ala Trp Trp Asn
 [2540] 65 70 75 80
 [2541] Tyr Ser Phe Thr Arg Leu Gly Gly Leu Thr Met Thr Asp Thr Asp Arg
 [2542] 85 90 95
 [2543] Trp Ala Ser Gln Glu Ala Val Arg Ser Thr Pro Gly Asn Ile Lys Leu
 [2544] 100 105 110
 [2545] Thr Ser Phe Ser Asp Arg Arg Asp Arg Ala Leu Phe Ser Glu Ala Tyr
 [2546] 115 120 125
 [2547] Glu Asp Pro Val Ser Gly Ile Phe Thr Gly Arg Ala Ser Val Gly Asn
 [2548] 130 135 140
 [2549] Pro Glu Phe Thr Gly Pro Ile Thr Tyr Ile Gly Gln Glu Glu Thr Gln
 [2550] 145 150 155 160
 [2551] Thr Asp Val Asp Leu Leu Lys Lys Gly Met Asn Ala Ala Gly Ala Thr
 [2552] 165 170 175
 [2553] Asp Gly Phe Val Ala Ala Leu Ser Pro Gly Ser Ala Ala Arg Leu Thr
 [2554] 180 185 190
 [2555] Asn Lys Phe Tyr Asp Thr Asp Glu Glu Val Val Ala Ala Cys Ala Asp
 [2556] 195 200 205
 [2557] Ala Leu Ser Gln Glu Tyr Lys Ile Ile Thr Asp Ala Gly Leu Thr Val
 [2558] 210 215 220
 [2559] Gln Leu Asp Ala Pro Asp Leu Ala Glu Ala Trp Asp Gln Ile Asn Pro
 [2560] 225 230 235 240
 [2561] Glu Pro Ser Val Lys Asp Tyr Leu Asp Trp Ile Gly Thr Arg Ile Asp

[2562]	245	250	255
[2563]	Ala Ile Asn Ser Ala Val Lys Gly Leu Pro Lys Glu Gln Thr Arg Leu		
[2564]	260	265	270
[2565]	His Ile Cys Trp Gly Ser Trp His Gly Pro His Val Thr Asp Ile Pro		
[2566]	275	280	285
[2567]	Phe Gly Asp Ile Ile Gly Glu Ile Leu Arg Ala Glu Val Gly Gly Phe		
[2568]	290	295	300
[2569]	Ser Phe Glu Gly Ala Ser Pro Arg His Ala His Glu Trp Arg Val Trp		
[2570]	305	310	315
[2571]	Glu Glu Asn Lys Leu Pro Glu Gly Ser Val Ile Tyr Pro Gly Val Val		
[2572]	325	330	335
[2573]	Ser His Ser Ile Asn Ala Val Glu His Pro Arg Leu Val Ala Asp Arg		
[2574]	340	345	350
[2575]	Ile Val Gln Phe Ala Lys Leu Val Gly Pro Glu Asn Val Ile Ala Ser		
[2576]	355	360	365
[2577]	Thr Asp Cys Gly Leu Gly Gly Arg Leu His Ser Gln Ile Ala Trp Ala		
[2578]	370	375	380
[2579]	Lys Leu Glu Ser Leu Val Glu Gly Ala Arg Ile Ala Ser Lys Glu Leu		
[2580]	385	390	395
[2581]	Phe		
[2582]	<210> 94		
[2583]	<211> 666		
[2584]	<212> DNA		
[2585]	<213> 智人		
[2586]	<400> 94		
[2587]	atgggtgaca ccaaggagca gcgcatacctg aaccacgtgc tgcagcatgc ggagcccggg	60	
[2588]	aacgcacaga gcgtgctgga ggccattgac acctactgcg agcagaagga gtgggccatg	120	
[2589]	aacgtgggcg acaagaaagg caagatcgtg gacgccgtga ttcaggagca ccagccctcc	180	
[2590]	gtgtgctctgg agctgggggc ctactgtggc tactcagctg tgcgcatggc ccgctgctg	240	
[2591]	tcaccagggg cgaggctcat caccatcgag atcaaccccg actgtgccgc catcaccag	300	
[2592]	cggatgggtgg atttcgctgg cgtgaaggac aaggtcacc ttgtggttgg agcgtcccag	360	
[2593]	gacatcatcc cccagctgaa gaagaagtat gatgtggaca cactggacat ggtcttcctc	420	
[2594]	gaccactgga aggaccggtg cctgccggac acgcttctct tggaggaatg tggcctgctg	480	
[2595]	cggaagggga cagtgtctact ggctgacaac gtgatctgcc caggtgcgcc agacttccta	540	
[2596]	gcacacgtgc gcgggagcag ctgctttgag tgcacacact accaatcgtt cctggaatac	600	
[2597]	agggaggtgg tggacggcct ggagaaggcc atctacaagg gccaggcag cgaagcaggg	660	
