(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 116897204 B
(45)授权公告日 2024.02.27

(21)申请号 202180002260.0
(22)申请日 2021.04.22
(65)同一申请的已公布的文献号
申请公布号 CN 116897204 A
(43)申请公布日 2023.10.17
(30)优先权数据
10-2021-0006131 2021.01.15 KR
(85)PCT国际申请进入国家阶段日
2021.08.25
(86)PCT国际申请的申请数据
PCT/KR2021/005074 2021.04.22
(87)PCT国际申请的公布数据
WO2022/154189 KO 2022.07.21
(83)生物保藏信息
KCCM12877P 2020.12.02
(73)专利权人 CJ第一制糖株式会社
地址 韩国首尔市
(72)发明人 裴智妍 沈世勋 李沚泫 李智惠 朴暠会 金孝珍 徐昌一
(74)专利代理机构 北京市中伦律师事务所 11410
专利代理师 钟锦舜 王奕勋
(51)Int.Cl.
C12N 9/10(2006.01)
(56)对比文件
JP 2019165635 A,2019.10.03
US 2015315551 A1,2015.11.05
NCBI Reference Sequence: WP_075723423.1.phytoene synthase [Corynebacterium stationis].GENBANK.2017,全文.
审查员 李娟娟

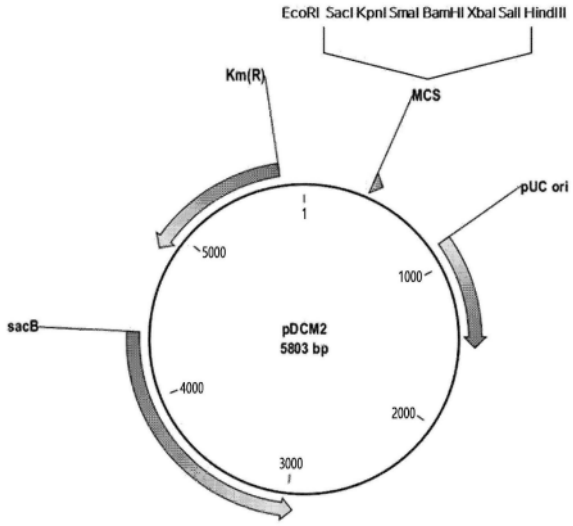
权利要求书1页 说明书14页
序列表7页 附图1页

(54)发明名称

新型八氢番茄红素合成酶变体及使用其生产XMP或GMP的方法

(57)摘要

本公开涉及一种新型八氢番茄红素合成酶变体、包含所述变体的停滞棒杆菌(Corynebacterium stationis)菌株以及使用所述菌株生产XMP或GMP的方法。



CN 116897204 B

1.一种八氢番茄红素合成酶变体,其由SEQ ID NO:1表示的氨基酸序列组成,其中作为对应于SEQ ID NO:3的位置133的氨基酸的丙氨酸被缬氨酸取代。

2.一种多核苷酸,其编码根据权利要求1所述的变体。

3.一种停滞棒杆菌(Corynebacterium stationis)菌株,其包含根据权利要求1所述的变体或编码所述变体的多核苷酸。

4.根据权利要求3所述的菌株,其与包含SEQ ID NO:3的多肽或编码所述多肽的多核苷酸的停滞棒杆菌相比,XMP生产能力增加。

5.根据权利要求4所述的菌株,其中XMP被转化成GMP。

6.一种用于生产XMP的方法,所述方法包括在培养基中培养包含根据权利要求1所述的变体或编码所述变体的多核苷酸的停滞棒杆菌菌株。

7.一种用于生产GMP的方法,所述方法包括:(a)根据权利要求6所述的方法生产XMP;以及(b)将步骤(a)中生产的XMP转化成GMP。

新型八氢番茄红素合成酶变体及使用其生产XMP或GMP的方法

技术领域

[0001] 本公开涉及一种新型八氢番茄红素合成酶变体、包含所述变体的停滞棒杆菌(Corynebacterium stationis)菌株以及使用所述菌株生产XMP或GMP的方法。

背景技术

[0002] 正在进行多种研究以开发出用于生产XMP、GMP和其他有用物质的高效微生物和发酵工艺技术。例如,主要使用目标物质特异性方法,在所述方法中编码参与XMP或GMP生物合成的酶的基因的表达增加,或在所述方法中除去了生物合成不需要的基因(EP 3722430 A1、US 2020-0347346 A1)。

[0003] 但是,随着对XMP和GMP的需求增加,仍然需要进行研究以有效增加XMP和GMP生产能力。

发明内容

[0004] 技术问题

[0005] 本发明人开发了一种新型八氢番茄红素合成酶变体、包含所述变体的停滞棒杆菌菌株以及使用所述菌株生产XMP或GMP的方法,从而完成本公开。

[0006] 技术方案

[0007] 本公开的一个目的为提供一种八氢番茄红素合成酶变体,其由SEQ ID NO:1表示的氨基酸序列组成,其中作为对应于氨基酸序列SEQ ID NO:3的位置133的氨基酸的丙氨酸被缬氨酸取代。

[0008] 本公开的另一个目的为提供一种编码本公开的变体的多核苷酸。

[0009] 本公开的另一个目的为提供一种停滞棒杆菌菌株,其包含本公开的变体或编码所述变体的多核苷酸并且具有XMP或GMP生产能力。

[0010] 本公开的另一个目的为提供一种用于生产XMP或GMP的方法,其包括在培养基中培养包含本公开的变体或编码所述变体的多核苷酸并且具有XMP或GMP生产能力的停滞棒杆菌菌株。

[0011] 有益效果

[0012] 在培养包含本公开的八氢番茄红素合成酶变体的停滞棒杆菌菌株的情况下,与具有未修饰多肽的现有微生物的情况相比,有可能以较高产率生产XMP或GMP。

附图说明

[0013] 图1为pDCM2质粒的示意图。

具体实施方式

[0014] 将如下详细地描述本公开。同时,本公开中所公开的每个描述和实施方案均可应用于其他描述和实施方案。换句话讲,本公开中所公开的各种要素的所有组合均属于本公

开的范围。此外,不可以认为本公开的范围受到以下所述的具体描述的限制。此外,在本说明书通篇,引用了许多论文和专利文件并指出了其引用文。所引用的论文和专利文件中所公开的全部内容以引用的方式并入本说明书中,以更清楚地描述本发明所属的技术领域的水平和本发明的内容。

[0015] 本公开的一个方面提供一种八氢番茄红素合成酶变体,其由SEQ ID NO:1表示的氨基酸序列组成,其中作为对应于氨基酸序列SEQ ID NO:3的位置133的氨基酸的丙氨酸被缬氨酸取代。

[0016] 本公开的变体可具有由SEQ ID NO:1表示的氨基酸序列,包含所述氨基酸序列或基本上由所述氨基酸序列组成。

[0017] 在本公开的变体中,在由SEQ ID NO:1表示的氨基酸序列中对应于基于氨基酸序列SEQ ID NO:3的位置133的氨基酸为缬氨酸,并且可包含与由SEQ ID NO:1表示的氨基酸序列有至少70%、75%、80%、85%、90%、95%、96%、97%、98%、99%、99.5%、99.7%或99.9%同源性或同一性的氨基酸序列。明显的是,具有缺失、修饰、取代、保守取代或添加了一些序列的氨基酸序列的变体也包括在本公开的范围内,只要氨基酸序列具有此类同源性或同一性并表现出对应于本公开的变体功效的功效。

[0018] 其实例包括具有在氨基酸序列的N末端和C末端和/或氨基酸序列内部的不改变本公开变体功能的序列添加或缺失、天然存在的突变、沉默突变或保守取代的变体。

[0019] “保守取代”意指一种氨基酸被具有类似结构和/或化学性质的另一种氨基酸取代。这种氨基酸取代可通常基于残基的极性、电荷、溶解度、疏水性、亲水性和/或两亲性性质的类似性而发生。通常,保守取代可能几乎不影响或不影响蛋白或多肽的活性。

[0020] 在本公开中,术语“变体”是指氨基酸序列与通过一个或多个氨基酸的保守取代和/或修饰而变异之前的变体的氨基酸序列不同但维持功能或性质的多肽。这种变体通常可以通过修饰多肽氨基酸序列的一个或多个氨基酸并评估修饰多肽的性质来鉴别。换句话讲,与变异之前多肽的能力相比,变体的能力可能增加、不变或降低。一些变体可包括除去了一个或多个部分诸如N末端前导序列或跨膜结构域的变体。其他变体可包括将N末端和/或C末端的一部分从成熟蛋白除去的变体。术语“变体”可与术语诸如修饰、修饰多肽、修饰蛋白、突变体、突变蛋白和趋异(divergent)互换使用,并且不限于此,只要其为以变化含义使用的术语即可。出于本公开的目的,变体可为包含由SEQ ID NO:1表示的氨基酸序列的多肽,其中丙氨酸为对应于氨基酸序列SEQ ID NO:3的位置133的氨基酸,被缬氨酸取代。

[0021] 变体可包括对多肽的性质和二级结构具有最小影响的胺基酸的缺失或添加。例如,共翻译或翻译后参与蛋白易位的信号(或前导)序列可与变体的N末端缀合。变体可与其他序列或接头缀合,以便被鉴别、纯化或合成。

[0022] 在本公开中,术语“同源性”或“同一性”意指两个给定氨基酸序列或碱基序列之间的类似性程度且可以百分比的形式表示。术语同源性和同一性常常可互换使用。

[0023] 保守多核苷酸或多肽的序列同源性或同一性由标准比对算法确定,并且可以一起使用由所用程序建立的默认空位罚分。基本上,同源或相同序列通常能够在中等或高度严格的条件下与序列整体或一部分杂交。明显的是,杂交还包括多核苷酸与包含通用密码子或考虑到密码子简并性的密码子的多核苷酸的杂交。

