



(12)发明专利

(10)授权公告号 CN 102939382 B

(45)授权公告日 2017.02.15

(21)申请号 201180026764.2
 (22)申请日 2011.06.03
 (65)同一申请的已公布的文献号
 申请公布号 CN 102939382 A
 (43)申请公布日 2013.02.20
 (30)优先权数据
 61/351,286 2010.06.03 US
 (85)PCT国际申请进入国家阶段日
 2012.11.29
 (86)PCT国际申请的申请数据
 PCT/US2011/039092 2011.06.03
 (87)PCT国际申请的公布数据
 W02011/153449 EN 2011.12.08
 (73)专利权人 丹尼斯科美国公司
 地址 美国加利福尼亚州
 (72)发明人 B·S·鲍尔 T·卡佩尔
 B·R·凯莱门
 (74)专利代理机构 北京市中咨律师事务所
 11247
 代理人 陈迎春 黄革生

(51)Int.Cl.
 C12N 15/81(2006.01)
 C12N 15/00(2006.01)
 C12N 5/00(2006.01)
 (56)对比文件
 CN 101421401 A,2009.04.29,全文.
 WO 2008113847 A2,2008.09.25,全文,尤其是权利要求1-9,说明书第4页第26行至第5页第34行,第5页最后一段,第8页第18-22行,第9页第30-35行,第11页第23-32行,第12页第10-17行.
 VERENA SEIDL ET AL..Trichoderma reesei:genetic approaches to improving strain efficiency.《BIOFUELS》.2010,348-350.

审查员 管冰

权利要求书2页 说明书52页
 序列表66页 附图10页

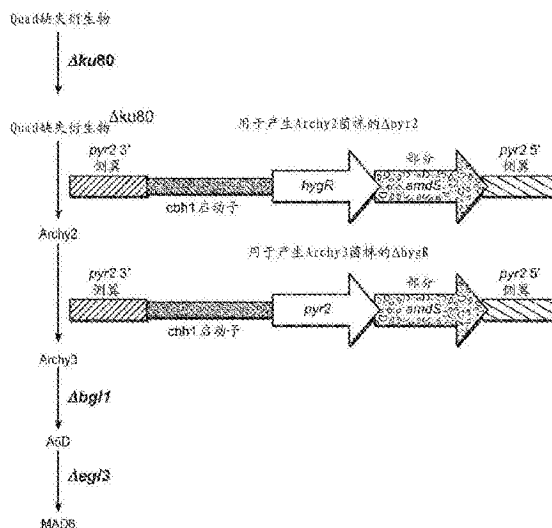
(54)发明名称

丝状真菌宿主菌株和DNA构建体以及它们的使用方法

(57)摘要

本发明涉及丝状真菌宿主菌株和用于其产生和使用的重组DNA构建体。所述丝状真菌宿主菌株特别可用于实现重组酶和变体的可靠表达。

用于产生筛选菌株的缺失载体



[接上页]

(56)对比文件

R storms et al..Plasmid vectors for protein production, gene expression and molecular manipulations in *Aspergillus niger*.《Plasmid》.2005,191-204.

NEUZA D S P CARVALHO ET AL..Expanding

the ku70 toolbox for filamentous fungi: establishment of complementation vectors and recipient strains for advanced gene analyses.《Applied Microbiology and Biotechnology》.2010,1463-1473.

1. 一种丝状真菌宿主细胞表达系统, 包含:

a. 真菌宿主细胞, 所述真菌宿主细胞在其染色体DNA中含有非同源重组(NHR)途径的一个或多个组分中的破坏、缺乏第一可选择功能的第一可选择标记、和可有效赋予第二可选择功能的第二可选择标记; 和

b. 核酸分子, 所述核酸分子含有(1)当引入所述真菌宿主细胞中时给所述在宿主细胞染色体上的第一可选择标记赋予所述第一可选择功能的序列;(2)可有效表达一种或多种目的基因的序列;(3)和与侧接所述可选择标记的序列具有实质同源性的序列;

其中所述同源序列引起同源重组事件, 所述同源重组事件导致有功能的第一可选择标记、导致所述第二可选择标记的去除和导致所述目的基因的表达。

2. 根据权利要求1所述的丝状真菌宿主细胞表达系统, 其中所述NHR途径的一个或多个组分选自ku80、ku70、rad50、mre11、xrs2、lig4和xrs。

3. 根据权利要求1所述的丝状真菌宿主细胞表达系统, 其中所述第一可选择标记和所述第二可选择标记是选自alsR、amdS、hygR、pyr2、pyr4、pyrG、sucA、博来霉素抗性标记、杀稻瘟素抗性标记、吡啶硫胺素抗性标记、氯嘧磺隆抗性标记、新霉素抗性标记、腺嘌呤途径基因、色氨酸途径基因和胸苷激酶的不同标记。

4. 根据权利要求1所述的丝状真菌宿主细胞表达系统, 其中所述真菌宿主细胞来自选自木霉属(*Trichoderma*)、青霉属(*Penicillium*)、曲霉属(*Aspergillus*)、腐质霉属(*Humicola*)、金孢子菌属(*Chrysosporium*)、镰刀菌属(*Fusarium*)、链孢霉属(*Neurospora*)和裸孢壳属(*Emericella*)的属。

5. 根据权利要求4所述的丝状真菌宿主细胞表达系统, 其中所述木霉是里氏木霉(*T. reesei*)。

6. 根据权利要求4所述的丝状真菌宿主细胞表达系统, 其中所述曲霉是黑曲霉(*A. niger*)。

7. 根据权利要求1所述的丝状真菌宿主细胞表达系统, 其中所述目的基因选自半纤维素酶、过氧化物酶、蛋白酶、纤维素酶、木聚糖酶、脂肪酶、磷脂酶、酯酶、角质酶、果胶酶、角蛋白酶、还原酶、氧化酶、酚氧化酶、脂加氧酶、木质素酶、支链淀粉酶、鞣酸酶、戊聚糖酶、malanase、 β -葡聚糖酶、阿拉伯糖苷酶、透明质酸酶、软骨素酶、漆酶、淀粉酶、葡糖淀粉酶及其混合物。