[2598]	ccttaa	666	
[2599]	<210> 95		
[2600]	<211> 675		
[2601]	<212> DNA		
[2602]	<213> 智人		
[2603]	<400> 95		

[2604]	catatggggg ataccaaaga acagcggatt ctgaatcatg ttttacagca cgctgaaccg	60
[2605]	ggcaatgctc agagcgtttt agaagcaata gatacctatt gtgaacagaa ggaatgggca	120
[2606]	atgaatgttg gcgataaaaa gggcaaaatt gttgatgcag ttatccagga acatcagccg	180
[2607]	agcgttttac ttgaactggg cgcatattgc ggttattccg cagttcggat ggcacggctg	240
[2608]	ctgtcccttg gcgctcgctt aattaccatt gaaattaatc cggattgcgc agcaattacc	300
[2609]	cagcggatgg ttgactttgc aggtgttaaa gataaagtca ctttggttgt cggcgctagc	360
[2610]	caggatatta ttccgcagct gaaaaaaaaa tacgatgttg ataccctgga tatggtcttt	420
[2611]	ttagatcatt ggaaagatcg gtatctgccc gatactctgt tattggaaga gtgcggctctg	480
[2612]	ctgcggaaag gcaccgttct actggcagat aatgttattt gtcctggggc tcctgatttt	540
[2613]	ctagctcatg ttcggggcag cagctgtttc gaatgtaccc attaccaatc gtttctggaa	600
[2614]	tatcggaag ttgttgatgg tctggaaaag gcaatttata aaggtcctgg tagcgaggct	660
[2615]	ggcccgtaag agctc	675
[2616]	<210>	96
[2617]	<211>	6557
[2618]	<212>	DNA
[2619]	<213>	人工序列
[2620]	<220>	
[2621]	<223>	pELAC载体
[2622]	<400>	96
[2623]	tggcgaatgg gacgcgccct gtagcggcgc attaacgcgc gcgggtgtgg tggttacgcg	60
[2624]	cagcgtgacc gctacacttg ccagcgccct agcgcgccgt ctttctgctt tcttcccttc	120
[2625]	ctttctgcc acgttcgccg gctttccccg tcaagctcta aatcgggggc tcccttttagg	180
[2626]	gttccgattt agtgctttac ggcacctga ccccaaaaaa cttgattagg gtgatggttc	240
[2627]	acgtagtggg ccatcgccct gatagacggg ttttcgccct ttgacgttgg agtccacgtt	300
[2628]	ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct cggctctattc	360
[2629]	ttttgattta taagggattt tgccgatttc ggcctattgg ttaaaaaatg agctgattta	420
[2630]	acaaaaattt aacgcgaatt ttaacaaaat attaacgttt acaatttcag gtggcacttt	480
[2631]	tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt caaatatgta	540
[2632]	tccgctcatg agacaataac cctgataaat gtttcaataa tattgaaaaa ggaagagtat	600
[2633]	gagtattcaa catttccgtg tcgcccttat tccctttttt gcggcatttt gccttctctg	660
[2634]	ttttgtcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg	720
[2635]	agtgggttac atcgaactgg atctcaacag cggtaaagatc cttgagagtt ttcgccccga	780
[2636]	agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcgg tattatcccc	840
[2637]	tattgacgcc gggcaagagc aactcggtcg ccgcatacac tattctcaga atgacttggt	900
[2638]	tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa gagaattatg	960
[2639]	cagtgtgcc ataaccatga gtgataacac tgcggccaac ttacttctga caacgatcgg	1020
[2640]	aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa ctgccttga	1080
[2641]	tcgttgggaa ccggagctga atgaagccat accaaacgac gacggtgaca ccacgatgcc	1140
[2642]	tgcagcaatg gcaacaacgt tgcgcaaact attactggc gaactactta ctctagcttc	1200
[2643]	ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac ttctgcgctc	1260
[2644]	ggcccttccg gctggctggt ttattgtctga taaatctgga gccggtgagc gtgggtctcg	1320