[0024] 任何两个多核苷酸或多肽序列是否具有同源性、类似性或同一性可以使用已知的

计算机算法诸如“FASTA”程序,例如使用如Pearson等人,(1988)Proc.Natl.Acad.Sci.USA 85:2444中的默认参数来确定。替代地,同源性、类似性或同一性可使用如EMBOSS包(EMBOSS:The European Molecular Biology Open Software Suite,Rice等人,2000,Trends Genet.16:276-277)(版本5.0.0或之后版本)(包括GCG程序包(Devereux,J.等人,Nucleic Acids Research 12:387(1984))的Needleman程序中所执行的Needleman-Wunsch算法(Needleman和Wunsch,1970,J.Mol.Biol.48:443-453)、BLASTP、BLASTN、FASTA(Atschul,S.F.等人,JMOLEC BIOL 215:403(1990);Guide to Huge Computers,Martin J.Bishop编,Academic Press,San Diego,1994;以及CARILLO等人(1988)SIAM J Applied Math 48:1073)来确定。例如,可使用美国国家生物技术信息中心的BLAST或ClustalW确定同源性、类似性或同一性。

[0025] 多核苷酸或多肽的同源性、类似性或同一性可如例如Smith和Waterman,Adv.Appl.Math(1981)2:482中所发表的通过使用例如GAP计算机程序诸如Needleman等人(1970),J Mol Biol.48:443比较序列信息来确定。概括地说,GAP程序可定义为通过两个序列中的较短序列中符号(即,核苷酸或氨基酸)的总数除以类似比对符号的数目所获得的值。GAP程序的默认参数可包括(1)二元比较矩阵(包括相同值为1且非同值为0)和Gribsko等人,(1986)Nucl.Acids Res.14:6745(或EDNAFULL(NCBI NUC4.4的EMBOSS版)取代矩阵)的加权比较矩阵,如Schwartz and Dayhoff编,Atlas Of Protein Sequence And Structure,National Biomedical Research Foundation,第353-358页(1979)中所公开;(2)每个空位的罚分为3.0并且每个空位的每个符号额外罚分0.10(或空位开放罚分为10,空位延伸罚分为0.5);以及(3)末端空位没有罚分。

[0026] 作为本公开的实例,本公开可表现出八氢番茄红素合成酶活性。本公开的变体可表现出活性,以便与表现出八氢番茄红素合成酶活性的野生型多肽相比,XMP或GMP生产能力增加。

[0027] 在本公开中,术语“八氢番茄红素合成酶”是指参与类胡萝卜素生物合成的多肽。的多肽。具体地说,本公开的八氢番茄红素合成酶可与CrtB2或PSY互换使用。在本公开中,八氢番茄红素合成酶的序列可获自已知数据库NCBI的GenBank。具体地说,八氢番茄红素合成酶可为由crtB2编码的表现出八氢番茄红素合成酶活性的多肽,但不限于此。

[0028] 在本公开中,术语“对应于”是指多肽或氨基酸残基中所列出的位置处的氨基酸残基与多肽中所列出的氨基酸残基类似、相同或同源。鉴别对应位置处的氨基酸可为确定称为特定序列的序列中的特定氨基酸。如本文所用,“对应区域”通常是指相关蛋白或参考蛋白中类似或对应的位置。

[0029] 例如,将任意氨基酸序列与SEQ ID NO:3比对,并且基于此,可以相对于SEQ ID NO:3的氨基酸残基和对应氨基酸残基的数值位置对氨基酸序列的每个氨基酸残基进行编号。例如,如本公开中所述的序列比对算法可通过与查询序列(也称为“参考序列”)比较来确定氨基酸位置或修饰诸如取代、插入或缺失发生的位置。

[0030] 针对此类比对,例如,可使用Needleman-Wunsch算法(Needleman和Wunsch,1970,J.Mol.Biol.48:443-453)、EMBOSS包的Needleman程序(EMBOSS:The European Molecular Biology Open Software Suite,Rice等人,2000,Trends Genet.16:276-277)等,但程序和算法不限于此,并且可适当使用本领域中已知的序列比对程序、成对序列比较算法等。

[0031] 本公开的另一个方面为提供一种编码本公开的变体的多核苷酸。

[0032] 在本公开中,术语“多核苷酸”为具有一定长度或更长的以核苷酸聚合物的形式的在其中核苷酸单体通过共价键连接成长链的DNA或RNA链,并且更具体地指编码变体的多核苷酸片段。

[0033] 编码本公开的变体的多核苷酸可包括编码由SEQ ID NO:1表示的氨基酸序列的碱基序列。作为本公开的实例,本公开的多核苷酸具有或包含序列SEQ ID NO:2。本公开的多核苷酸可由序列SEQ ID NO:2组成或基本上由其组成。

[0034] 在本公开的多核苷酸中,可在编码区中进行各种修饰,只要本公开的变体的氨基酸序列不由于密码子简并性或旨在表达本公开的变体的生物体中优选的密码子而变化即可。具体地说,本公开的多核苷酸具有或包含与序列SEQ ID NO:2有70%或更多、75%或更多、80%或更多、85%或更多、90%或更多、95%或更多、96%或更多、97%或更多、98%或更多但小于100%同源性或同一性的碱基序列,或可由与序列SEQ ID NO:2有70%或更多、75%或更多、80%或更多、85%或更多、90%或更多、95%或更多、96%或更多、97%或更多、98%或更多但小于100%同源性或同一性的碱基序列或基本上由其组成。在此,在具有同源性或同一性的序列中,编码对应于SEQ ID NO:1的位置133的氨基酸的密码子可为编码缬氨酸的密码子。

[0035] 本公开的多核苷酸可包含探针,其可由已知基因序列制备,例如没有限制的序列,只要其为可在严格条件下与本公开的多核苷酸序列的整体或一部分的互补序列杂交的序列即可。“严格条件”意指实现多核苷酸之间的特异性杂交的条件。这些条件明确描述于文件中(参见J.Sambrook等人,Molecular Cloning,ALaboratory Manual,第2版,Cold Spring Harbor Laboratory press,Cold Spring Harbor,New York,1989;F.M.Ausubel等人,Current Protocols in Molecular Biology,John Wiley&Sons,Inc.,New York,9.50-9.51,11.7-11.8)。其实例包括具有较高同源性或同一性的多核苷酸(即具有70%或更多、75%或更多、80%或更多、85%或更多、90%或更多、95%或更多、96%或更多、97%或更多、98%或更多或99%或更多同源性或同一性的多核苷酸)彼此不杂交的条件,或在等效于60℃、1×SSC、0.1%SDS,具体地60℃、0.1×SSC、0.1%SDS,更具体地68℃、0.1×SSC、0.1%SDS的盐浓度和温度(其为针对常见Southern杂交的洗涤条件)下进行洗涤一次、具体地两次至三次的条件。

[0036] 杂交要求两个核酸具有互补序列,但碱基之间允许根据杂交的严格性而存在错配。术语“互补”用于描述能够彼此杂交的核苷酸碱基之间的关系。例如,至于DNA,腺嘌呤与胸腺嘧啶互补,并且胞嘧啶与鸟嘌呤互补。因此,本公开的多核苷酸可还包含基本上类似的核酸序列以及与整个序列互补的分离的核酸片段。

[0037] 特别说,与本公开的多核苷酸具有同源性或同一性的多核苷酸可使用在T_m值55℃下包括杂交步骤的杂交条件和上文所述条件来检测。T_m值可为60℃、63℃或65℃,但不限于此,且可以由本领域的技术人员根据目的进行适当调整。

[0038] 杂交多核苷酸的适当严格性取决于多核苷酸的长度和互补程度,并且变量是本领域中熟知的(例如,J.Sambrook等人,Molecular Cloning,A Laboratory Manual,第2版,Cold Spring Harbor Laboratory press,Cold Spring Harbor,New York,1989)。

[0039] 本公开的另一个方面为提供一种载体,其包含本公开的多核苷酸。载体可为用于

在宿主细胞中表达多核苷酸的表达载体,但不限于此。

[0040] 本公开的“载体”可包括DNA构建体,其包含编码感兴趣的多肽的多核苷酸序列,所述序列可操作性地连接至合适的表达控制区(或表达控制序列),以使得感兴趣的多肽可在合适的宿主中表达。表达控制区可含有能够起始转录的启动子、用于控制转录的任何操纵子序列、编码合适的mRNA核糖体结合位点的序列以及控制转录和翻译终止的序列。可将载体转化到合适的宿主细胞中,然后独立于宿主基因组复制和起作用,或者可以将其整合到基因组本身中。

[0041] 本公开中所用的载体没有特别限制,可使用本领域中已知的任何载体。通常使用的载体的实例包括天然或重组质粒、粘粒、病毒和噬菌体。例如,可以将pWE15、M13、MBL3、MBL4、IXII、ASHII、APII、t10、t11、Charon4A、Charon21A等用作噬菌体载体或粘粒载体,并且可以将pDZ系统、pBR系统、pUC系统、pBluescript II系统、pGEM系统、pTZ系统、pCL系统、pET系统等用作质粒载体。具体地说,可以使用pDZ、pDC、pDCM2、pACYC177、pACYC184、pCL、pECCG117、pUC19、pBR322、pMW118和pCC1BAC载体等。

[0042] 例如,可以将编码感兴趣的多肽的多核苷酸通过载体插入到染色体中,以用于细胞内染色体插入。将多核苷酸插入到染色体中可以通过本领域中已知的任何方法(例如,同源重组)进行,但不限于此。载体可以进一步含有选择标记物,以用于确认染色体插入。选择标记物用于选择与载体转化的细胞,即,用于确认感兴趣的核酸分子的插入,并且可以使用赋予可选择表型诸如抗药性、营养缺陷型、细胞毒剂抗性或表面多肽的表达的标记物。在用选择剂处理的环境中,只有表达选择标记物的细胞存活或表现出其他表型性状,因此可以选择转化细胞。

