(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 105555947 A
(43)申请公布日 2016.05.04

(21)申请号 201480051856.X
(22)申请日 2014.09.25
(30)优先权数据
14/037666 2013.09.26 US
(85)PCT国际申请进入国家阶段日
2016.03.21
(86)PCT国际申请的申请数据
PCT/US2014/057386 2014.09.25
(87)PCT国际申请的公布数据
WO2015/048243 EN 2015.04.02
(71)申请人 纳幕尔杜邦公司
地址 美国特拉华州威尔明顿
(72)发明人 J. 杨
(74)专利代理机构 中国专利代理(香港)有限公司 72001
代理人 唐华东 黄希贵
(51)Int.Cl.
C12N 1/22(2006.01)
C12P 7/10(2006.01)

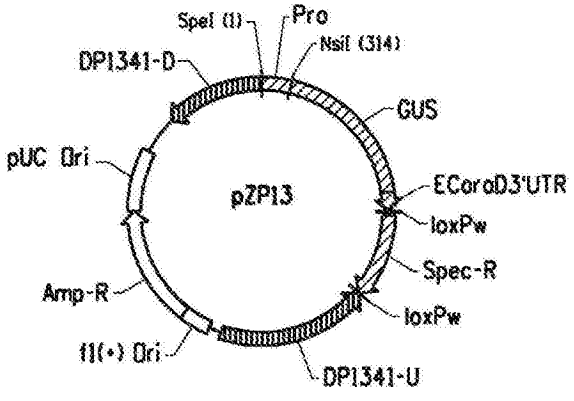
权利要求书1页 说明书10页
序列表15页 附图1页

(54)发明名称

高表达发酵单胞菌启动子

(57)摘要

本发明创建了合成的衍生启动子,其用于发酵单胞菌(Zymomonas)、发酵杆菌(Zymobacter)以及相关的细菌中嵌合基因的表达。该启动子具有运动发酵单胞菌(Z.mobilis)甘油醛-3-磷酸脱氢酶基因启动子的位置90处C到T的变化。具有此变化的启动子与其所源自的天然启动子相比具有更高的表达并且可用于针对表达编码区或调控RNA以获得高水平表达的蛋白质或调控RNA而进行的基因工程。



CN 105555947 A

1.一种合成的核酸分子，所述合成的核酸分子包含具有在天然运动发酵单胞菌(Z.mobilis)甘油醛-3-磷酸脱氢酶基因启动子的位置90处T对C的碱基取代的运动发酵单胞菌甘油醛-3-磷酸脱氢酶基因启动子，所述天然运动发酵单胞菌甘油醛-3-磷酸脱氢酶基因启动子的序列示于SEQ ID 1中。

2.一种合成的核酸分子，所述合成的核酸分子包含选自SEQ ID NO:2、4、6和7的运动发酵单胞菌甘油醛-3-磷酸脱氢酶基因启动子序列。

3.一种嵌合基因，所述嵌合基因包含可操作地连接至异源核酸分子的权利要求1或2所述的合成的核酸分子。

4.根据权利要求3所述的嵌合基因，其中所述异源核酸分子编码蛋白质或肽。

5.根据权利要求3所述的嵌合基因，其中所述异源核酸分子编码选自反义RNA、核糖酶和干扰RNA的调控RNA分子。

6.一种载体，所述载体包含权利要求1或2所述的合成的核酸分子。

7.一种转化细菌细胞的方法，所述细菌细胞选自发酵单胞菌(Zymomonas)细胞和发酵杆菌(Zymobacter)细胞，所述方法包括将权利要求1或2所述的合成的核酸分子引入到所述细胞中。

8.根据权利要求7所述的方法，其中引入包括将权利要求1或2所述的合成的核酸分子整合到所述细胞的基因组中或者保持在所述细胞内稳定复制的质粒上。

9.一种重组细菌菌株，所述重组细菌菌株包含权利要求1或2所述的合成的核酸，其中所述细菌菌株选自发酵单胞菌和发酵杆菌。

10.根据权利要求9所述的重组细菌菌株，其中所述分离的核酸分子为嵌合基因的启动子。

11.根据权利要求10所述的重组细菌菌株，其中所述嵌合基因包含分离的核酸分子，所述分离的核酸分子编码选自下列的酶：木糖异构酶、木酮糖激酶、转酮醇酶、转醛醇酶、L-阿拉伯糖异构酶、L-核酮糖激酶、L-核酮糖-5-磷酸-4-差向异构酶、核糖-5-磷酸异构酶和核酮糖-磷酸-3-差向异构酶。

12.一种用于生产改善的合成甘油醛-3-磷酸脱氢酶基因启动子的方法，所述方法包括：

a)从来自发酵单胞菌和发酵杆菌的甘油醛-3-磷酸脱氢酶基因分离包含甘油醛-3-磷酸脱氢酶基因启动子区域的核酸片段；以及

b)将在位置90处T对C的碱基取代引入到a)的分离的片段中，从而产生改善的合成甘油醛-3-磷酸脱氢酶基因启动子。

13.一种由权利要求12所述的方法制备的改善的合成甘油醛-3-磷酸脱氢酶基因启动子。

高表达发酵单胞菌启动子

技术领域

[0001] 本发明涉及微生物学和基因工程领域。更具体地，鉴定了用于引导细菌中嵌合基因表达的新型启动子。

背景技术

[0002] 微生物的乙醇生产提供了化石燃料的替代性能源且因此成为当前研究的重要领域。运动发酵单胞菌(Zymomonas mobilis)是可用于商业乙醇生产的天然产乙醇生物。然而，运动发酵单胞菌天然情况下并不利用木糖，其为水解木质纤维素物质中主要的戊糖。人们希望产乙醇生物利用木糖以充分利用木质纤维素生物质材料水解物中的碳水化合物底物。

[0003] 运动发酵单胞菌和天然情况下不利用木糖的其它产乙醇生物可通过引入基因来进行针对木糖的基因工程，所述基因编码1)木糖异构酶，其催化木糖至木酮糖的转化；2)木酮糖激酶，其使木酮糖磷酸化以形成木酮糖-5-磷酸；3)转酮醇酶；和4)转醛醇酶。针对木糖代谢对运动发酵单胞菌菌株的工程化(US 5,514,583,US 5,712,133,US 6,566,107,US 5,726,053,Feldmann等人(1992)Appl Microbiol Biotechnol 38:354-361,Zhang等人(1995)Science 267:240-243)，以及对棕榈发酵杆菌(Zymobacter palmae)菌株的工程化(Yanase等人(2007)Appl.Environ.Mirobiol.73:2592-2599)已有成功。然而，工程化的菌株在木糖上的生长以及产生乙醇通常不如葡萄糖一样良好。

[0004] 对这种工程化，已通过在运动发酵单胞菌细胞中有活性的启动子来表达编码用于木糖代谢的异源蛋白质的基因，所述启动子通常为运动发酵单胞菌甘油醛-3-磷酸脱氢酶基因的启动子或运动发酵单胞菌烯醇酶基因的启动子。针对木糖利用进行工程化的菌株已经通过在木糖培养基上连续传代而进行驯化，产生了具有改善的木糖利用的菌株，如US 7223575和US 7741119中所述。US 7989206和US 7998722公开了具有改善活性的运动发酵单胞菌甘油醛-3-磷酸脱氢酶基因启动子以及其在发酵单胞菌(Zymomonas)中表达木糖异构酶用于木糖利用的用途。

[0005] 对于另外的强启动子依然有需求，所述另外的强启动子可用于对发酵单胞菌以及其它产乙醇生物的基因工程以表达赋予改善的木糖利用的基因。所述启动子还可用于其它基因的表达。

发明内容

[0006] 本发明提供了具有改善活性的合成启动子，其为运动发酵单胞菌甘油醛-3-磷酸脱氢酶基因启动子的衍生物。

[0007] 因此，本发明提供了合成的核酸分子，其包含具有在位置90处T对C的碱基取代的运动发酵单胞菌甘油醛-3-磷酸脱氢酶基因启动子；其中所述位置编号属于SEQ ID NO:1。与天然运动发酵单胞菌甘油醛-3-磷酸脱氢酶基因启动子相比，这一碱基取代具有增加所述启动子强度的效果。

[0008] 在其它方面，本发明提供了包含合成的核酸分子的嵌合基因以及包含其的载体，所述合成的核酸分子具有上述碱基取代。

[0009] 在另外的方面，本发明提供了转化细菌细胞的方法，所述细菌细胞选自发酵单胞菌细胞和发酵杆菌(Zymobacter)细胞，所述方法包括将本发明的合成核酸分子引入到所述细胞中。

[0010] 在另一方面，本发明提供了包含本发明的合成核酸分子的重组细菌菌株，所述重组细菌菌株选自发酵单胞菌和发酵杆菌。

[0011] 在另一个实施例中，本发明提供用于生产改善的合成甘油醛-3-磷酸脱氢酶基因启动子的方法，该方法包括：

[0012] a)从来自发酵单胞菌和发酵杆菌的甘油醛-3-磷酸脱氢酶基因分离包含甘油醛-3-磷酸脱氢酶基因启动子区域的核酸片段；

[0013] b)将在位置90处T对C的碱基取代引入到a)的分离片段，从而产生改善的合成甘油醛-3-磷酸脱氢酶基因启动子。

[0014] 附图说明和序列描述

[0015] 图1示出了pZP13的质粒图谱。

[0016] 根据下面的详细描述以及构成本申请一部分的附带的序列描述可更充分地理解本发明。

[0017] 下列序列符合37C.F.R.1.821-1.825(“对含核苷酸序列和/或氨基酸序列公开内容的要求-序列规则”)并且符合世界知识产权组织(WIPO)标准ST.25(2009)及EPO和PCT的序列表要求(规则5.2和49.5(a-bis)，以及行政性指示的208节和附录C)。用于核苷酸和氨基酸序列数据的符号和格式遵循在37C.F.R.§1.822中示出的规定。

[0018] SEQ ID NO:1为运动发酵单胞菌CP4菌株的天然ZmPgap的核苷酸序列。[4236中#1]

[0019] SEQ ID NO:2为在位置90处具有C到T变化的运动发酵单胞菌CP4菌株的ZmPgap的核苷酸序列。

[0020] SEQ ID NO:3为运动发酵单胞菌的ZM4菌株的天然ZmPgap的核苷酸序列。[4236中#2]

[0021] SEQ ID NO:4为在位置90处具有C到T变化的运动发酵单胞菌ZM4菌株的ZmPgap的核苷酸序列。

[0022] SEQ ID NO:5为pZB4的ZmPgap的核苷酸序列。

[0023] [4236中T缺失#3]

[0024] SEQ ID NO:6为在位置90处具有C到T变化的pZB4的ZmPgap的核苷酸序列。

[0025] SEQ ID NO:7为运动发酵单胞菌ZM4菌株的ZmPgap的核苷酸序列，其在位置285处具有T缺失(如在pZB4中)并且在位置90处具有C到T的变化。

[0026] SEQ ID NO:8为1,951-bp DP1341-U片段的核苷酸序列。

[0027] SEQ ID NO:9为1,255-bp DP1341-U片段的核苷酸序列。

[0028] SEQ ID NO:10为GUS编码序列的核苷酸序列。

[0029] SEQ ID NO:11为大肠杆菌(E.coli)araD 3′UTR的核苷酸序列。

[0030] SEQ ID NO:12为Spec-R表达盒的核苷酸序列。

[0031] SEQ ID NO:13为pZP1332的核苷酸序列。

[0032] SEQ ID NO:14为pZP1337的核苷酸序列。

[0033] SEQ ID NO:15为loxPw片段的核苷酸序列，其位于所述Spec-R表达盒的5′。

[0034] SEQ ID NO:16为loxPw片段的核苷酸序列，其位于所述Spec-R表达盒的3′。

[0035] 包括loxPw位点和限制位点的侧翼loxPw序列为SEQ ID NO:15和16。

具体实施方式

[0036] 可使用下列定义阐释权利要求书和说明书：

[0037] 如本文所使用，术语“包含”、“包含的”、“包括”、“包括的”、“具有”、“具有的”、“含有”或“含有的”，或者它们的任何其它变型旨在涵盖非排他性的包括。例如，包含元素列表的组合物、混合物、工艺、方法、制品或设备不必仅限于那些元素，而可以包括其他未明确列出的元素，或此类组合物、混合物、工艺、方法、制品或设备固有的元素。此外，除非明确指明相反，“或”是指包含性的“或”而非排他性的“或”。例如，条件A或B满足下列任一项：A为真实的(或存在的)且B为虚假的(或不存在的)，A为虚假的(或不存在的)且B为真实的(或存在的)，以及A和B均为真实的(或存在的)。