8. 根据权利要求1所述的丝状真菌宿主细胞表达系统, 其中所述目的基因选自乙酰酯酶、氨肽酶、淀粉酶、阿拉伯糖酶、阿拉伯呋喃糖苷酶、羧肽酶、过氧化氢酶、纤维素酶、几丁质酶、凝乳酶、角质酶、脱氧核糖核酸酶、差向异构酶、酯酶、 α -半乳糖苷酶、 β -半乳糖苷酶、 α -葡聚糖酶、葡聚糖裂解酶、内切- β -葡聚糖酶、葡糖淀粉酶、葡糖氧化酶、 α -葡糖苷酶、 β -葡糖苷酶、葡糖醛酸糖苷酶、半纤维素酶、己糖氧化酶、水解酶、转化酶、异构酶、漆酶、脂肪酶、裂解酶、甘露糖苷酶、氧化酶、氧化还原酶、果胶酸裂解酶、果胶乙酰酯酶、果胶解聚酶、果胶甲基酯酶、果胶裂解酶、过氧化物酶、酚氧化酶、植酸酶、多聚半乳糖醛酸酶、蛋白酶、鼠李糖半乳糖醛酸酶、核糖核酸酶、奇异果甜蛋白、转移酶、转运蛋白、转谷氨酰胺酶、木聚糖酶、己糖氧化酶及其组合。

9. 根据权利要求1所述的丝状真菌宿主细胞表达系统, 其中所述目的基因选自肽激素、生长因子、凝血因子、趋化因子、细胞因子、淋巴因子、抗体、受体、粘附分子、微生物抗原及

其片段。

丝状真菌宿主菌株和DNA构建体以及它们的使用方法

[0001] 相关专利申请的交叉引用

[0002] 本专利申请要求于2010年6月3日提交的美国临时申请61/351,286的优先权,所述美国临时申请的公开内容以引用方式并入本文。

技术领域

[0003] 本发明涉及丝状真菌宿主菌株和用于其产生和使用的重组DNA构建体。丝状真菌宿主菌株特别可用于以可靠或较不可变的方式表达目的蛋白质,且用于有效筛选编码重组蛋白的DNA文库。

背景技术

[0004] 丝状真菌宿主细胞菌株已被工程改造为表达各种蛋白质。随后,任选在纯化后,这些蛋白质可以用于各种工业、学术或其他应用中。表达过程通常会是无法预测的。仅极少数(如果有的话)所制备的转化体实际上产生目的酶,这并非罕见情况。异源基因(例如非天然基因,或以与天然形式不同的形式存在的天然基因)的表达的可变性可由于与其核酸和/或氨基酸序列无关的因素而出现。例如,非同源整合在丝状真菌中占优势。因此,表达载体随机整合到基因组内,可能导致对转化体间表达水平的位置效应。此外,可能生成不稳定转化体,使得有必要对转化体进行进一步筛选,以获得稳定转化体。可变性还可通过由于多核原生质体的转化所导致的异核体的生成而出现。因此,以减低的可变性生产目的酶的可靠方法具有明确优点。

[0005] 对于某些工业应用,由真菌宿主菌株生产的这些蛋白质通常被工程改造的,以获得新的期望特征或不同水平的某些特征。在这些情况下,通常使用现有丝状真菌宿主细胞菌株来筛选编码变体蛋白质的DNA文库。表达效力和/或水平的可变性使得难以将给定变体的特征与另一变体的特征进行比较。因此,如果变体可以被可靠地表达(如果它们可以由特定宿主细胞表达的话),并且如果变体可以以较不可变的水平表达而使得它们的特征可以更容易地评估和比较,则明确存在特定的优点。

[0006] 虽然端粒的、染色体外的复制载体可以用作基因组整合的替代物,但这种方法并不消除转化体之间的表达水平的可变性。因此,若有手段来减少丝状真菌宿主菌株基因表达中的序列依赖性差异,将给本领域带来好处。

发明内容

[0007] 本发明涉及丝状真菌宿主菌株和用于其产生和使用的重组DNA构建体。丝状真菌宿主菌株能可靠地产生转化体且表达酶,其中表达水平的可变性减低。丝状真菌宿主菌株可用于有效筛选编码重组蛋白的DNA文库。

[0008] 特别地,本发明提供丝状真菌宿主细胞表达系统,其包含:a)真菌宿主细胞,所述真菌宿主细胞在其染色体DNA中含有非同源重组(NHR)途径的一个或多个组分的破坏、缺乏第一可选择功能的第一可选择标记的部分、和可有效赋予第二可选择功能的第二可选择标

记;和b)核酸分子,所述核酸分子含有当被引入该真菌宿主细胞中时给该第一可选择标记赋予第一可选择功能的序列、可有效表达一种或多种目的基因或变体目的基因的序列、和与侧接染色体可选择标记的序列具有实质同源性的序列;其中所述同源序列引起同源重组事件,所述同源重组事件导致有功能的第一可选择标记、导致第二可选择标记的去除和导致目的基因或变体目的基因的表达。在一些实施方案中,NHR途径的一个或多个组分包括ku80、ku70、rad50、mre11、xrs2、lig4和xrs2中的一者或多者。在某些实施方案中,在b)中引入真菌宿主细胞中的核酸分子可以是非天然的分子或者以对于真菌宿主细胞非天然的形式存在的天然分子。