[2645]	cggtatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag ttatctacac	1380

[2646]	gacggggagt caggcaacta tggatgaacg aaatagacag atcgctgaga taggtgcctc	1440
[2647]	actgattaag cattggtaac tgtcagacca agtttactca tatatacttt agattgattt	1500
[2648]	aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata atctcatgac	1560
[2649]	caaaatccct taacgtgagt tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa	1620
[2650]	aggatcttct tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa caaaaaaacc	1680
[2651]	accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt ttccgaaggt	1740
[2652]	aactggcttc agcagagcgc agataccaaa tactgtcctt ctagtgtagc cgtagttagg	1800
[2653]	ccaccacttc aagaactctg tagcaccgcc tacatactc gctctgctaa tctgttacc	1860
[2654]	agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa gacgatatgt	1920
[2655]	accggataag gcgcagcggc cgggctgaac ggggggttcg tgcacacagc ccagcttgga	1980
[2656]	gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa gcgccacgct	2040
[2657]	tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc agggtcggaa caggagagcg	2100
[2658]	cacgaggag cttccagggg gaaacgctg gtatctttat agtcctgtcg ggtttcgcca	2160
[2659]	cctctgactt gagcgtcgat ttttgtgatg ctctgcaggg gggcggagcc tatggaaaaa	2220
[2660]	cgccagcaac gcggcctttt tacggttcct ggccttttgc tggccttttg ctcacatgtt	2280
[2661]	ctttctgcg ttatccccg attctgtgga taaccgtatt accgcctttg agtgagctga	2340
[2662]	taccgtcgc cgcagccgaa cgaccgagcg cagcagtgca gtgagcgagg aagcggaaga	2400
[2663]	gcgcctgatg cggatatttc tccttacgca tctgtcgggt atttcacacc gcatatatgg	2460
[2664]	tgcactctca gtacaactctg ctctgatgcc gcatagttaa gccagtatac actccgctat	2520
[2665]	cgctacgtga ctgggtcatg gctgcgcccc gacaccgcc aacaccgct gacgcgccct	2580
[2666]	gacgggcttg tctgtcccg gcatccgctt acagacaagc tgtgaccgtc tccgggagct	2640
[2667]	gcatgtgtca gaggttttca ccgtcatcac cgaaacgcgc gaggcagctg cggtaaagct	2700
[2668]	catcagcgtg gtctggaagc gattcacaga tgtctgcctg ttcacccg cccagctcgt	2760
[2669]	tgagtttctc cagaagcgtt aatgtctggc ttctgataaa gcgggccatg ttaaggcg	2820
[2670]	tttttctctg tttggtcact gatgcctccg tgtaagggg atttctgttc atgggggtaa	2880
[2671]	tgataccgat gaaacgagag aggatgtcga cgatacgggt tactgatgat gaacatgccc	2940
[2672]	ggttactgga acgttgtgag ggtaaacaaac tggcgggatg gatgcggcgg gaccagagaa	3000
[2673]	aaatcactca gggtaaatgc cagcgcttcg ttaatacaga tgtaggtgtt ccacagggt	3060
[2674]	gccagcagca tctgcgatg cagatccgga acataatggt gcagggcgct gacttccgcg	3120
[2675]	tttccagact ttacgaaaca cggaaccga agaccattca tgttggtgct caggtcgcag	3180
[2676]	acgttttgca gcagcagtcg cttcacgttc gctcgcgtat cggtgattca ttctgctaac	3240
[2677]	cagtaaggca accccgccag cctagccggg tcctcaacga caggagcacg atcatgcgca	3300
[2678]	cccgtggggc cgccatgccg gcgataatgg cctgcttctc gccgaaacgt ttggtggcgg	3360
[2679]	gaccagtgac gaaggcttga gcgaggcgt gcaagattcc gaataccgca agcgacaggc	3420
[2680]	cgatcatcgt cgcgtccag cgaaagcggc cctcgccgaa aatgaccag agcgtgccg	3480