[0038] 此外，涉及元素或组分的实例(即出现)的数目在本发明元素或组分前的不定冠词“一个”或“一种”旨在为非限制性的。因此，应将“一个”或“一种”理解为包括一个(种)或至少一个(种)，并且元素或组分的词语单数形式也包括复数指代，除非有数字明显表示单数。

[0039] 如本文所用，术语“发明”或“本发明”是非限制性术语，并且不旨在意指本发明的任何单独实施例，而是涵盖如本说明书和权利要求所述的所有可能的实施例。

[0040] 如本文所使用，修饰本发明的成分或反应物的量使用的术语“约”是指可发生的用数字表示的量的变化，例如：在真实世界中，通过用于制备浓缩物或使用溶液的一般测量和液体处理操作；通过这些操作中非故意的误差；通过用于制备组合物或执行方法的成分的制造、来源或纯度中的差异；等等。术语“约”还包括由于相对于由特定起始混合物所得的组合物的不同平衡条件而不同的量。无论是否由术语“约”来修饰，权利要求包括量的等同量。在一个实施例中，术语“约”指在报告数值10％以内，优选地在报告数值5％以内。

[0041] 术语“木糖异构酶”是指催化D-木糖和D-木酮糖的互变的酶。木糖异构酶(XI)属于归类为EC 5.3.1.5的酶的组。

[0042] 术语“木酮糖激酶”指磷酸化木酮糖形成木酮糖-5-磷酸的酶。

[0043] 术语“转酮醇酶”和“转醛醇酶”指戊糖磷酸化途径中的两种酶，其将木酮糖-5-磷酸转化成连接戊糖代谢与Entner-Douderoff糖酵解途径的中间体，该途径使木糖代谢成乙醇。

[0044] 术语“基因”指表达特定蛋白质或功能性RNA分子的核酸片段，其可任选地包括位于编码序列之前(5′非编码序列)和之后的调控序列(3′非编码序列)。“天然基因”或“野生型基因”指具有其自身调控序列的天然存在的基因。“嵌合基因”指不是天然基因的任何基因，包含在天然情况下不是一起存在的调控序列和编码序列。因此，嵌合基因可包括源于不同来源的调控序列和编码序列，或者包括源于同一来源但以不同于天然存在的方式排列的调控序列和编码序列。“内源基因”是指在生物基因组中处于其天然位置的天然基因。“外来基因”指正常情况下不存在于宿主生物中的基因，但是它通过基因转移引入到宿主生物中。外来基因可包含插入到非天然生物内的天然基因，或嵌合基因。

[0045] 术语“启动子”指能够控制编码序列或功能性RNA表达的DNA序列。一般来讲，编码序列位于启动子序列的3′端。启动子可整个源于天然基因，或者由源于天然存在的不同启动子的不同元件构成，或者甚至包含合成的DNA片段。本领域内的技术人员应当理解，不同的启动子可以在不同的组织或细胞类型中，或者在不同的发育阶段，或者响应不同的环境条件而引导基因的表达。通常将在大多数细胞类型中、在大多数情况下引起基因表达的启动子称为“组成型启动子”。

[0046] 如本文所使用，术语“表达”指衍生自基因的编码(mRNA)或功能性RNA的转录和稳定积聚。表达也可指将mRNA翻译成多肽。“过表达”指在转基因生物中产生的基因产物超出在正常生物或未转化生物中产生的基因产物的水平。

[0047] 如本文所用，术语“转化”指将核酸片段转移至宿主生物内，导致在基因上稳定遗传。所转移的核酸可以是宿主细胞中保留的质粒形式，或者某些转移的核酸可以被整合到所述宿主细胞基因组中。含有所转移的核酸片段的宿主生物被称为“转基因”或“重组”或“转化”生物。

[0048] 如本文所使用，术语“质粒”和“载体”是指通常携带有不属于细胞中心代谢的部分的基因的染色体外元件，并且通常是环状双链DNA分子的形式。这类元件可以是源自任何来源的自主复制序列、基因组整合序列、噬菌体或单链或双链DNA或RNA的核苷酸序列(线性或环状)，其中多个核苷酸序列已连接或重组进入独特构建体中，该独特构建体能够将所选基因产物的启动子片段和DNA序列与适当的3′非翻译序列一起引入细胞中。

[0049] 术语“可操作地连接”指单核酸片段上的核酸序列的关联，使得其中一个核酸序列的功能受到另一个核酸序列的影响。例如，当启动子能够影响编码序列的表达(即，所述编码序列受到启动子的转录控制)时，则该启动子与该编码序列可操作地连接。编码序列可以按有义或反义的取向可操作地连接至调控序列。

[0050] 术语“选择性标记”指一种鉴定因子，通常是抗生素或化学药品抗性基因，该因子能基于标记基因的效应，即，对抗生素的抗性进行选择，其中所述效应用于追踪遗传的受关注核酸和/或用于鉴定遗传了受关注核酸的细胞或生物。

[0051] 如本文所用，术语“密码子简并性”指允许核苷酸序列在不影响所编码的多肽的氨基酸序列的情况下发生变化的遗传密码的性质。技术人员将充分意识到：特定宿主细胞在使用核苷酸密码子以确定给定氨基酸时所表现出的“密码子偏好性”。因此，当合成基因用以改善在宿主细胞中的表达时，希望对基因进行设计，使得其密码子使用频率接近该宿主细胞优选的密码子使用频率。

[0052] 术语“密码子优化的”在其涉及用于转化不同宿主的核酸分子的基因或编码区时，是指在不改变由DNA编码的多肽的情况下，改变核酸分子的基因或编码区中的密码子以反映宿主生物通常的密码子使用。

[0053] 术语“异源”指非天然存在于受关注的位置。例如异源基因指宿主生物中非天然存在的但是通过转基因被引入到宿主生物中的基因。例如，存在于嵌合基因中的异源核酸分子是与其它嵌合基因片段相关的非天然存在核酸分子，例如具有彼此非天然相关的编码区和启动子片段的核酸分子。

[0054] 如本文所使用，“合成的核酸分子”是人造或者调控的RNA或DNA的聚合物，它是单链-或双链的，任选地包含合成的、非天然-的或改变的核苷酸碱基。DNA聚合物形式的合成

的核酸分子可由一种或多种cDNA、基因组DNA或合成DNA的片段构成。

[0055] 术语“运动发酵单胞菌甘油醛-3-磷酸脱氢酶基因启动子”和“ZmPgap”指带有启动子活性的核酸分子，其具有天然存在于所述运动发酵单胞菌基因中的甘油醛-3-磷酸脱氢酶编码区上游的核苷酸序列。这些术语指运动发酵单胞菌菌株的启动子，例如CP4和ZM4菌株(分别为SEQ ID NO:1和2)并且指在序列和/或长度上的变体，其在基本上并无不同的水平上引导表达，例如pZB4的ZmPgap(SEQ ID NO:3)。

[0056] 本文使用的标准重组DNA和分子克隆技术是本领域所熟知的，并且在如下文献中有更全面的描述:Sambrook,J.,Fritsch,E.F.和Maniatis,T.,Molecular Cloning:A Laboratory Manual,第2版;Cold Spring Harbor Laboratory:Cold Spring Harbor,New York,1989(下文称为“Maniatis”);和Silhavy,T.J.,Bennan,M.L.和Enquist,L.W.,Experiments with Gene Fusions;Cold Spring Harbor Laboratory:Cold Spring Harbor,New York,1984;以及Ausubel,F.M等人,Current Protocols in Molecular Biology,由Greene Publishing and Wiley-Interscience出版,1987。

[0057] 本发明涉及用于在发酵单胞菌、发酵杆菌以及相关的细菌中嵌合基因表达的合成衍生启动子。所述启动子与其所衍生来源的天然启动子相比具有更高的表达，并且可用于针对在发酵单胞菌、发酵杆菌以及相关的细菌中表达编码区或调控RNA以获得高水平表达的蛋白质或调控RNA而进行的基因工程。例如，可将本启动子用于基因工程来获得改善的发酵单胞菌细胞木糖利用，其通过表达木糖利用途径的酶的编码区而进行，例如木糖异构酶。

[0058] 改善的甘油醛-3-磷酸脱氢酶基因启动子

[0059] 本启动子是来自所述运动发酵单胞菌甘油醛-3-磷酸脱氢酶基因(ZmPgap)的启动子的衍生物，其在ZmPgap的位置90处中具有T对C的取代，位置90是指SEQ ID NO:1中的核苷酸+90。具有这一核苷酸变化的启动子与天然的ZmPgap相比具有提高的活性。

[0060] 在一个实施例中，本启动子具有SEQ IN NO:2的核苷酸序列，除了位置90处的T之外其与SEQ ID NO:1相同。

[0061] 在多个实施例中，就SEQ ID NO:1的核苷酸而言，本启动子是在位置90处具有T对C的取代的任意ZmPgap。ZmPgap的序列不是单个序列，而是可在序列上有一些对启动子功能没有实质影响的变异。对启动子功能没有实质影响意味着所述启动子序列引导了基本类似于通过存在于天然运动发酵单胞菌菌株中ZmPgap引导的表达水平的表达水平。序列的变异可天然存在于运动发酵单胞菌不同分离株或者菌株间，例如就SEQ ID NO:1的核苷酸而言在位置276处上所述CP4与ZM4的ZmPgap间的差异，该位置上在CP4中为A并且在ZM4中为G(分别SEQ ID NO:1和3)。在一个实施例中，本启动子具有所述ZM4菌株的ZmPgap的序列，其在位置90处具有T对C的取代(SEQ ID NO:4)。

[0062] 此外，在ZmPgap序列中还可有对启动子功能没有实质影响的一种或多种插入或缺失，例如存在于pZB4载体中的。就SEQ ID NO:1的核苷酸而言，所述pZB4载体的ZmPgap在位置285处缺失T，产生了SEQ ID NO:5。如在所述CP4菌株中该启动子在位置276处具有A。在一个实施例中，本启动子具有所述pZB4载体的ZmPgap的序列，其在位置90处具有T对C的取代(SEQ ID NO:6)。

[0063] 一种或多种插入或缺失可影响相对于SEQ ID NO:1的位置编号，这取决于所述插入或缺失是否存在于所述启动子5’端与SEQ ID NO:1中位置90处之间。因此，C到T的位置变

化可通过邻近序列进行最佳识别。所述C到T的变化的邻近序列为:CAGGGACGACAAT,其中粗体和带下划线的C是在本衍生启动子中变为T的核苷酸。本领域的技术人员应易于识别在该核苷酸处从C到T的变化(也称为T对C的取代),该处就SEQ ID NO:1而言称为位置90。

[0064] 在另一个实施例中,本启动子在位置90处具有T对C的取代,在位置285处具有缺失T(如在pZB4中)并且在位置276处具有G(如所述ZM4菌株)(SEQ ID NO:7)。在位置90处具有T对C的取代、在位置285处具有缺失T(如在pZB4中)并且在位置276处具有A(如所述CP4菌株)的启动子与SEQ ID NO:6相同。所有位置是相对SEQ ID NO:1而言。

[0065] 因此,所述启动子的示例包括SEQ ID NO:2、4、6和7。包括上文所给定的上文中在位置90处具有T对C的取代的ZmPgap在本文称为PgapU或P_{gapU}。

[0066] 制备改善的PgapU

[0067] 可通过本领域的技术人员已知的任意方法将所描述的位置90处取代引入ZmPgap核酸分子中。例如,可合成具有所述突变以及附近DNA序列的寡核苷酸并将其克隆进入更大的启动子DNA片段,以取代不含所述位置90处取代的片段。可合成包含所述核苷酸变化以及一些相邻启动子序列的引物并将其用于PCR来制备所述启动子片段。可将整个启动子DNA片段合成为连接在一起的多个寡核苷酸。可使用定点突变来引入所述核苷酸取代。