[0043] 在本公开中,术语“转化”意指将包含编码靶多肽的多核苷酸的载体引入到宿主细胞或微生物中,以使得由多核苷酸编码的多肽可以在宿主细胞中表达。转化的多核苷酸可以通过插入到宿主细胞的染色体中来定位,或定位于染色体外,只要其可以在宿主细胞中表达即可。多核苷酸包含编码感兴趣的多肽的DNA和/或RNA。多核苷酸可以以任何形式引入,只要其可以引入到宿主细胞并表达即可。例如,可以将多核苷酸以表达盒的形式引入到宿主细胞,表达盒为包含自我表达所需的所有元件的基因构建体。表达盒通常可以包含与多核苷酸可操作地连接的启动子、转录终止信号、核糖体结合位点和翻译终止信号。表达盒可为能够自我复制的表达载体的形式。多核苷酸可以以其自身的形式引入到宿主细胞中并且可操作地连接至在宿主细胞中表达所需的序列,但不限于此。

[0044] 在上文中,术语“可操作地连接”意指多核苷酸序列与启动子序列功能性地连接,启动子序列起始并介导编码本公开的感兴趣变体的多核苷酸的转录。

[0045] 本公开的另一个方面为提供一种停滞棒杆菌菌株,其包含本公开的变体或本公开的多核苷酸。

[0046] 本公开的菌株可以包含本公开的修饰多肽、编码所述多肽的多核苷酸或包含本公开的多核苷酸的载体。

[0047] 在本公开中,“菌株(或微生物)”包括所有野生型微生物或天然或人工基因修饰的微生物,并且其可为特定机制由于外部基因的插入或内源性基因的活性增强或失活而削弱或加强的微生物,或者其可为包含针对感兴趣的多肽、蛋白或产物的遗传修饰的微生物。

[0048] 本公开的菌株可为:包含本公开的变体、本公开的多核苷酸或包含本公开的多核

苷酸的载体中的任何一种或多种的菌株;经修饰以表达本公开的变体或本公开的多核苷酸的菌株;表达本公开的变体或本公开的多核苷酸的菌株(例如,重组菌株);或表现出本公开的变体的活性的菌株(例如,重组菌株),但不限于此。

[0049] 本公开的菌株可为具有5′-黄苷单磷酸酯(XMP)或5′-鸟苷单磷酸酯(GMP)生产能力的菌株。

[0050] 本公开的菌株可为天然具有八氢番茄红素合成酶或XMP或GMP生产能力的微生物,或者将本公开的变体或编码变体的多核苷酸(或包含多核苷酸的载体)引入到不具有八氢番茄红素合成酶或XMP或GMP生产能力的亲本菌株中和/或对亲本菌株赋予XMP或GMP生产能力的微生物,但不限于此。

[0051] 例如,本公开的菌株为用包含本公开的多核苷酸或编码本公开的变体的多核苷酸的载体转化并表达本公开的变体的细胞或微生物。出于本公开的目的,本公开的菌株可包括包含本公开的变体并且可生产XMP或GMP的所有微生物。例如,本公开的菌株可为重组菌株,在重组菌株中将编码本公开的变体的多核苷酸引入到天然野生型微生物或生产XMP或GMP的微生物中,以因此表达八氢番茄红素合成酶变体并具有增加的XMP或GMP生产能力。具有增加的XMP或GMP生产能力的重组菌株可为与天然野生型微生物或八氢番茄红素合成酶未修饰微生物(即,表达野生型八氢番茄红素合成酶(SEQ ID NO:3)的微生物或不表达修饰修饰的(SEQ ID NO:1)蛋白的微生物)相比具有增加的XMP或GMP生产能力的微生物,但不限于此。例如,本公开的具有增加的XMP或GMP生产能力的菌株可为与包含多肽SEQ ID NO:3或编码所述多肽的多核苷酸的停滞棒杆菌相比具有增加的XMP或GMP生产能力的微生物,但不限于此。例如,八氢番茄红素合成酶未修饰微生物作为用于比较XMP或GMP生产能力的增加的靶菌株,可为ATCC6872或CJX1664菌株(KCCM12285P,KR 10-1950141B1),但不限于此。

[0052] 例如,与变异之前的亲本菌株或未修饰微生物的XMP或GMP生产能力相比,具有增加的生产能力的重组菌株的XMP或GMP生产能力可增加约1%或更多、特别地约1%或更多、约2.5%或更多、约5%或更多、约6%或更多、约7%或更多、约8%或更多、约9%或更多、约10%或更多、约10.5%或更多、约11%或更多、约11.5%或更多、约12%或更多、约12.5%或更多、约13%或更多、约13.5%或更多、约14%或更多、或约14.5%或更多(上限不进行具体限制并且可为例如约200%或更少、约150%或更少、约100%或更少、约50%或更少、约45%或更少、约40%或更少、约35%或更少、约30%或更少、约25%或更少、约20%或更少、或约15%或更少),但增加的量不限于此,只要生产能力与变化之前的亲本菌株或未修饰微生物的生产能力相比具有增加量的+值即可。在另一个实例中,与变异之前的亲本菌株或未修饰微生物的XMP或GMP生产能力相比,具有增加的生产能力的重组菌株的XMP或GMP生产能力可增加约1.1倍或更多、约1.12倍或更多、约1.13倍或更多、或约1.14倍或更多(上限不进行具体限制并且可为例如约10倍或更少、约5倍或更少、约3倍或更少或约2倍或更少)。术语“约”为包括±0.5、±0.4、±0.3、±0.2、±0.1等所有的范围,包括等于或类似于在术语“约”之后的值的范围中的所有值,但不限于此。

[0053] 在本公开中,“未修饰微生物”不排除包含可在微生物中天然存在的突变的菌株,并且可为野生型菌株或天然菌株本身,或者可为通过因天然或人工因素所致的遗传变异而性状改变之前的菌株。例如,未修饰微生物可为未引入或尚未引入本说明书中所述的八氢番茄红素合成酶变体的菌株。术语“未修饰微生物”可与“修饰之前的菌株”、“修饰之前的微

生物”、“未变异菌株”、“未修饰菌株”、“未变异微生物”或“参考微生物”互换使用。

[0054] 在本公开的另一个实例中,本公开的微生物可为停滞棒杆菌、生乳棒杆菌(Corynebacterium crudilactis)、沙漠棒杆菌(Corynebacterium deserti)、有效棒杆菌(Corynebacterium efficiens)、帚石南棒杆菌(Corynebacterium callunae)、谷氨酸棒杆菌(Corynebacterium glutamicum)、单一棒杆菌(Corynebacterium singulare)、耐盐棒杆菌(Corynebacterium halotolerans)、纹带棒杆菌(Corynebacterium striatum)、产氨棒杆菌(Corynebacterium ammoniagenes)、污染棒杆菌(Corynebacterium pollutisoli)、亚胺棒杆菌(Corynebacterium imitans)、testudinoris棒杆菌或微黄棒杆菌(Corynebacterium flavescens)。

[0055] 在本公开中,多肽的“弱化”包括与内源性活性相比活性降低或活性缺失的两种情况。弱化可与诸如失活、缺陷、下调、降低、减小和衰减的术语互换使用。

[0056] 弱化还可以包括:与微生物最初拥有的多肽的活性相比,由于编码动态的多核苷酸的变异等而多肽本身的活性降低或消除的情况;与天然菌株相比,由于编码多肽的多核苷酸的基因的表达的抑制或翻译成多肽的抑制而细胞中总体多肽活性水平和/或浓度(表达水平)较低的情况;多核苷酸根本不表达的情况;和/或即使当多核苷酸表达,也不表现出多肽活性的情况。“内源性活性”意指性状改变之前的亲本菌株或通过因天然或人工因素所致的遗传变异而性状改变时的野生型或未修饰微生物最初拥有的特定多肽的活性。内源性活性可与“修饰之前的活性”互换使用。与内源性活性相比,多肽的活性“失活、缺陷、降低、下调、减少或衰弱”的事实意指与性状改变之前的亲本菌株或野生型或未修饰微生物最初拥有的特定多肽的活性相比,多肽的活性降低。

[0057] 此类多肽活性弱化可通过本领域中已知的任何方法进行,但是方法不限于此,并且弱化可以通过应用本领域中熟知的各种方法来实现(例如,Nakashima N.等人,Bacterial cellular engineering by genome editing and gene silencing.Int JMol Sci.2014;15(2):2773-2793;Sambrook等人,Molecular Cloning 2012)。

[0058] 特别地,本公开中多肽活性的弱化可为:

[0059] 1)缺失编码多肽的基因的全部或部分;

[0060] 2)修饰表达控制区(或表达控制序列)以降低编码多肽的基因的表达;

[0061] 3)修饰构成多肽的氨基酸序列以消除或弱化多肽活性(例如,氨基酸序列中一个或多个氨基酸的缺失/取代/添加);

[0062] 4)修饰编码多肽的基因序列以消除或弱化多肽活性(例如,经过修饰以消除或弱化多肽活性的编码多肽的多肽基因的核酸碱基序列中一个或多个核酸碱基的缺失/取代/添加);

[0063] 5)修饰编码多肽的基因转录物或编码5′-UTR区的碱基序列的起始密码子;

[0064] 6)引入与编码多肽的基因的转录物互补结合的反义寡核苷酸(例如,反义RNA);

[0065] 7)在编码多肽的基因的SD序列(Shine-Dalgarno sequence)之前添加与SD序列互补的序列以形成不可连接核糖体的二级结构;