[2681]	gcacctgtcc tacgagttgc atgataaaga agacagtc atagtcggcg acgatatgca	3540
[2682]	tgccccgcgc ccaccggaag gagctgactg ggttgaaggc tctcaaggc atcggtcgag	3600
[2683]	atcccggtgc ctaatgagtg agctaactta cattaattgc gttgcgtca ctgcccgtt	3660
[2684]	tccagtcggg aaacctgtcg tgccagctgc attaatgaat cggccaacgc gcggggagag	3720
[2685]	gcggtttgcg tattgggcgc cagggtggtt tttcttttca ccagtgaac gggcaacagc	3780
[2686]	tgattgccct tcaccgctg gccctgagag agttgcagca agcgggtccac gctggtttgc	3840
[2687]	cccagcaggc gaaaatcctg tttgatggtg gttaacggcg ggatataaca tgagctgtct	3900

[2688]	tcggtatcgt cgtatccac taccgagata tccgcaccaa cgcgagccc ggactcggta	3960
[2689]	atggcgcgca ttgcgcccag cgccatctga tcgttggcaa ccagcatcgc agtgggaacg	4020
[2690]	atgccctcat tcagcatttg catggtttgt tgaaaaccgg acatggcact ccagtcgcct	4080
[2691]	tcccgttccg ctatcggtg aatttgattg cgagtgaatg atttatgcca gccagccaga	4140
[2692]	cgagacgcg ccgagacaga acttaattggg cccgctaaca gcgagattg ctggtgacct	4200
[2693]	aatgcgacca gatgctccac gccagtcgc gtaccgtctt catgggagaa aataatactg	4260
[2694]	ttgatgggtg tctggtcaga gacatcaaga aataacgccg gaacattagt gcaggcagct	4320
[2695]	tccacagcaa tggcatcctg gtcattccagc ggatagttaa tgatcagccc actgacgcgt	4380
[2696]	tgcgcgagaa gattgtgcac cgccgcttta caggtctcga cgccgcttcg ttctaccatc	4440
[2697]	gacaccacca cgctggcacc cagttgatcg gcgagagatt taatcgccgc gacaatttgc	4500
[2698]	gacggcgctg gcagggccag actggaggtg gcaacgcaa tcagcaacga ctgtttgccc	4560
[2699]	gccagtgtgt gtgccacgcg gttgggaatg taattcagct ccgccatcgc cgtttccact	4620
[2700]	ttttcccgcg ttttcgaga aacgtggtg gcctggttca ccacgcggga aacggtctga	4680
[2701]	taagagacac cggcactctc tgcgacatcg tataacgtta ctggtttcac attcaccacc	4740
[2702]	ctgaattgac tctcttccgg gcgctatcat gccataccgc gaaaggtttt gcgccattcg	4800
[2703]	atggtgtccg ggatctcgac gctctccctt atgcgactcc tgcattagga agcagcccag	4860
[2704]	tagtaggttg aggccgttga gcaccgccgc cgcaaggaat ggtgcatgca aggagatggc	4920
[2705]	gcccacagct cccccggcca cggggcctgc caccataccc acgccgaaac aagcgtcat	4980
[2706]	gagcccgaag tggcgagccc gatcttcccc atcggtagatg tcggcgatat aggcgccagc	5040
[2707]	aaccgcacct gtggcgccgg tgatgccggc cagcatgcgt ccggcgtaga ggatcgagat	5100
[2708]	ctgcgggcag tgagcgcaac gcaattaatg tgagttagct cactcattag gcaccccagg	5160
[2709]	ctttacactt tatgcttccg gctcgtataa tgtgtggaat tgtgagcgga taacaatttc	5220
[2710]	acacaggatc tagatttaag aaggagatat acatatggcc gaagaaggta aacttgtaat	5280
[2711]	ctggattaac ggcgataaag gctataacgg tctcgtgaa gtcggtaaga aattcgagaa	5340
[2712]	agataccgga attaaagtca ccgttgagca tccggataaa ctggaagaga aattcccaca	5400
[2713]	ggttgccgca acaggcgatg gccctgacat tatcttctgg gcacacgacc gctttggtgg	5460
[2714]	ctacgtcaa tctggcctgt tggctgaaat caccgccgac aaagcgttcc aggacaagct	5520
[2715]	gtatccgttt acctgggatg ccgtacgtta caacggcaag ctgattgctt acccgatcgc	5580
[2716]	tggtgaagcg ttatcgtga ttataacaa agatctgctg ccgaacccgc caaaaacctg	5640
[2717]	ggaagagatc ccggcgttg ataaagaact gaaagcgaag ggtaagagcg cgtgatgtt	5700