[0068] 嵌合基因以及载体中改善的Pgap,向细菌细胞中的引入

[0069] 可将本启动子可操作地连接至在细菌细胞中表达的异源核酸分子,形成本发明另一方面的嵌合核酸分子或嵌合基因。所述嵌合基因的设计以及构建是本领域技术人员熟知的。嵌合基因通常包括启动子,待表达的异源核酸分子以及3′终止控制区。终止控制区可源于多种基因并且通常取自对目标宿主细胞而言是天然的基因。所述可操作地连接的异源核酸分子可以是任意核酸分子,其在细菌细胞中期望进行表达,包括例如,蛋白质或肽的编码区、或者用于功能RNA表达的核酸。功能RNA包括例如,反义RNA、核糖酶和干扰RNA。此外,还可构建操作元,其包含本文所述的启动子以及通过所述启动子表达的多个编码区。

[0070] 本文所述的启动子可用于嵌合基因,其用以在属于发酵单胞菌或发酵杆菌的细菌中进行表达。可使用嵌合基因进行任意蛋白质的表达,所述蛋白质参与发酵单胞菌或发酵杆菌产物的生成。例如,可通过具有这些启动子的嵌合基因来表达参与氨基酸(诸如丙氨酸)或山梨醇或木糖醇合成的一种或多种酶。可在不利用木糖的天然发酵单胞菌或发酵杆菌菌株中或者在木糖-利用菌株中表达所述嵌合基因。还可使用本文所述的启动子进行涉及木糖代谢或者另一代谢途径的酶的表达。

[0071] 本文所述的嵌合基因通常在载体中构建或转移到载体中用于进一步的操作。载体是本领域熟知的。某些载体能够在广泛的宿主细菌中复制并可通过接合进行转移。pRK404以及三种相关载体:pRK437、pRK442和pRK442(H)的完整并且带注释的序列是可用的。已证明这些衍生物是在革兰氏阴性菌中进行的遗传操作中的有用工具(Scott等人,Plasmid 50(1):74-79(2003))。

[0072] 其它熟知的载体可用于不同的目标宿主细胞。对于在发酵单胞菌中的表达而言尤其有用的是兼而能在大肠杆菌和发酵单胞菌中复制的载体,例如在US 5,514,583中描述的pZB188。载体可包括在细胞中稳定自主复制的质粒以及携带待整合到细菌基因组中的构建体的质粒。用于DNA整合的质粒可包括转座子、与靶细菌基因组同源的核酸序列区或其它支持整合的序列。其它类型的载体还可是转座体,例如,其使用可通过EPICENTRE®市售可

用的系统而产生。如何为所需靶宿主和所需功能选择适当的载体是众所周知的。

[0073] 还可在不具有可操作地连接用于表达的核酸分子的情况下将本文所述的启动子构建到载体中并且整合到邻近内源编码区来替换在细菌基因组中的内源启动子，或者添加启动子，例如添加于操作元内的编码区。染色体启动子的替换可使用以下方法完成，例如，Yuan等人(Metab.Eng.(2006)8:79-90)和White等人(Can.J.Microbiol.(2007)53:56-62)所描述的方法。

[0074] 可通过熟知的方法将包含本文所述的启动子的载体引入到细菌细胞中，例如使用冻融转化、钙介导转化、电穿孔或者接合。

[0075] 使用PgapU进行异源核酸分子的表达

[0076] 可使用本文所述的改善的Pgap获得提高水平的嵌合基因表达。如本文实例3中所示，与通过天然Pgap所表达蛋白质的活性相比，包含可操作地连接至编码区的本启动子的嵌合基因在细胞提取物中给出了提高超过4倍的所表达蛋白质的活性。

[0077] 在多个实施例中，编码提高木糖或阿拉伯糖利用的蛋白质的一种或多种序列可作为包含本启动子的嵌合基因的一部分来表达。可在包含木糖异构酶、木酮糖激酶转酮醇酶和转醛醇酶的木糖利用途径中表达编码区。这些酶，它们的编码序列，以及它们参与木糖利用途径是已知的。具体地，如US 7,998,722(其以引用方式并入本文)中公开的任意木糖异构酶是可通过将其编码序列可操作地连接至本启动子来表达的木糖异构酶。

[0078] 另外，对于阿拉伯糖的利用，可表达的编码区包括L-阿拉伯糖异构酶(其将L-阿拉伯糖转化成L-核酮糖)、L-核酮糖激酶(其将L-核酮糖转化成L-核酮糖-5-磷酸)以及L-核酮糖-5-磷酸-4-差向异构酶(其将L-核酮糖-5-磷酸转化成D-木酮糖)，如US 5,843,760中公开，其以引用方式并入本文。此外，可表达编码提高木糖利用的其它蛋白质的序列，例如核糖-5-磷酸异构酶(RPI)，其催化核酮糖-5-磷酸与核糖-5-磷酸转化的并且属于归类为EC 5.3.1.6(US 20120156746中公开，其以引用方式并入本文)的酶组，以及核酮糖-磷酸-3-差向异构酶(RPE)，其催化D-核酮糖-5-磷酸与D-木酮糖-5-磷酸转化的并且属于归类为EC 5.1.3.1(US 20130157331中公开，其以引用方式并入本文)的酶组。

[0079] 在一个实施例中，本启动子用于表达编码木糖异构酶的异源核苷酸分子，其用于引导与来自所述ZmPgap的表达相比而提高的木糖异构酶的表达。所述改善的Pgap和木糖异构酶编码区形成嵌合基因，其通常还包括3′终止控制区。终止控制区可源于多种基因并且通常取自对目标宿主细胞而言是天然的基因。嵌合基因的构建在本领域是熟知的。

[0080] 可在嵌合基因中使用任意木糖异构酶编码区以通过改善的PgapU来表达木糖异构酶。木糖异构酶属于EC5.3.1.5组。US 7,998,722和US 20110318801中公开了可使用的合适的木糖异构酶蛋白质以及编码序列的示例，其以引用的方式并入全文。

[0081] 实例

[0082] 本发明将在以下的实例中进一步阐述。应该理解，这些实例尽管说明了本发明的优选实例，但仅是以例证的方式给出的。通过上述论述和这些实例，本领域的技术人员可确定本发明的必要特征，并且在不脱离本发明的实质和范围内的前提下，可对本发明进行各种变化和修改以适应多种用途和条件。

[0083] 基本方法

[0084] 缩写的含义如下："kb"是指千碱基，"bp"指碱基对，"nt"是指核苷酸，"hr"是指小

时，“min”是指分钟，“sec”是指秒钟，“d”是指天数，“L”是指升，“ml”或“mL”是指毫升，“μL”是指微升，“μg”是指微克，“ng”是指纳克，“mg”是指毫克，“mM”是指毫摩，“μM”是指微摩，“nm”是指纳米，“μmol”是指微摩尔，“pmol”是指皮摩尔，“XI”是指木糖异构酶，“nt”是指核苷酸。

[0085] 本文使用的标准重组DNA和分子克隆技术是本领域所熟知的，并且在如下文献中有更全面的描述：Sambrook，J.，Fritsch，E.F.和Maniatis，T.，Molecular Cloning：A Laboratory Manual，第2版，Cold Spring Harbor Laboratory：Cold Spring Harbor，NY(1989)(下文称为“Maniatis”)和Silhavy，T.J.，Bennan，M.L.和Enquist，L.W.，Experiments with Gene Fusions，Cold Spring Harbor Laboratory：Cold Spring Harbor，NY(1984)；和Ausubel，F.M.等人，Current Protocols in Molecular Biology，由Greene Publishing Assoc.and Wiley-Interscience出版，Hoboken，NJ(1987)

[0086] 实例1

[0087] 启动子-GUS测试整合载体的构建

[0088] 为了将DNA整合到所述运动发酵单胞菌基因中，构建所述pZP13双交(DCO)自杀载体。图1示出它的示意图。该质粒具有pBluescript的骨架，包含大肠杆菌复制位点但是没有运动发酵单胞菌复制位点。因此，它不可以在运动发酵单胞菌中繁殖。在pZP13中靶定基因组重组的两个片段DP1341-U和DP1341-D以PCR通过运动发酵单胞菌基因组DNA合成。所提及的序列在所述运动发酵单胞菌基因序列中：Seo等人(2005)Nat.Biotechnol.23：63-68。NCBI参考号：NC_006526.2。1，951-bp DP1341-U片段(SEQ ID NO：8)包括ZMO1250编码序列的前217bp(从nt-1至nt-217)以及1，734-bp的上游序列。1，255-bp DP1341-D片段(SEQ ID NO：9)包括ZMO1250编码序列的后704bp(从nt-218至nt-921)以及551-bp的下游序列。这两个片段介导位于它们之间的转基因整合到所述运动发酵单胞菌基因组中介于nt-217与nt-218间的ZMO1250编码序列。ZMO1250是位于从1，355，129至1，356，049的运动发酵单胞菌基因序列的开放阅读框，其编码306-aa的假定蛋白。

[0089] pZP13还经构造以包含嵌合GUS表达盒，其位于所述靶定片段间，其用作测定测试启动子强度的报告基因。其由以SpeI和NsiI位点界定的测试启动子序列(Pro)、1，815-bpβ-葡糖苷酸酶编码序列(GUS；SEQ ID NO：10)和166-bp大肠杆菌araD 3’UTR(ECaraD 3’UTR；SEQ ID NO：11)构成。位于所述靶定片段间的还有1，014-bp Spec-R表达盒(SEQ ID NO：12)，其用于使用奇放线菌素选择转化的运动发酵单胞菌细胞，侧接loxPw片段。包括loxPw位点和限制位点的侧翼loxPw序列是SEQ ID NO：15和16。

[0090] 用于表达所述GUS编码区的测试启动子是所述运动发酵单胞菌甘油醛-3-磷酸脱氢酶基因(P_{gap})以及该启动子的衍生物，P_{gapU}。P_{gap}是强305-bp糖酵解启动子(SEQ ID NO：1)，在此前已用于在重组运动发酵单胞菌菌株中表达木糖利用基因。P_{gapU}是305-bp P_{gap}(SEQ ID NO：2)的衍生物，其在所述启动子序列的位置90处具有从C到T的变化。包含驱动GUS表达的这些启动子的pZP13自杀载体命名为包含P_{gap}的pZP1332(SEQ ID NO：13)以及包含P_{gapU}的pZP1337(SEQ ID NO：14)。

[0091] 实例2

[0092] 启动子-GUS表达盒的整合

[0093] 将所述测试启动子-GUS嵌合基因整合到所述野生型运动发酵单胞菌菌株ZW1

(ATCC#31821)的基因组中。通过30℃下以150rpm振荡在MRM3G5(1％酵母提取物,15mM KH_2PO_4,4mM $MgSO_4$和50g/L葡萄糖)中过夜生长ZW1来制备感受态细胞。使用Shimadzu UV-1200分光光度计(Kyoto,Japan)测量所述OD_{600}值。收集细胞并将其重悬在新鲜培养基中至OD_{600}值为0.05。细胞在同样条件下生长至早中对数期(OD_{600}接近0.5),然后收集并用冰冷的水洗涤2次并且使用冰冷的10％甘油洗涤一次。收集所得到的感受态细胞并重悬在冰冷的10％甘油中至OD_{600}值接近100。因为运动发酵单胞菌的转化需要非甲基化的DNA,所以首先将pZP1332和pZP1337转化进入大肠杆菌SCS110感受态细胞中(Stratagene,La Jolla,CA)。对于每次转化,将所转化细胞的一个菌落在10mL LB-Amp100(含100mg/L氨苄青霉素的LB液体培养基)中于37℃生长过夜。使用QIAprep Spin DNA Miniprep Kit(Qiagen,Valencia,CA)从每10mL培养物制备DNA。