[0009] 基因缺失可以通过使用缺失质粒来实现。例如,可以将待缺失或破坏的所需基因插入质粒中。然后在适当的限制性酶位点(在所需的基因编码区的内部)切割该缺失质粒,并且用可选择标记替换基因编码序列或其部分。来自待缺失或破坏的基因的基因座的侧翼DNA序列(优选在约0.5至2.0kb之间)保留在可选择标记基因的任一侧上。合适的缺失质粒将通常具有存在于其中的独特限制性酶位点,以使得含有该缺失基因的片段(包括侧翼DNA序列)和可选择标记基因能够作为单一线性小片被去除。缺失质粒还可以通过使用PCR扩增所需侧翼区和可选择标记来构建,在扩增的片段的末端具有限制性酶位点以促进片段的连接。作为另外一种选择,可通过指定适当的侧翼DNA和可选择标记序列来从头合成缺失质粒。

[0010] 在一些实施方案中,所述第一和第二可选择标记是不同的标记。在一些实施方案中,所述第一和第二可选择标记独立地选自alsR、amdS、hygR、pyr2、pyr4、pyrG、sucA、博来霉素抗性标记、杀稻瘟素抗性标记、吡啶硫胺素抗性标记、氯嘧磺隆抗性标记、新霉素抗性标记、腺嘌呤途径基因、色氨酸途径基因和胸苷激酶。在一些实施方案中,所述同源序列中的至少一个在pyr2序列的上游或下游。在一些实施方案中,所述同源序列在pyr2序列的上游和下游。在其他实施方案中,所述同源序列包含可有效表达一种或多种目的基因或一种或多种变体目的基因的序列和对第一可选择标记赋予第一可选择功能的序列。在一些实施方案中,丝状真菌宿主细胞是选自木霉属(*Trichoderma*)、青霉属(*Penicillium*)、曲霉属(*Aspergillus*)、腐质霉属(*Humicola*)、金孢子菌属(*Chrysosporium*)、镰刀菌属(*Fusarium*)和裸孢壳属(*Emericella*)的属中的种。在一些实施方案中,木霉是里氏木霉(*T.reesei*),而在其他实施方案中,曲霉是黑曲霉(*A.niger*)。

[0011] 本发明提供丝状真菌宿主细胞表达系统,其中所述目的基因或变体目的基因选自半纤维素酶、过氧化物酶、蛋白酶、纤维素酶、木聚糖酶、脂肪酶、磷脂酶、酯酶、角质酶(cutinase)、果胶酶、角蛋白酶、还原酶、氧化酶、酚氧化酶、脂加氧酶、木质素酶、支链淀粉酶、鞣酸酶、戊聚糖酶、malanase、 β -葡聚糖酶、阿拉伯糖苷酶、透明质酸酶、软骨素酶、漆酶、淀粉酶、葡糖淀粉酶及其混合物。目的基因或变体基因的非限制性例子编码:涉及淀粉代谢的蛋白质或酶、涉及糖原代谢的蛋白质或酶、乙酰酯酶、氨肽酶、淀粉酶、阿拉伯糖酶、阿拉伯呋喃糖苷酶、羧肽酶、过氧化氢酶、纤维素酶、几丁质酶、凝乳酶、角质酶、脱氧核糖核酸酶、差向异构酶、酯酶、 α -半乳糖苷酶、 β -半乳糖苷酶、 α -葡聚糖酶、葡聚糖裂解酶、内切- β -葡聚糖酶、葡糖淀粉酶、葡萄糖氧化酶、 α -葡糖苷酶、 β -葡糖苷酶、葡糖醛酸糖苷酶、半纤维素酶、己糖氧化酶、水解酶、转化酶、异构酶、漆酶、脂肪酶、裂解酶、甘露糖苷酶、氧化酶、氧化还原酶、果胶酸裂解酶、果胶乙酰酯酶、果胶解聚酶、果胶甲基酯酶、果胶裂解酶

(pectinolytic enzyme)、过氧化物酶、酚氧化酶、植酸酶、多聚半乳糖醛酸酶、蛋白酶、鼠李糖半乳糖醛酸酶、核糖核酸酶、奇异果甜蛋白、转移酶、转运蛋白、转谷氨酰胺酶、木聚糖酶、己糖氧化酶(D-己糖:O₂-氧化还原酶, EC 1.1.3.5)、其变体及其组合。在一些实施方案中,目的基因或变体目的基因编码选自下述的多肽:肽激素、生长因子、凝血因子、趋化因子、细胞因子、淋巴因子、抗体、受体、粘附分子和微生物抗原(例如HBV表面抗原、HPV E7等)及其变体(例如片段)。

[0012] 此外,本发明提供在丝状真菌宿主细胞中表达目的基因或变体目的基因的方法,所述方法包括:将对第一可选择标记赋予第一可选择功能的核酸分子引入丝状真菌宿主细胞中,所述核酸分子可以是非天然或天然(但以非天然形式存在)分子;使宿主细胞生长;以及选择具有第一可选择功能但缺乏第二可选择功能的宿主细胞。在一些实施方案中,如此实现的表达比使用本领域的常规方法实现的那些表达更可靠。在一些实施方案中,所述方法进一步包括测定目的基因或变体目的基因的表达,和/或测定由目的基因或由变体目的基因编码的多肽的生物化学功能。

[0013] 本文引用的所有专利、专利申请、文献、核苷酸和蛋白质序列数据库登记号和文章以引用方式整体并入本文。

附图说明

[0014] 下述的图和表意欲举例说明而不是限制本文公开说明书或权利要求的范围和内容。

[0015] 图1提供了示出从quad缺失衍生菌株衍生MAD6宿主菌株的示意图。

[0016] 图2提供了里氏木霉ku80缺失盒的示意图。

[0017] 图3提供了用于制备Archy2菌株的pyr2缺失盒的示意图。

[0018] 图4提供了用于制备Archy3菌株的hygR缺失盒的示意图。

[0019] 图5提供了里氏木霉bg11缺失盒的示意图。

[0020] 图6提供了里氏木霉eg13缺失盒的示意图。

[0021] 图7提供了用于表达cre重组酶的里氏木霉端粒质粒载体的示意图。

[0022] 图8示出了由于目的基因或变体目的基因(GOI)盒的多核苷酸的引入,pyr2可选择标记的失活和amdS可选择标记的活化,其中在这个实例中GOI编码CBH2变体。

