



(12)发明专利

(10)授权公告号 CN 104080918 B

(45)授权公告日 2017.02.22

(21)申请号 201380006973.X
 (22)申请日 2013.02.19
 (65)同一申请的已公布的文献号
 申请公布号 CN 104080918 A
 (43)申请公布日 2014.10.01
 (30)优先权数据
 2012-039129 2012.02.24 JP
 (85)PCT国际申请进入国家阶段日
 2014.07.28
 (86)PCT国际申请的申请数据
 PCT/JP2013/053958 2013.02.19
 (87)PCT国际申请的公布数据
 W02013/125509 JA 2013.08.29
 (73)专利权人 旭化成株式会社
 地址 日本东京都
 (72)发明人 小西一诚 今津晋一
 (74)专利代理机构 北京三友知识产权代理有限公司 11127
 代理人 丁香兰 李洋

(51)Int.Cl.
C12P 7/58(2006.01)
C12N 1/15(2006.01)
C12N 1/19(2006.01)
C12N 1/21(2006.01)
C12N 5/10(2006.01)
C12N 15/09(2006.01)
 (56)对比文件
 CN 102066552 A,2011.05.18,
 Moon TS等.Production of glucaric acid
 from a synthetic pathway in recombinant
 Escherichia coil.《Appl Environ
 Microbiol.》.2009,第75卷(第3期),
 Tae SeokMoona等.Use of modular,
 syntheticscaffoldsforimprovedproductionof
 glucaricacid in engineered E.coli.
 《Metabolic Engineering》.2010,(第12期),第
 298-305页.

审查员 刘新蕾

权利要求书1页 说明书12页
序列表12页 附图4页

(54)发明名称

葡萄糖二酸的制造方法

(57)摘要

本发明的课题在于提供具有显著改善的葡萄糖二酸生产能力的转化体的制作方法和利用该转化体有效生产葡萄糖二酸的方法。在本发明的葡萄糖二酸的制造方法中,向保有肌醇-1-磷酸合成酶基因、肌醇单磷酸酶基因、肌肉肌醇加氧酶基因和糖醛酸脱氢酶基因的转化体中导入对功能性肌醇单磷酸酶的过剩生产或肌醇单磷酸酶的活化进行诱导的基因重组或变异。

1. 一种葡萄糖二酸的制造方法,其包括下述工序:

1) 准备转化体的工序,该转化体保有肌醇-1-磷酸合成酶基因、肌醇单磷酸酶基因、肌肉肌醇加氧酶基因和糖醛酸脱氢酶基因,来源于大肠杆菌,并具有对该转化体内的功能性肌醇单磷酸酶的过剩生产进行诱导的基因重组,所述转化体的特征在于保有包含编码肌醇-1-磷酸合成酶的核酸的表达盒、包含编码肌醇单磷酸酶基因的核酸的表达盒、包含编码肌肉肌醇加氧酶的核酸的表达盒和包含编码糖醛酸脱氢酶的核酸的表达盒的4种表达盒;

2) 在适于所述转化体的生长和/或维持的条件下,使该转化体与能够由该转化体转换为葡萄糖二酸的碳源进行接触的工序;以及

3) 从在所述2)中得到的培养物中分离出葡萄糖二酸或者葡萄糖二酸盐的工序。

2. 如权利要求1所述的制造方法,其中,所述碳源包括可在所述转化体内转换为葡萄糖-6-磷酸的化合物。

3. 如权利要求2所述的制造方法,其中,所述碳源为选自由D-葡萄糖、蔗糖、寡糖、多糖、淀粉、纤维素、米糠、废糖蜜以及含有D-葡萄糖的生物物质组成的组中的1种以上。

4. 一种转化体,该转化体保有肌醇-1-磷酸合成酶基因、肌醇单磷酸酶基因、肌肉肌醇加氧酶基因和糖醛酸脱氢酶基因,来源于大肠杆菌,并且具有对该转化体内的功能性肌醇单磷酸酶的过剩生产进行诱导的基因重组,所述转化体的特征在于保有包含编码肌醇-1-磷酸合成酶的核酸的表达盒、包含编码肌醇单磷酸酶基因的核酸的表达盒、包含编码肌肉肌醇加氧酶的核酸的表达盒和包含编码糖醛酸脱氢酶的核酸的表达盒的4种表达盒。

葡萄糖二酸的制造方法

技术领域

[0001] 本发明涉及基因重组技术在葡萄糖二酸的制造中的应用。

背景技术

[0002] 葡萄糖二酸(四羟基己二酸)是在植物和哺乳动物中早已发现的化合物。

[0003] 近年来,在美国国家再生能源研究所关于需要由生物物质制作的高附加值化学品的报告(非专利文献1)中举出了葡萄糖二酸作为前12名以内的化合物。在该报告中作为可以以葡萄糖二酸为原料制备出的葡萄糖二酸衍生物还示例出了葡糖二酸- γ -内酯、葡糖二酸- δ -内酯或葡糖二酸二内酯等内酯类(可期待其作为溶剂的用途);多羟基聚酰胺类(可期待其作为新型尼龙的用途)等。此外,该报告中还推测可利用淀粉的硝酸氧化反应或碱性漂白剂存在下的催化氧化反应作为已知的葡萄糖二酸制造方法。

[0004] 进一步地,最近在专利文献1中公开了可进行葡萄糖二酸的生物合成的转化体。即,在该专利文献中,分别编码肌肉肌醇-1-磷酸合成酶(Ino1)、肌肉肌醇加氧酶(MIOX)和糖醛酸脱氢酶(udh)的3个基因被转染至大肠杆菌宿主中。这样得到的转化体在培养基内以0.72g/L~1.13g/L的浓度生产出了葡萄糖二酸。但是,据专利文献1的发明人推测,在该专利文献的转化体中无需导入肌醇单磷酸酶(suhB)基因。

[0005] 即,在以葡萄糖为底物进行葡萄糖二酸的生物合成的通路中,理论上需要下述5种活性:

[0006] 活性1:由适当的碳源生成葡萄糖-6-磷酸的活性;

[0007] 活性2:将葡萄糖-6-磷酸转换为肌肉肌醇-1-磷酸的活性、即肌醇-1-磷酸合成酶活性;

[0008] 活性3:将肌肉肌醇-1-磷酸转换为肌肉肌醇的活性、即以肌肉肌醇-1-磷酸为底物的磷酸酶活性;

[0009] 活性4:将肌肉肌醇转换为葡萄糖醛酸的活性、即肌肉肌醇加氧酶活性;和

[0010] 活性5:将葡萄糖醛酸转换为葡萄糖二酸的活性、即糖醛酸脱氢酶活性。

[0011] 但是,实际上,由于活性1的产物葡萄糖-6-磷酸为原核微生物普遍生成的代谢中间体,因而无需对原核微生物赋予该活性。

[0012] 此外,关于活性3,已知也有不少微生物株表达内源性肌醇单磷酸酶、或者具有能够以肌肉肌醇-1-磷酸为底物的通用单磷酸酶活性。因而,在专利文献1的转化体中没有导入肌醇单磷酸酶基因。

[0013] 并且,在专利文献1中,基于所制作的转化体的代谢分析,得出在用于进行葡萄糖二酸的生物合成的转化体中无需导入肌醇单磷酸酶基因的结论。即,在专利文献1中记载了下述内容:“应当注意的是,我们并没有过量表达suhB基因或同源磷酸酶。但是,培养产物中未检测到肌肉肌醇-1-磷酸,而另一方面,却有肌肉肌醇的积聚。因此,我们得出结论:磷酸酶活性并不限制通过该通路的代谢的流量(flux)。”(第33页、第2~5行)。

[0014] 因而并不存在将肌醇单磷酸酶基因导入到用于葡萄糖二酸的生物合成的转化体

中的明确动机。

[0015] 现有技术文献

[0016] 专利文献

[0017] 专利文献1:W02009/145838号小册子

[0018] 非专利文献

[0019] 非专利文献1:Top Value Added Chemicals from Biomass Volume I-Results of Screening for Potential Candidates from Sugars and Synthesis Gas””、<http://www1.eere.energy.gov/biomass/pdfs/35523.pdf>、T.Werpy和G.Peterson编、2004年8月发行

发明内容

[0020] 发明所要解决的课题

[0021] 本发明所要解决的课题在于具有显著改善的葡萄糖二酸生产能力的转化体的制造及其应用。

[0022] 解决课题的手段

[0023] 如上所述,即使是在公开了能够生物合成葡萄糖二酸的转化体的专利文献1中,也并未向该转化体中导入肌醇单磷酸酶基因,并且该活性也并未引起特别的关注。

[0024] 但是,与专利文献1的预想相反,本发明人发现,肌醇单磷酸酶活性在用于葡萄糖二酸生物合成的转化体中具有重要作用。特别值得吃惊的是,通过增强肌醇单磷酸酶活性,这样的转化体的葡萄糖二酸生产能力提高了数十~百倍。

[0025] 因而,本发明的第一方面在于:

[0026] (1) 一种葡萄糖二酸的制造方法,其包括下述工序:

[0027] 1) 准备转化体的工序,该转化体保有肌醇-1-磷酸合成酶基因、肌醇单磷酸酶基因、肌肉肌醇加氧酶基因和糖醛酸脱氢酶基因,并具有对该转化体内的功能性肌醇单磷酸酶的过剩生产或肌醇单磷酸酶的活化进行诱导的基因重组或变异;

[0028] 2) 在适于上述转化体的生长和/或维持的条件下,使该转化体与能够由该转化体转换为葡萄糖二酸的碳源进行接触的工序;以及

[0029] 3) 从在上述2)中得到的培养物中分离出葡萄糖二酸或者葡萄糖二酸盐的工序。

[0030] 更特定地说,上述制造方法为使用转化体的葡萄糖二酸的制造方法,该转化体保有肌醇-1-磷酸合成酶基因、肌醇单磷酸酶基因、肌肉肌醇加氧酶基因和糖醛酸脱氢酶基因,其特征在于,上述转化体具有对功能性肌醇单磷酸酶的过剩生产或肌醇单磷酸酶的活化进行诱导的基因重组或变异。

[0031] 本发明葡萄糖二酸的发酵生产中,使用下述碳源作为培养基材是适宜的,该基材含有适于生成作为肌醇-1-磷酸合成酶(上述活性2)的底物的葡萄糖-6-磷酸的化合物。因而,本发明的适宜方式为:

[0032] (2) 如上述(1)所述的制造方法,其中,上述碳源包括可在上述转化体内转换为葡萄糖-6-磷酸的化合物;和

[0033] (3) 如上述(2)所述的制造方法,其中,上述碳源为选自由D-葡萄糖、蔗糖、寡糖、多糖、淀粉、纤维素、米糠、废糖蜜以及含有D-葡萄糖的生物物质组成的组中的1种以上。

[0034] 对于以大肠杆菌为代表的原核微生物来说,由于其迅速的生长能力以及发酵管理的容易性,因而从工业发酵生产的方面考虑是极富魅力的,同时从应用基因重组技术时的实际成绩、可靠的安全性的方面考虑也具有优点。此外,对于不具有从葡萄糖经肌肉肌醇来生物合成葡萄糖二酸的通路的多数原核微生物来说,通过使用与基因重组技术联合的合成生物学方法,具有容易调节葡萄糖二酸生产率的优点。特别是大肠杆菌等原核微生物宿主不具有对作为葡萄糖二酸生物合成通路中间体的肌肉肌醇的同化能力(分解能力),因而更容易应用合成生物学的方法。因而,本发明的适宜方式包括:

[0035] (4)如上述(1)至(3)的任一项所述的制造方法,其特征在于,上述转化体来源于不具有肌肉肌醇同化能力的微生物;和

[0036] (5)如上述(1)至(4)的任一项所述的制造方法,其中,上述转化体来源于选自由大肠杆菌、芽胞杆菌属细菌、棒状杆菌属细菌、发酵单胞菌属细菌组成的组中的细菌。

[0037] 无论宿主微生物是否具有内源性肌醇单磷酸酶活性,均可通过肌醇单磷酸酶在该细胞内的过剩生产来增强该细胞的肌醇单磷酸酶活性。可应用各种公知的技术使肌醇单磷酸酶在细胞中过剩生产。因而,本发明包括下述方式:

[0038] (6)如上述(1)至(5)的任一项所述的制造方法,其中,上述肌醇单磷酸酶的过剩生产是通过对上述转化体进行以下操作而被诱导的:

[0039] a) 导入外来肌醇单磷酸酶基因、

[0040] b) 增加内源性肌醇单磷酸酶基因的拷贝数、

[0041] c) 在内源性肌醇单磷酸酶基因的调整区域导入变异、

[0042] d) 利用高表达诱导性外来调整区域来置换内源性肌醇单磷酸酶基因的调整区域、或者

[0043] e) 使内源性肌醇单磷酸酶基因的调整区域缺失;和

[0044] (7)如上述(6)所述的制造方法,其中,上述肌醇单磷酸酶的过剩生产是通过对上述转化体导入外来肌醇单磷酸酶基因而被诱导的。

[0045] 此外,在宿主细胞具有内源性肌醇单磷酸酶基因的情况下,也可利用下述方式来增强该细胞的肌醇单磷酸酶活性。

[0046] (8)如上述(1)至(5)的任一项所述的制造方法,其中,上述肌醇单磷酸酶的活化是通过对上述转化体进行以下操作而被诱导的:

[0047] f) 在内源性肌醇单磷酸酶基因中导入变异、

[0048] g) 置换内源性肌醇单磷酸酶基因的一部分或全部、

[0049] h) 使内源性肌醇单磷酸酶基因的一部分缺失、

[0050] i) 使降低肌醇单磷酸酶活性的其它蛋白质减少、

[0051] j) 使降低肌醇单磷酸酶活性的化合物的生成减少。

[0052] 此外,本发明还谋求用于上述葡萄糖二酸的制造方法的转化体。因而,本发明的第二方面在于:

[0053] (9)一种转化体,该转化体保有肌醇-1-磷酸合成酶基因、肌醇单磷酸酶基因、肌肉肌醇加氧酶基因和糖醛酸脱氢酶基因,其中,该转化体具有对该转化体内的功能性肌醇单磷酸酶的过剩生产或肌醇单磷酸酶的活化进行诱导的基因重组或变异。

[0054] 更特定地说,上述转化体为保有肌醇-1-磷酸合成酶基因、肌醇单磷酸酶基因、肌

肉肌醇加氧酶基因和糖醛酸脱氢酶基因的转化体,其特征在于,其具有对功能性肌醇单磷酸酶的过剩生产或肌醇单磷酸酶的活化进行诱导的基因重组或变异。

[0055] 此外,本发明第一方面中记载的方式也适用于本发明的第二方面。这些方式为:

[0056] (10) 如上述(9)所述的转化体,其特征在于,上述转化体来源于不具有肌肉肌醇同化能力的微生物;

[0057] (11) 如上述(9)或(10)所述的转化体,其中,上述转化体来源于选自由大肠杆菌、芽胞杆菌属细菌、棒状杆菌属细菌、发酵单胞菌属细菌组成的组中的细菌;

[0058] (12) 如上述(9)至(11)的任一项所述的转化体,其中,上述肌醇单磷酸酶的过剩生产是通过对上述转化体进行以下操作而被诱导的:

[0059] a) 导入外来肌醇单磷酸酶基因、

[0060] b) 增加内源性肌醇单磷酸酶基因的拷贝数、

[0061] c) 在内源性肌醇单磷酸酶基因的调整区域导入变异、

[0062] d) 利用高表达诱导性外来调整区域来置换内源性肌醇单磷酸酶基因的调整区域、或者

[0063] e) 使内源性肌醇单磷酸酶基因的调整区域缺失;

[0064] (13) 如上述(12)所述的转化体,其中,上述肌醇单磷酸酶的过剩生产是通过对上述转化体导入外来肌醇单磷酸酶基因而被诱导的;和

[0065] (14) 如上述(9)至(11)的任一项所述的转化体,其中,上述肌醇单磷酸酶的活化是通过对上述转化体进行以下操作而被诱导的:

[0066] f) 在内源性肌醇单磷酸酶基因中导入变异、

[0067] g) 置换内源性肌醇单磷酸酶基因的一部分或全部、

[0068] h) 使内源性肌醇单磷酸酶基因的一部分缺失、

[0069] i) 使降低肌醇单磷酸酶活性的其它蛋白质减少、

[0070] j) 使降低肌醇单磷酸酶活性的化合物的生成减少。

[0071] 发明的效果

[0072] 根据本发明,能够利用微生物培养技术实现葡萄糖二酸工业生产的效率化。

附图说明

[0073] 图1示出了IN01基因的编码区域(序列编号1)。

[0074] 图2示出了suhB基因的编码区域(序列编号3)。

[0075] 图3示出了miox基因的编码区域(序列编号5)。

[0076] 图4示出了udh基因的编码区域(序列编号7)。

具体实施方式

[0077] 本发明的课题通过增强保有肌醇-1-磷酸合成酶基因、肌醇单磷酸酶基因、肌肉肌醇加氧酶基因和糖醛酸脱氢酶基因的转化体中的肌醇单磷酸酶活性而得到解决。

[0078] 本发明的转化体可使用各种宿主微生物细胞来制作。特别是在以原核微生物为宿主时,由于能够在该宿主细胞内新构建葡萄糖二酸生物合成通路(即,无现有的内源性通路的影响),因而在合成生物学方法的应用中是极富魅力的。所示例出的原核微生物为埃希氏

菌、假单胞菌、芽胞杆菌、地芽胞杆菌、甲烷单胞菌、甲基芽孢杆菌、嗜甲基菌、精肱杆菌、甲基球菌、棒状杆菌、短杆菌、发酵单胞菌和李斯特氏菌属的细菌。在工业发酵生产中适宜的原核微生物的非限定性示例包括大肠杆菌、芽胞杆菌属细菌、棒状杆菌属细菌、发酵单胞菌属细菌。大肠杆菌由于其迅速的生长能力和发酵管理的容易性的原因，为特别优选的本发明的宿主微生物的示例。

[0079] 此外，能够作为本发明的宿主细胞使用的细胞株可以为通常含义下的野生型，或者也可以为营养要求性变异株、抗生素抗性变异株。进一步地，能够作为本发明的宿主细胞使用的细胞株可以已经被转化从而使之具有与上述变异相关的各种标记基因。这些变异或基因能够提供对于本发明转化体的制作·维持·管理有益的性质。优选可通过使用对于氯霉素、氨苄青霉素、卡那霉素、四环素等抗生素显示出耐性的菌株来简便地进行本发明葡萄糖二酸的生产。

[0080] 在面向合成生物学的本发明中，为了在宿主细胞中构建新的葡萄糖二酸生物合成通路，在宿主细胞不表达内源性肌醇-1-磷酸合成酶的情况下，导入外来肌醇-1-磷酸合成酶基因。需要说明的是，在本说明书中，使用“外来”或“外来性”这一术语所表示的含义是，在转化前的宿主微生物不具有要由本发明导入的基因的情况下；在实质上不表达由该基因编码的酶的情况下；以及在该酶的氨基酸序列由不同的基因编码、但在转化后不表达相匹配的内源性酶活性的情况下，向宿主中导入基于本发明的基因或核酸序列。

[0081] 肌醇-1-磷酸合成酶基因是已知的（例如，基因库登录号Nos. AB032073、AF056325、AF071103、AF078915、AF120146、AF207640、AF284065、BC111160、L23520、U32511），它们均可用于本发明的目的。特别是具有序列编号1所表示的编码区域核苷酸序列的肌醇-1-磷酸合成酶基因能够在本发明中适当地使用。但是，能够在本发明中使用的肌醇-1-磷酸合成酶基因并不限于上述基因，其可来源于其它生物、或者也可人工合成，只要在上述宿主微生物细胞内能够实质性地表达肌醇-1-磷酸合成酶活性即可。

[0082] 从而，对于能够用于本发明目的的肌醇-1-磷酸合成酶基因，只要在上述宿主微生物细胞内能够实质性地表达肌醇-1-磷酸合成酶活性，可以具有能够在自然界发生的全部变异、或者具有人工导入的变异和修饰。例如，已知在编码特定氨基酸的各种密码子中存在冗余密码子（redundancy）。因此，在本发明中也可利用最终可翻译成相同氨基酸的代替密码子。即，由于基因编码的简并，因而编码某一特定氨基酸时可使用2个以上的密码子，因此氨基酸序列可利用任意的1组类似的DNA低聚核苷酸进行编码得到。虽然该组中仅唯一的成员与天然型酶的基因序列相同，但即使是有错配的DNA低聚核苷酸，在适当的严谨条件下（例如，3xSSC、68℃杂交、2xSSC、0.1% SDS和68℃清洗）也能够与天然型序列杂交，能够对编码天然型序列的DNA进行鉴定、分离，进一步还可将这样的基因用于本发明中。特别是已知大部分生物优选使用特定密码子（最佳密码子）的子集（Gene, Vol. 105, pp. 61-72, 1991等），因而根据宿主微生物进行“密码子最佳化”也是有用的。

[0083] 并且，在本发明中，也可通过将肌醇-1-磷酸合成酶基因以“表达盒”的形式导入到宿主微生物细胞内来更稳定地得到高水平的肌醇-1-磷酸合成酶活性，这对本领域技术人员来说是可以理解的。在本说明书中，“表达盒”是指含有与表达对象的核酸或表达对象的基因功能性结合的对于转录和翻译进行调节的核酸序列的核苷酸。代表性地，本发明的表达盒以功能性结合的状态在编码序列的5'上游含有启动子序列、在3'下游含有终止子序

列、根据情况进一步含有通常的调节元件,在这样的情况下,表达对象的核酸或表达对象的基因被“可表达地导入”到宿主微生物中。

[0084] 关于启动子,不论是结构性启动子还是调节启动子,均被定义为使RNA聚合酶与DNA结合、引发RNA合成的DNA序列。强启动子是指高频引发mRNA合成的启动子,也适于用于本发明中。可根据该宿主细胞的性质等使用lac系、trp系、TAC或TRC系、噬菌体的主要操纵子和启动子区域、fd包覆蛋白质的调节区域、针对糖酵解系酶(例如,3-磷酸甘油酸激酶、甘油醛-3-磷酸脱氢酶)、谷氨酸脱羧酶A、丝氨酸羟甲基转移酶的启动子等。除了启动子和终止子序列以外,可作为其它调节元件的示例举出的序列为选择标记、扩增信号、复制起点等。适宜的调节序列例如记载于“Gene Expression Technology:Methods in Enzymology 185”,Academic Press (1990)中。