[0094] 将约1μg非甲基化的质粒DNA与50μL的ZW1感受态细胞在1MM电穿孔杯(VWR,West Chester,PA)中混合。使用BT720Transporater Plus(BTX-Genetronics,San Diego,CA)在2.0KV下将所述质粒DNA电穿孔进入所述细胞中。使转化细胞于30℃下在1mL MMG5培养基(50g/L葡萄糖,10g/L酵母提取物,5g/L胰蛋白胨,2.5g/L $(NH_4)_2SO_4$,2g/L K_2HPO_4和1mM $MgSO_4$)中复苏4小时,并且于30℃下在带有AnaeroPack(Mitsubishi Gas Chemical,New York,NY)的厌氧广口瓶内的MMG5-Spec250板(具有250mg/L的奇放线菌素和15g/L的琼脂的MMG5)上生长3天。获得所述抗奇放线菌素的克隆并且划线接种到新鲜的MMG5-Spec250板上,并且在如上所述的同样条件下生长,这表明所述嵌合GUS/Spec-R转基因构建体已整合到ZW1的基因组中。为了进一步确定所述GUS表达盒的整合和表达,在14mL带盖的Falcon管中将转化子接种到5mL MRM3G5-Spec200(具有200mg/L奇放线菌素的MRM3G5)并且于30℃、150rpm振荡条件下过夜生长。将100μL细胞培养液离心沉淀于1.5mL微量离心管中。使用1mL 0.1M pH 7.0磷酸盐缓冲液洗涤所述收集的细胞并且然后重悬于100μL GUS染色液(1mL溶液包括500μL 0.2M磷酸钠缓冲液,pH 7.0,20μL 0.5M EDTA,pH 8.0,10μL 10％triton X-100,20μL 50mg/mL X-Glue二甲基甲酰胺溶液,450μL水)。通过在37℃下过夜孵育所述悬浮液中进行颜色反应。通过移除所述GUS溶液并且加入100μL 70％无水乙醇来终止该反应。根据标志性蓝色测定,很多转化子具有GUS染色阳性反应。因此,这些转化子不仅有嵌合GUS/Spec-R的构建体的整合,还同时表达GUS和Spec-R转基因。这些转化子命名为ZW1-ZP1332和ZW1-ZP1337。

[0095] 实例3

[0096] 通过GUS活性分析进行启动子的比较

[0097] 为了测量GUS活性,包含ZW1-ZP1332(ZW1-ZP1332#11和#12)或者ZW1-ZP1337(ZW1-ZP1337#1和#8)的两种转化子在2mL MRM3G5-Spec200、30℃、150rpm振荡条件下过夜培养。然后将50μL培养液加入5mL新鲜MRM3G5-Spec200并且于14mL带盖的Falcon管中于30℃、150rpm振荡条件下生长16小时直到OD_{600}接近6。通过离心收集细胞,用1mL GUS分析基准缓冲液(50mM $NaPO_4$缓冲液,pH 7.0,1mM EDTA)洗涤两次,重悬浮于1mL蛋白质提取缓冲液中(包含5％甘油的GUS分析基准缓冲液)。将各细胞悬浮液置于含裂解基质珠(BIO101,La Jolla,CA)的BIO101 2mL带盖的管中并且在第5步中在BIO101 Fastprep FP1200上20秒振荡3次。在微型离心机中离心沉淀细胞碎片和珠,并且收集上清作为蛋白质提取物。

[0098] 使用Pierce Coomassie Protein Assay Solution(Thermo Fisher Scientific

Inc.,Rockford,IL)来测定所述提取物中的蛋白质浓度。在所述分析中,将10μL标准BSA溶液(包括0.025、0.05、0.1和0.2μg/μL)以及适当稀释的蛋白质提取物与200μL的Protein Assay Solution在96孔板中混合,并且在室温下孵育5分钟。在Victor III(Perkin Elmer,Waltham,MA)上测量OD_{595}并且通过所述设备根据所得的标准曲线计算蛋白质浓度。表1是在所有提取液中蛋白质浓度的总结。

[0099] 表1细胞提取液中的蛋白质浓度

[0100]

启动子	ZW1-ZP 菌株	蛋白质浓度(ng/μL)
P_{gap}	1332-11	798
	1332-12	539

[0101]

P_{gapU}	1337-1	705
	1337-8	1,000

[0102] 根据GUS方案来调整用于GUS活力分析的方案:使用所述GUS作为基因表达的报告基因(SeanR.Gallagher编辑,Academic Press,Inc.,1992)。将其改进以用于96孔板。在所述分析中,使用所述蛋白质提取缓冲液(具有5%甘油的GUS分析基准缓冲液)将所述ZW1-ZP1332#11以及#12蛋白质提取液稀释20倍并且将所述ZW1-ZP1337#1以及#8稀释80倍。通过将3,930μL GUS分析基准缓冲液,20μL 1M DTT和50μL 100mM p-硝基苯基-β-D-葡糖苷酸(PNPG)混合来制备新鲜GUS分析工作缓冲液并且将其预热到37℃。通过将200μL 0.4M Na_2CO_3终止溶液加入到每个孔中来预备Falcon96孔平底板。在0.2mL PCR管中通过加入25μL稀释的蛋白质提取液和100μL预热的GUS工作缓冲液来组配GUS分析反应,并且随后在37℃条件下孵育。在15、30、45和60分钟,从所述反应管中移出25μL反应液并且立即在预备好的96孔板中与200μL终止溶液混合。对每个蛋白质提取液进行三组平行反应。向孔中加入20μL GUS分析工作缓冲液与5μL蛋白质提取缓冲液的混合物作为空白对照。反应完全后,在Victor III读板机中扫描所述96孔板来测定各孔中的OD_{405}。从各反应中减去所述空白对照的读数。将反映经β-葡萄糖醛酸酶所释放的4-硝基酚的平均OD_{405}相对于反应时间来作图,产生指示每分钟OD_{405}变化的斜率S(OD_{405}/分钟)。最终,基于所述反应斜率S以及加入到反应中的蛋白量来计算反应速度R(OD_{405}/ug蛋白质/分钟)。在下表2中给出各个蛋白质提取物的反应斜率(S)以及反应速度(R)。使用平均反应速度来计算比活,其指示相对于P_{gap}的启动子强度。结果证明所述启动子P_{gapU}具有的强度比所述启动子P_{gap}高约4.4倍。

[0103] 表2.P_{ganU}以及P_{gap}的相对启动子强度

[0104]

启动子	ZW1-ZP 菌株	S(OD_{405}/分钟)	R(OD_{405}/μg 蛋白质/分钟)	比活
P_{gap}	1332-11	0.0065	0.033	100%
	1332-12	0.0052	0.039	
P_{gapU}	1337-1	0.0065	0.144	442%
	1337-8	0.0113	0.174	

序列表

<110> E I du Pont de Nemours and Company

<120> 高表达发酵单胞菌启动子

<130> CL6026

<160> 16

<170> PatentIn版本3.5

<210> 1
<211> 305
<212> DNA
<213> 运动发酵单胞菌（Zymomonas mobilis）

<400> 1
gttcgatcaa caacccgaat cctatcgtaa tgatgttttg cccgatcagc ctcaatcgac 60

aattttacgc gtttcgatcg aagcagggac gacaattggc tgggaacggt atactggaat 120

aaatggtctt cgttatggta ttgatgtttt tggtgcatcg gccccggcga atgatctata 180

tgctcatttc ggcttgaccg cagtcggcat cacgaacaag gtgttggccg cgatcgccgg 240

taagtcggca cgttaaaaaa tagctatgga atataatagc tacttaataa gttaggagaa 300

taaac 305

<210> 2
<211> 305
<212> DNA
<213> 人工序列

<220>
[0001] <223> 在位置90处具有C到T的碱基变化的Pgap

<400> 2
gttcgatcaa caacccgaat cctatcgtaa tgatgttttg cccgatcagc ctcaatcgac 60

aattttacgc gtttcgatcg aagcagggat gacaattggc tgggaacggt atactggaat 120

aaatggtctt cgttatggta ttgatgtttt tggtgcatcg gccccggcga atgatctata 180

tgctcatttc ggcttgaccg cagtcggcat cacgaacaag gtgttggccg cgatcgccgg 240

taagtcggca cgttaaaaaa tagctatgga atataatagc tacttaataa gttaggagaa 300

taaac 305

<210> 3
<211> 305
<212> DNA
<213> 运动发酵单胞菌（Zymomonas mobilis）

<400> 3
gttcgatcaa caacccgaat cctatcgtaa tgatgttttg cccgatcagc ctcaatcgac 60

aattttacgc gtttcgatcg aagcagggac gacaattggc tgggaacggt atactggaat 120

aaatggtctt cgttatggta ttgatgtttt tggtgcatcg gccccggcga atgatctata 180

tgctcatttc ggcttgaccg cagtcggcat cacgaacaag gtgttggccg cgatcgccgg 240

taagtcggca cgttaaaaaa tagctatgga atatagtagc tacttaataa gttaggagaa 300

taaac 305

<210> 4
<211> 305
<212> DNA
<213> 人工序列

<220>
<223> 在ZM4序列中位置90处具有C到T的变化的Pgap

<400> 4
gttcgatcaa caacccgaat cctatcgtaa tgatgttttg cccgatcagc ctcaatcgac 60

aattttacgc gtttcgatcg aagcagggat gacaattggc tgggaacggt atactggaat 120

aaatggtctt cgttatggta ttgatgtttt tggtgcatcg gccccggcga atgatctata 180

tgctcatttc ggcttgaccg cagtcggcat cacgaacaag gtgttggccg cgatcgccgg 240

taagtcggca cgttaaaaaa tagctatgga atatagtagc tacttaataa gttaggagaa 300

taaac 305

<210> 5
<211> 304
<212> DNA
<213> 人工序列

<220>
<223> 在pZB4中的Pgap - 在位置285处具有T缺失

<400> 5
gttcgatcaa caacccgaat cctatcgtaa tgatgttttg cccgatcagc ctcaatcgac 60

aattttacgc gtttcgatcg aagcagggac gacaattggc tgggaacggt atactggaat 120

aaatggtctt cgttatggta ttgatgtttt tggtgcatcg gccccggcga atgatctata 180

tgctcatttc ggcttgaccg cagtcggcat cacgaacaag gtgttggccg cgatcgccgg 240

taagtcggca cgttaaaaaa tagctatgga atataatagc tactaataag ttaggagaat 300

aaac 304

[0002]