[0023] 图9示出了如实施例2中所述的pENTR/D-TOPO载体。

[0024] 图10示出了如实施例2中所述的pTrex3gM载体。

[0025] 图11示出了如实施例2中所述的Fv43B表达载体pTrex3gM-Fv43B,。

[0026] 图12示出了如实施例2中所述的Fv43C表达载体pTrex3gM-Fv43C,。

[0027] 图13是表征使用由fv43B转化的里氏木霉quad缺失克隆所表达的Fv43B的SDS-PAGE照片。相对于总蛋白质的蛋白质百分比依照实施例2进行定量测定,并且在相应泳道下方列出。

[0028] 图14是表征使用由fv43C转化的里氏木霉quad缺失克隆所表达的Fv43C的SDS-PAGE照片。相对于加载的总蛋白质的蛋白质百分比依照实施例2进行定量测定,并且在相应泳道下方列出。

[0029] 图15是表征使用MAD6构建体所表达的Fv43B和Fv43C的SDS-PAGE照片。相对于总蛋

白质的蛋白质百分比依照实施例2进行定量测定,并且在相应泳道下方列出。

[0030] 图16A是如实施例3中所述检查CBH2变体的4个SDS-PAGE照片。

[0031] 图16B示出如实施例3中所述的CBH2变体的平均表达。

具体实施方式

[0032] 本发明涉及丝状真菌宿主菌株和用于其产生和使用的重组DNA构建体。与本领域已知的其他表达方法相比较,丝状真菌宿主菌株可以用于以更高的可靠性和/或更低的表达水平可变性提供目的基因或变体在这些宿主中的表达。在一个具体的实施方案中,丝状真菌宿主菌株可用于有效筛选编码重组蛋白的DNA文库。

[0033] 本文描述的方法以改善的可靠性表达目的蛋白质或目的变体。在这个意义上,术语“改善的可靠性”反映在(1)至少60%(例如至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%或至少99%)的转化体是稳定转化体;或者至少60%(例如至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%或至少99%)的转化体如所意图以超过本底表达水平表达目的蛋白或变体;和(2)目的蛋白或变体以变化小于60%(例如小于55%、小于50%、小于45%、小于40%、小于35%、小于30%、小于25%、小于20%、小于15%、小于10%、小于5%或小于2%)的表达水平表达,其中术语“表达水平变化”定义为最高和最低表达水平之间的差值除以最高表达水平和本底表达水平之间的差值,所述表达水平都用相同构建体和相同目的基因或变体来测定。

[0034] 应当了解,前述一般描述和以下详细描述都仅是示例性和解释性的,并不限制本文所述的组合物和方法。在本专利申请中,除非另有特别说明,否则所用的单数包括复数。除非另有说明,否则所用的“或”意指“和/或”。同样,术语“包含”、“含有”、“具有”和“包括”并非意图进行限制。本文提及的所有专利和出版物,包括在这些专利和出版物中公开的所有氨基酸和核苷酸序列明确地以引用方式并入。本文提供的标题并不对本发明的各个方面或实施方案构成限制,可通过参考整篇说明书了解本发明的各个方面或实施例。因此,通过参考整篇说明书,本文的术语得到更全面的定义。

[0035] 除非本文另外定义,否则本文使用的所有技术和科学术语具有本发明所属领域的普通技术人员通常理解的相同含义。Singleton等人,DICTIONARY OF MICROBIOLOGY AND MOLECULAR BIOLOGY(微生物学和分子生物学词典),第2版,纽约州的约翰威立父子出版公司(John Wiley and Sons,New York)(1994年),和Hale&Marham,THE HARPER COLLINS DICTIONARY OF BIOLOGY(哈珀柯林斯生物学辞典),纽约州的Harper Perennial,(1991年)为技术人员提供了本发明中使用的许多术语的一般词典。虽然任何与本文所述的那些方法和材料相似或等同的方法和材料可用于实施或测试本发明,但描述的是优选的方法和材料。数值范围包括限定该范围的数值。除非另有说明,否则分别地,核酸从左至右以5'至3'方向书写;氨基酸序列从左至右以氨基至羧基方向书写。关于本领域定义和术语,从业者可具体参考Sambrook等人,MOLECULAR CLONING:A LABORATORY MANUAL(分子克隆实验指南)(第二版),纽约州普莱恩维尔市的冷泉港出版社(Cold Spring Harbor Press,Plainview,N.Y.),1989年,和Ausubel FM等人,Current Protocols in Molecular Biology(分子生物学实验手册),纽约州纽约市的约翰&威立父子出版公司,1993年。应当了解,本公开内容不限于所描述的具体方法、方案和试剂,因为它们可变化。