[0066] 8)添加在编码多肽的鸡眼序列的开放阅读框(ORF)的3′末端的相反方向上转录的启动子(逆转录工程化,RTE);或

[0067] 9)选自1)至8)的两种或更多种的组合,但不特别限于此。

[0068] 例如:

[0069] 1)缺失编码多肽的基因的一部分或全部可除去染色体中编码感兴趣的内源性多肽的整个多核苷酸、用缺失一些核苷酸的多肽置换、或用标记物基因置换。

[0070] 2)表达控制区(或表达控制序列)的修饰可为缺失、插入、非保守或保守取代,或由于其组合而发生表达控制区(或表达控制序列)的变异,或用表现出较弱活性的序列置换。表达控制区含有启动子、操纵子序列、编码核糖体结合位点的序列和控制转录和翻译终止的序列,但不限于此。

[0071] 5)修饰编码多肽的基因转录物或编码5′-UTR区的碱基序列的起始密码子可为例如用与内源性起始密码子相比具有较低多肽表达速率的编码另一种起始密码子的碱基序列取代,但不限于此。

[0072] 3)和4)氨基酸序列或多核苷酸序列的修饰可为缺失、插入或非保守或保守取代多肽的氨基酸序列或编码多肽的核苷酸序列,或者序列中因其组合所致而发生变异,或者用经修饰以表现出较弱活性的氨基酸序列或多核苷酸序列或经修饰以失活的氨基酸序列或多核苷酸序列置换,以使得多肽的活性弱化,但不限于此。例如,基因表达可以通过将变异引入到多核苷酸序列中并形成密码子来抑制或弱化,但修饰不限于此。

[0073] 6)对于与编码多肽的基因的转录物互补结合的反义寡核苷酸(例如,反义RNA)的引入,可参考文件,例如Weintraub,H.等人,Antisense-RNA as a molecular tool for genetic analysis,Reviews-Trends in Genetics,第1卷(1)1986。

[0074] 7)在编码多肽的基因的SD序列之前添加与SD序列互补的序列以形成不可连接核糖体的二级结构可为使mRNA不能翻译或减慢mRNA翻译速率。

[0075] 8)添加在编码多肽的鸡眼序列的开放阅读框(ORF)的3′末端的相反方向上转录的启动子(逆转录工程化,RTE)可为通过制备与编码多肽的基因的转录物互补的反义核苷酸来弱化活性。

[0076] 在本公开中,术语多肽活性的“增强”意指与内源性活性相比,多肽的活性增加。增强可与术语诸如活化、上调、过表达和增加互换使用。在此,活化、增强、上调、过表达和增加包括表现出最初没有的活性和表现出与内源性活性或修饰之前的活性相比改善的活性。“内源性活性”意指性状改变之前的亲本菌株或通过因天然或人工因素所致的遗传变异而性状改变时的未修饰微生物最初拥有的特定多肽的活性。这可与“修饰之前的活性”互换使用。与内源性活性相比,多肽的活性“增强”、“上调”、“过表达”或“增加”的事实意指与性状改变之前的亲本菌株或未修饰微生物最初拥有的特定多肽的活性和/或浓度(表达水平)相比,多肽的活性改善。

[0077] 增强可以通过引入外来多肽或增强多肽的活性和/或浓度(表达水平)来实现。多肽活性的增强可以通过多肽的活性程度和表达水平或者由多肽产生的产物的量的增加来确认。

[0078] 对于多肽活性的增强,可应用本领域中熟知的各种方法,并且方法不受限制,只要感兴趣的多肽的活性与修饰之前微生物的活性相比增强即可。具体地说,可使用本领域的技术人员熟知的基因工程和/或蛋白质工程,其为分子生物学的例行方法,但方法不限于此(例如,Sitnicka等人,Functional Analysis of Genes.Advances in Cell Biology.2010,第2卷.1-16;Sambrook等人,Molecular Cloning 2012等)。

[0079] 特别地,本公开的多肽的活性的增强可为:

[0080] 1)编码多肽的多核苷酸的细胞内拷贝数增加;

[0081] 2)用表现出强活性的序列置换编码多肽的染色体上的基因表达调控区;

[0082] 3)修饰编码多肽的基因转录物或编码5′-UTR区的碱基序列的起始密码子;

[0083] 4)修饰多肽的氨基酸序列以增强多肽的活性;

[0084] 5)修饰编码多肽的多核苷酸以增强多肽的活性(例如,修饰多肽基因的多核苷酸序列以编码经修饰以增强多肽的活性的多肽);

[0085] 6)引入表现出多肽的活性的外来多肽或编码多肽的外来多核苷酸;

[0086] 7)密码子优化编码多肽的多核苷酸;

[0087] 8)分析多肽的三级结构以进行选择并修饰或化学修饰暴露位点;或

[0088] 9)选自1)至8)的两种或更多种的组合,但不特别限于此。

[0089] 更特别地:

[0090] 1)编码多肽的多核苷酸的细胞内拷贝数的增加可以通过引入到载体的宿主细胞中来实现,所述载体可被复制并独立于宿主起作用并且可操作地连接编码多肽的多核苷酸。替代地,增加可通过将编码多肽的多核苷酸的一个拷贝或两个或更多个拷贝引入到宿主细胞的染色体中来实现。引入到染色体中可以通过将能够将多核苷酸插入到宿主细胞的染色体中的载体引入宿主细胞中来进行,但不限于此。载体如上文所述。

[0091] 2)用表现出强活性的序列置换编码多肽的染色体上的基因表达控制区(或表达控制序列)可为例如缺失、插入、非保守或保守取代,或由于其组合而发生序列的变异,或用表现出较强活性的序列置换,以使得表达控制区的活性进一步增强。表达控制区没有特别限制,但可含有启动子、操纵子序列、编码核糖体结合位点的序列和控制转录和翻译终止的序列等。例如,置换可为用强启动子置换初始启动子,但不限于此。

[0092] 已知强启动子的实例包括CJ1至CJ7启动子(US 7662943 B2)、lac启动子、trp启动子、trc启动子、tac启动子、λ噬菌体PR启动子、PL启动子、tet启动子、gapA启动子、SPL7启动子、SPL13(sm3)启动子(US 10584338 B2)、O2启动子(US 10273491 B2)、tkt启动子和yccA启动子,但不限于此。

[0093] 3)修饰编码多肽的基因转录物或编码5′-UTR区的碱基序列的起始密码子可为用与内源性起始密码子相比具有较高多肽表达速率的编码另一种起始密码子的碱基序列取代,但不限于此。

[0094] 4)和5)氨基酸序列或多核苷酸序列的修饰可为缺失、插入、非保守或保守取代多肽的氨基酸序列或编码多肽的核苷酸序列,或者序列中因其组合所致而发生变异,或者用经修饰以表现出较强活性的氨基酸序列或多核苷酸序列或经修饰以更具活性的氨基酸序列或多核苷酸序列置换,以使得多肽的活性增强,但不限于此。置换可具体地通过将多核苷酸通过同源重组插入到染色体中来进行,但不限于此。此处所用的载体可以进一步含有选择标记物,以用于确认染色体插入。选择标记物如上文所述。

[0095] 6)引入表现出多肽的活性的外来多肽可为将编码表现出与多肽的活性相同或类似的活性的多肽的外来多核苷酸引入到宿主细胞中。外来多核苷酸的来源或序列不受限制,只要其表现出与多肽的活性相同或类似的活性即可。引入可以通过适当地选择本领域的技术人员已知的转化方法来进行。当引入的多核苷酸在宿主细胞中表达时,可以产生多

肽,并且其活性可以增加。

[0096] 7)密码子优化编码多肽的多核苷酸可为密码子优化内源性多核苷酸以便增加宿主细胞中的转录或翻译或者密码子优化外来多核苷酸以便在宿主细胞中进行优化的转录或翻译。

[0097] 8)分析多肽的三级结构以选择并修饰或化学修饰暴露位点可为例如通过将待分析的多肽的序列信息与储存已知蛋白的序列信息的数据库比较来根据序列的类似性程度确定模板蛋白候选物,基于此确认结构并选择和修饰或化学修饰待修饰或化学修饰的暴露部分。

[0098] 多肽活性的此类增强可为基于野生型微生物菌株或修饰之前的微生物菌株中表达的多肽的活性或浓度的对应多肽的活性或浓度表达水平的增加或者由多肽产生的产物的量的增加,但不限于此。

[0099] 在本公开的微生物中,多肽的部分或整个修饰可以通过以下来诱导:(a)使用在微生物中插入染色体的载体进行同源重组或使用工程核酸酶(例如,CRISPR-Cas9)进行基因组编辑;和/或(b)用光诸如紫外线和辐射和/或化学品进行处理,但不限于此。用于修饰基因的一部分或全部的方法可以包括使用DNA重组技术的方法。例如,通过将含有与感兴趣的基因同源的核苷酸序列的核苷酸序列或载体引入微生物中以引起同源重组,可以缺失基因的一部分或全部。引入的核苷酸序列或载体可包含显性选择标记物,但不限于此。

[0100] 在本公开的微生物中,变体、多核苷酸、XMP和GMP等如其他方面所述。

[0101] 本公开的另一个方面提供一种用于生产XMP或GMP的方法,其包括在培养基中培养包含本公开的变体或本公开的多核苷酸的停滞棒杆菌菌株。

[0102] 用于生产本公开的XMP或GMP的方法可以包括在培养基中培养包含本公开的变体、或本公开的多核苷酸、或本公开的载体的停滞棒杆菌菌株。

[0103] 在本公开中,术语“培养物”意指使本公开的停滞棒杆菌菌株在适当控制的环境条件下生长。本公开的培养过程可以根据本领域已知的合适的培养基和培养条件进行。本领域技术人员可以根据选择的菌株容易地调整和使用这种培养过程。具体地说,培养物可以是分批型、连续型和/或分批补料型,但不限于此。