<210> 6
<211> 304
<212> DNA
<213> 人工序列

<220>
<223> 在位置90处具有C到T的变化的pZB4 Pgap

<400> 6
gttcgatcaa caacccgaat cctatcgtaa tgatgttttg cccgatcagc ctcaatcgac 60

aattttacgc gtttcgatcg aagcagggat gacaattggc tgggaacggt atactggaat 120

aaatggtctt cgttatggta ttgatgtttt tggtgcatcg gccccggcga atgatctata 180

tgctcatttc ggcttgaccg cagtcggcat cacgaacaag gtgttggccg cgatcgccgg 240

taagtcggca cgttaaaaaa tagctatgga atataatagc tactaataag ttaggagaat 300

aaac 304

<210> 7
<211> 304
<212> DNA
<213> 人工序列

<220>
<223> ZM4的Pgap，其在位置285处具有T缺失并且在位置90处具有C到T的变化

<400> 7
gttcgatcaa caacccgaat cctatcgtaa tgatgttttg cccgatcagc ctcaatcgac 60

aattttacgc gtttcgatcg aagcagggat gacaattggc tgggaacggt atactggaat 120

aaatggtctt cgttatggta ttgatgtttt tggtgcatcg gccccggcga atgatctata 180

tgctcatttc ggcttgaccg cagtcggcat cacgaacaag gtgttggccg cgatcgccgg 240

taagtcggca cgttaaaaaa tagctatgga atatagtagc tactaataag ttaggagaat 300

aaac 304

<210> 8
<211> 1951
<212> DNA
<213> 运动发酵单胞菌（Zymomonas mobilis）

<400> 8
aactgccacc acccataaat cccagcgggt ttattttaaa aataaaacaa atgataacca 60

aaacgatcac gccgccgata ccgaaacggc caacgacaaa gccgagtaag ccgctaatca 120

ggttaaaccc gccgccaccg ccgttattac caccgctaaa gtcttcacca cgttcgtctt 180

cgacgttatc gctgcttcgg tagtcatcaa gccgcattat gcaccttccc aattttaccc 240

gcaggcctat ttataggccg acagcgggtt ttatctatag ctttcatttt tttgtgtttt 300

tttgaaaaaa aacgcttgca gaacacatag ctgattgtta tacgcgcttc acacccaacg 360

cggtgattga catagtgttc cctgatagct cagcggtaga gcgtccgact gttaatcggc 420

aggtcgtagg ttcgaatcct actcggggag ccatttcaac gatagtcctt atagttttat 480

tttctcttta aaaatttatc ttatttccga tcaatcggtt atatttcgcc tgatagattc 540

aagacattct gtattgggat tgtctgaata tatttcgggt agatctgttt aaagtggttt 600

ttttgtccac tttcctaatt aggttatttc cataacgtga cccaatagcg tggccgtgag 660

cgcttggctg ccgcggaata gggatatttt ctttttcga gataccgcca gccattaagc 720

[0003] cattaagcca ttaagccatt aagccattaa gccattaagc cattaagcca ttaagatatc 780

gctactcatt ctttcggcaa ggaaaaaggt gattagtcgt tatttggggt aataaatgtt 840

tatcctcgga tcgagatttc ttgaataaag aaggacgagt ggttttcttc atttgagcct 900

tttttctttt ttgctttttc tatttcgctg gaaacacccc ttttcaccta ttccagatag 960

attgtccccc ttcgcttttt taaagcattg atcgtagaaa tcactccgat ggtcgtatga 1020

cagcgtttcc tttggaaagc ggggttttac agcctttatc aatagggtca ggatgcggga 1080

atcatcctct atatggagaa gtatgatgtc ccctatttcg ataccctcta tccgtctgaa 1140

tgacggcaat gatttgcctg cggtaggctt tggcacctat aaactcaatg gttcggccgg 1200

tgtttccgat atcgttagcg cgattaaagt cggataccgc ttgctcgatt ctgcctttaa 1260

ttatgagaat gaaggtgctg taggcgaagc tgttcgcgag gccggtattg cccgtgataa 1320

attgcggatc gtgtccaaat taccggggcg acatcatcat tttgaagagg ccattgcaac 1380

ggtcgaggag tcactttatc gggcacagct tgattattat gatctctatc tgattcattg 1440

gcctaatccg agcaaagatc tctatgtcga ggcgtggcag gccttgattg aagcccgtaa 1500

aaaagggtta atccgttcga ttggtgtctg taacttttta ccagaacatc tggaaaggct 1560

gattaaagag acgggtgtaa cgccggttgt caatcaggtg gaattgcatc cctattttcc 1620

gcaggaagag cagcgggcat gggataaagc ccatggcatt gttaccgagt catggagtcc 1680

tttgggacgt gccagcaagt tgctacagga cgacacgatt aaaaaaattg cggatcgctt 1740

gggtaagtct atcccgcaag ttattttgcg ttggcatgtc cagcttggag cgataccgat 1800

tcctaaagct tctagcaaag agcgtcagat tgaaaatcta tccttgtttg atttcgagct 1860

gtcgccgcag gatgtagaaa taatcgcgac tttagcacgg cctgatggac gcttggctga 1920

ccaagaccct gcccgctatg aagaatttta a 1951

<210> 9
<211> 1255
<212> DNA
<213> 运动发酵单胞菌（Zymomonas mobilis）

<400> 9
aataatcccg acagcagttg cattatgatg ggttaaatgc accagaacta gccagtcctg 60

tgcaatgcgt tgcatccatg tcccgatatt ggacacgaac attccactta tccaaaggcg 120

ataattgaag ctgttcagtg accggaaggt tccatggaaa agtttttca tcagaggctc 180

gctctatttt tacatgccga gcagaagcgc cttgatccgc tagcatccca tattgggagg 240

atgtttttgt gcgccagaaa agcggaaaaa atatctacga tcatagacgg tcgtcagatt 300

ttccggtgat tggcttgttc aagctatgcg tcaagagcgg gatatcaagc gtcatgataa 360

gaaaaaggat acaggataga cgcttttgaa aaaggccatc gctgtatcct aagagtattt 420

tagttgataa tttttactaa ataaggcgta taaccttatg cttttacggt aaaatattta 480

ttatttttta ttgaccgaat atgtgttttt cgctttcgca aagacatgaa taccatttca 540

tctcaacaga atttaactat caatatcatg atggaaggta tcgcaaacgg ctggatcgcc 600

gctatcaatc ccgcgtttga gccagcgcat tctttccgca gatgtgccat gagtgaaact 660

ttcaggaacg gggcgcatcc ctgcggcttt ctctaaggta tcgtcgccaa tagcctgcgc 720

tgcccgcata ccttcttcga catcgccctc ttctatccgg tcgcggttat gcgctgccca 780

[0004] gactccggca tagcaatcgg cctgtaattc cattctcacc tgtagggcat tagcggcagc 840

tttgccgaca ttagcttgtt gttcgcggat ttggtcggat aagcccatta aatcttgcac 900

atgatgcccg acttcatggg cgataacata ggcttgggca aaatcaccgc tggcaggaag 960

cggttgacca gttcattata gaaatctgta tcgagataaa cgcgctgatc tgctgggcag 1020

taaaaaggcc ccatcgccga ttctgctgca ccgcagcctg aacgacctct ttgactataa 1080

aagaccaatg ttgtcggttt atatttttta ccagcatcag cgaaaatctt gccccatgtg 1140

tcttcggtac tagccaagac acggcaggag aaaaggcttg cggcgttaag ttggcaggcc 1200

tgttgtgccg aacgaacaga atgcacctga tgaggtgcca tcgcgggttg tgtcg 1255

<210> 10
<211> 1815
<212> DNA
<213> 大肠杆菌（Escherichia coli）

<400> 10
atgcatgtac gtcctgtaga aaccccaacc cgtgaaatca aaaaactcga cggcctgtgg 60

gcattcagtc tggatcgcga aaactgtgga attgatcagc gttggtggga aagcgcgtta 120

caagaaagcc gggcaattgc tgtgccaggc agttttaacg atcagttcgc cgatgcagat 180

attcgtaatt atgcgggcaa cgtctggtat cagcgcgaag tctttatacc gaaaggttgg 240

gcaggccagc gtatcgtgct gcgtttcgat gcggtcactc attacggcaa agtgtgggtc 300

aataatcagg aagtgatgga gcatcagggc ggctatacgc catttgaagc cgatgtcacg 360

ccgtatgtta ttgccgggaa aagtgtacgt atcaccgttt gtgtgaacaa cgaactgaac 420

tggcagacta tcccgccggg aatggtgatt accgacgaaa acggcaagaa aaagcagtct 480

tacttccatg atttctttaa ctatgccgga atccatcgca gcgtaatgct ctacaccacg 540

ccgaacacct gggtggacga tatcaccgtg gtgacgcatg tcgcgcaaga ctgtaaccac 600

gcgtctgttg actggcaggt ggtggccaat ggtgatgtca gcgttgaact gcgtgatgcg 660

gatcaacagg tggttgcaac tggacaaggc actagcggga ctttgcaagt ggtgaatccg 720

cacctctggc aaccgggtga aggttatctc tatgaactgt gcgtcacagc caaaagccag 780

acagagtgtg atatctaccc gcttcgcgtc ggcatccggt cagtggcagt gaagggcgaa 840

cagttcctga ttaaccacaa accgttctac tttactggct ttggtcgtca tgaagatgcg 900

gacttgcgtg gcaaaggatt cgataacgtg ctgatggtgc acgaccacgc attaatggac 960

tggattgggg ccaactccta ccgtacctcg cattaccctt acgctgaaga gatgctcgac 1020

tgggcagatg aacatggcat cgtggtgatt gatgaaactg ctgctgtcgg ctttaacctc 1080

tctttaggca ttggtttcga agcgggcaac aagccgaaag aactgtacag cgaagaggca 1140

gtcaacgggg aaactcagca agcgcactta caggcgatta aagagctgat agcgcgtgac 1200

aaaaaccacc caagcgtggt gatgtggagt attgccaacg aaccggatac ccgtccgcaa 1260

ggtgcacggg aatatttcgc gccactggcg gaagcaacgc gtaaactcga cccgacgcgt 1320

ccgatcacct gcgtcaatgt aatgttctgc gacgctcaca ccgataccat cagcgatctc 1380

tttgatgtgc tgtgcctgaa ccgttattac ggatggtatg tccaaagcgg cgatttggaa 1440

acggcagaga aggtactgga aaaagaactt ctggcctggc aggagaaact gcatcagccg 1500

attatcatca ccgaatacgg cgtggatacg ttagccgggc tgcactcaat gtacaccgac 1560

atgtggagtg aagagtatca gtgtgcatgg ctggatatgt atcaccgcgt ctttgatcgc 1620

[0005]

gtcagcgccg tcgtcggtga acaggtatgg aatttcgccg attttgcgac ctcgcaaggc 1680

atattgcgcg ttggcggtaa caagaaaggg atcttcactc gcgaccgcaa accgaagtcg 1740

gcggctttc tgctgcaaaa acgctggact ggcatgaact tcggtgaaaa accgcagcag 1800

ggaggcaaac aatga 1815

<210> 11
<211> 166
<212> DNA
<213> 大肠杆菌（Escherichia coli）

<400> 11
tgactgtata aaaccacagc caatcaaacg aaaccaggct atactcaagc ctggtttttt 60

gatggatttt cagcgtggcg caggcaggtt ttatcttaac ccgacactgg cgggacaccc 120

cgcaagggac agaagtctcc ttctggctgg cgacggacaa cgggcc 166

<210> 12
<211> 1014
<212> DNA
<213> 人工序列

<220>
<223> Spec-R表达盒

<400> 12
agcacaggat gacgcctaac aattcattca agccgacacc gcttcgcggc gcggcttaat 60

tcaggagtta aacatcatga gggaagcggt gatcgccgaa gtatcgactc aactatcaga 120

ggtagttggc gtcatcgagc gccatctcga accgacgttg ctggccgtac atttgtacgg 180

ctccgcagtg gatggcggcc tgaagccaca cagtgatatt gatttgctgg ttacggtgac 240

tgtaaggctt gatgaaacaa cgcggcgagc tttgatcaac gacctttttgg aaacttcggc 300

ttcccctgga gagagcgaga ttctccgcgc tgtagaagtc accattgttg tgcacgacga 360

catcattccg tggcgttatc cagctaagcg cgaactgcaa tttggagaat ggcagcgcaa 420

tgacattctt gcaggtatct tcgagccagc cacgatcgac attgatctgg ctatcttgct 480

gacaaaagca agagaacata gcgttgcctt ggtaggtcca gcggcggagg aactctttga 540

tccggttcct gaacaggatc tatttgaggc gctaaatgaa accttaacgc tatggaactc 600

gccgcccgac tgggctggcg atgagcgaaa tgtagtgctt acgttgtccc gcatttggta 660

cagcgcagta accggcaaaa tcgcgccgaa ggatgtcgct gccgactggg caatggagcg 720

cctgccggcc cagtatcagc ccgtcatact tgaagctagg caggcttatc ttggacaaga 780

agatcgcttg gcctcgcgcg cagatcagtt ggaagaattt gttcactacg tgaaaggcga 840

gatcaccaag gtagtcggca aataatgtct aacaattcgt tcaagccgac gccgcttcgc 900

ggcgcggctt aactcaagcg ttagagagct ggggaagact atgcgcgatc tgttgaaggt 960

ggttctaagc ctcgtacttg cgatggcatc ggggcaggca cttgctgacc tgcc 1014

<210> 13
<211> 9512
<212> DNA
<213> 人工序列

<220>
<223> 构建的质粒

<400> 13

[0006]

ctagttcgat caacaacccg aatcctatcg taatgatgtt ttgcccgatc agcctcaatc 60

gacaatttta cgcgtttcga tcgaagcagg gacgacaatt ggctgggaac ggtatactgg 120

aataaatggt cttcgttatg gtattgatgt ttttggtgca tcggccccgg cgaatgatct 180

atatgctcat ttcggcttga ccgcagtcgg catcacgaac aaggtgttgg ccgcgatcgc 240

cggtaagtcg gcacgttaaa aaatagctat ggaatataat agctacttaa taagttagga 300

gaataaacat gcatgtacgt cctgtagaaa ccccaacccg tgaaatcaaa aaactcgacg 360

gcctgtgggc attcagtctg gatcgcgaaa actgtggaat tgatcagcgt tggtgggaaa 420

gcgcgttaca agaaagccgg gcaattgctg tgccaggcag ttttaacgat cagttcgccg 480

atgcagatat tcgtaattat gcgggcaacg tctggtatca gcgcgaagtc tttataccga 540

aaggttgggc aggccagcgt atcgtgctgc gtttcgatgc ggtcactcat tacggcaaag 600

tgtgggtcaa taatcaggaa gtgatggagc atcagggcgg ctatacgcca tttgaagccg 660

atgtcacgcc gtatgttatt gccgggaaaa gtgtacgtat caccgtttgt gtgaacaacg 720

aactgaactg gcagactatc ccgccgggaa tggtgattac cgacgaaaac ggcaagaaaa 780

agcagtctta cttccatgat ttctttaact atgccggaat ccatcgcagc gtaatgctct 840

acaccacgcc gaacacctgg gtggacgata tcaccgtggt gacgcatgtc gcgcaagact 900

gtaaccacgc gtctgttgac tggcaggtgg tggccaatgg tgatgtcagc gttgaactgc 960

gtgatgcgga tcaacaggtg gttgcaactg gacaaggcac tagcgggact ttgcaagtgg 1020

tgaatccgca cctctggcaa ccgggtgaag gttatctcta tgaactgtgc gtcacagcca 1080

aaagccagac agagtgtgat atctacccgc ttcgcgtcgg catccggtca gtggcagtga 1140

agggcgaaca gttcctgatt aaccacaaac cgttctactt tactggcttt ggtcgtcatg 1200

[0007]