[0036] 1. 定义

[0037] 以下术语通过参考整篇说明书得到更全面的定义。

[0038] 如本文所用,术语“多肽”是指由通过肽键连接的氨基酸残基的单链构成的化合物或分子。如本文所用,术语“蛋白质”可与术语“多肽”同义。

[0039] “变体”意指通过向C和N末端的任一者或二者添加一个或多个氨基酸、在氨基酸序列中的一个位点或多个不同位点处置换一个或多个氨基酸、或在蛋白质的任一或两个末端处或在氨基酸序列中的一个或多个位点处缺失一个或多个氨基酸而从前体蛋白质(例如天然蛋白质)衍生的蛋白质。目的蛋白(例如由“目的基因”编码的)的变体或“目的变体”(例如由“变体目的基因”编码的)的制备可以通过本领域已知的任何手段进行。例如,通过以下方式制备目的变体:修饰编码天然蛋白质的DNA序列(例如目的基因),将经修饰的DNA序列转化到合适宿主中,并表达经修饰的DNA序列以形成目的变体。在一个非限制性实例中,目的纤维素酶的变体可以通过本领域已知的任何手段进行。例如,通过以下方式制备纤维素酶变体:修饰编码其天然的(或天然存在的)相对物(counterpart)的DNA序列,将经修饰的DNA序列转化到合适宿主中,并表达经修饰的DNA序列以形成变体纤维素酶。本发明的变体目的酶包括这样的多肽:与天然目的酶的氨基酸序列相比较,所述多肽包含改变的氨基酸序列。在某些实施方案中,变体目的酶可保留天然目的酶的一些特征,但同时具有某些相对于天然目的酶而言改变的特征。例如,本发明的变体纤维素酶包括这样的肽:与前体酶氨基酸序列相比较,所述肽包含改变的氨基酸序列,其中所述变体纤维素酶保留前体酶的特征性纤维素分解性质,但在某个具体方面可以具有改变的特性。例如,变体纤维素酶可以具有增加的最适pH或增加的温度或氧化稳定性或者减低的与非纤维素材料的亲和力或结合,但将保留其特征性纤维素分解活性。

[0040] 在一个非限制性实例中,设想到根据本发明的目的变体可以衍生自编码变体的核苷酸序列,其中所表达的变体的功能活性得到保留。例如,纤维素酶变体可以衍生自编码纤维素酶变体的DNA片段,其中所表达的变体的纤维素酶活性得到保留。在一些实施方案中,编码纤维素酶的DNA片段可进一步包括在5'端或3'端附着至纤维素酶DNA序列的编码铰链或接头的DNA序列或其部分,其中所编码的纤维素酶结构域的功能活性得到保留。术语“变体”和“衍生物”在本文中可互换使用。

[0041] “等价残基”可以通过测定前体或参考酶的三级结构水平上的同源性进行定义,所述前体或参考酶的三级结构已通过X射线晶体学进行测定。例如,等价残基定义为这样的残基:在比对后,纤维素酶和红褐肉座菌(*Hypocrea jecorina*)CBH2的特定氨基酸残基的两个或更多个主链原子的原子坐标(N对N,CA对CA,C对C和O对O)相差在0.13nm内且优选在0.1nm内。比对是在使最佳模型定向和定位以给出所讨论的酶和前体/参考酶的非氢蛋白质原子的原子坐标的最大限度重叠后实现的。例如,合适的模型包括对在可获得的最分辨率下的实验衍射数据给出最低R因子的晶体学模型。参见例如,美国专利申请No. 2006/0205042。

[0042] 与前体或参考酶的特定残基功能上类似的等价残基定义为那些可呈这样的构象的残基,该构象使得它们改变、修饰或促成酶的结构、底物结合或以预定方式进行的催化。例如,红褐肉座菌CBH2的等价残基是纤维素酶的那些呈这样的构象的氨基酸,该构象使得它们改变、修饰或促成蛋白质结构、底物结合或以限定方式进行的催化。在一些实施方案中,等价残基可以是那些占据类似位置达这样的程度的残基:尽管给定残基的主链原子可

能不满足在占据同源位置的基础上的等价标准,但残基的超过一个(例如2、3个或更多个)侧链原子的原子坐标与前体/参考酶的相应侧链原子相差在短距离内(例如在约0.02nm内、在约0.05nm内、在约0.08nm内、在约0.10nm内、在约0.12nm内、在约0.13nm内、在约0.14nm内、在约0.15nm内、在约0.17nm内、在约0.18nm内、在约0.20nm内、在约0.25nm内等)。例如,已通过X射线晶体学获得三级结构的纤维素酶可以适当地包含等价残基,其中残基的至少两个侧链原子的原子坐标与红褐肉座菌CBH2的相应侧链原子相差在0.13nm内,即使给定残基的主链原子不满足在占据同源位置的基础上的等价标准。红褐肉座菌CBH2的晶体结构显示于Zou等人(1999)Structure(结构)7(9):1035-45中。

[0043] 术语“核酸分子”包括RNA、DNA和cDNA分子。应当理解,由于遗传密码的简并性,可以产生多个编码给定蛋白质和/或其变体的核苷酸序列。本发明设想了编码变体目的酶的每种可能的变体核苷酸序列,考虑到遗传密码的简并性,所有这些变体核苷酸序列都是可能的。例如,由于遗传密码的简并性,多个核苷酸序列可以编码纤维素酶例如CBH2和/或其变体,其可以通过本文描述的方法或过程产生。

[0044] “异源”核酸构建体或序列具有对于它在其中表达的细胞而言并非天然的或并非以天然形式存在的序列部分。就控制序列而言的异源是指这样的控制序列(即启动子或增强子),对于其目前调节表达的另一基因,在自然界中其是不起到这个调节作用的。一般来讲,异源核酸序列对其所在的细胞或基因组部分不是内源的,而是已通过感染、转染、转化、微注射、电穿孔等添加至细胞。“异源”核酸构建体可含有与天然细胞中发现的控制序列/DNA编码序列组合相同或不同的控制序列/DNA编码序列组合。

[0045] 如本文所用,术语“载体”是指设计用于在不同宿主细胞之间转移的核酸构建体。“表达载体”是指能够将异源DNA片段掺入外来细胞中并在该细胞中表达异源DNA片段的载体。许多原核和真核表达载体是可商购获得的。适当表达载体的选择在本领域技术人员知识范围之内。

[0046] 因此,“表达盒”或“表达载体”是通过重组或合成方式生成的核酸构建体,其具有一系列能让特定核酸在靶细胞中转录的指定核酸元件。可将重组表达盒掺入质粒、染色体、线粒体DNA、质体DNA、病毒或核酸片段中。通常,表达载体的重组表达盒部分包括待转录的核酸序列和启动子等序列。