[0085] 上述说明的表达盒嵌入到例如由质粒、噬菌体、转座子、IS元件、噬粒、粘粒、或者线状或环状DNA等构成的载体中而被插入到宿主微生物中。优选质粒和噬菌体。这些载体在宿主微生物中可自体复制,也可通过染色体复制。适宜的质粒例如为:大肠杆菌的pLG338、pACYC184、pBR322、pUC18、pUC19、pKC30、pRep4、pHS1、pKK223-3、pDHE19.2、pHS2、pPLc236、pMBL24、pLG200、pUR290、pIN-III113-B1、 λ gt11或pBdCI;杆菌的pUB110、pC194或pBD214;棒状杆菌属的pSA77或pAJ667等。除它们以外,能够使用的质粒等还记载于“Cloning Vectors”,Elsevier,1985中。表达盒向载体中的导入可通过包括利用适当的限制酶切割、克隆化以及连接的惯用方法进行。

[0086] 在如上所述构建出本发明的具有表达盒的载体后,作为将该载体导入到宿主微生物中时能够适用的方法,例如使用共沉淀、原生质体融合、电穿孔、逆转录病毒转染等惯用的克隆化法和转染法。它们的示例记载于“分子生物学的最新实验计划(Current Protocols in Molecular Biology)”,F.Ausubel等,Publ.Wiley Interscience,New York,1997;或Sambrook等,“分子克隆:实验室手册”,第2版,Cold Spring Harbor Laboratory,Cold Spring Harbor Laboratory Press,Cold Spring Harbor,NY,1989中。

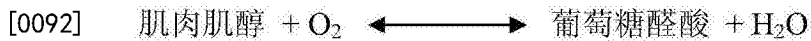
[0087] 本发明人令人吃惊地发现,在不具有内源性葡萄糖二酸生物合成通路的宿主微生物内导入葡萄糖二酸生物合成通路而得到的转化体中,肌醇单磷酸酶活性具有重大作用。如上所述,在迄今为止的研究中,均未对肌醇单磷酸酶活性予以特别关注。但是,出人意料地,通过增强肌醇单磷酸酶活性,这样的转化体的葡萄糖二酸生产能力大幅提高。

[0088] 因而,本发明的一个方式包括:在保有肌醇-1-磷酸合成酶基因、肌醇单磷酸酶基因、肌肉肌醇加氧酶基因和糖醛酸脱氢酶基因的转化体中进行肌醇单磷酸酶的过剩生产。

[0089] 对于本发明中谋求的肌醇单磷酸酶,除了对肌醇-1-磷酸显示出高底物特异性外,还显示出了可作用于广泛的底物的磷酸单酯水解酶活性,从而包括实质上能够水解肌醇-1-磷酸的蛋白质。作为代表性的肌醇单磷酸酶,例如已知有肌醇-1-单磷酸酶,大量生物来源的该基因(suhB基因)在基因库登录号Nos.ZP__04619988、YP__001451848等中公布。特别是大肠杆菌来源的suhB基因(序列编号3:AAC75586(MG1655))的使用在以大肠杆菌为宿主细胞的情况下较为方便。

[0090] 本发明的转化微生物所应具有的下述生物活性为肌肉肌醇加氧酶活性。该酶代表性地通过下述反应将肌肉肌醇转换为葡萄糖醛酸。

[0091] [化1]



[0093] 各种肌肉肌醇加氧酶基因是已知的,且可加以利用。例如,在W02002/074926号小册子中公开了隐球菌和人来源的肌肉肌醇加氧酶基因及其异源表达。此外,专利文献1所公开的肌肉肌醇加氧酶基因也可用于本发明中。进一步地,例如已知有被赋予了下述基因库登录号的来源于多种生物的肌肉肌醇加氧酶基因,它们可用于本发明中。

[0094] ACCESSION No. AY738258 (智人肌肉肌醇加氧酶 (MIOX))

[0095] ACCESSION No. NM101319 (拟南芥肌醇加氧酶1 (MIOX1))

[0096] ACCESSION No. NM001101065 (牛肌肉肌醇加氧酶 (MIOX))

[0097] ACCESSION No. NM001030266 (斑马鱼肌肉肌醇加氧酶 (miox))

[0098] ACCESSION No. NM214102 (野猪肌肉肌醇加氧酶 (MIOX))

[0099] ACCESSION No. AY064416 (智人肌肉肌醇加氧酶 (MIOX))

[0100] ACCESSION No. NM001247664 (番茄肌肉肌醇加氧酶 (MIOX))

[0101] ACCESSION No. XM630762 (盘基网柄菌AX4肌醇加氧酶 (miox))

[0102] ACCESSION No. NM145771 (褐鼠肌肉肌醇加氧酶 (Miox))

[0103] ACCESSION No. NM017584 (智人肌肉肌醇加氧酶 (MIOX))

[0104] ACCESSION No. NM001131282 (苏门答腊猩猩肌肉肌醇加氧酶 (MIOX))

[0105] 具有序列编号5所表示的编码区域核苷酸序列的miox基因的使用是特别方便的。

[0106] 本发明的转化微生物所应具有的最后一种生物活性为糖醛酸脱氢酶活性。该酶代表性地在NAD⁺的存在下通过下述反应将葡萄糖醛酸转换为葡萄糖二酸。

[0107] [化2]



[0109] 各种糖醛酸脱氢酶基因是已知的且可加以利用。例如,专利文献1所记载的绿脓杆菌或农杆菌属细菌的糖醛酸脱氢酶也可在本发明中进行使用。进一步地,例如已知有被赋予了下述基因库登录号的udh基因,它们可用于本发明中。

[0110] ACCESSION No. BK006462 (根癌农杆菌str. C58糖醛酸脱氢酶 (udh) 基因)

[0111] ACCESSION No. EU377538 (丁香假单胞菌番茄致病变种str. DC3000糖醛酸脱氢酶 (udh) 基因)

[0112] 具有序列编号7所表示的编码区域核苷酸序列的udh基因的使用是特别方便的。

[0113] 本领域技术人员容易理解,在肌醇-1-磷酸合成酶基因中记载的变异、修饰和密码子最佳化以及表达盒、启动子等调节序列和质粒等以及基于此的转化的说明对于本发明的肌醇单磷酸酶基因、肌肉肌醇加氧酶基因和糖醛酸脱氢酶基因也是全部适用的。因而,本发明的转化体可以保有包含编码肌醇-1-磷酸合成酶的核酸的表达盒、包含编码肌醇单磷酸酶基因的核酸的表达盒、包含编码肌肉肌醇加氧酶的核酸的表达盒和包含编码糖醛酸脱氢酶的核酸的表达盒这4个表达盒。本发明的适宜的转化体保有包含具有序列编号1所表示的核苷酸序列的核酸的表达盒、包含具有序列编号3所表示的核苷酸序列的核酸的表达盒、包含具有序列编号5所表示的核苷酸序列的核酸的表达盒以及包含具有序列编号7所表示的核苷酸序列的核酸的表达盒。

[0114] 上述4个表达盒可配置在1个载体上被转染至宿主微生物中。或者可以将配置有其中任意2个以上的表达盒的载体与配置有剩下的表达盒的载体共转染至宿主微生物中,还

可以将配置有各个表达盒的4个载体共转染至宿主微生物中。进一步地,也可以将上述4个表达盒中的任意一个以上嵌入到宿主微生物的基因组中,剩下的表达盒以质粒的形式存在于该转化微生物内。例如,也可将配置有包含编码肌肉肌醇加氧酶的核酸的表达盒和包含编码糖醛酸脱氢酶的核酸的表达盒的质粒转染到大肠杆菌AKC-018株(于2011年10月25日以FERM P-22181保藏于独立行政法人产品评价技术基础机构专利生物保藏中心。国际保藏编号:FERM BP-11514)中,该大肠杆菌AKC-018株在染色体上具有包含编码肌醇-1-磷酸合成酶的核酸(INO1)的表达盒和包含编码肌醇单磷酸酶的核酸(suhB)的表达盒这两种表达盒。

[0115] 此外还认为有多种微生物细胞表达本发明中谋求的肌醇单磷酸酶活性(即具有编码肌醇单磷酸酶活性的内源性基因)。因而,本发明中的肌醇单磷酸酶的过剩生产可通过下述操作进行诱导:增加内源性肌醇单磷酸酶基因的拷贝数;在内源性肌醇单磷酸酶基因的调整区域导入变异;利用高表达诱导性外来调整区域来置换内源性肌醇单磷酸酶基因的调整区域;和使内源性肌醇单磷酸酶基因的调整区域缺失。具体地说,为了实现肌醇单磷酸酶的过剩表达,可通过进行下述操作来实现:利用含有内源性肌醇单磷酸酶基因、或者含有在该内源性基因的编码区域添加了适宜的调整区域的表达盒的构建体对上述宿主微生物进行转化,使该转化体内该肌醇单磷酸酶基因的拷贝数与原来的宿主细胞相比有实质性的增加;或者利用公知的基因重组技术对于具有内源性肌醇单磷酸酶基因的原来的宿主细胞实施染色体的变异、添加和缺失;或者使用诱变剂等染色体中随机导入变异。可使用公知的SDS-PAGE分析法等对肌醇单磷酸酶的过剩生产进行确认。

[0116] 进一步地,用于增强肌醇单磷酸酶活性的本发明的其它方式包括在宿主微生物细胞中诱导肌醇单磷酸酶的活化。为了完成该目的,可示例出下述这样的方法:1)在内源性肌醇单磷酸酶基因中导入变异;2)置换内源性肌醇单磷酸酶基因的一部分或全部;3)使内源性肌醇单磷酸酶基因的一部分缺失;4)使降低肌醇单磷酸酶活性的其它蛋白质减少;和/或5)使降低肌醇单磷酸酶活性的化合物的生成减少。

[0117] 关于上述用于增强肌醇单磷酸酶活性的1)~5)的方法,具体地说,可在对肌醇单磷酸酶的基因实施变异、添加或者缺失后对该基因所编码的肌醇单磷酸酶的活性进行评价,从而得到肌醇单磷酸酶活性得到增强的肌醇单磷酸酶。

[0118] 对于上述得到的转化体,为了进行本发明的葡萄糖二酸的生产,可在适于上述转化体的生长和/或维持的条件下进行培养和维持。来源于各种宿主微生物细胞的转化体所用的适宜的培养基组成、培养条件、培养时间对于本领域技术人员为已知的。

[0119] 培养基可以为含有1种以上的碳源、氮源、无机盐、维生素以及必要时的微量元素或维生素等微量成分的天然、半合成、合成培养基。但是,不消说所使用的培养基必须适当满足所要培养的转化体的营养要求。进一步地,为了使上述转化体与能够由该转化体转换为葡萄糖二酸的碳源接触,本发明的培养基需要含有最终能够用作葡萄糖二酸生产的底物的碳源、即需要含有可在转化体内转换为葡萄糖-6-磷酸的化合物。碳源可以为D-葡萄糖、蔗糖、寡糖、多糖、淀粉、纤维素、米糠、废糖蜜,进一步可为含有D-葡萄糖的生物物质。作为适宜的生物物质,可示例出玉米分解液或纤维素分解液。此外,在转化体表达有用的附加性状的情况下、例如在具有针对抗生素的耐性标记的情况下,培养基也可以含有相应的抗生素。由此使发酵中的杂菌所致的污染风险降低。

[0120] 在宿主微生物无法同化上述纤维素或多糖类碳源的情况下,可通过实施在该宿主微生物中导入外来基因等公知的基因工程方法使其适应于使用这些碳源的葡萄糖二酸的生产中。作为外来基因,例如可以举出纤维酶基因或淀粉酶基因等。