aagatgcgga cttgcgtggc aaaggattcg ataacgtgct gatggtgcac gaccacgcat 1260

taatggactg gattggggcc aactcctacc gtacctcgca ttacccttac gctgaagaga 1320

tgctcgactg ggcagatgaa catggcatcg tggtgattga tgaaactgct gctgtcggct 1380

ttaacctctc tttaggcatt ggtttcgaag cgggcaacaa gccgaaagaa ctgtacagcg 1440

aagaggcagt caacggggaa actcagcaag cgcacttaca ggcgattaaa gagctgatag 1500

cgcgtgacaa aaaccaccca agcgtggtga tgtggagtat tgccaacgaa ccggataccc 1560

gtccgcaagg tgcacgggaa tatttcgcgc cactggcgga agcaacgcgt aaactcgacc 1620

cgacgcgtcc gatcacctgc gtcaatgtaa tgttctgcga cgctcacacc gataccatca 1680

gcgatctctt tgatgtgctg tgcctgaacc gttattacgg atggtatgtc caaagcggcg 1740

atttggaaac ggcagagaag gtactggaaa aagaacttct ggcctggcag gagaaactgc 1800

atcagccgat tatcatcacc gaatacggcg tggatacgtt agccgggctg cactcaatgt 1860

acaccgacat gtggagtgaa gagtatcagt gtgcatggct ggatatgtat caccgcgtct 1920

ttgatcgcgt cagcgccgtc gtcggtgaac aggtatggaa tttcgccgat tttgcgacct 1980

cgcaaggcat attgcgcgtt ggcggtaaca agaaagggat cttcactcgc gaccgcaaac 2040

cgaagtcggc ggcttttctg ctgcaaaaac gctggactgg catgaacttc ggtgaaaaac 2100

cgcagcaggg aggcaaacaa tgatctagat gactgtataa aaccacagcc aatcaaacga 2160

aaccaggcta tactcaagcc tggttttttg atggatttcc agcgtggcgc aggcaggttt 2220

tatcttaacc cgacactggc gggacacccc gcaagggaca gaagtctcct tctggctggc 2280

gacggacaac gggccaagct tggaagggcc ggccaagctt gaattcgcga tcgcataact 2340

tcgtataatg tatgctatac gaagttatgc ggccgcagca caggatgacg cctaacaatt 2400

cattcaagcc gacaccgctt cgcggcgcgg cttaattcag gagttaaaca tcatgaggga 2460

agcggtgatc gccgaagtat cgactcaact atcagaggta gttggcgtca tcgagcgcca 2520

tctcgaaccg acgttgctgg ccgtacattt gtacggctcc gcagtggatg gcggcctgaa 2580

gccacacagt gatattgatt tgctggttac ggtgactgta aggcttgatg aaacaacgcg 2640

gcgagctttg atcaacgacc ttttggaaac ttcggcttcc cctggagaga gcgagattct 2700

ccgcgctgta gaagtcacca ttgttgtgca cgacgacatc attccgtggc gttatccagc 2760

taagcgcgaa ctgcaatttg gagaatggca gcgcaatgac attcttgcag gtatcttcga 2820

gccagccacg atcgacattg atctggctat cttgctgaca aaagcaagag aacatagcgt 2880

tgccttggta ggtccagcgg cggaggaact ctttgatccg gttcctgaac aggatctatt 2940

tgaggcgcta aatgaaacct taacgctatg gaactcgccg cccgactggg ctggcgatga 3000

gcgaaatgta gtgcttacgt tgtcccgcat ttggtacagc gcagtaaccg gcaaaatcgc 3060

gccgaaggat gtcgctgccg actgggcaat ggagcgcctg ccggcccagt atcagcccgt 3120

catacttgaa gctaggcagg cttatcttgg acaagaagat cgcttggcct cgcgcgcaga 3180

tcagttggaa gaatttgttc actacgtgaa aggcgagatc accaaggtag tcggcaaata 3240

atgtctaaca attcgttcaa gccgacgccg cttcgcggcg cggcttaact caagcgttag 3300

agagctgggg aagactatgc gcgatctgtt gaaggtggtt ctaagcctcg tacttgcgat 3360

ggcatcgggg caggcacttg ctgacctgcc ttaattaaat aacttcgtat aatgtatgct 3420

atacgaagtt atggccggcc aactgccacc acccataaat cccagcgggt ttattttaaa 3480

aataaaacaa atgataacca aaacgatcac gccgccgata ccgaaacggc caacgacaaa 3540

gccgagtaag ccgctaatca ggttaaaccc gccgccaccg ccgttattac caccgctaaa 3600

gtcttcacca cgttcgtctt cgacgttatc gctgcttcgg tagtcatcaa gccgcattat 3660

gcaccttccc aattttaccc gcaggcctat ttataggccg acagcgggtt ttatctatag 3720

ctttcatttt tttgtgtttt tttgaaaaaa aacgcttgca gaacacatag ctgattgtta 3780

tacgcgcttc acacccaacg cggtgattga catagtgttc cctgatagct cagcggtaga 3840

gcgtccgact gttaatcggc aggtcgtagg ttcgaatcct actcggggag ccatttcaac 3900

gatagtcctt atagttttat tttctcttta aaaatttatc ttatttccga tcaatcggtt 3960

atatttcgcc tgatagattc aagacattct gtattgggat tgtctgaata tatttcgggt 4020

agatctgttt aaagtggttt ttttgtccac tttcctaatt aggttatttc cataacgtga 4080

cccaatagcg tggccgtgag cgcttggctg ccgcggaata gggatatttt ctttttcga 4140

gataccgcca gccattaagc cattaagcca ttaagccatt aagccattaa gccattaagc 4200

cattaagcca ttaagatatc gctactcatt ctttcggcaa ggaaaaaggt gattagtcgt 4260

tatttggggt aataaatgtt tatcctcgga tcgagatttc ttgaataaag aaggacgagt 4320

ggttttcttc atttgagcct tttttctttt ttgcttttc tatttcgctg gaaacacccc 4380

ttttcaccta ttccagatag attgtcccct ttcgcttttt taaagcattg atcgtagaaa 4440

tcactccgat ggtcgtatga cagcgtttc tttggaaagc ggggttttac agcctttatc 4500

aatagggtca ggatggggga atcatcctct atatggagaa gtatgatgtc ccctatttcg 4560

ataccctcta tccgtctgaa tgacggcaat gatttgcctg cggtaggctt tggcacctat 4620

[0008]

aaactcaatg gttcggccgg tgtttccgat atcgttagcg cgattaaagt cggataccgc 4680

ttgctcgatt ctgcctttaa ttatgagaat gaaggtgctg taggcgaagc tgttcgcgag 4740

gccggtattg cccgtgataa attgcggatc gtgtccaaat taccggggcg acatcatcat 4800

tttgaagagg ccattgcaac ggtcgaggag tcactttatc gggcacagct tgattattat 4860

gatctctatc tgattcattg gcctaatccg agcaaagatc tctatgtcga ggcgtggcag 4920

gccttgattg aagcccgtaa aaaagggtta atccgttcga ttggtgtctg taacttttta 4980

ccagaacatc tggaaaggct gattaaagag acgggtgtaa cgccggttgt caatcaggtg 5040

gaattgcatc cctattttcc gcaggaagag cagcgggcat gggataaagc ccatggcatt 5100

gttaccgagt catggagtcc tttgggacgt gccagcaagt tgctacagga cgacacgatt 5160

aaaaaaattg cggatcgctt gggtaagtct atcccgcaag ttattttgcg ttggcatgtc 5220

cagcttggag cgataccgat tcctaaagct tctagcaaag agcgtcagat tgaaaatcta 5280

tccttgtttg atttcgagct gtcgccgcag gatgtagaaa taatcgcgac tttagcacgg 5340

cctgatggac gcttggctga ccaagaccct gcccgctatg aagaatttta aggtacccaa 5400

ttcgccctat agtgagtcgt attacgcgcg ctcactggcc gtcgttttac aacgtcgtga 5460

ctgggaaaac cctggcgtta cccaacttaa tcgccttgca gcacatcccc ctttcgccag 5520

ctggcgtaat agcgaagagg cccgcaccga tcgcccttcc caacagttgc gcagcctgaa 5580

tggcgaatgg gacgcgccct gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg 5640

cagcgtgacc gctacacttg ccagcgccct agcgcccgct cctttcgctt tcttcccttc 5700

ctttctcgcc acgttcgccg gctttccccg tcaagctcta aatcgggggc tccctttagg 5760

[0009]

gttccgattt agtgctttac ggcacctcga ccccaaaaaa cttgattagg gtgatggttc 5820
acgtagtggg ccatcgccct gatagacggt ttttcgccct ttgacgttgg agtccacgtt 5880
ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct cggtctattc 5940
ttttgattta taagggattt tgccgatttc ggcctattgg ttaaaaaatg agctgattta 6000
acaaaaattt aacgcgaatt ttaacaaaat attaacgctt acaatttagg tggcactttt 6060
cggggaaatg tgcgcggaac ccctatttgt ttatttttct aaatacattc aaatatgtat 6120
ccgctcatga gacaataacc ctgataaatg cttcaataat attgaaaaag gaagagtatg 6180
agtattcaac atttccgtgt cgcccttatt cccttttttg cggcattttg ccttcctgtt 6240
tttgctcacc cagaaacgct ggtgaaagta aaagatgctg aagatcagtt gggtgcacga 6300
gtgggttaca tcgaactgga tctcaacagc ggtaagatcc ttgagagttt tcgccccgaa 6360
gaacgttttc caatgatgag cacttttaaa gttctgctat gtggcgcggt attatcccgt 6420
attgacgccg ggcaagagca actcggtcgc cgcatacact attctcagaa tgacttggtt 6480
gagtactcac cagtcacaga aaagcatctt acggatggca tgacagtaag agaattatgc 6540
agtgctgcca taaccatgag tgataacact gcggccaact tacttctgac aacgatcgga 6600
ggaccgaagg agctaaccgc ttttttgcac aacatggggg atcatgtaac tcgccttgat 6660
cgttgggaac cggagctgaa tgaagccata ccaaacgacg agcgtgacac cacgatgcct 6720
gtagcaatgg caacaacgtt gcgcaaacta ttaactggcg aactacttac tctagcttcc 6780
cggcaacaat taatagactg gatggaggcg gataaagttg caggaccact tctgcgctcg 6840
gcccttccgg ctggctggtt tattgctgat aaatctggag ccggtgagcg tgggtctcgc 6900
ggtatcattg cagcactggg gccagatggt aagccctccc gtatcgtagt tatctacacg 6960
acggggagtc aggcaactat ggatgaacga aatagacaga tcgctgagat aggtgcctca 7020
ctgattaagc attggtaact gtcagaccaa gtttactcat atatacttta gattgattta 7080
aaacttcatt tttaatttaa aaggatctag gtgaagatcc tttttgataa tctcatgacc 7140
aaaatccctt aacgtgagtt ttcgttccac tgagcgtcag accccgtaga aaagatcaaa 7200
ggatcttctt gagatccttt ttttctgcgc gtaatctgct gcttgcaaac aaaaaaacca 7260
ccgctaccag cggtggtttg tttgccggat caagagctac caactctttt tccgaaggta 7320
actggcttca gcagagcgca gataccaaat actgtccttc tagtgtagcc gtagttaggc 7380
caccacttca agaactctgt agcaccgcct acatacctcg ctctgctaat cctgttacca 7440
gtggctgctg ccagtggcga taagtcgtgt cttaccgggt tggactcaag acgatagtta 7500
ccggataagg cgcagcggtc gggctgaacg gggggttcgt gcacacagcc cagcttggag 7560
cgaacgacct acaccgaact gagataccta cagcgtgagc tatgagaaag cgccacgctt 7620
cccgaaggga gaaaggcgga caggtatccg gtaagcggca gggtcggaac aggagagcgc 7680
acgagggagc ttccaggggg aaacgcctgg tatctttata gtcctgtcgg gtttcgccac 7740
ctctgacttg agcgtcgatt tttgtgatgc tcgtcagggg ggcggagcct atggaaaaac 7800
gccagcaacg cggccttttt acggttcctg gccttttgct ggccttttgc tcacatgttc 7860
tttcctgcgt tatcccctga ttctgtggat aaccgtatta ccgcctttga gtgagctgat 7920
accgctcgcc gcagccgaac gaccgagcgc agcgagtcag tgagcgagga agcggaagag 7980
cgcccaatac gcaaaccgcc tctccccgcg cgttggccga ttcattaatg cagctggcac 8040