[0047] 如本文所用,术语“质粒”是指用作克隆载体的环状双链(ds)DNA构建体,并且其在许多细菌和一些真核生物中形成染色体外自我复制遗传元件。

[0048] 如本文所用,术语“编码可选择标记的核苷酸序列”是指能够在细胞中表达的核苷酸序列,并且可选择标记的表达使得含有所表达基因的细胞能够在相应选择性试剂存在下或在相应选择性生长条件下生长。

[0049] 如本文所用,术语“启动子”是指起到指导下游基因转录的作用的核酸序列。启动子一般将适于靶基因在其中被表达的宿主细胞。启动子连同其他转录和翻译调控核酸序列(也称为“控制序列”)通常用于表达给定基因。一般来讲,转录和翻译调控核酸序列包括但不限于启动子序列、核糖体结合位点、转录起始和终止序列、翻译起始和终止序列以及增强子或激活序列。

[0050] 如本文所定义,“嵌合基因构建体”是指可由不同基因的部分(包括调控元件)构成的非天然基因(即已被引入到宿主中的基因或不以其在宿主中的天然形式存在的基因)。用

于转化宿主细胞的嵌合基因构建体通常由这样的转录调控区(启动子)组成,所述转录调控区有效连接至蛋白编码序列,或者,在可选择标记嵌合基因中,其有效连接至编码例如给被转化的细胞赋予抗生素抗性的蛋白质的可选择标记基因。用于转化到宿主细胞中的本发明典型嵌合基因包含组成型或诱导型转录调控区、蛋白编码序列和终止子序列。如果期望分泌靶蛋白,则嵌合基因构建体还可包含编码信号肽的第二DNA序列。例如,本文描述的某些构建体,例如Archy 3里氏木霉菌株是嵌合基因构建体。

[0051] 核酸当被置于与另一核酸序列发生功能关系时,该核酸是“有效连接的”。例如,编码分泌前导区的DNA与编码某多肽的DNA有效连接,如果该编码分泌前导区的DNA作为参与该多肽的分泌的前蛋白表达的话;启动子或增强子与某编码序列有效连接,如果该启动子或增强子影响该序列的转录的话;或者,核糖体结合位点与某编码序列有效连接,如果该核糖体结合位点被定位为有利于翻译的话。一般来讲,“有效连接”是指被连接的DNA序列是连续的,并且在分泌前导区的情况下,被连接的DNA序列是连续的并在阅读框中。然而,增强子不必是连续的。连接是通过在便利的限制性位点处的连接来实现。如果不存在此类位点,那么根据常规做法使用合成的寡核苷酸衔接子、接头或PCR引物。

[0052] 当选择标记具有完全选择功能时,那么该选择标记在本文中被称为是“可操作的”。

[0053] 如本文所用,术语“基因”是指DNA的参与产生多肽链的区段,其可包括或不包括在编码区之前和之后的区域,例如5'非翻译(5' UTR)或“前导”序列和3' UTR或“尾随”序列以及各个编码区段(外显子)之间的间插序列(内含子)。

[0054] 一般来讲,编码目的变体的核酸分子将在中等至高度严格性条件下与野生型前体/参考序列杂交。例如,编码变体纤维素酶(如CBH2)的核酸分子将在中等至高度严格性条件下与野生型序列(如本文提供的SEQ ID NO:7)杂交。然而,在某些实施方案中,编码目的酶的核苷酸序列可以反映实质上不同的密码子使用,但仍旧编码相同目的酶。例如,依照特定密码子被宿主利用的频率(参见例如,Te'o等人,FEMS Microbiology Letters(欧洲微生物学会联合会微生物学快报),190:13-19,2000年,描述了用于在丝状真菌中表达的基因的最佳化),编码序列可以进行修饰,以利于目的酶或变体在特定原核或真核表达系统中的更强表达。

[0055] 如果某核酸序列与参考核酸序列在中等至高度严格性杂交和洗涤条件下彼此特异性杂交,那么认为这两个序列“可选择性杂交”。杂交条件基于核酸结合复合物或探针的解链温度(T_m)。例如,“最大限度严格性”通常在约 $T_m-5^\circ\text{C}$ (比探针的 T_m 低 5°C)下发生;“高严格性”在比 T_m 低约5至约 10°C 下发生;“中等”或“中间严格性”在比探针的 T_m 低约10至约 20°C 下发生;“低严格性”在比 T_m 低约20至约 25°C 下发生。在功能上,可以使用最大限度严格性条件来鉴定与杂交探针具有严格的同一性或接近严格的同一性的序列;而高严格性条件用于鉴定与探针具有约80%或更高序列同一性的序列。

[0056] 中等和高严格性杂交条件是本领域众所周知的(参见例如,Sambrook,等人,1989年,第9和11章,以及Ausubel,F.M.,等人,1993年,将它们明确地以引用方式并入本文)。高严格性条件的一个例子包括在约 42°C 下,在50%甲酰胺、5xSSC、5xDenhardt's溶液、0.5% SDS和100 $\mu\text{g}/\text{mL}$ 变性载体DNA中杂交,然后在室温下在2xSSC和0.5% SDS中洗涤两次,并且在 42°C 下在0.1xSSC和0.5% SDS中再洗涤两次。

[0057] 当用来指涉例如细胞或核酸、蛋白质或载体时,术语“重组”表示该细胞、核酸、蛋白质或载体已通过引入异源核酸或蛋白质或者改变天然核酸或蛋白质而被修饰,或者表示该细胞衍生自如此修饰的细胞。因此,例如,重组细胞表达在天然(非重组)形式的该细胞内未发现的基因,或者表达原本会异常表达、低表达或根本不表达的天然基因。

[0058] 如本文所用,指涉细胞的术语“转化的”、“稳定转化的”或“转基因的”意指该细胞有非天然(或未以其天然形式存在)的核酸序列整合到其基因组内或作为附加型质粒存在,该附加型质粒经过多代都得到维持。