[0121] 培养可以为分批式也可以为连续式。并且,在任一情况下,均可为在适当的培养时刻补给所追加的上述碳源等的形式。进一步地,培养应该在维持适宜的温度、氧浓度、pH等的条件下继续进行。来源于一般的微生物宿主细胞的转化体的适宜培养温度通常为15℃~45℃、优选为25℃~37℃的范围。宿主微生物为好氧性的情况下,为了确保发酵中的适当的氧浓度,需要进行振荡(烧瓶培养等)、搅拌/通气(发酵罐培养等)。这些培养条件可由本领域技术人员容易地设定。

[0122] 本领域技术人员可通过公知方法的组合由上述培养物中精制出葡萄糖二酸。例如,在专利文献1中具体记载了对于该目的有用的葡萄糖二酸的检测和定量方法。

[0123] 被提供了上述说明的本领域技术人员可充分实施本发明。下面出于进一步说明的目的提供实施例,因而本发明并不限于该实施例。需要说明的是,在本说明书中,只要不特别声明,核苷酸序列被记载为从5'向3'方向。

[0124] 实施例

[0125] 实施例1:质粒的构建

[0126] 1-a) 肌醇单磷酸酶表达盒

[0127] 在LB培养基中(2ml)中于37℃对大肠杆菌株W3110(NBRC 12713)进行振荡培养。培养终止后,从培养液中回收菌体,使用Nucleo Spin Tissue(产品名、MACHERY-NAGEL社制造)提取基因组DNA。将所提取的基因组DNA用于模板,利用下述引物进行PCR扩增(PrimeSTAR Max DNA聚合酶(产品名、Takara-Bio制造)反应条件:98℃10sec,55℃5sec,72℃20sec,28循环),克隆suhB基因的编码区域(序列编号3)。

[0128] [化3]

[0129] 正向:atgcatccgatgctgaac(序列编号9)

[0130] 反向:ttaacgcttcagagcgctcg(序列编号10)

[0131] 所得到的suhB编码区域可转录地插入到下述序列的启动子的下游。

[0132] [化4]

[0133] 启动子:gtcgtttttctgcttaggattttgttatttaaatgaagcctgtaatgccttgcttccattgcggataaatcctacttttttattgccttcaataaatttaaggagttc(序列编号11)

[0134] 即,在质粒pNFP-A51(于2011年10月25日以FERM P-22182保藏在独立行政法人产品评价技术基础机构专利生物保藏中心。国际保藏编号:FERM BP-11515)的多克隆位点插入终止子序列和上述启动子序列。在所导入的启动子序列的下游连接上述克隆的suhB编码区域,构建pNFP-A54。将所构建的pNFP-A54利用氯化钙法(参照羊土社遗传子工学实验ノ一ト(基因工程实验笔记)上田村隆明著)转染到大肠杆菌AKC-016株(于2011年4月20日以FERM P-22104保藏在独立行政法人产品评价技术基础机构专利微生物保藏中心。国际保藏编号:FERM BP-11512)中。通过SDS-PAGE确认到肌醇单磷酸酶在该大肠杆菌的可溶性级分中的高表达。

[0135] 1-b) 肌醇-1-磷酸合成酶表达盒

[0136] 从酿酒酵母的培养液中回收菌体,使用Nucleo Spin Tissue(产品名、MACHERY-

NAGEL社制造)提取基因组DNA。将所提取的基因组DNA用于模板,利用下述引物进行PCR扩增(PrimeSTAR Max DNA聚合酶(产品名、Taraka-Bio制造)反应条件:98°C10sec,55°C5sec,72°C20sec,28循环),克隆INO1基因的编码区域(序列编号1)。

[0137] [化5]

[0138] 正向:atgacagaagataatattgct(序列编号12)

[0139] 反向:ttacaacaatctctcttcg(序列编号13)

[0140] 所得到的ino1编码区域可转录地插入到下述序列的启动子的下游。

[0141] [化6]

[0142] 启动子:ctcaagcccaaaggaagagtggaggcagtcagtcgcgtaatgcttaggcacaggattgattgtcgcgaatgattgacacgattccgcttgacgctgcgtaagggtttgtaattttacaggcaacctttattcactaa caaatagctggtggaa(序列编号14)

[0143] 即,在上述质粒pNFP-A51的多克隆位点插入终止子序列和上述启动子序列。在所导入的启动子序列的下游连接上述克隆的ino1编码区域,构建pNFP-D78。将所构建的pNFP-D78利用氯化钙法(参照羊土社遗传子工学实验ノート上田村隆明著)转染到大肠杆菌AKC-016株(于2011年4月20日以FERM P-22104保藏在独立行政法人产品评价技术基础机构专利微生物保藏中心。国际保藏编号:FERM BP-11512)中。通过SDS-PAGE确认到肌醇-1-磷酸合成酶在该大肠杆菌的可溶性级分中的高表达。

[0144] 1-c) 肌肉肌醇加氧酶表达盒

[0145] 肌肉肌醇加氧酶(miox)基因如下得到:通过人工合成制作具有序列编号5的核苷酸序列的DNA,以该DNA为模板,利用下述引物进行PCR(PrimeSTAR Max DNA聚合酶(产品名、Taraka-Bio制造)反应条件:98°C10sec,55°C5sec,72°C20sec,28循环),从而得到该肌肉肌醇加氧酶(miox)基因。

[0146] [化7]

[0147] 正向:atgaaagttgatgttggtcct(序列编号15)

[0148] 反向:ttaccaggacagggtgcc(序列编号16)

[0149] 将所得到的miox编码区域可转录地插入到序列编号11的启动子的下游。即,在上述pNFP-A51的多克隆位点插入终止子序列和上述启动子序列。在所导入的启动子序列的下游连接上述克隆的miox编码区域,构建pNFP-H26。将所构建的pNFP-H26利用氯化钙法(参照羊土社遗传子工学实验ノート上田村隆明著)转染到大肠杆菌株FERM P-22104中。通过SDS-PAGE确认到肌肉肌醇加氧酶在该大肠杆菌的可溶性级分中的高表达。

[0150] 1-d) 糖醛酸脱氢酶表达盒

[0151] 糖醛酸脱氢酶(udh)基因如下得到:通过人工合成制作具有序列编号7的核苷酸序列的DNA,以该DNA为模板,利用下述引物进行PCR(PrimeSTAR Max DNA聚合酶(产品名、Taraka-Bio制造)反应条件:98°C10sec,55°C5sec,72°C20sec,28循环),从而得到该糖醛酸脱氢酶(udh)基因。

[0152] [化8]

[0153] 正向:atgaccactacccccttcaat(序列编号17)

[0154] 反向:tcagttgaacgggccgg(序列编号18)

[0155] 将所得到的udh编码区域可转录地插入到序列编号11的启动子的下游。即,在上述

pNFP-A51的多克隆位点插入终止子序列和上述启动子序列。在所导入的启动子序列的下游连接上述克隆的udh编码区域,构建pNFP-H45。将所构建的pNFP-H45利用氯化钙法(参照羊土社遗传子工学实验ノ一ト上田村隆明著)转染至大肠杆菌株FERM P-22104中。通过SDS-PAGE确认到糖醛酸脱氢酶在该大肠杆菌的可溶性级分中的高表达。

[0156] 1-e) 用于转化的质粒的构建

[0157] 将上述制作的pNFP-D78利用SalI消化,进行末端平滑化和5'末端脱磷酸化。将pNFP-A54中的suhB表达盒克隆化,与pNFP-D78连接。得到了与pNFP-D78中的IN01表达盒顺方向地连接有suhB表达盒的pNFP-G22。接下来,将pNFP-G22利用SalI消化,进行末端平滑化和5'末端脱磷酸化。将实施例1中制作的pNFP-H26中的miox表达盒和pNFP-H45中的udh表达盒克隆化,将两个表达盒与pNFP-G22连接。得到了与pNFP-G22中的IN01表达盒和suhB表达盒顺方向地连接有miox表达盒和udh表达盒的本发明的质粒。

[0158] 实施例2:

[0159] 2-a) 利用经含有表达盒的质粒转染而得到的转化体、使用罐状培养槽所进行的葡萄糖二酸的生产

[0160] 使用氯化钙法(参照羊土社遗传子工学实验ノ一ト上田村隆明著),将按照上述步骤构建的本发明的质粒转染至大肠杆菌AKC-016株(于2011年4月20日以FERM P-22104保藏在独立行政法人产品评价技术基础机构专利微生物保藏中心。国际保藏编号:FERM BP-11512)中。

[0161] 将所得到的转化体在含有氨苄青霉素(100mg/L)的LB微板上于37℃培养一天,形成菌落。将含有氨苄青霉素(100mg/L)的LB培养基30mL装入150mL容量的烧瓶中,利用白金耳从上述微板进行菌落的植菌,在37℃以180rpm进行3~5小时的培养,直至OD(600nm)为0.5左右,将其作为主培养所用的前培养液。

[0162] 在1000mL容量的罐状培养装置(丸菱生物工程社制造)中加入300mL含有10g/L的葡萄糖与100mg/L氨苄青霉素的合成培养基(表1),添加6mL的前培养液进行主培养(使用罐状培养装置的葡萄糖二酸生产试验)。培养条件如下:培养温度32℃;培养pH 6.0[下限];碱添加28%(重量/容量)氨水;搅拌850rpm;通气1vvm。适当添加作为原料的葡萄糖添料溶液(表2),使培养液中的葡萄糖浓度为0g/L~5g/L。

[0163] [表1]

[0164] 合成培养基组成

	KH_2PO_4	13.3 g
	$(\text{NH}_4)_2\text{HPO}_4$	4 g
	$\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$	1.2 g
	EDTA·2Na	8.4 mg
	$\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$	2.5 mg
	$\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$	15 mg
[0165]	$\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$	1.5 mg
	H_3BO_3	3 mg
	$\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$	2.5 mg
	$\text{Zn}(\text{CH}_3\text{COO})_2 \cdot 2\text{H}_2\text{O}$	13 mg
	$\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$	100 mg
	总量	1 L

[0166] 使用8N KOH,将pH调整为6.3

[0167] [表2]

[0168] 葡萄糖添料溶液

	葡萄糖	700	g
	MgSO ₄ ·7H ₂ O	20	g
	EDTA·2Na	13	mg
	CoCl ₂ ·6H ₂ O	5	mg
	MnCl ₂ ·4H ₂ O	29	mg
[0169]	CuCl ₂ ·2H ₂ O	4	mg
	H ₃ BO ₃	5	mg
	Na ₂ MoO ₄ ·2H ₂ O	4	mg
	Zn(CH ₃ COO) ₂ ·2H ₂ O	21	mg
	FeCl ₃ ·6H ₂ O	41	mg
	总量	1	L

[0170] 将上述培养液在4℃进行10,000g×10分钟的离心分离,回收上清,测定培养上清的葡萄糖二酸浓度。具体地说,将Shim-Pak SCR-H(保护柱)和Shim-Pak SCR-101H(均为商品名、岛津GLC社制造)联结,进行HPLC分析(检测器:RI、柱温度:40℃、流速:1mL/min、移动层:0.1%甲酸),从而对培养上清中的葡萄糖二酸浓度进行定量。

[0171] 其结果,在基于本发明的保有肌醇-1-磷酸合成酶基因、肌醇单磷酸酶基因、肌肉肌醇加氧酶基因和糖醛酸脱氢酶基因的转化体中,通过增强肌醇单磷酸酶活性,在该转化体的培养上清中生产了约73g/L(培养时间68小时)的葡萄糖二酸。

[0172] 参考例:

[0173] 制作不过剩生产肌醇单磷酸酶的转化体,使用该肌醇单磷酸酶非增强株,除此以外,按照上述实施例2进行葡萄糖二酸生产试验,结果在培养时间68小时仅生产了0.26g/L的葡萄糖二酸。

[0174] 关于本说明书中提及的质粒和微生物,在它们被记载为保藏的情况下,均保藏于(保藏机关的名称)“IPOD独立行政法人产品评价技术基础机构专利微生物保藏中心(IPOD, NITE)”;(保藏机关的地址)“日本邮政编码305-8566茨城县筑波市东1丁目1番地1中央第6”。

[0175] 工业实用性

[0176] 本发明可用于葡萄糖二酸的工业发酵生产。

序列表

<110> 旭化成化学株式会社
 <120> 葡萄糖二酸的制造方法
 <130> F112222-W0
 <150> 日本专利申请2012-39129
 <151> 2012-02-24
 <160> 18
 <170> PatentIn version 3.5
 <210> 1
 <211> 1602
 <212> DNA
 <213> 酿酒酵母
 <220>
 <221> CDS
 <222> (1)..(1602)
 <400> 1
 atg aca gaa gat aat att gct cca atc acc tcc gtt aaa gta gtt acc 48
 Met Thr Glu Asp Asn Ile Ala Pro Ile Thr Ser Val Lys Val Val Thr
 1 5 10 15
 gag aag tgc acg tac aag gac aac gag ctg ctc acc aag tac agc tac 96
 Asp Lys Cys Thr Tyr Lys Asp Asn Glu Leu Leu Thr Lys Tyr Ser Tyr
 20 25 30
 gaa aat gct gta gtt acg aag aca gct agt ggc cgc ttc gat gtc acg 144
 Glu Asn Ala Val Val Thr Lys Thr Ala Ser Gly Arg Phe Asp Val Thr
 35 40 45
 [0001] ccc act gtt caa gac tac gtg ttc aaa ctt gac tta aaa aag ccg gaa 192
 Pro Thr Val Gln Asp Tyr Val Phe Lys Leu Asp Leu Lys Lys Pro Glu
 50 55 60
 aaa cta gga att atg ctc att ggg tta ggt ggc aac aat ggc tcc acc 240
 Lys Leu Gly Ile Met Leu Ile Gly Leu Gly Gly Asn Asn Gly Ser Thr
 65 70 75 80
 tta gtg gcc tcg gta ttg gcg aat aag cac aat gtg gag ttt caa act 288
 Leu Val Ala Ser Val Leu Ala Asn Lys His Asn Val Glu Phe Gln Thr
 85 90 95
 aag gaa ggc gtt aag caa cca aac tac ttc ggc tcc atg act caa tgt 336
 Lys Glu Gly Val Lys Gln Pro Asn Tyr Phe Gly Ser Met Thr Gln Cys
 100 105 110
 tct acc ttg aaa ctg ggt gtc gat gcg gag ggg aat gac gtt tat gct 384
 Ser Thr Leu Lys Leu Gly Val Asp Ala Glu Gly Asn Asp Val Tyr Ala
 115 120 125
 cct ttt aac tct ctg ttg ccc atg gtt agc cca aac gac ttt gtc gtc 432
 Pro Phe Asn Ser Leu Leu Pro Met Val Ser Pro Asn Asp Phe Val Val
 130 135 140
 tct ggt tgg gac atc aat aac gca gat cta tac gaa gct atg cag aga 480
 Ser Gly Trp Asp Ile Asn Asn Ala Asp Leu Tyr Glu Ala Met Gln Arg
 145 150 155 160
 agt cag gtt ctc gaa tat gat ctg caa caa cgc ttg aag gcg aag atg 528
 Ser Gln Val Leu Glu Tyr Asp Leu Gln Gln Arg Leu Lys Ala Lys Met
 165 170 175
 tcc ttg gtg aag cct ctt cct tcc att tac tac cct gat ttc att gea 576
 Ser Leu Val Lys Pro Leu Pro Ser Ile Tyr Tyr Pro Asp Phe Ile Ala
 180 185 190
 gct aat caa gat gag aga gcc aat aac tgc atc aat ttg gat gaa aaa 624
 Ala Asn Gln Asp Glu Arg Ala Asn Asn Cys Ile Asn Leu Asp Glu Lys

	195	200	205	
	ggc aac gta acc acg agg ggt aag tgg gcc cat ctg caa cgc atc aga Gly Asn Val Thr Thr Arg Gly Lys Trp Ala His Leu Gln Arg Ile Arg 210	215	220	672
	cgc gat att cag aat ttc aaa gaa gaa aac gcc ctt gat aaa gta atc Arg Asp Ile Gln Asn Phe Lys Glu Glu Asn Ala Leu Asp Lys Val Ile 225	230	235	720
	gtt ctt tgg act gca aat act gag agg tca gta gaa gta tct cct ggt Val Leu Trp Thr Ala Asn Thr Glu Arg Ser Val Glu Val Ser Pro Gly 245	250	255	768
	gtt aat gac acc atg gaa aac ctc ttg cag tct att aag aat gac cat Val Asn Asp Thr Met Glu Asn Leu Leu Gln Ser Ile Lys Asn Asp His 260	265	270	816
	gaa gag att gct cct tcc acg atc ttt gca gca gca tct atc ttg gaa Glu Glu Ile Ala Pro Ser Thr Ile Phe Ala Ala Ala Ser Ile Leu Glu 275	280	285	864
	ggt gtc ccc tat att aat ggt tca ccg cag aat act ttt gtt ccc ggc Gly Val Pro Tyr Ile Asn Gly Ser Pro Gln Asn Thr Phe Val Pro Gly 290	295	300	912
	ttg gtt cag ctg gct gag cat gag ggt aca ttc att gcg gga gac gat Leu Val Gln Leu Ala Glu His Glu Gly Thr Phe Ile Ala Gly Asp Asp 305	310	315	960
	ctc aag tcg gga caa acc aag ttg aag tct gtt ctg gcc cag ttc tta Leu Lys Ser Gly Gln Thr Lys Leu Lys Ser Val Leu Ala Gln Phe Leu 325	330	335	1008
	gtg gat gca ggt att aaa ccg gtc tcc att gca tcc tat aac cat tta Val Asp Ala Gly Ile Lys Pro Val Ser Ile Ala Ser Tyr Asn His Leu 340	345	350	1056
[0002]	ggc aat aat gac ggt tat aac tta tct gct cca aaa caa ttt agg tct Gly Asn Asn Asp Gly Tyr Asn Leu Ser Ala Pro Lys Gln Phe Arg Ser 355	360	365	1104
	aag gag att tcc aaa agt tct gtc ata gat gac atc atc gcg tct aat Lys Glu Ile Ser Lys Ser Ser Val Ile Asp Asp Ile Ile Ala Ser Asn 370	375	380	1152
	gat atc ttg tac aat gat aaa ctg ggt aaa aaa gtt gac cac tgc att Asp Ile Leu Tyr Asn Asp Lys Leu Gly Lys Lys Val Asp His Cys Ile 385	390	395	1200
	gtc att aaa tat atg aag ccc gtc ggg gac tca aaa gtg gca atg gac Val Ile Lys Tyr Met Lys Pro Val Gly Asp Ser Lys Val Ala Met Asp 405	410	415	1248
	gag tat tac agt gag ttg atg tta ggt ggc cat aac cgg att tcc att Glu Tyr Tyr Ser Glu Leu Met Leu Gly Gly His Asn Arg Ile Ser Ile 420	425	430	1296
	cac aat gtt tgc gaa gat tct tta ctg gct acg ccc ttg atc atc gat His Asn Val Cys Glu Asp Ser Leu Leu Ala Thr Pro Leu Ile Ile Asp 435	440	445	1344
	ctt tta gtc atg act gag ttt tgt aca aga gtg tcc tat aag aag gtg Leu Leu Val Met Thr Glu Phe Cys Thr Arg Val Ser Tyr Lys Lys Val 450	455	460	1392
	gac cca gtt aaa gaa gat gct ggc aaa ttt gag aac ttt tat cca gtt Asp Pro Val Lys Glu Asp Ala Gly Lys Phe Glu Asn Phe Tyr Pro Val 465	470	475	1440
	tta acc ttc ttg agt tac tgg tta aaa gct cca tta aca aga cca gga Leu Thr Phe Leu Ser Tyr Trp Leu Lys Ala Pro Leu Thr Arg Pro Gly 485	490	495	1488
	ttt cac ccg gtg aat ggc tta aac aag caa aga acc gcc tta gaa aat Phe His Pro Val Asn Gly Leu Asn Lys Gln Arg Thr Ala Leu Glu Asn			1536

500	505	510	
ttt tta aga ttg ttg att gga ttg cct tct caa aac gaa cta aga ttc			1584
Phe Leu Arg Leu Leu Ile Gly Leu Pro Ser Gln Asn Glu Leu Arg Phe			
515	520	525	
gaa gag aga ttg ttg taa			1602
Glu Glu Arg Leu Leu			
530			
<210> 2			
<211> 533			
<212> PRT			
<213> 酿酒酵母			
<400> 2			
Met Thr Glu Asp Asn Ile Ala Pro Ile Thr Ser Val Lys Val Val Thr			
1	5	10	15
Asp Lys Cys Thr Tyr Lys Asp Asn Glu Leu Leu Thr Lys Tyr Ser Tyr			
20	25	30	
Glu Asn Ala Val Val Thr Lys Thr Ala Ser Gly Arg Phe Asp Val Thr			
35	40	45	
Pro Thr Val Gln Asp Tyr Val Phe Lys Leu Asp Leu Lys Lys Pro Glu			
50	55	60	
Lys Leu Gly Ile Met Leu Ile Gly Leu Gly Gly Asn Asn Gly Ser Thr			
65	70	75	80
Leu Val Ala Ser Val Leu Ala Asn Lys His Asn Val Glu Phe Gln Thr			
85	90	95	
Lys Glu Gly Val Lys Gln Pro Asn Tyr Phe Gly Ser Met Thr Gln Cys			
100	105	110	
Ser Thr Leu Lys Leu Gly Val Asp Ala Glu Gly Asn Asp Val Tyr Ala			
115	120	125	
Pro Phe Asn Ser Leu Leu Pro Met Val Ser Pro Asn Asp Phe Val Val			
130	135	140	
Ser Gly Trp Asp Ile Asn Asn Ala Asp Leu Tyr Glu Ala Met Gln Arg			
145	150	155	160
Ser Gln Val Leu Glu Tyr Asp Leu Gln Gln Arg Leu Lys Ala Lys Met			
165	170	175	
Ser Leu Val Lys Pro Leu Pro Ser Ile Tyr Tyr Pro Asp Phe Ile Ala			
180	185	190	
Ala Asn Gln Asp Glu Arg Ala Asn Asn Cys Ile Asn Leu Asp Glu Lys			
195	200	205	
Gly Asn Val Thr Thr Arg Gly Lys Trp Ala His Leu Gln Arg Ile Arg			
210	215	220	
Arg Asp Ile Gln Asn Phe Lys Glu Glu Asn Ala Leu Asp Lys Val Ile			
225	230	235	240

[0003]