[0104] 在本公开中,术语“培养基”意指含有培养本公开的停滞棒杆菌菌株所需的营养物质作为主要组分的混合物质,并且培养基供应存活和发展不可缺少的营养物质、生长因子等,包括水。具体地说,作为用于培养本公开的停滞棒杆菌菌株的培养基和其他培养条件,可以没有特别限制地使用任何一种,只要其为用于微生物的普通培养的培养基即可。本公开的停滞棒杆菌菌株可以在有氧条件下控制温度、pH等的同时在含有适当碳源、氮源、磷源、无机化合物、氨基酸和/或维生素等的普通培养基中培养。

[0105] 特别地,棒杆菌属菌株的培养基可见于美国微生物学会的文件“Manual of Methods for General Bacteriology”(Washington,D.C.,USA,1981)中。

[0106] 在本公开中,碳源包括:碳水化合物,诸如葡萄糖、蔗糖(saccharose)、乳糖、果糖、蔗糖(sucrose)和麦芽糖;糖醇,诸如甘露醇和山梨糖醇;有机酸,诸如丙酮酸、乳酸和柠檬酸;氨基酸,诸如谷氨酸、蛋氨酸和赖氨酸;等。可以使用天然有机营养物质,诸如淀粉水解物、糖蜜、赤糖糊、米糠、木薯、甘蔗渣和玉米浆。具体地说,可以使用碳水化合物,诸如葡萄糖和经过灭菌的预处理糖蜜(即,转化为还原糖的糖蜜),并且可以以各种方式使用适当量

的其他碳源而没有限制。这些碳源可以单独使用或以两种或更多种的组合使用,但不限于此。

[0107] 作为氮源,可以使用:无机氮源,诸如氨、硫酸铵、氯化铵、乙酸铵、磷酸铵、碳酸铵和硝酸铵;和有机氮源,诸如氨基酸(诸如谷氨酸、蛋氨酸和谷氨酰胺)、蛋白胨、NZ-胺、肉提取物、酵母提取物、麦芽提取物、玉米浆、酪蛋白水解物、鱼或其分解产物以及脱脂豆饼或其分解产物。这些氮源可以单独使用或以两种或更多种的组合使用,但不限于此。

[0108] 磷源可包括磷酸二氢钾、磷酸氢二钾或与其对应的含钠盐。作为无机化合物,可以使用氯化钠、氯化钙、氯化铁、硫酸镁、硫酸铁、硫酸锰、碳酸钙等。除这些之外,还可以含有氨基酸、维生素和/或合适的前体等。这些组分或前体可以分批或连续地添加到培养基中,但添加方式不限于此。

[0109] 在培养本公开的停滞棒杆菌菌株期间,可以通过以适当方式向培养基中添加化合物诸如氢氧化铵、氢氧化钾、氨、磷酸、硫酸等来调节培养基的pH。在培养期间,可以通过使用消泡剂诸如脂肪酸聚乙二醇酯来抑制起泡。可向培养基中注入氧气或含氧气体以维持培养基的好氧状态,或可不注入气体或可以注入氮气、氢气或二氧化碳气体以维持培养基的厌氧和微氧状态,但用于维持状态的方法不限于此。

[0110] 在本公开的培养中,培养温度可维持在20℃至45℃,具体地25℃至40℃,并且菌株可培养约10至160小时,但培养条件不限于此。

[0111] 通过本公开的培养物生产的XMP或GMP可以分泌到培养基中或可以保留在细胞中。

[0112] 本公开的用于生产XMP或GMP的方法可还包括制备本公开的停滞棒杆菌菌株的步骤、制备用于菌株培养的培养基的步骤或这些的组合(以任何顺序),例如,在培养步骤之前。

[0113] 本公开的用于生产XMP或GMP的方法可还包括从根据培养物的培养基(进行培养的培养基)或从停滞棒杆菌菌株中回收XMP或GMP的步骤。在培养步骤之后可还包括回收步骤。

[0114] 回收可为根据培养本公开的用于培养微生物的方法例如分批、连续或分批补料方法,通过本领域已知的合适方法来收集感兴趣的XMP或GMP。例如,离心、过滤、用结晶蛋白质沉淀剂处理(盐析)、提取、超声分解、超滤、渗析、各种形式的色谱法诸如分子筛色谱法(凝胶过滤)、吸附色谱法、离子交换色谱法和亲和亲和色谱法、HPLC或其组合。可以通过本领域已知的合适方法从培养基或微生物中回收感兴趣的XMP或GMP。

[0115] 本公开的用于生产XMP或GMP的方法还可包括纯化步骤。纯化可以通过本领域已知的合适方法进行。在一个实例中,当本公开的用于生产XMP或GMP的方法包括回收步骤和纯化步骤时,回收步骤和纯化步骤可以不分顺序地连续或不连续地进行,或者可以同时或通过合并成一个步骤来进行,但进行步骤的方式不限于此。

[0116] 特别地,本公开的用于生产GMP生产方法还可以包括将XMP转化为GMP的步骤。在本公开的用于生产GMP的方法中,可在培养步骤或回收步骤之后还包括转化步骤。转化步骤可以通过本领域已知的合适方法进行。例如,可以使用棒状微生物、大肠杆菌或5′-黄苷酸脱氨基酶(KR 10-0655902B1)进行转化,但进行转化的方式不限于此。

[0117] 在本公开的方法中,变体、多核苷酸、载体、菌株、XMP和GMP等如其他方面所述。

[0118] 本公开的另一个方面为提供一种用于XMP或GMP生产的组合物,其包含有包含本公开变体的停滞棒杆菌菌株、编码所述变体的多核苷酸、包含所述多核苷酸的载体或本公开

的多核苷酸;一种培养基,在其中培养了此停滞棒杆菌菌株;或其中两个或多个的组合。

[0119] 本公开的组合物还可包含常用于XMP或GMP生产的组合物中的任意合适的赋形剂。此类赋形剂可为例如防腐剂、润湿剂、分散剂、悬浮剂、缓冲剂、稳定剂或等渗剂,但不限于此。

[0120] 在本公开的组合物中,变体、多核苷酸、载体、菌株、培养基、XMP和GMP等如其他方面所述。

[0121] 实施例

[0122] 在下文中,将通过实施例更详细地描述本公开。然而,以下实施例仅为用于说明本公开的优选实施例,且因此不旨在将本公开的范围限制于此。同时,本说明书中未描述的技术事项可以被本公开的技术领域或类似技术领域的技术人员充分理解并容易地实施。

[0123] 实施例1:质粒的构建

[0124] 使用BIONICS Co.,Ltd的基因合成服务设计并合成了用于在棒杆菌染色体中插入和置换基因的质粒(pDCM2,图1,SEQ ID NO:11)。通过参考众所周知的关于sacB系统的论文(Gene,145(1994)69-73),将质粒设计成含有限制性酶位点以容易地用于克隆。因此合成的pDCM2质粒具有以下性质:

[0125] 1)pDCM2质粒具有仅在大肠杆菌中起作用的复制起点,因此自我复制在大肠杆菌中是有可能的,但在棒杆菌中不可能。

[0126] 2)pDCM2质粒具有卡那霉素抗性基因作为选择标记物。

[0127] 3)pDCM2质粒具有Levan蔗糖基因(sacB)作为二级阳性选择标记物。

[0128] 4)最终构建的菌株中没有来源于pDCM2质粒的基因信息。

[0129] 实施例2.微生物中八氢番茄红素合成酶变体表达的载体的构建

[0130] 为了探究变体(A133V;SEQ ID NO:1)(其中八氢番茄红素合成酶氨基酸序列(SEQ ID NO:3)的位置133处的丙氨酸被缬氨酸取代)对XMP生产的作用,如下构建了用于构建表达变体的菌株的载体。

[0131] 使用野生型停滞棒杆菌ATCC6872的gDNA(基因组DNA)作为模板,并且一对具有序列SEQ ID NO:5和6的引物以及一对具有序列SEQ ID NO:7和8的引物进行PCR。使用上文获得的两个片段的混合物作为模板和一对具有序列SEQ ID NO:5和SEQ ID NO:8的引物再次进行重叠PCR以获得片段。PCR如下进行:在94℃下变性5分钟;在94℃下30秒、在55℃下30秒、且在72℃下1分30秒,重复30次;并且在72℃下5分钟。pDCM2载体用smaI处理,并且将上文获得的PCR产物融合克隆到其中。使用In-Fusion®HD克隆试剂盒(Clontech)进行融合克隆。所得质粒命名为pDCM2-crtB2(A133V)。

[0132] 实施例3:对表达八氢番茄红素合成酶变体的微生物的XMP生产能力的评估

[0133] 3-1.用于表达八氢番茄红素合成酶变体的菌株的构建

[0134] 将实施例2中构建的载体转化到停滞棒杆菌CJX1664菌株(KCCM12285P,KR 10-1950141B1)中。

[0135] 使用SEQ ID NO:9和10,从发生同源重组的菌株中选择引入变体的菌株。选择的菌株命名为CJX1664_crtB2_A133V。

[0136] 3-2.表达八氢番茄红素合成酶变体的菌株的XMP生产能力的比较

[0137] 通过对实施例3-1中构建的每种菌株和对照亲本菌株的烧瓶发酵效价评估来分析

XMP生产能力。

[0138] 首先,将每种菌株接种到含有2mL种子培养基的直径18mm的试管中,并在30℃下振荡培养24小时以用作种子培养溶液。向含有32mL以下生产培养基(24mL主培养基+8mL单独灭菌培养基)的250mL角挡板烧瓶(corner-baffle flask)中,接种0.7mL种子培养溶液并在30℃下以170rpm振荡培养75小时。培养完成之后,通过HPLC测量XMP生产能力。每种测试菌株的培养溶液中的XMP浓度以及XMP浓度的增加率如以下表1所示。