[0059] 如本文所用,术语“表达”是指基于基因的核酸序列而产生多肽的过程。该过程包括转录和翻译。

[0060] 在将核酸序列插入细胞中的语境下的术语“引入的”,是指“转染”或“转化”或“转导”,并且包括指涉将核酸序列掺入真核或原核细胞内,在其中核酸序列可被掺入细胞的基因组(例如,染色体、质粒、质体或线粒体DNA)中、被转变成自主复制子或被瞬时表达(例如,转染的mRNA)。

[0061] 术语“目的蛋白或变体的表达”是指目的基因或目的变体的转录和翻译,其产物包括前体RNA、mRNA、多肽、翻译后加工的多肽及其衍生物。例如,“CBH2表达”是指cbh2基因或其变体的转录和翻译,其产物包括前体RNA、mRNA、多肽、翻译后加工的多肽及其衍生物,包括来自相关菌种例如康宁木霉(*Trichoderma koningii*)、红褐肉座菌(也称为长梗木霉(*Trichoderma longibrachiatum*)、里氏木霉或绿色木霉(*Trichoderma viride*))和施氏肉座菌(*Hypocrea schweinitzii*)的CBH2。表达水平可以通过各种已知方法进行测定,包括例如用于目的蛋白或变体的蛋白质印迹、用于目的基因或变体基因的mRNA的RNA印迹分析和逆转录酶聚合酶链反应(RT-PCR)测定、和在合适底物上进行酶促活性测定。举例说,用于CBH2表达的测定包括如以下所述的用于CBH2蛋白质的蛋白质印迹、用于cbh2mRNA的RNA印迹分析和逆转录酶聚合酶链反应(RT-PCR)测定、和磷酸膨胀纤维素酶(PASC)和对羟基苯甲酸酰肼(PAHBAH)测定:(a)PASC:(Karlsson,J.等人(2001年),*Eur. J. Biochem*(欧洲生物化学杂志),268,6498-6507,Wood,T.(1988年)于*Methods in Enzymology*(酶学方法),第160卷。*Biomass Parta Cellulose and Hemicellulose*(生物量部分纤维素和半纤维素)(Wood,W.&Kellog,S.编辑),第19-25页,美国加利福尼亚州圣地亚哥市的美国学术出版社(Academic Press,San Diego,Calif.,USA))和(b)PAHBAH:(Lever,M.(1972年)*Anal. Biochem.*(分析生物化学),47,273,Blakeney,A.B.&Mutton,L.L.(1980年)*J. Sci. Food&Agriculture*(食品科学与农业杂志),31,889,Henry,R.J.(1984年)*J. of the Institute of Brewing*(酿造学会志),90,37)。

[0062] 术语“宿主细胞”是指这样的细胞,其含有载体,且支持表达构建体的复制和/或转录或转录和翻译(表达)。在本发明中使用的宿主细胞可以是原核细胞如大肠杆菌(*E. coli*)细胞,或真核细胞如酵母、植物、昆虫、两栖动物或哺乳动物细胞。在某些实施方案中,宿主细胞适当地是丝状真菌细胞。

[0063] 术语“丝状真菌”意指本领域技术人员所了解的任何和所有丝状真菌。优选的真菌选自曲霉属、木霉属、镰刀菌属、金孢子菌属、青霉属、腐质霉属、链孢霉属(*Neurospora*)、或其另外的有性形式,例如裸孢壳属、肉座菌属(*Hypocrea*)。目前已证实无性工业真菌里氏木霉是子囊菌纲红褐肉座菌的克隆衍生物(参见,Kuhls等人,PNAS(美国国家科学院院刊),

93:7755-7760,1996年)。

[0064] 许多微生物能产生水解纤维素的酶,包括木腐真菌木霉属,堆肥菌高温单孢菌属(*Thermomonospora*)、芽孢杆菌属(*Bacillus*)和纤维单孢菌属(*Cellulomonas*);链霉菌属(*Streptomyces*);和真菌腐质霉属、曲霉属和镰刀菌属。

[0065] 如本文所用,术语“分离的”或“纯化的”是指从与其天然结合的至少一种组分中移出的核酸或氨基酸。

[0066] “丝状真菌”包括亚门真菌门(*Eumycota*)和卵菌门(*Oomycota*)的所有丝状形式。例如,丝状真菌包括但不限于枝顶孢属(*Acremonium*)、曲霉属、裸孢壳属、镰刀菌属、腐质霉属、毛霉属(*Mucor*)、毁丝霉属(*Myceliophthora*)、链孢霉属、柱霉属(*Scytalidium*)、草根霉属(*Thielavia*)、弯颈霉属(*Tolyposcladium*)或木霉属的种。在一些实施方案中,丝状真菌可以是棘孢曲霉(*Aspergillus aculeatus*)、泡盛曲霉(*Aspergillus awamori*)、臭曲霉(*Aspergillus foetidus*)、日本曲霉(*Aspergillus japonicus*)、构果曲霉(*Aspergillus nidulans*)、黑曲霉或米曲霉(*Aspergillus oryzae*)。在一些实施方案中,丝状真菌是拟杆镰刀菌(*Fusarium bactridioides*)、纹枯镰刀菌(*Fusarium cerealis*)、克地镰刀菌(*Fusarium crookwellense*)、黄色镰刀菌(*Fusarium culmorum*)、禾谷镰刀菌(*Fusarium graminearum*)、禾赤镰刀菌(*Fusarium graminum*)、异孢镰刀菌(*Fusarium heterosporum*)、荆条镰刀菌(*Fusarium negundi*)、尖孢镰刀菌(*Fusarium oxysporum*)、网状镰刀菌(*Fusarium reticulatum*)、粉红镰刀菌(*Fusarium roseum*)、接骨木镰刀菌(*Fusarium sambucinum*)、肤色镰刀菌(*Fusarium sarcochromum*)、拟枝孢镰刀菌(*Fusarium sporotrichioides*)、硫色镰刀菌(*Fusarium sulphureum*)、簇囊镰刀菌(*Fusarium torulosum*)、拟丝孢镰刀菌(*Fusarium trichothecioides*)或镶片镰刀菌(*Fusarium venenatum*)。在一些实施方案中,丝状真菌是特异腐质霉(*Humicola insolens*)、柔毛腐质霉(*Humicola lanuginosa*)、米赫毛霉(*Mucor miehei*)、嗜热毁丝霉(*Myceliophthora thermophila*)、粗糙链孢霉(*Neurospora crassa*)、嗜热柱霉菌(*Scytalidium thermophilum*)或土生梭孢壳(*Thielavia terrestris*)。在一些实施方案中,丝状真菌是哈茨木霉(*Trichoderma harzianum*)、康宁木霉、长梗木霉、里氏木霉,例如,RL-P37(Sheir-Neiss等人,Appl.Microbiol.Biotechnology(应用微生物学与生物技术),20(1984年),第46-53页;Montenecourt B.S.,Can.,1-20,1987年),QM9414(ATCC No.26921),NRRL 15709,ATCC 13631、56764、56466、56767或绿色木霉例如ATCC 32098和32086。在一些实施方案中,丝状真菌是里氏木霉RutC30,其作为里氏木霉ATCC 56765得自美国典型培养物保藏中心。关于这点,在一些实施方案中,本发明提供本文描述的任何一种丝状真菌的全细胞肉汤制备物。