Val Leu Trp Thr Ala Asn Thr Glu Arg Ser Val Glu Val Ser Pro Gly
 245 250 255
 Val Asn Asp Thr Met Glu Asn Leu Leu Gln Ser Ile Lys Asn Asp His
 260 265 270
 Glu Glu Ile Ala Pro Ser Thr Ile Phe Ala Ala Ala Ser Ile Leu Glu
 275 280 285
 Gly Val Pro Tyr Ile Asn Gly Ser Pro Gln Asn Thr Phe Val Pro Gly
 290 295 300
 Leu Val Gln Leu Ala Glu His Glu Gly Thr Phe Ile Ala Gly Asp Asp
 305 310 315 320
 Leu Lys Ser Gly Gln Thr Lys Leu Lys Ser Val Leu Ala Gln Phe Leu
 325 330 335
 Val Asp Ala Gly Ile Lys Pro Val Ser Ile Ala Ser Tyr Asn His Leu
 340 345 350
 Gly Asn Asn Asp Gly Tyr Asn Leu Ser Ala Pro Lys Gln Phe Arg Ser
 355 360 365
 Lys Glu Ile Ser Lys Ser Ser Val Ile Asp Asp Ile Ile Ala Ser Asn
 370 375 380
 [0004] Asp Ile Leu Tyr Asn Asp Lys Leu Gly Lys Lys Val Asp His Cys Ile
 385 390 395 400
 Val Ile Lys Tyr Met Lys Pro Val Gly Asp Ser Lys Val Ala Met Asp
 405 410 415
 Glu Tyr Tyr Ser Glu Leu Met Leu Gly Gly His Asn Arg Ile Ser Ile
 420 425 430
 His Asn Val Cys Glu Asp Ser Leu Leu Ala Thr Pro Leu Ile Ile Asp
 435 440 445
 Leu Leu Val Met Thr Glu Phe Cys Thr Arg Val Ser Tyr Lys Lys Val
 450 455 460
 Asp Pro Val Lys Glu Asp Ala Gly Lys Phe Glu Asn Phe Tyr Pro Val
 465 470 475 480
 Leu Thr Phe Leu Ser Tyr Trp Leu Lys Ala Pro Leu Thr Arg Pro Gly
 485 490 495
 Phe His Pro Val Asn Gly Leu Asn Lys Gln Arg Thr Ala Leu Glu Asn
 500 505 510
 Phe Leu Arg Leu Leu Ile Gly Leu Pro Ser Gln Asn Glu Leu Arg Phe
 515 520 525
 Glu Glu Arg Leu Leu
 530

<210>	3		
<211>	804		
<212>	DNA		
<213>	大肠杆菌		
<220>			
<221>	CDS		
<222>	(1)..(804)		
<400>	3		
	atg cat ccg atg ctg aac atc gcc gtg cgc gca gcg cgc aag gcg ggt	48	
	Met His Pro Met Leu Asn Ile Ala Val Arg Ala Ala Arg Lys Ala Gly		
	1 5 10 15		
	aal tia att gcc aaa aac tat gaa acc ccg gac gct gta gaa gcg agc	96	
	Asn Leu Ile Ala Lys Asn Tyr Glu Thr Pro Asp Ala Val Glu Ala Ser		
	20 25 30		
	cag aaa ggc agt aac gat ttc gtg acc aac gta gat aaa gct gcc gaa	144	
	Gln Lys Gly Ser Asn Asp Phe Val Thr Asn Val Asp Lys Ala Ala Glu		
	35 40 45		
	gcg gtg att atc gac acg att cgt aaa tct tac cca cag cac acc atc	192	
	Ala Val Ile Ile Asp Thr Ile Arg Lys Ser Tyr Pro Gln His Thr Ile		
	50 55 60		
	atc acc gaa gaa agc ggt gaa ctt gaa ggt act gat cag gat gtt caa	240	
	Ile Thr Glu Glu Ser Gly Glu Leu Glu Gly Thr Asp Gln Asp Val Gln		
	65 70 75 80		
	tgg gtt atc gat cca ctg gat ggc act acc aac ttt atc aaa cgt ctg	288	
	Trp Val Ile Asp Pro Leu Asp Gly Thr Thr Asn Phe Ile Lys Arg Leu		
	85 90 95		
[0005]	ccg cac ttc cgc gta tct atc gct gtt cgt atc aaa ggc cgc acc gaa	336	
	Pro His Phe Ala Val Ser Ile Ala Val Arg Ile Lys Gly Arg Thr Glu		
	100 105 110		
	gtt gct gtg gta tac gat cct atg cgt aac gaa ctg ttc acc gcc act	384	
	Val Ala Val Val Tyr Asp Pro Met Arg Asn Glu Leu Phe Thr Ala Thr		
	115 120 125		
	cgc ggt cag ggc gca cag ctg aac ggc tac cga ctg cgc ggc agc acc	432	
	Arg Gly Gln Gly Ala Gln Leu Asn Gly Tyr Arg Leu Arg Gly Ser Thr		
	130 135 140		
	gct cgc gat ctc gac ggt act att ctg gcg acc ggc ttc ccg ttc aaa	480	
	Ala Arg Asp Leu Asp Gly Thr Ile Leu Ala Thr Gly Phe Pro Phe Lys		
	145 150 155 160		
	gca aaa cag tac gcc act acc tac atc aac atc gtc ggc aaa ctg ttc	528	
	Ala Lys Gln Tyr Ala Thr Thr Tyr Ile Asn Ile Val Gly Lys Leu Phe		
	165 170 175		
	aac gaa tgt gca gac ttc cgt cgt acc ggt tct gcg gcg ctg gat ctg	576	
	Asn Glu Cys Ala Asp Phe Arg Arg Thr Gly Ser Ala Ala Leu Asp Leu		
	180 185 190		
	gct tac gtc gct gcg ggt cgt gtt gac ggt ttc ttt gaa atc ggt ctg	624	
	Ala Tyr Val Ala Ala Gly Arg Val Asp Gly Phe Phe Glu Ile Gly Leu		
	195 200 205		
	cgc ccg tgg gac ttc gcc gca ggc gag ctg ctg gtt cgt gaa gcg ggc	672	
	Arg Pro Trp Asp Phe Ala Ala Gly Glu Leu Leu Val Arg Glu Ala Gly		
	210 215 220		
	ggc atc gtc agc gac ttc acc ggt ggt cat aac tac atg ctg acc ggt	720	
	Gly Ile Val Ser Asp Phe Thr Gly Gly His Asn Tyr Met Leu Thr Gly		
	225 230 235 240		
	aac atc gtt gct ggt aac ccg cgc gtt gtt aaa gcc atg ctg gcg aac	768	
	Asn Ile Val Ala Gly Asn Pro Arg Val Val Lys Ala Met Leu Ala Asn		
	245 250 255		

atg cgt gac gag tta agc gac gct ctg aag cgt taa 804
 Met Arg Asp Glu Leu Ser Asp Ala Leu Lys Arg
 260 265

<210> 4
 <211> 267
 <212> PRT
 <213> 大肠杆菌

<400> 4

Met His Pro Met Leu Asn Ile Ala Val Arg Ala Ala Arg Lys Ala Gly
 I 5 10 15

Asn Leu Ile Ala Lys Asn Tyr Glu Thr Pro Asp Ala Val Glu Ala Ser
 20 25 30

Gln Lys Gly Ser Asn Asp Phe Val Thr Asn Val Asp Lys Ala Ala Glu
 35 40 45

Ala Val Ile Ile Asp Thr Ile Arg Lys Ser Tyr Pro Gln His Thr Ile
 50 55 60

Ile Thr Glu Glu Ser Gly Glu Leu Glu Gly Thr Asp Gln Asp Val Gln
 65 70 75 80

Trp Val Ile Asp Pro Leu Asp Gly Thr Thr Asn Phe Ile Lys Arg Leu
 85 90 95

[0006]

Pro His Phe Ala Val Ser Ile Ala Val Arg Ile Lys Gly Arg Thr Glu
 100 105 110

Val Ala Val Val Tyr Asp Pro Met Arg Asn Glu Leu Phe Thr Ala Thr
 115 120 125

Arg Gly Gln Gly Ala Gln Leu Asn Gly Tyr Arg Leu Arg Gly Ser Thr
 130 135 140

Ala Arg Asp Leu Asp Gly Thr Ile Leu Ala Thr Gly Phe Pro Phe Lys
 145 150 155 160

Ala Lys Gln Tyr Ala Thr Thr Tyr Ile Asn Ile Val Gly Lys Leu Phe
 165 170 175

Asn Glu Cys Ala Asp Phe Arg Arg Thr Gly Ser Ala Ala Leu Asp Leu
 180 185 190

Ala Tyr Val Ala Ala Gly Arg Val Asp Gly Phe Phe Glu Ile Gly Leu
 195 200 205

Arg Pro Trp Asp Phe Ala Ala Gly Glu Leu Leu Val Arg Glu Ala Gly
 210 215 220

Gly Ile Val Ser Asp Phe Thr Gly Gly His Asn Tyr Met Leu Thr Gly
 225 230 235 240

Asn Ile Val Ala Gly Asn Pro Arg Val Val Lys Ala Met Leu Ala Asn
 245 250 255

Met Arg Asp Glu Leu Ser Asp Ala Leu Lys Arg
260 265

<210> 5
<211> 858
<212> DNA
<213> 人工序列

<220>
<223> miox基因编码区域

<220>
<221> CDS
<222> (1).. (858)

<400> 5
atg aaa gtt gat gtt ggt cct gat ccg tct ctg gtt tat cgt cca gac 48
Met Lys Val Asp Val Gly Pro Asp Pro Ser Leu Val Tyr Arg Pro Asp
1 5 10 15
gtg gat ccg gag atg gca aag agc aaa gac tct ttc cgt aac tac act 96
Val Asp Pro Glu Met Ala Lys Ser Lys Asp Ser Phe Arg Asn Tyr Thr
20 25 30
tct ggt ccg ctg ctg gat cgc gtt ttc acg acc tac aaa ctg atg cac 144
Ser Gly Pro Leu Leu Asp Arg Val Phe Thr Thr Tyr Lys Leu Met His
35 40 45
acc cat cag acc gtt gat ttc gtg agc cgc aag cgc atc cag tat ggt 192
Thr His Gln Thr Val Asp Phe Val Ser Arg Lys Arg Ile Gln Tyr Gly
50 55 60
tct ttc tct tac aaa aag atg acc att atg gag gct gtt ggt atg ctg 240
Ser Phe Ser Tyr Lys Lys Met Thr Ile Met Glu Ala Val Gly Met Leu
65 70 75 80
gat gac ctg gtt gac gaa agc gac cct gac gtt gac ttc cct aac tct 288
Asp Asp Leu Val Asp Glu Ser Asp Pro Asp Val Asp Phe Pro Asn Ser
85 90 95
ttc cat gca ttt cag acc gcc gaa ggt att cgt aaa gct cat cct gat 336
Phe His Ala Phe Gln Thr Ala Glu Gly Ile Arg Lys Ala His Pro Asp
100 105 110
aaa gat tgg ttc cac ctg gtc ggt ctg ctg cac gac ctg ggt aag atc 384
Lys Asp Trp Phe His Leu Val Gly Leu Leu His Asp Leu Gly Lys Ile
115 120 125
atg gcg ctg tgg ggc gaa cca caa tgg gcg gta gtg ggc gat act ttc 432
Met Ala Leu Trp Gly Glu Pro Gln Trp Ala Val Val Gly Asp Thr Phe
130 135 140
ccg gtg ggc tgc cgc cca cag gca tcc gtg gtc ttc tgc gac tct acc 480
Pro Val Gly Cys Arg Pro Gln Ala Ser Val Val Phe Cys Asp Ser Thr
145 150 155 160
ttc cag gat aac ccg gat ctg cag gat ccg cgc tat tcc acc gaa ctg 528
Phe Gln Asp Asn Pro Asp Leu Gln Asp Pro Arg Tyr Ser Thr Glu Leu
165 170 175
ggc atg tac cag ccg cat tgc ggc ctg gag aac gtt ctg atg tct tgg 576
Gly Met Tyr Gln Pro His Cys Gly Leu Glu Asn Val Leu Met Ser Trp
180 185 190
ggt cac gac gag tac ctg tat cag atg atg aaa ttc aac aaa ttc tcc 624
Gly His Asp Glu Tyr Leu Tyr Gln Met Met Lys Phe Asn Lys Phe Ser
195 200 205
ctg ccg tcc gag gca ttt tac atg atc cgt ttt cac tcc ttc tac ccg 672
Leu Pro Ser Glu Ala Phe Tyr Met Ile Arg Phe His Ser Phe Tyr Pro
210 215 220
tgg cat acc ggt ggc gat tat cgt cag ctg tgc tct cag cag gat ctg 720