[0139] <种子培养基(pH 7.5)>

[0140] 葡萄糖1%、蛋白胨1%、牛肉提取物1%、酵母提取物1%、氯化钠0.25%、腺嘌呤100mg/L和鸟嘌呤100mg/L(基于1升蒸馏水)

[0141] <XMP烧瓶生产培养基(主培养基)>

[0142] 葡萄糖40g/L、硫酸镁10g/L、氯化钙100mg/L、硫酸铁20mg/L、硫酸锰10mg/L、硫酸锌10mg/L、硫酸铜0.8mg/L、组氨酸20mg/L、胱氨酸15mg/L、β-丙氨酸15mg/L、生物素100μg/L、硫胺素5mg/L、腺嘌呤50mg/L、鸟嘌呤25mg/L、烟酸15mg/L并且pH 7.0

[0143] <XMP烧瓶生产培养基(单独灭菌培养基)>

[0144] 磷酸二氢钾18g/L、磷酸一氢钾42g/L、尿素7g/L和硫酸铵5g/L

[0145] 将实验重复3次,并且其分析结果的平均值呈现于以下表1中。

[0146] [表1]

[0147] XMP生产能力的比较

[0148]

菌株	XMP浓度(g/L)	XMP浓度的增加率(%)
CJX1664	4.72	-
CJX1664_crtB2_A133V	5.40	14.4

[0149] 如表1所呈现,与对照组相比,CJX1664_crtB2_A133V菌株表现出XMP生产能力增加了14.4%。

[0150] CJX1664_crtB2_A133V命名为CN02-2071,在2020年12月2日保藏在韩国微生物培养中心(根据布达佩斯条约的保藏机构),登录号为KCCM12877P。

[0151] 根据上文描述,本公开所属技术领域的技术人员将理解,在不改变技术精神或基本特征的情况下,可以以其他具体形式实施本公开。就这一点而言,应当理解,上文所述的实施例在所有方面都是示例性的,而不是限制性的。本公开的范围应被解释为包括来源于以下将描述的权利要求的含义和范围而不是上文所述的详细描述的所有变化或修改形式以及其等效概念。

国际承认用于专利程序的微生物保藏
布达佩斯条约

国 际 表 格

接收人：CJ 第一制糖株式会社
CJ 第一制糖中心
东湖路 330
首尔中区 100-400
韩国

由本页底部标示的
国际保藏单位
根据第 7 条第 1 款颁发的
收到原始保藏材料的证明

[0152]

I. 微生物鉴定	
保藏人提供的鉴定参考： **停滞棒杆菌 CN02-2071**	国际保藏单位授予的登录号： **KCCM12877P**
II. 科学描述和/或提议的分类学命名	
以上 I 中鉴定的微生物具有： □ 科学描述 □ 提议的分类学命名 （在适用时用叉号标记）	
III. 接收和受理	
I 中鉴定的微生物由本国际保藏机构于 2020 年 12 月 2 日（原始保藏日）[1] 收到并接受。	
IV. 收到转化请求	
本国际保藏机构已于 （原始保藏日）收到以上 I 中鉴定的微生物并于 （转化请求的接受日）收到将原始保藏转化为根据布达佩斯条约的保藏的请求。	
V. 国际保藏机构	
名称：韩国培养物保藏中心 地址：Yurim B/D 45，Hongjenae-2ga-gil 西大门区 首尔 03641 韩国	本国际保藏机构全权代表人签名 日期：2020 年 12 月 2 日

[1] 如果条款 6.4(d) 适用，该日期即是已经获得国际保藏机构状态的日期。

BP/4 表 （单页）

[0001] <110> CJ第一制糖株式会社
[0002] <120> 新型八氢番茄红素合成酶变体及使用其生产XMP或GMP的方法
[0003] <130> OPA21127
[0004] <150> KR 10-2021-0006131
[0005] <151> 2021-01-15
[0006] <160> 11
[0007] <170> KoPatentIn 3.0
[0008] <210> 1
[0009] <211> 283
[0010] <212> PRT
[0011] <213> 人工序列
[0012] <220>
[0013] <223> CRTB2_A133V
[0014] <400> 1
[0015] Met Asn Ser Pro Thr Arg Arg Ala Leu Ile Glu Thr Tyr Asp Lys Thr
[0016] 1 5 10 15
[0017] Ala Arg Arg Ala Ser Gly Lys Ile Ile Ser Gly Tyr Ser Ser Ser Phe
[0018] 20 25 30
[0019] Ser Leu Ala Thr Arg Leu Leu Ala Glu Pro Ala Arg Thr Asp Ile Arg
[0020] 35 40 45
[0021] His Leu Tyr Ala Val Val Arg Ile Ala Asp Glu Ile Val Asp Gly Thr
[0022] 50 55 60
[0023] Ala Leu Glu Val Gly Ala Asn Pro Ala Lys Leu Leu Asn Glu Tyr Glu
[0024] 65 70 75 80
[0025] Glu Gln Val Arg Lys Ala Ala Glu Gln Lys Phe His Thr Asp Pro Val
[0026] 85 90 95
[0027] Met His Ala Phe Ala Met Thr Ala Arg Arg Cys Asn Phe Lys Asp Glu
[0028] 100 105 110
[0029] His Leu Arg Ala Phe Phe Thr Ser Met Arg Arg Asp Leu Asp Ser Ser
[0030] 115 120 125
[0031] Ala His Asp Ala Val Ser Phe Asp Ser Tyr Val Tyr Gly Ser Ala Glu
[0032] 130 135 140
[0033] Val Ile Gly Leu Leu Cys Leu Asp Val Phe Leu Ala Asp Arg Lys Thr
[0034] 145 150 155 160
[0035] Gly Ala Lys Glu Arg Glu Gln Leu Asp Ala Gly Ala Arg His Leu Gly
[0036] 165 170 175
[0037] Ala Ala Phe Gln Lys Ile Asn Phe Leu Arg Asp Arg His Glu Asp Phe
[0038] 180 185 190
[0039] Glu Ile Leu Gly Arg Glu Tyr Phe Pro Glu Gly Thr Thr Leu Glu Ala
[0040] 195 200 205
[0041] Ile Leu Ser Asp Ile Arg Asp Asp Ile Glu Ile Ala Arg Thr Val Ile