gacaggtttc ccgactggaa agcgggcagt gagcgcaacg caattaatgt gagttagctc 8100

actcattagg caccccaggc tttacacttt atgcttccgg ctcgtatgtt gtgtggaatt 8160

gtgagcggat aacaatttca cacaggaaac agctatgacc atgattacgc caagcgcgca 8220

attaaccctc actaaaggga acaaaagctg gagctcaata atcccgacag cagttgcatt 8280

atgatgggtt aaatgcacca gaactagcca gtcctgtgca atgcgttgca tccatgtccc 8340

gatattggac acgaacattc cacttatcca aaggcgataa ttgaagctgt tcagtgaccg 8400

gaaggttcca tggaaaagtt ttttcatcag aggctcgctc tatttttaca tgccgagcag 8460

aagcgccttg atccgctagc atcccatatt gggaggatgt ttttgtgcgc cagaaaagcg 8520

gaaaaaatat ctacgatcat agacggtcgt cagattttcc ggtgattggc ttgttcaagc 8580

tatgcgtcaa gagcgggata tcaagcgtca tgataagaaa aaggatacag gatagacgct 8640

tttgaaaaag gccatcgctg tatcctaaga gtattttagt tgataatttt tactaaataa 8700

ggcgtataac cttatgcttt tacggtaaaa tatttattat tttttattga ccgaatatgt 8760

gtttttcgct ttcgcaaaga catgaatacc atttcatctc aacagaattt aactatcaat 8820

atcatgatgg aaggtatcgc aaacggctgg atcgccgcta tcaatcccgc gtttgagcca 8880

gcgcattctt tccgcagatg tgccatgagt gaaactttca ggaacggggc gcatccctgc 8940

ggctttctct aaggtatcgt cgccaatagc ctgcgctgcc cgcatacctt cttcgacatc 9000

gccctcttct atccggtcgc ggttatgcgc tgcccagact ccggcatagc aatcggcctg 9060

taattccatt ctcacctgta gggcattagc ggcagctttg ccgacattag cttgttgttc 9120

gcggatttgg tcggataagc ccattaaatc ttgcacatga tgcccgactt catgggcgat 9180

[0010]

aacataggct tgggcaaaat caccgctggc aggaagcggt tgaccagttc attatagaaa 9240

tctgtatcga gataaacgcg ctgatctgct gggcagtaaa aaggccccat cgccgattct 9300

gctgcaccgc agcctgaacg acctctttga ctataaaaga ccaatgttgt cggtttatat 9360

tttttaccag catcagcgaa aatcttgccc catgtgtctt cggtactagc caagacacgg 9420

caggagaaaa ggcttgcggc gttaagttgg caggcctgtt gtgccgaacg aacagaatgc 9480

acctgatgag gtgccatcgc gggttgtgtc ga 9512

<210> 14
<211> 9512
<212> DNA
<213> 人工序列

<220>
<223> 构建的质粒

<400> 14
ctagttcgat caacaacccg aatcctatcg taatgatgtt ttgcccgatc agcctcaatc 60

gacaatttta cgcgtttcga tcgaagcagg gatgacaatt ggctgggaac ggtatactgg 120

aataaatggt cttcgttatg gtattgatgt ttttggtgca tcggccccgg cgaatgatct 180

atatgctcat ttcggcttga ccgcagtcgg catcacgaac aaggtgttgg ccgcgatcgc 240

cggtaagtcg gcacgttaaa aaatagctat ggaatataat agctacttaa taagttagga 300

gaataaacat gcatgtacgt cctgtagaaa ccccaacccg tgaaatcaaa aaactcgacg 360

gcctgtgggc attcagtctg gatcgcgaaa actgtggaat tgatcagcgt tggtgggaaa 420

gcgcgttaca agaaagccgg gcaattgctg tgccaggcag ttttaacgat cagttcgccg 480

[0011]

atgcagatat tcgtaattat gcgggcaacg tctggtatca gcgcgaagtc tttataccga 540
aaggttgggc aggccagcgt atcgtgctgc gtttcgatgc ggtcactcat tacggcaaag 600
tgtgggtcaa taatcaggaa gtgatggagc atcagggcgg ctatacgcca tttgaagccg 660
atgtcacgcc gtatgttatt gccgggaaaa gtgtacgtat caccgtttgt gtgaacaacg 720
aactgaactg gcagactatc ccgccgggaa tggtgattac cgacgaaaac ggcaagaaaa 780
agcagtctta cttccatgat ttctttaact atgccggaat ccatcgcagc gtaatgctct 840
acaccacgcc gaacacctgg gtggacgata tcaccgtggt gacgcatgtc gcgcaagact 900
gtaaccacgc gtctgttgac tggcaggtgg tggccaatgg tgatgtcagc gttgaactgc 960
gtgatgcgga tcaacaggtg gttgcaactg gacaaggcac tagcgggact ttgcaagtgg 1020
tgaatccgca cctctggcaa ccgggtgaag gttatctcta tgaactgtgc gtcacagcca 1080
aaagccagac agagtgtgat atctacccgc ttcgcgtcgg catccggtca gtggcagtga 1140
agggcgaaca gttcctgatt aaccacaaac cgttctactt tactggcttt ggtcgtcatg 1200
aagatgcgga cttgcgtggc aaaggattcg ataacgtgct gatggtgcac gaccacgcat 1260
taatggactg gattggggcc aactcctacc gtacctcgca ttacccttac gctgaagaga 1320
tgctcgactg ggcagatgaa catggcatcg tggtgattga tgaaactgct gctgtcggct 1380
ttaacctctc tttaggcatt ggtttcgaag cgggcaacaa gccgaaagaa ctgtacagcg 1440
aagaggcagt caacggggaa actcagcaag cgcacttaca ggcgattaaa gagctgatag 1500
cgcgtgacaa aaaccaccca agcgtggtga tgtggagtat tgccaacgaa ccggataccc 1560
gtccgcaagg tgcacgggaa tatttcgcgc cactggcgga agcaacgcgt aaactcgacc 1620
cgacgcgtcc gatcacctgc gtcaatgtaa tgttctgcga cgctcacacc gataccatca 1680
gcgatctctt tgatgtgctg tgcctgaacc gttattacgg atggtatgtc caaagcggcg 1740
atttggaaac ggcagagaag gtactggaaa aagaacttct ggcctggcag gagaaactgc 1800
atcagccgat tatcatcacc gaatacggcg tggatacgtt agccgggctg cactcaatgt 1860
acaccgacat gtggagtgaa gagtatcagt gtgcatggct ggatatgtat caccgcgtct 1920
ttgatcgcgt cagcgccgtc gtcggtgaac aggtatggaa tttcgccgat tttgcgacct 1980
cgcaaggcat attgcgcgtt ggcggtaaca agaaagggat cttcactcgc gaccgcaaac 2040
cgaagtcggc ggcttttctg ctgcaaaaac gctggactgg catgaacttc ggtgaaaaac 2100
cgcagcaggg aggcaaacaa tgatctagat gactgtataa aaccacagcc aatcaaacga 2160
aaccaggcta tactcaagcc tggtttttg atggatttc agcgtggcgc aggcaggttt 2220
tatcttaacc cgacactggc gggacacccc gcaagggaca gaagtctcct tctggctggc 2280
gacggacaac gggccaagct tggaagggcc ggccaagctt gaattcgcga tcgcataact 2340
tcgtataatg tatgctatac gaagttatgc ggccgcagca caggatgacg cctaacaatt 2400
cattcaagcc gacaccgctt cgcggcgcgg cttaattcag gagttaaaca tcatgaggga 2460
agcggtgatc gccgaagtat cgactcaact atcagaggta gttggcgtca tcgagcgcca 2520
tctcgaaccg acgttgctgg ccgtacattt gtacggctcc gcagtggatg gcggcctgaa 2580
gccacacagt gatattgatt tgctggttac ggtgactgta aggcttgatg aaacaacgcg 2640
gcgagctttg atcaacgacc ttttggaaac ttcggcttcc cctggagaga gcgagattct 2700
ccgcgctgta gaagtcacca ttgttgtgca cgacgacatc attccgtggc gttatccagc 2760

[0012]

taagcgcgaa ctgcaatttg gagaatggca gcgcaatgac attcttgcag gtatcttcga 2820
gccagccacg atcgacattg atctggctat cttgctgaca aaagcaagag aacatagcgt 2880
tgccttggta ggtccagcgg cggaggaact ctttgatccg gttcctgaac aggatctatt 2940
tgaggcgcta aatgaaacct taacgctatg gaactcgccg cccgactggg ctggcgatga 3000
gcgaaatgta gtgcttacgt tgtcccgcat ttggtacagc gcagtaaccg gcaaaatcgc 3060
gccgaaggat gtcgctgccg actgggcaat ggagcgcctg ccggcccagt atcagcccgt 3120
catacttgaa gctaggcagg cttatcttgg acaagaagat cgcttggcct cgcgcgcaga 3180
tcagttggaa gaatttgttc actacgtgaa aggcgagatc accaaggtag tcggcaaata 3240
atgtctaaca attcgttcaa gccgacgccg cttcgcggcg cggcttaact caagcgttag 3300
agagctgggg aagactatgc gcgatctgtt gaaggtggtt ctaagcctcg tacttgcgat 3360
ggcatcgggg caggcacttg ctgacctgcc ttaattaaat aacttcgtat aatgtatgct 3420
atacgaagtt atggccggcc aactgccacc acccataaat cccagcgggt ttattttaaa 3480
aataaaacaa atgataacca aaacgatcac gccgccgata ccgaaacggc caacgacaaa 3540
gccgagtaag ccgctaatca ggttaaaccc gccgccaccg ccgttattac caccgctaaa 3600
gtcttcacca cgttcgtctt cgacgttatc gctgcttcgg tagtcatcaa gccgcattat 3660
gcaccttccc aattttaccc gcaggcctat ttataggccg acagcgggtt ttatctatag 3720
ctttcatttt tttgtgtttt tttgaaaaaa aacgcttgca gaacacatag ctgattgtta 3780
tacgcgcttc acacccaacg cggtgattga catagtgttc cctgatagct cagcggtaga 3840
gcgtccgact gttaatcggc aggtcgtagg ttcgaatcct actcggggag ccatttcaac 3900
gatagtcctt atagttttat tttctcttta aaaatttatc ttatttccga tcaatcggtt 3960
atatttcgcc tgatagattc aagacattct gtattgggat tgtctgaata tatttcgggt 4020
agatctgttt aaagtggttt ttttgtccac tttcctaatt aggttatttc cataacgtga 4080
cccaatagcg tggccgtgag cgcttggctg ccgcggaata gggatatttt ctttttcga 4140
gataccgcca gccattaagc cattaagcca ttaagccatt aagccattaa gccattaagc 4200
cattaagcca ttaagatatc gctactcatt ctttcggcaa ggaaaaaggt gattagtcgt 4260
tatttggggt aataaatgtt tatcctcgga tcgagatttc ttgaataaag aaggacgagt 4320
ggttttcttc atttgagcct ttttttcttt ttgcttttttc tatttcgctg gaaacacccc 4380
ttttcaccta ttccagatag attgtcccct ttcgcttttt taaagcattg atcgtagaaa 4440
tcactccgat ggtcgtatga cagcgttttc tttggaaagc ggggttttac agcctttatc 4500
aatagggtca ggatgggggga atcatcctct atatggagaa gtatgatgtc ccctatttcg 4560
ataccctcta tccgtctgaa tgacggcaat gatttgcctg cggtaggctt tggcacctat 4620
aaactcaatg gttcggccgg tgtttccgat atcgttagcg cgattaaagt cggataccgc 4680
ttgctcgatt ctgcctttaa ttatgagaat gaaggtgctg taggcgaagc tgttcgcgag 4740
gccggtattg cccgtgataa attgcggatc gtgtccaaat taccgggggcg acatcatcat 4800
tttgaagagg ccattgcaac ggtcgaggag tcactttatc gggcacagct tgattattat 4860
gatctctatc tgattcattg gcctaatccg agcaaagatc tctatgtcga ggcgtggcag 4920
gccttgattg aagcccgtaa aaaagggtta atccgttcga ttggtgtctg taacttttta 4980
ccagaacatc tggaaaggct gattaaagag acgggtgtaa cgccggttgt caatcaggtg 5040