[0007]

Trp His Thr Gly Gly Asp Tyr Arg Gln Leu Cys Ser Gln Gln Asp Leu
 225 230 235 240
 gat atg ctg ccg tgg gtt cag gag ttc aat aaa ttc gac ctg tat acc 768
 Asp Met Leu Pro Trp Val Gln Glu Phe Asn Lys Phe Asp Leu Tyr Thr
 245 250 255
 aaa tgc cct gac ctg cca gat gtt gaa tcc ctg cgc cca tac tac caa 816
 Lys Cys Pro Asp Leu Pro Asp Val Glu Ser Leu Arg Pro Tyr Tyr Gln
 260 265 270
 ggc ctg att gac aaa tac tgc ccg ggc acc ctg tcc tgg taa 858
 Gly Leu Ile Asp Lys Tyr Cys Pro Gly Thr Leu Ser Trp
 275 280 285

 <210> 6
 <211> 285
 <212> PRT
 <213> 人工序列

 <220>
 <223> 合成结构

 <400> 6

 Met Lys Val Asp Val Gly Pro Asp Pro Ser Leu Val Tyr Arg Pro Asp
 1 5 10 15

 Val Asp Pro Glu Met Ala Lys Ser Lys Asp Ser Phe Arg Asn Tyr Thr
 20 25 30

 Ser Gly Pro Leu Leu Asp Arg Val Phe Thr Thr Tyr Lys Leu Met His
 35 40 45

 [0008] Thr His Gln Thr Val Asp Phe Val Ser Arg Lys Arg Ile Gln Tyr Gly
 50 55 60

 Ser Phe Ser Tyr Lys Lys Met Thr Ile Met Glu Ala Val Gly Met Leu
 65 70 75 80

 Asp Asp Leu Val Asp Glu Ser Asp Pro Asp Val Asp Phe Pro Asn Ser
 85 90 95

 Phe His Ala Phe Gln Thr Ala Glu Gly Ile Arg Lys Ala His Pro Asp
 100 105 110

 Lys Asp Trp Phe His Leu Val Gly Leu Leu His Asp Leu Gly Lys Ile
 115 120 125

 Met Ala Leu Trp Gly Glu Pro Gln Trp Ala Val Val Gly Asp Thr Phe
 130 135 140

 Pro Val Gly Cys Arg Pro Gln Ala Ser Val Val Phe Cys Asp Ser Thr
 145 150 155 160

 Phe Gln Asp Asn Pro Asp Leu Gln Asp Pro Arg Tyr Ser Thr Glu Leu
 165 170 175

 Gly Met Tyr Gln Pro His Cys Gly Leu Glu Asn Val Leu Met Ser Trp
 180 185 190

 Gly His Asp Glu Tyr Leu Tyr Gln Met Met Lys Phe Asn Lys Phe Ser
 195 200 205

Leu Pro Ser Glu Ala Phe Tyr Met Ile Arg Phe His Ser Phe Tyr Pro
 210 215 220
 Trp His Thr Gly Gly Asp Tyr Arg Gln Leu Cys Ser Gln Gln Asp Leu
 225 230 235 240
 Asp Met Leu Pro Trp Val Gln Glu Phe Asn Lys Phe Asp Leu Tyr Thr
 245 250 255
 Lys Cys Pro Asp Leu Pro Asp Val Glu Ser Leu Arg Pro Tyr Tyr Gln
 260 265 270
 Gly Leu Ile Asp Lys Tyr Cys Pro Gly Thr Leu Ser Trp
 275 280 285

<210> 7
 <211> 807
 <212> DNA
 <213> 人工序列

<220>
 <223> udh基因编码区域

<220>
 <221> CDS
 <222> (1)..(807)

[0009] <400> 7
 atg acc act acc ccc ttc aat cgc ctg ctg ctc acc gga gcc gea ggc 48
 Met Thr Thr Thr Pro Phe Asn Arg Leu Leu Thr Gly Ala Ala Gly
 1 5 10 15
 ggc ctg ggc aag gtc ett cgc gaa cgc ctg aaa ggc tac gcc gag gtc 96
 Gly Leu Gly Lys Val Leu Arg Glu Arg Leu Lys Gly Tyr Ala Glu Val
 20 25 30
 ctg cgc ctg tct gac atc agc ccc atg gcc ccg gcc gcg ggc ccg cat 144
 Leu Arg Leu Ser Asp Ile Ser Pro Met Ala Pro Ala Ala Gly Pro His
 35 40 45
 gaa gaa gtc att acc tgt gac ctg gcc gac aag gct gcg gtg cat acc 192
 Glu Glu Val Ile Thr Cys Asp Leu Ala Asp Lys Ala Ala Val His Thr
 50 55 60
 ctg gtc gag ggc gta gac gcc atc atc cac ttt ggc ggg gtt tct acc 240
 Leu Val Glu Gly Val Asp Ala Ile Ile His Phe Gly Gly Val Ser Thr
 65 70 75 80
 gaa cac gcc ttc gaa gag att ctc ggc ccc aat atc tgc gcc gtg ttc 288
 Glu His Ala Phe Glu Glu Ile Leu Gly Pro Asn Ile Cys Gly Val Phe
 85 90 95
 cac gtg tac gag gcg gcg cgc aag cac ggg gtc aag cgc atc atc ttc 336
 His Val Tyr Glu Ala Ala Arg Lys His Gly Val Lys Arg Ile Ile Phe
 100 105 110
 gcc agc tcc aac cac acc atc ggt ttc tat cgc cag gat gag cgc atc 384
 Ala Ser Ser Asn His Thr Ile Gly Phe Tyr Arg Gln Asp Glu Arg Ile
 115 120 125
 gac gct cac gcg ccg cgc cgg ccc gac agc tat tac ggg ctg tcc aag 432
 Asp Ala His Ala Pro Arg Arg Pro Asp Ser Tyr Tyr Gly Leu Ser Lys
 130 135 140
 tgc tac ggc gaa gat gtg gcc agc ttc tac ttt gac cgc tac gcc atc 480
 Cys Tyr Gly Glu Asp Val Ala Ser Phe Tyr Phe Asp Arg Tyr Gly Ile
 145 150 155 160

gag acc gtc agc att cgc atc ggc tgc tgc ttc ccg cag cca cag aac 528
 Glu Thr Val Ser Ile Arg Ile Gly Ser Ser Phe Pro Gln Pro Gln Asn
 165 170 175

ctg cgc atg ctc tgc acc tgg ctc agt tac gac gac ctg gtg cag tgg 576
 Leu Arg Met Leu Cys Thr Trp Leu Ser Tyr Asp Asp Leu Val Gln Leu
 180 185 190

atc gaa cgc ggg ctg ttc acc ccc ggg gtt ggc cac acc atc gtc tac 624
 Ile Glu Arg Gly Leu Phe Thr Pro Gly Val Gly His Thr Ile Val Tyr
 195 200 205

ggc gcc tcc gac aat cgc acc gtg tgg tgg gac aac cgc cat gcc gcc 672
 Gly Ala Ser Asp Asn Arg Thr Val Trp Trp Asp Asn Arg His Ala Ala
 210 215 220

cac ctg ggc tat gta ccc aag gac agc tcc gaa acc ttc cgc gca gcc 720
 His Leu Gly Tyr Val Pro Lys Asp Ser Ser Glu Thr Phe Arg Ala Ala
 225 230 235 240

gtg gag gcc caa ccg gca ccc gcc gcc gat gac ccg agc atg gtc tac 768
 Val Glu Ala Gln Pro Ala Pro Ala Ala Asp Asp Pro Ser Met Val Tyr
 245 250 255

cag ggc ggc gct ttc gcc gtg gcc gcc ccg ttc aac tga 807
 Gln Gly Gly Ala Phe Ala Val Ala Gly Pro Phe Asn
 260 265

<210> 8
 <211> 268
 <212> PRT
 <213> 人工序列

<220>
 <223> 合成结构

[0010]

<400> 8
 Met Thr Thr Thr Pro Phe Asn Arg Leu Leu Leu Thr Gly Ala Ala Gly
 1 5 10 15

Gly Leu Gly Lys Val Leu Arg Glu Arg Leu Lys Gly Tyr Ala Glu Val
 20 25 30

Leu Arg Leu Ser Asp Ile Ser Pro Met Ala Pro Ala Ala Gly Pro His
 35 40 45

Glu Glu Val Ile Thr Cys Asp Leu Ala Asp Lys Ala Ala Val His Thr
 50 55 60

Leu Val Glu Gly Val Asp Ala Ile Ile His Phe Gly Gly Val Ser Thr
 65 70 75 80

Glu His Ala Phe Glu Glu Ile Leu Gly Pro Asn Ile Cys Gly Val Phe
 85 90 95

His Val Tyr Glu Ala Ala Arg Lys His Gly Val Lys Arg Ile Ile Phe
 100 105 110

Ala Ser Ser Asn His Thr Ile Gly Phe Tyr Arg Gln Asp Glu Arg Ile
 115 120 125

Asp Ala His Ala Pro Arg Arg Pro Asp Ser Tyr Tyr Gly Leu Ser Lys
 130 135 140

Cys Tyr Gly Glu Asp Val Ala Ser Phe Tyr Phe Asp Arg Tyr Gly Ile

	145	150	155	160	
	Glu Thr Val Ser Ile Arg Ile Gly Ser Ser Phe Pro Gln Pro Gln Asn	165	170	175	
	Leu Arg Met Leu Cys Thr Trp Leu Ser Tyr Asp Asp Leu Val Gln Leu	180	185	190	
	Ile Glu Arg Gly Leu Phe Thr Pro Gly Val Gly His Thr Ile Val Tyr	195	200	205	
	Gly Ala Ser Asp Asn Arg Thr Val Trp Trp Asp Asn Arg His Ala Ala	210	215	220	
	His Leu Gly Tyr Val Pro Lys Asp Ser Ser Glu Thr Phe Arg Ala Ala	225	230	235	240
	Val Glu Ala Gln Pro Ala Pro Ala Ala Asp Asp Pro Ser Met Val Tyr	245	250	255	
	Gln Gly Gly Ala Phe Ala Val Ala Gly Pro Phe Asn	260	265		
	<210> 9				
	<211> 18				
	<212> DNA				
	<213> 人工序列				
	<220>				
[0011]	<223> suhB编码区域正向PCR引物				
	<400> 9				
	atgcatccga tgctgaac				18
	<210> 10				
	<211> 19				
	<212> DNA				
	<213> 人工序列				
	<220>				
	<223> suhB编码区域反向PCR引物				
	<400> 10				
	ttaacgcttc agagcgtcg				19
	<210> 11				
	<211> 109				
	<212> DNA				
	<213> 人工序列				
	<220>				
	<223> suhB编码区域启动子				
	<400> 11				
	gtcgtttttc tgcttaggat tttgttattt aaattaagcc tgtaatgcct tgcttcatt				60
	gcggataaat cctacitttt tattgccttc aaataaattt aaggagttc				109
	<210> 12				
	<211> 22				
	<212> DNA				
	<213> 人工序列				
	<220>				
	<223> IN01编码区域正向PCR引物				