[0042] 210 215 220

[0043] Pro Gly Leu Pro Leu Arg Ala Arg Ala Gly Val Leu Ala Ala Thr Gly

[0044] 225 230 235 240

[0045] Ile Phe Glu Glu Leu Thr Ser Gln Leu Ala Asp Ser Arg Ser Asp Lys

[0046] 245 250 255

[0047] Thr Arg Ile Ser Val Pro Arg Arg Lys Lys Leu Gly Ile Ala Val Arg

[0048] 260 265 270

[0049] Ala Ala Val Thr Ala Leu Lys Gly Asn Asn Lys

[0050] 275 280

[0051] <210> 2

[0052] <211> 852

[0053] <212> DNA

[0054] <213> 人工序列

[0055] <220>

[0056] <223> crtB2_C398T

[0057] <400> 2

[0058] atgaattcac ccacgaggcg cgcgctgatt gagacctacg ataagactgc tcgtcgtgcg 60

[0059] tctggcaaaa ttatctcggg atattcctcc agtttctcgc ttgcaacgcg cctgctagct 120

[0060] gaacccgccc gaacagatat ccgccatctt tacgcagttg tccgtatagc ggacgaaatt 180

[0061] gtcgatggca ctgcacttga agtcggcgcc aatccggcca agctactcaa cgagtacgag 240

[0062] gaacaggttc gaaaggcagc agaacaaaaa tttcatacgg accctgtcat gcatgctttt 300

[0063] gccatgactg cacggcgctg caatttcaaa gacgagcacc tcagagcatt tttcacatcc 360

[0064] atgcgccgcg acttggattc ctccgctcat gatgctgtta gctttgattc atatgtctac 420

[0065] ggctcagcgg aagtcattgg attgctgtgc ctggatgtgt ttctcgcaga tagaaaaact 480

[0066] ggcgcgaagg aaagagaaca gctcgatgca ggagcaagac accttggtgc cgcgtttcag 540

[0067] aagatcaatt tccttagaga ccggcacgaa gattttgaaa tcctaggaag agagtacttt 600

[0068] cccgagggaa ctactctcga agctattctg tccgatatcc gagatgacat cgaaatcgct 660

[0069] cggaccgtca tcccggggct tcctttgaga gctcgagcag gagtattagc tgcgaccggc 720

[0070] atattcgaag aactaacatc acaattggca gatagccgat cggataagac gcgaattagt 780

[0071] gttccacgcc ggaagaagct cggaattgcg gtacgcgcag ctgtgacagc attgaaaggg 840

[0072] aacaataagt ga 852

[0073] <210> 3

[0074] <211> 283

[0075] <212> PRT

[0076] <213> 人工序列

[0077] <220>

[0078] <223> CRTB2

[0079] <400> 3

[0080] Met Asn Ser Pro Thr Arg Arg Ala Leu Ile Glu Thr Tyr Asp Lys Thr

[0081] 1 5 10 15

[0082] Ala Arg Arg Ala Ser Gly Lys Ile Ile Ser Gly Tyr Ser Ser Ser Phe

[0083] 20 25 30

[0084] Ser Leu Ala Thr Arg Leu Leu Ala Glu Pro Ala Arg Thr Asp Ile Arg
[0085] 35 40 45
[0086] His Leu Tyr Ala Val Val Arg Ile Ala Asp Glu Ile Val Asp Gly Thr
[0087] 50 55 60
[0088] Ala Leu Glu Val Gly Ala Asn Pro Ala Lys Leu Leu Asn Glu Tyr Glu
[0089] 65 70 75 80
[0090] Glu Gln Val Arg Lys Ala Ala Glu Gln Lys Phe His Thr Asp Pro Val
[0091] 85 90 95
[0092] Met His Ala Phe Ala Met Thr Ala Arg Arg Cys Asn Phe Lys Asp Glu
[0093] 100 105 110
[0094] His Leu Arg Ala Phe Phe Thr Ser Met Arg Arg Asp Leu Asp Ser Ser
[0095] 115 120 125
[0096] Ala His Asp Ala Ala Ser Phe Asp Ser Tyr Val Tyr Gly Ser Ala Glu
[0097] 130 135 140
[0098] Val Ile Gly Leu Leu Cys Leu Asp Val Phe Leu Ala Asp Arg Lys Thr
[0099] 145 150 155 160
[0100] Gly Ala Lys Glu Arg Glu Gln Leu Asp Ala Gly Ala Arg His Leu Gly
[0101] 165 170 175
[0102] Ala Ala Phe Gln Lys Ile Asn Phe Leu Arg Asp Arg His Glu Asp Phe
[0103] 180 185 190
[0104] Glu Ile Leu Gly Arg Glu Tyr Phe Pro Glu Gly Thr Thr Leu Glu Ala
[0105] 195 200 205
[0106] Ile Leu Ser Asp Ile Arg Asp Asp Ile Glu Ile Ala Arg Thr Val Ile
[0107] 210 215 220
[0108] Pro Gly Leu Pro Leu Arg Ala Arg Ala Gly Val Leu Ala Ala Thr Gly
[0109] 225 230 235 240
[0110] Ile Phe Glu Glu Leu Thr Ser Gln Leu Ala Asp Ser Arg Ser Asp Lys
[0111] 245 250 255
[0112] Thr Arg Ile Ser Val Pro Arg Arg Lys Lys Leu Gly Ile Ala Val Arg
[0113] 260 265 270
[0114] Ala Ala Val Thr Ala Leu Lys Gly Asn Asn Lys
[0115] 275 280
[0116] <210> 4
[0117] <211> 852
[0118] <212> DNA
[0119] <213> 人工序列
[0120] <220>
[0121] <223> crtB2
[0122] <400> 4
[0123] atgaattcac ccacgaggcg cgcgctgatt gagacctacg ataagactgc tcgtcgtgcg 60
[0124] tctggcaaaa ttatctcggg atattcctcc agtttctcgc ttgcaacgcg cctgctagct 120
[0125] gaacccgccc gaacagatat ccgccatctt tacgcagttg tccgtatagc ggacgaaatt 180