[0013]

gaattgcatc cctattttcc gcaggaagag cagcgggcat gggataaagc ccatggcatt 5100
gttaccgagt catggagtcc tttgggacgt gccagcaagt tgctacagga cgacacgatt 5160
aaaaaaattg cggatcgctt gggtaagtct atcccgcaag ttattttgcg ttggcatgtc 5220
cagcttggag cgataccgat tcctaaagct tctagcaaag agcgtcagat tgaaaatcta 5280
tccttgtttg atttcgagct gtcgccgcag gatgtagaaa taatcgcgac tttagcacgg 5340
cctgatggac gcttggctga ccaagaccct gcccgctatg aagaatttta aggtacccaa 5400
ttcgccctat agtgagtcgt attacgcgcg ctcactggcc gtcgttttac aacgtcgtga 5460
ctgggaaaac cctggcgtta cccaacttaa tcgccttgca gcacatcccc ctttcgccag 5520
ctggcgtaat agcgaagagg cccgcaccga tcgcccttcc caacagttgc gcagcctgaa 5580
tggcgaatgg gacgcgccct gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg 5640
cagcgtgacc gctacacttg ccagcgccct agcgcccgct cctttcgctt tcttcccttc 5700
ctttctcgcc acgttcgccg gctttccccg tcaagctcta aatcgggggc tccctttagg 5760
gttccgattt agtgctttac ggcacctcga ccccaaaaaa cttgattagg gtgatggttc 5820
acgtagtggg ccatcgccct gatagacggt ttttcgccct ttgacgttgg agtccacgtt 5880
ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct cggtctattc 5940
ttttgattta taagggattt tgccgatttc ggcctattgg ttaaaaaatg agctgattta 6000
acaaaaattt aacgcgaatt ttaacaaaat attaacgctt acaatttagg tggcactttt 6060
cggggaaatg tgcgcggaac ccctatttgt ttatttttct aaatacattc aaatatgtat 6120
ccgctcatga gacaataacc ctgataaatg cttcaataat attgaaaaag gaagagtatg 6180
agtattcaac atttccgtgt cgcccttatt ccctttttg cggcattttg ccttcctgtt 6240
tttgctcacc cagaaacgct ggtgaaagta aaagatgctg aagatcagtt gggtgcacga 6300
gtgggttaca tcgaactgga tctcaacagc ggtaagatcc ttgagagttt tcgccccgaa 6360
gaacgttttc caatgatgag cacttttaaa gttctgctat gtggcgcggt attatcccgt 6420
attgacgccg ggcaagagca actcggtcgc cgcatacact attctcagaa tgacttggtt 6480
gagtactcac cagtcacaga aaagcatctt acggatggca tgacagtaag agaattatgc 6540
agtgctgcca taaccatgag tgataacact gcggccaact tacttctgac aacgatcgga 6600
ggaccgaagg agctaaccgc ttttttgcac aacatggggg atcatgtaac tcgccttgat 6660
cgttgggaac cggagctgaa tgaagccata ccaaacgacg agcgtgacac cacgatgcct 6720
gtagcaatgg caacaacgtt gcgcaaacta ttaactggcg aactacttac tctagcttcc 6780
cggcaacaat taatagactg gatggaggcg gataaagttg caggaccact tctgcgctcg 6840
gcccttccgg ctggctggtt tattgctgat aaatctggag ccggtgagcg tgggtctcgc 6900
ggtatcattg cagcactggg gccagatggt aagccctccc gtatcgtagt tatctacacg 6960
acggggagtc aggcaactat ggatgaacga aatagacaga tcgctgagat aggtgcctca 7020
ctgattaagc attggtaact gtcagaccaa gtttactcat atatacttta gattgattta 7080
aaacttcatt tttaatttaa aaggatctag gtgaagatcc tttttgataa tctcatgacc 7140
aaaatccctt aacgtgagtt ttcgttccac tgagcgtcag accccgtaga aaagatcaaa 7200
ggatcttctt gagatccttt ttttctgcgc gtaatctgct gcttgcaaac aaaaaaacca 7260
ccgctaccag cggtggtttg tttgccggat caagagctac caactctttt tccgaaggta 7320

actggcttca gcagagcgca gataccaaat actgtccttc tagtgtagcc gtagttaggc 7380

caccacttca agaactctgt agcaccgcct acatacctcg ctctgctaat cctgttacca 7440

gtggctgctg ccagtggcga taagtcgtgt cttaccgggt tggactcaag acgatagtta 7500

ccggataagg cgcagcggtc gggctgaacg gggggttcgt gcacacagcc cagcttggag 7560

cgaacgacct acaccgaact gagataccta cagcgtgagc tatgagaaag cgccacgctt 7620

cccgaaggga gaaaggcgga caggtatccg gtaagcggca gggtcggaac aggagagcgc 7680

acgagggagc ttccaggggg aaacgcctgg tatctttata gtcctgtcgg gtttcgccac 7740

ctctgacttg agcgtcgatt tttgtgatgc tcgtcagggg ggcggagcct atggaaaaac 7800

gccagcaacg cggccttttt acggttcctg gccttttgct ggccttttgc tcacatgttc 7860

tttcctgcgt tatcccctga ttctgtggat aaccgtatta ccgcctttga gtgagctgat 7920

accgctcgcc gcagccgaac gaccgagcgc agcgagtcag tgagcgagga agcggaagag 7980

cgcccaatac gcaaaccgcc tctccccgcg cgttggccga ttcattaatg cagctggcac 8040

gacaggtttc ccgactggaa agcgggcagt gagcgcaacg caattaatgt gagttagctc 8100

actcattagg caccccaggc tttacacttt atgcttccgg ctcgtatgtt gtgtggaatt 8160

gtgagcggat aacaatttca cacaggaaac agctatgacc atgattacgc caagcgcgca 8220

attaaccctc actaaaggga acaaaagctg gagctcaata atcccgacag cagttgcatt 8280

atgatgggtt aaatgcacca gaactagcca gtcctgtgca atgcgttgca tccatgtccc 8340

gatattggac acgaacattc cacttatcca aaggcgataa ttgaagctgt tcagtgaccg 8400

[0014] gaaggttcca tggaaaagtt ttttcatcag aggctcgctc tatttttaca tgccgagcag 8460

aagcgccttg atccgctagc atcccatatt gggaggatgt ttttgtgcgc cagaaaagcg 8520

gaaaaaatat ctacgatcat agacggtcgt cagattttcc ggtgattggc ttgttcaagc 8580

tatgcgtcaa gagcgggata tcaagcgtca tgataagaaa aaggatacag gatagacgct 8640

tttgaaaaag gccatcgctg tatcctaaga gtattttagt tgataatttt tactaaataa 8700

ggcgtataac cttatgcttt tacggtaaaa tatttattat tttttattga ccgaatatgt 8760

gtttttcgct ttcgcaaaga catgaatacc atttcatctc aacagaattt aactatcaat 8820

atcatgatgg aaggtatcgc aaacggctgg atcgccgcta tcaatcccgc gtttgagcca 8880

gcgcattctt tccgcagatg tgccatgagt gaaactttca ggaacggggc gcatccctgc 8940

ggctttctct aaggtatcgt cgccaatagc ctgcgctgcc cgcatacctt cttcgacatc 9000

gccctcttct atccggtcgc ggttatgcgc tgcccagact ccggcatagc aatcggcctg 9060

taattccatt ctcacctgta gggcattagc ggcagctttg ccgacattag cttgttgttc 9120

gcggatttgg tcggataagc ccattaaatc ttgcacatga tgcccgactt catgggcgat 9180

aacataggct tgggcaaaat caccgctggc aggaagcggt tgaccagttc attatagaaa 9240

tctgtatcga gataaacgcg ctgatctgct gggcagtaaa aaggccccat cgccgattct 9300

gctgcaccgc agcctgaacg acctctttga ctataaaaga ccaatgttgt cggtttatat 9360

tttttaccag catcagcgaa aatcttgccc catgtgtctt cggtactagc caagacacgg 9420

caggagaaaa ggcttgcggc gttaagttgg caggcctgtt gtgccgaacg aacagaatgc 9480

acctgatgag gtgccatcgc gggttgtgtc ga 9512

<210> 15
<211> 64
<212> DNA
<213> 人工序列

<220>
<223> 含有loxPw位点和NotI位点的寡聚物

<400> 15
ataacttcgt ataatgtatg cttattgaag catattacat acgaatacga agttatgcgg 60

ccgc 64

[0015]

<210> 16
<211> 64
<212> DNA
<213> 人工序列

<220>
<223> 含有loxPw位点和PacI位点的Dna

<400> 16
ttaattaaat aacttcgtat aatgtatgct tattgaagca tattacatac gaatacgaag 60

ttat 64

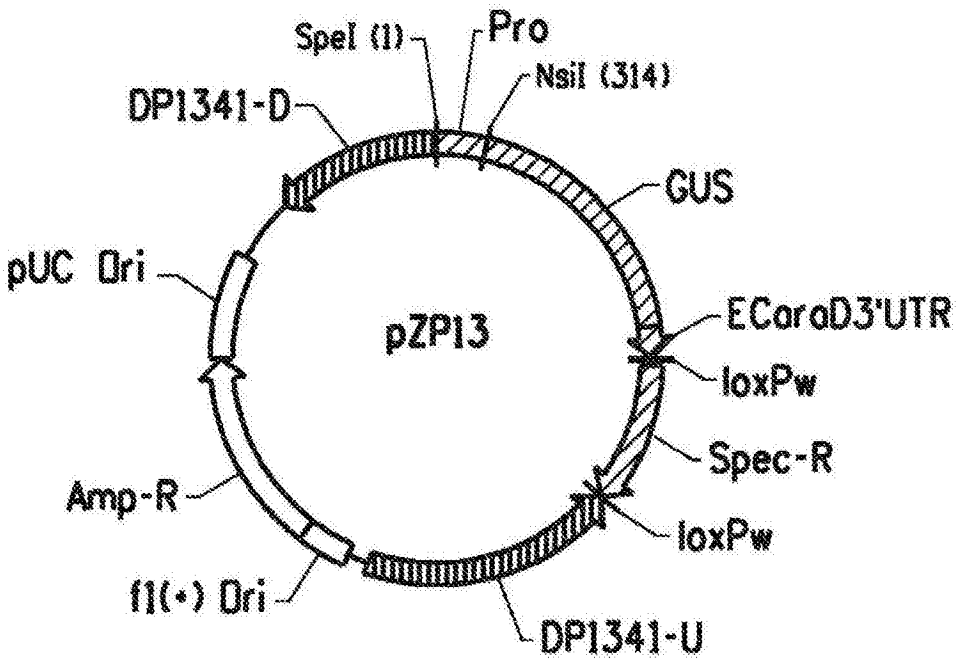

SpeI (1)
Pro
NsiI (314)
DP1341-D
GUS
pUC Ori
pZP13
ECaroD3'UTR
loxPw
Spec-R
Amp-R
loxPw
f1(+) Ori
DP1341-U

图1