	<400> 12 atgacagaag ataatatgac tc	22
	<210> 13 <211> 19 <212> DNA <213> 人工序列	
	<220> <223> IN01编码区域反向PCR引物	
	<400> 13 ttacaacaat ctctcttcg	19
	<210> 14 <211> 155 <212> DNA <213> 人工序列	
	<220> <223> IN01编码序列启动子	
	<400> 14 ctcaagccca aaggaagagt gaggcgagtc agtcgcgtaa tgcttaggca caggattgat	60
	ttgtcgcaat gattgacacg attccgcttg acgetgcgta aggtttttgt aattttacag	120
	gcaacctttt attcactaac aaatagctgg tggaa	155
	<210> 15 <211> 22 <212> DNA <213> 人工序列	
[0012]	<220> <223> miox编码区域正向PCR引物	
	<400> 15 atgaaagttg atgttggccc tg	22
	<210> 16 <211> 18 <212> DNA <213> 人工序列	
	<220> <223> Miox编码区域反向PCR引物	
	<400> 16 ttaccaggac aggtgccc	18
	<210> 17 <211> 21 <212> DNA <213> 人工序列	
	<220> <223> Udh编码区域正向PCR引物	
	<400> 17 atgaccacta ccccccttaa t	21
	<210> 18 <211> 17 <212> DNA <213> 人工序列	
	<220> <223> Udh编码区域反向PCR引物	
[0013]	<400> 18 tcagttgaac gggccgg	17

ATGACAGAAG ATAATATTGC TCCAATCACC TCCGTAAAAG TAGTTACCGA CAAGTGCACG
TACAAGGACA ACGAGCTGCT CACCAAGTAC AGCTACGAAA ATGCTGTAGT TACGAAGACA
GCTAGTGGCC GCTTCGATGT CACGCCACT GTTCAAGACT ACGTGTCAA ACTTGACTTA
AAAAAGCCGG AAAAAGTAGG AATTATGCTC ATTGGGTTAG GTGGCAACAA TGGCTCCACC
TTAGTGGCCT CGGTATTGGC GAATAAGCAC AATGTGGAGT TTCAAACATA GGAAGGCGTT
AAGCAACCAA ACTACTCGG CTCCATGACT CAATGTTCTA CCTTGAAACT GGGTGTCCAT
GCGGAGGGGA ATGACGTTTA TGCTCCTTTT AACTCTCTGT TGCCCATGGT TAGCCCAAAC
GACTTTGTCG TCTCTGGTTG GGACATCAAT AACGCAGATC TATACGAAGC TATGCAGAGA
AGTCAGGTTT TCGAATATGA TCTGCAACAA CGCTTGAAGG CGAAGATGTC CTTGGTGAAG
CCTCTTCTT CCATTTACTA CCTGATTTT ATTGCAGCTA ATCAAGATGA GAGAGCCAAT
AACTGCATCA ATTTGGATGA AAAAGGCAAC GTAACCACGA GGGGTAAGTG GGGCCATCTG
CAACGCATCA GACGCGATAT TCAGAAATTC AAAGAAGAAA ACGCCCTTGA TAAAGTAATC
GTTCTTTGGA CTGCAAATAC TGAGAGGTC GTAGAAGTAT CTCTGGTGT TAATGACACC
ATGAAAAACC TCTTGCAGTC TATTAAGAAT GACCATGAAG AGATTGCTCC TTCCACGATC
TTTGCAGCAG CATCTATCTT GGAAGGTGTC CCCTATATTA ATGGTTCACC GCAGAATACT
TTTGTCCCG GCTTGGTTCA GCTGGCTGAG CATGAGGGTA CATTCAATTGC GGGAGACGAT
CTCAAGTCGG GACAAACCAA GTTGAAGTCT GTTCTGGCCC AGTTCTTAGT GGATGCAGGT
ATTAAACCGG TCTCCATTGC ATCCTATAAC CATTTAGGCA ATAATGACGG TTATAACTTA
TCTGCTCCAA AACAAATTTAG GTCTAAGGAG ATTTCCAAAA GTTCTGTCAT AGATGACATC
ATCGGTCTA ATGATATCTT GTACAATGAT AACTGGGTA AAAAAGTTGA CCACTGCATT
GTCATTAAT ATATGAAGCC CGTCGGGGAC TCAAAAGTGG CAATGGACGA GTATTACAGT
GAGTTGATGT TAGGTGGCCA TAACCGGATT TCCATTCACA ATGTTTGGCA AGATTCTTTA
CTGGCTACGC CCTTGATCAT CGATCTTTTA GTCATGACTG AGTTTTGTAC AAGAGTGTCC
TATAAGAAGG TGGACCCAGT TAAAGAAGAT GCTGGCAAAT TTGAGAACTT TTATCCAGTT
TTAACCTTCT TGAGTACTG GTTAAAAGCT CCATTAACAA GACCAGGATT TCACCCGGTG
AATGGCTTAA ACAAGCAAAG AACCGCTTA GAAAATTTTT TAAGATTGTT GATTGGATTG
CCTTCTCAA ACGAACTAAG ATTCGAAGAG AGATTGTTGT AA (SEQ ID NO 1)

图1

atgcatccgatgctgaacatcgccgtgcgcgccagcgcgcaaggcgggtaatttaattgcc
aaaaactatgaaaccccggaacgctgtagaagcgcagccagaaaggcagtaacgatttcgtg
accaacgtagataaagctgccgaagcgggtgattatcgacacgattcgtaaattacc
cagcacaccatcatcaccgaagaaagcgggtgaacttgaaggctactgatcaggatgttcaa
tgggttatcgatccactggatggcactaccaactttatcaaacgtctgcccacttcg
gtatctatcgctgttcgtatcaaaggccgcaccgaagttgctgtggtatacgatcctatg
cgtaacgaactgttcaccgccactcgcggtcagggcgcacagctgaacggctaccgactg
cgcggcagcaccgctcgcgatctcgacgggtactattctggcgaccggttcccgttcaaa
gcaaacagtagccactacctacatcaacatcgteggcaaacgtttcaacgaatgtgca
gacttcgctcgtaaccggttctgcgggcgtggatctggcttacgtcgcctgcggtcgtgtt
gacggtttctttgaaatcggctctgcgcccgtgggacttcgcccagggcagctgctggtt
cgtgaagcgggcggcatcgtcagcgaactcaccggtggtcataactacatgctgaccggt
aacatcgttgctggtaacccgcgcggttgtaaagccatgctggcgaacatgcgtgacgag
ttaagcagcgtctgaagcgttaa (SEQ ID NO 3)

图2

```

      10      20      30      40      50      60
ATGAAAGTTG ATGTTGGTCC TGATCCGTCT CTGGTTTATC GTCCAGACGT GGATCCGGAG

      70      80      90      100     110     120
ATGGCAAAGA GCAAAGACTC TTTCCGTAAC TACACTTCTG GTCCGCTGCT GGATCGCGTT

      130     140     150     160     170     180
TTCACGACCT ACAAACCTGAT GCACACCCAT CAGACCGTTG ATTTTCGTGAG CCGCAAGCGC

      190     200     210     220     230     240
ATCCAGTATG GTTCTTTCTC TTACAAAAAG ATGACCATTA TGGAGGCTGT TGGTATGCTG

      250     260     270     280     290     300
GATGACCTGG TTGACGAAAG CGACCCTGAC GTTGACTICC CTAACTCTTT CCATGCATTT

      310     320     330     340     350     360
CAGACCGCCG AAGGTATTCT TAAAGCTCAT CCTGATAAAG ATTGGTTCCA CCTGGTCGGT

      370     380     390     400     410     420
CTGCTGCACG ACCTGGGTAA GATCATGGCG CTGTGGGGCG AACCACAATG GCGGTTAGTG

      430     440     450     460     470     480
GGCGATACTT TCCCGGTGGG CTGCCGCCCA CAGGCATCCG TGGTCTTCTG CGACTCTACC

      490     500     510     520     530     540
TTCCAGGATA ACCCGGATCT GCAGGATCCG CGCTATTCCA CCGAACTGGG CATGTACCAG

      550     560     570     580     590     600
CCGCATTGCG GCCTGGAGAA CGTCTGTATG TCTTGGGGTC ACGACGAGTA CCTGTATCAG

      610     620     630     640     650     660
ATGATGAAAT TCAACAAATT CTCCTGCCG TCCGAGGCAT TTTACATGAT CCGTTTTAC

      670     680     690     700     710     720
TCCTTCTACC CGTGGCATA CCGTGGCGAT TATCGTCAGC TGTGCTCTCA GCAGGATCTG

      730     740     750     760     770     780
GATATGCTGC CGTGGGTTCA GGAGTTCAAT AAATTCGACC TGTATACCAA ATGCCCTGAC

      790     800     810     820     830     840
CTGCCAGATG TTGAATCCCT GCGCCCATAC TACCAAGGCC TGATTGACAA AACTGCCCC

      850     860
GGCACCCGTG CCTGGTAA (SEQ ID NO 5)

```

图3

```

      10      20      30      40      50      60
atgaccaacta cccccttcaa tcgecttgetg etcaccggag ccgcagggcgg cctgggcaag

      70      80      90      100     110     120
gtccttcgcg aacgcctgaa aggctacgcc gaggtcctgc gcctgtctga catcageccc

      130     140     150     160     170     180
atggccccgg ccgcggggccc gcatgaagaa gtcattacct gtgacctggc cgacaaggct

      190     200     210     220     230     240
gcgggtgcata ccttggtcga gggcgtagac gccatcatcc actttggcgg ggtttctacc

      250     260     270     280     290     300
gaacacgcct tcgaagagat tctcgcccc aatatctgcg gcgtgttcca cgtgtacgag

      310     320     330     340     350     360
goggcgcgca agcacggggt caagcgcate atcttcgcca gctccaacca caccatcggt

      370     380     390     400     410     420
ttctatcgcc aggatgagcg catcgacget cacgcgccgc gccggcccga cagetattac

      430     440     450     460     470     480
gggctgtcca agtgetacgg cgaagatgtg gccagcttct actttgaccg ctacggcate

      490     500     510     520     530     540
gagaccgtca gcattegcct eggetcgtcg ticcgcagc cacagaacct gcgcattctc

      550     560     570     580     590     600
tgcaectggc tcagttaaga cgacctggtg cagttgateg aacgcgggct gttcaccccc

      610     620     630     640     650     660
ggggttggcc acaccatcgt ctacggcgcc tccgacaate gcaccgtgtg gtgggacaac

      670     680     690     700     710     720
cgccatgccg cgcacctggg ctatgtacce aaggacagct ccgaaacctt ccgocgagcc

      730     740     750     760     770     780
gtggaggccc aaccggcacc cgccgccgat gaccogagca tggctotacca gggcggcget

      790     800     810
ttcgccgtgg ccggcccgtt caactga (SEQ ID NO 7)

```

图4