[2718]	caacctgcaa gaacctact tcacctggcc gctgattgct gctgacgggg gttatgcgtt	5760
[2719]	caagtatgaa aacggcaagt acgacattaa agacgtgggc gtggataacg ctggcgcgaa	5820
[2720]	agcgggtctg accttccctg ttgacctgat taaaaacaaa cacatgaatg cagacaccga	5880
[2721]	ttactccatc gcagaagctg cttttaataa aggcgaaaca gcgatgacca tcaacggccc	5940
[2722]	gtgggcatgg tccaacatcg acaccagcaa agtgaattat ggtgtaacgg tactgccgac	6000
[2723]	cttcaagggt caaccatcca aaccgttcgt tggcgtgctg agcgagagta ttaacggcg	6060
[2724]	cagtccgaac aaagagctgg caaaagagtt cctcgaaaac tatctgctga ctgatgaagg	6120
[2725]	tctggaagcg gttaataaag acaaaccgct ggggtgccgtg gcgctgaagt cttacagga	6180
[2726]	agagtggcg aaagatccac gtattgccgc cactatggaa aacgcccaga aaggtgaaat	6240
[2727]	catgccgaac atccgcgaga tgtccgcttt ctggtatgcc gtgcgtactg cggtgatcaa	6300
[2728]	cgccgccagc ggtcgtcaga ctgtcgatga agccctgaaa gacgcgcaga ctaaggatcc	6360
[2729]	gaattcgagc tccgtcgaca agcttgcggc cgcactcgag caccaccacc accaccactg	6420

[2730]	agatccggct gctaacaag cccgaaagga agctgagttg gctgctgcca ccgctgagca 6480
[2731]	ataactagca taacccttg gggcctctaa acgggtcttg aggggttttt tgctgaaagg 6540
[2732]	aggaactata tccggat 6557
[2733]	<210> 97
[2734]	<211> 22
[2735]	<212> DNA
[2736]	<213> 人工序列
[2737]	<220>
[2738]	<223> 引物
[2739]	<400> 97
[2740]	gtgagttagc tcactcatta gg 22
[2741]	<210> 98
[2742]	<211> 22
[2743]	<212> DNA
[2744]	<213> 人工序列
[2745]	<220>
[2746]	<223> 引物
[2747]	<400> 98
[2748]	cgattggtaa tgggtacatt cg 22
[2749]	<210> 99
[2750]	<211> 20
[2751]	<212> DNA
[2752]	<213> 人工序列
[2753]	<220>
[2754]	<223> 引物
[2755]	<400> 99
[2756]	ggatctcagt ggtggtggtg 20
[2757]	<210> 100
[2758]	<211> 44
[2759]	<212> DNA
[2760]	<213> 人工序列
[2761]	<220>
[2762]	<223> 引物
[2763]	<400> 100
[2764]	ccaagcttgc atgccaaatt cctgtgaatt agctgattta gtac 44
[2765]	<210> 101
[2766]	<211> 38
[2767]	<212> DNA
[2768]	<213> 人工序列
[2769]	<220>
[2770]	<223> 引物
[2771]	<400> 101

[2772] tttggtatcc cccatagagg cgaaggctcc ttgaatag 38
[2773] <210> 102
[2774] <211> 22
[2775] <212> DNA
[2776] <213> 人工序列
[2777] <220>
[2778] <223> 引物
[2779] <400> 102
[2780] atgggggata ccaaagaaca gc 22
[2781] <210> 103
[2782] <211> 38
[2783] <212> DNA
[2784] <213> 人工序列
[2785] <220>
[2786] <223> 引物
[2787] <400> 103
[2788] cggtaccggg ggatcatgct agttattgct cagcgggtg 38
[2789] <210> 104
[2790] <211> 20
[2791] <212> DNA
[2792] <213> 人工序列
[2793] <220>
[2794] <223> 引物
[2795] <400> 104
[2796] tgacatccca ttcggtgaca 20
[2797] <210> 105
[2798] <211> 20
[2799] <212> DNA
[2800] <213> 人工序列
[2801] <220>
[2802] <223> 引物
[2803] <400> 105
[2804] ttcttcccat acacgccact 20
[2805] <210> 106
[2806] <211> 20
[2807] <212> DNA
[2808] <213> 人工序列
[2809] <220>
[2810] <223> 引物
[2811] <400> 106
[2812] aggtggggat gacgtcaaatt 20
[2813] <210> 107

-
- [2814] <211> 20
[2815] <212> DNA
[2816] <213> 人工序列
[2817] <220>
[2818] <223> 引物
[2819] <400> 107
[2820] gaactgaggc cggctttaag 20