[0126] gtcgatggca ctgcacttga agtcggcgcc aatccggcca agctactcaa cgagtacgag 240

[0127] gaacaggttc gaaaggcagc agaacaaaaa tttcatacgg accctgtcat gcatgctttt 300

[0128] gccatgactg cacggcgctg caatttcaaa gacgagcacc tcagagcatt tttcacatcc 360

[0129] atgcgccgcg acttggattc ctccgctcat gatgctgcta gctttgattc atatgtctac 420

[0130] ggctcagcgg aagtcattgg attgctgtgc ctggatgtgt ttctcgcaga tagaaaaact 480

[0131] ggcgcgaagg aaagagaaca gctcgatgca ggagcaagac accttggtgc cgcgtttcag 540

[0132] aagatcaatt tccttagaga ccggcacgaa gattttgaaa tcctaggaag agagtacttt 600

[0133] cccgagggaa ctactctcga agctattctg tccgatatcc gagatgacat cgaaatcgct 660

[0134] cggaccgtca tcccggggct tcctttgaga gctcgagcag gagtattagc tgcgaccggc 720

[0135] atattcgaag aactaacatc acaattggca gatagccgat cggataagac gcgaattagt 780

[0136] gttccacgcc ggaagaagct cggaattgcg gtacgcgcag ctgtgacagc attgaaaggg 840

[0137] aacaataagt ga 852

[0138] <210> 5

[0139] <211> 35

[0140] <212> DNA

[0141] <213> 人工序列

[0142] <220>

[0143] <223> 引物

[0144] <400> 5

[0145] tcgagctcgg tacccgtgag gatgcctaac cgaaa 35

[0146] <210> 6

[0147] <211> 41

[0148] <212> DNA

[0149] <213> 人工序列

[0150] <220>

[0151] <223> 引物

[0152] <400> 6

[0153] catatgaatc aaagctaaca gcatcatgag cggaggaatc c 41

[0154] <210> 7

[0155] <211> 41

[0156] <212> DNA

[0157] <213> 人工序列

[0158] <220>

[0159] <223> 引物

[0160] <400> 7

[0161] ggattcctcc gctcatgatg ctgttagctt tgattcatat g 41

[0162] <210> 8

[0163] <211> 35

[0164] <212> DNA

[0165] <213> 人工序列

[0166] <220>

[0167] <223> 引物

[0168] <400> 8

[0169] ctctagagga tcccctgttc cctttcaatg ctgtc 35

[0170] <210> 9

[0171] <211> 20

[0172] <212> DNA

[0173] <213> 人工序列

[0174] <220>

[0175] <223> 引物-5F

[0176] <400> 9

[0177] gtgaggatgc ctaaccgaaa 20

[0178] <210> 10

[0179] <211> 20

[0180] <212> DNA

[0181] <213> 人工序列

[0182] <220>

[0183] <223> 引物

[0184] <400> 10

[0185] tgttcccttt caatgctgtc 20

[0186] <210> 11

[0187] <211> 5803

[0188] <212> DNA

[0189] <213> 人工序列

[0190] <220>

[0191] <223> pDCM2

[0192] <400> 11

[0193] gttcgcttgc tgtccataaa accgcccagt ctagctatcg ccatgtaagc ccactgcaag 60

[0194] ctacctgctt tctctttgcg cttgcgtttt cccttgtcca gatagcccag tagctgacat 120

[0195] tcatccgggg tcagcaccgt ttctgcggac tggctttcta cgtgttccgc ttcctttagc 180

[0196] agcccttgcg ccctgagtgc ttgcggcagc gtgaagctag cttttatcgc cattcgccat 240

[0197] tcaggctgcg caactgttgg gaagggcgat cggtgcgggc ctcttcgcta ttacgccagc 300

[0198] tggcgaaagg gggatgtgct gcaaggcgat taagttgggt aacgccaggg ttttcccagt 360

[0199] cacgacgttg taaaacgacg gccagtgaat tcgagctcgg tacccgggga tcctctagag 420

[0200] tcgacctgca ggcatgcaag cttggcgtaa tcatggtcat agctgtttcc tgtgtgaaat 480

[0201] tgttatccgc tcacaattcc acacaacata cgagccggaa gcataaagtg taaagcctgg 540

[0202] ggtgcctaat gagtgagcta actcacatta attgcgttgc gctcactgcc cgctttccag 600

[0203] tcgggaaacc tgtcgtgcca gctgcattaa tgaatcggcc aacgcgcggg gagaggcggt 660

[0204] ttgcgtattg ggcgctcttc cgcttcctcg ctcactgact cgctgcgctc ggtcgttcgg 720

[0205] ctgcggcgag cggtatcagc tcactcaaag gcggtaatac ggttatccac agaatcaggg 780

[0206] gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag 840

[0207] gccgcgttgc tggcgttttt ccataggctc cgccccccctg acgagcatca caaaaatcga 900

[0208] cgctcaagtc agaggtggcg aaacccgaca ggactataaa gataccaggc gtttcccccct 960

[0209] ggaagctccc tcgtgcgctc tcctgttccg accctgccgc ttaccggata cctgtccgcc 1020

[0210] tttctccctt cgggaagcgt ggcgctttct caatgctcac gctgtaggta tctcagttcg 1080
[0211] gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac cccccgttca gcccgaccgc 1140
[0212] tgcgccttat ccggtaacta tcgtcttgag tccaacccgg taagacacga cttatcgcca 1200
[0213] ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag 1260
[0214] ttcttgaagt ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgct 1320
[0215] ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg caaacaaacc 1380
[0216] accgctggta gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag aaaaaaagga 1440
[0217] tctcaagaag atcctttgat cttttctacg gggtctgacg ctcagtggaa cgaaaactca 1500
[0218] cgttaaggga ttttggtcat gagattatca aaaaggatct tcacctagat ccttttgggg 1560
[0219] tgggcgaaga actccagcat gagatccccg cgctggagga tcatccagcc ctgatagaaa 1620
[0220] cagaagccac tggagcacct caaaaacacc atcatacact aaatcagtaa gttggcagca 1680
[0221] tcacccgacg cactttgcgc cgaataaata cctgtgacgg aagatcactt cgcagaataa 1740
[0222] ataaatcctg gtgtccctgt tgataccggg aagccctggg ccaacttttg gcgaaaatga 1800
[0223] gacgttgatc ggcacgtaag aggttccaac tttcaccata atgaaataag atcactaccg 1860
[0224] ggcgtatttt ttgagttatc gagattttca ggagctgata gaaacagaag ccactggagc 1920
[0225] acctcaaaaa caccatcata cactaaatca gtaagttggc agcatcaccc gacgcacttt 1980
[0226] gcgccgaata aatacctgtg acggaagatc acttcgcaga ataaataaat cctggtgtcc 2040
[0227] ctgttgatac cgggaagccc tgggccaact tttggcgaaa atgagacgtt gatcggcacg 2100
[0228] taagaggttc caactttcac cataatgaaa taagatcact accgggcgta ttttttgagt 2160
[0229] tatcgagatt ttcaggagct ctttggcatc gtctctcgcc tgtcccctca gttcagtaat 2220
[0230] ttcctgcatt tgcctgtttc cagtcggtag atattccaca aaacagcagg gaagcagcgc 2280
[0231] ttttccgctg cataaccctg cttcggggtc attatagcga ttttttcggt atatccatcc 2340
[0232] tttttcgcac gatatacagg attttgccaa agggttcgtg tagactttcc ttggtgtatc 2400
[0233] caacggcgtc agccgggcag gataggtgaa gtaggcccac ccgcgagcgg gtgttccttc 2460
[0234] ttcactgtcc cttattcgca cctggcggtg ctcaacggga atcctgctct gcgaggctgg 2520
[0235] ccggctaccg ccggcgtaac agatgagggc aagcggatgg ctgatgaaac caagccaacc 2580
[0236] aggaagggca gcccacctat caaggtgtac tgccttccag acgaacgaag agcgattgag 2640
[0237] gaaaaggcgg cggcggccgg catgagcctg tcggcctacc tgctggccgt cggccagggc 2700
[0238] tacaaaatca cgggcgtcgt ggactatgag cacgtccgcg agggcgtccc ggaaaacgat 2760
[0239] tccgaagccc aacctttcat agaaggcggc ggtggaatcg aaatctcgtg atggcaggtt 2820
[0240] gggcgtcgct tggtcggtca tttcgaaaaa ggttaggaat acggttagcc atttgcctgc 2880
[0241] ttttatatag ttcantatgg gattcacctt tatgttgata agaaataaaa gaaaatgcca 2940
[0242] ataggatatc ggcattttct tttgcgtttt tatttgttaa ctgttaattg tccttgttca 3000
[0243] aggatgctgt ctttgacaac agatgttttc ttgcctttga tgttcagcag gaagctcggc 3060
[0244] gcaaacgttg attgtttgtc tgcgtagaat cctctgtttg tcatatagct tgtaatcacg 3120
[0245] acattgtttc ctttcgcttg aggtacagcg aagtgtgagt aagtaaaggt tacatcgtta 3180
[0246] ggcggatcaa gatccatttt taacacaagg ccagttttgt tcagcggctt gtatgggcca 3240
[0247] gttaaagaat tagaaacata accaagcatg taaatatcgt tagacgtaat gccgtcaatc 3300
[0248] gtcatttttg atccgcggga gtcagtgaac aggtaccatt tgccgttcat tttaaagacg 3360
[0249] ttcgcgcgtt caatttcatc tgttactgtg ttagatgcaa tcagcggttt catcactttt 3420
[0250] ttcagtgtgt aatcatcgtt tagctcaatc ataccgagag cgccgtttgc taactcagcc 3480
[0251] gtgcgttttt tatcgctttg cagaagtttt tgactttctt gacggaagaa tgatgtgctt 3540

[0252] ttgccatagt atgctttgtt aaataaagat tcttcgcctt ggtagccatc ttcagttcca 3600
[0253] gtgtttgctt caaatactaa gtatttgtgg cctttatctt ctacgtagtg aggatctctc 3660
[0254] agcgtatggt tgtcgcctga gctgtagttg ccttcatcga tgaactgctg tacattttga 3720
[0255] tacgtttttc cgtcaccgtc aaagattgat ttataatcct ctacaccgtt gatgttcaaa 3780
[0256] gagctgtctg atgctgatac gttaacttgt gcagttgtca gtgtttgttt gccgtaatgt 3840
[0257] ttaccggaga aatcagtgta gaataaacgg atttttccgt cagatgtaaa tgtggctgaa 3900
[0258] cctgaccatt cttgtgtttg gtcttttagg atagaatcat ttgcatcgaa tttgtcgctg 3960
[0259] tctttaaaga cgcggccagc gtttttccag ctgtcaatag aagtttcgcc gacttttttga 4020
[0260] tagaacatgt aaatcgatgt gtcatccgca tttttaggat ctccggctaa tgcaaagacg 4080
[0261] atgtggtagc cgtgatagtt tgcgacagtg ccgtcagcgt tttgtaatgg ccagctgtcc 4140
[0262] caaacgtcca ggccttttgc agaagagata tttttaattg tggacgaatc aaattcagaa 4200
[0263] acttgatatt tttcattttt ttgctgttca gggatttgca gcatatcatg gcgtgtaata 4260
[0264] tgggaaatgc cgtatgtttc cttatatggc ttttggttcg tttctttcgc aaacgcttga 4320
[0265] gttgcgcctc ctgccagcag tgcggtagta aaggttaata ctgttgcttg ttttgcaaac 4380
[0266] tttttgatgt tcatcgttca tgtctccttt tttatgtact gtgttagcgg tctgcttctt 4440
[0267] ccagccctcc tgtttgaaga tggcaagtta gttacgcaca ataaaaaaag acctaaaata 4500
[0268] tgtaaggggt gacgccaaag tatacacttt gcccttaca cattttaggt cttgcctgct 4560
[0269] ttatcagtaa caaacccgcg cgatttactt ttcgacctca ttctattaga ctctcgtttg 4620
[0270] gattgcaact ggtctatttt cctcttttgt ttgatagaaa atcataaaag gatttgcaga 4680
[0271] ctacgggcct aaagaactaa aaaatctatc tgtttctttt cattctctgt atttttata 4740
[0272] gtttctgttg catgggcata aagttgcctt tttaatcaca attcagaaaa tatcataata 4800
[0273] tctcatttca ctaaataata gtgaacggca ggtatatgtg atgggttaaa aaggatcacc 4860
[0274] ccagagtccc gctcagaaga actcgtcaag aaggcgatag aaggcgatgc gctgcgaatc 4920
[0275] gggagcggcg ataccgtaaa gcacgaggaa gcggtcagcc cattcgccgc caagctcttc 4980
[0276] agcaatatca cgggtagcca acgctatgtc ctgatagcgg tccgccacac ccagccggcc 5040
[0277] acagtcgatg aatccagaaa agcggccatt ttccaccatg atattcggca agcaggcatc 5100
[0278] gccatgggtc acgacgagat cctcgccgtc gggcatccgc gccttgagcc tggcgaacag 5160
[0279] ttcggctggc gcgagcccct gatgctcttc gtccagatca tcctgatcga caagaccggc 5220
[0280] ttccatccga gtacgtgctc gctcgatgcg atgtttcgct tggtggtcga atgggcaggt 5280
[0281] agccggatca agcgtatgca gccgccgcat tgcatcagcc atgatggata ctttctcggc 5340
[0282] aggagcaagg tgagatgaca ggagatcctg ccccggcact tcgcccaata gcagccagtc 5400
[0283] ccttcccgct tcagtgacaa cgtcgagaca gctgcgcaag gaacgcccgt cgtggccagc 5460
[0284] cacgatagcc gcgctgcctc gtcttggagt tcattcaggg caccggacag gtcggtcttg 5520
[0285] acaaaaagaa ccgggcgccc ctgcgctgac agccggaaca cggcggcatc agagcagccg 5580
[0286] attgtctgtt gtgcccagtc atagccgaat agcctctcca cccaagcggc cggagaacct 5640
[0287] gcgtgcaatc catcttgttc aatcatgcga aacgatcctc atcctgtctc ttgatcagat 5700
[0288] cttgatcccc tgcgccatca gatccttggc ggcaagaaag ccatccagtt tactttgcag 5760
[0289] ggcttcccaa ccttaccaga gggcgcccca gctggcaatt ccg 5803

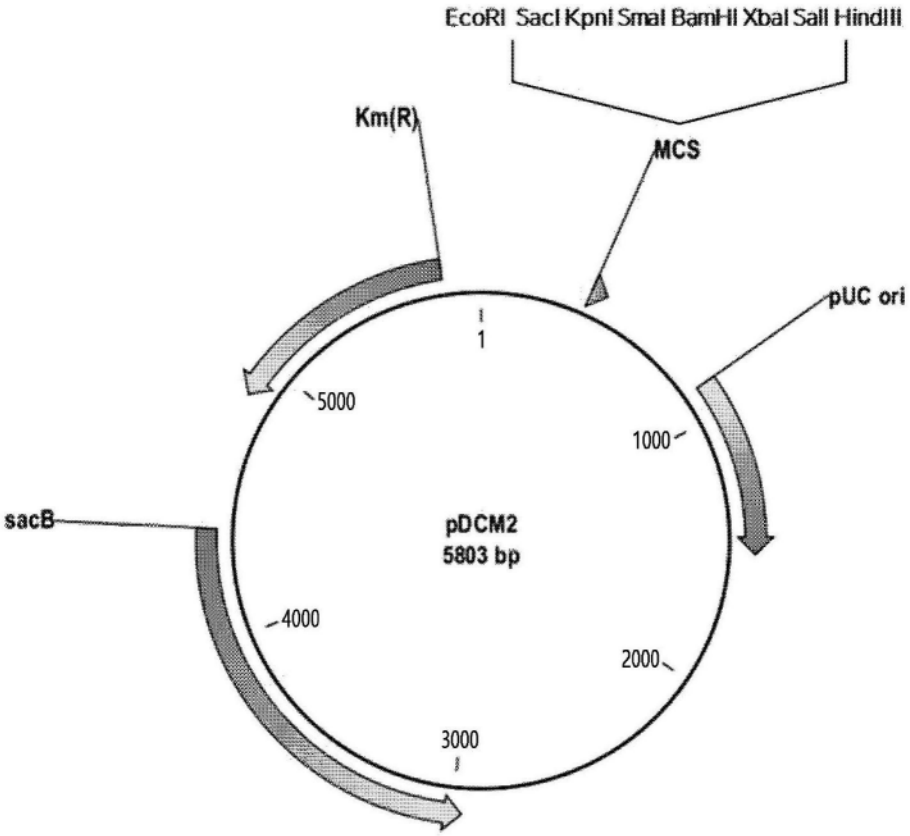

EcoRI SacI KpnI SmaI BamHI XbaI SalI HindIII
Km(R)
MCS
pUC ori
1
5000
1000
sacB
pDCM2
5803 bp
4000
2000
3000

图1