[0067] 一般来讲,在适合于产生酶的细胞培养基中培养微生物。使用本领域已知的程序,在包含碳源和氮源以及无机盐的合适的营养培养基中进行培养。适于生长和产酶的合适培养基、温度范围和其他条件在本领域是已知的。作为一个非限制性实例,里氏木霉产纤维素的正常温度范围是24℃至28℃。

[0068] 一般来讲,“全细胞肉汤制备物”经发酵产生后不进行回收和/或纯化或者进行最低限度的回收和/或纯化即加以使用。例如,一旦目的酶或变体(或超过一种目的酶或变体)被细胞分泌到细胞培养基中,就将含有目的酶或变体的细胞培养基加以使用。在一些实施

方案中,全细胞肉汤制备物包含发酵材料的未分级内容物,所述内容物包括细胞培养基、胞外酶和细胞在内。作为另外一种选择,全细胞肉汤制备物可通过任何便利的方法来处理,例如,通过沉淀、离心、亲和力、过滤或本领域已知的任何其他方法。在一些实施方案中,例如,可将全细胞肉汤制备物浓缩且随后在不进行进一步的纯化的情况下使用。在一些实施方案中,全细胞肉汤制备物包含降低细胞活力或杀死细胞的化学试剂。在一些实施方案中,使用本领域已知的方法对细胞进行裂解或透化处理。例如,目的纤维素酶或变体(例如CBH2或其变体Fv43B,Fv43A)可以由细胞分泌到细胞培养基中,所述细胞培养基含有纤维素酶或变体。细胞培养基可以作为全细胞肉汤制备物使用。

[0069] 2. 分子生物学

[0070] 在某些实施方案中,本发明提供目的酶或变体的表达。在某些实施方案中,将编码目的酶或变体的基因置于在丝状真菌中起作用的启动子控制下。在这个实施方案的一个实例中,本文提供表达处于在丝状真菌中起作用的合适启动子控制下的变体cbh2基因的方法。可以应用在重组遗传学领域中的已知技术(参见例如,Sambrook等人,Molecular Cloning,A Laboratory Manual(分子克隆实验指南),第2版,1989年;Kriegler,Gene Transfer and Expression:A Laboratory Manual(基因转移和表达实验指南),1990年;和Ausubel等人,编辑,CURRENT PROTOCOLS IN MOLECULAR BIOLOGY(分子生物学实验手册),纽约州的格林出版社和威立出版公司(Greene Publishing and Wiley-Interscience,New York),1994年)。

[0071] 3. 重组蛋白的表达

[0072] 本发明的方法涉及被工程改造来表达重组蛋白的宿主细胞,而不将表达方法限制于任何具体方法。目的重组蛋白或变体优选从细胞分泌。本发明提供已用包含蛋白编码核酸序列的表达载体转导、转化或转染的宿主细胞。培养条件(例如温度、pH等)是以前在转导、转化或转染之前用于亲本宿主细胞的那些条件,并且对于本领域技术人员将显而易见。

[0073] 在一种方法中,用表达载体转染丝状真菌细胞或酵母细胞,所述表达载体具有在宿主细胞系中起作用、与编码目的蛋白或变体的DNA区段有效连接的启动子或生物学活性的启动子片段或一种或多种(例如一系列)增强子,从而使目的蛋白或变体在细胞系中被表达。

[0074] A. 核酸构建体/表达载体。

[0075] 编码目的蛋白或变体的天然或合成多核苷酸片段可以掺入嵌合构建体或载体中,所述嵌合构建体或载体能够被引入丝状真菌或酵母细胞中且在该丝状真菌或酵母细胞中复制。本文所公开的载体和方法适用于在宿主细胞中表达蛋白质或变体。任何载体都可以使用,只要它能够在它被引入的细胞中复制并且存活。大量合适的载体和启动子是本领域技术人员已知的,并且许多是可商购获得的。克隆和表达载体在以下文献中也有描述: Sambrook等人,1989;Ausubel F M等人,1989和Strathern等人,The Molecular Biology of the Yeast Saccharomyces(酵母属的分子生物学),1981,所述每篇文献明确地以引用方式并入本文。适用于真菌的表达载体在van den Hondel,C.A.M.J.J.等人(1991年)于: Bennett,J.W.和Lasure,L.L.(编辑)More Gene Manipulations in Fungi(真菌中的更多基因操作),美国学术出版社(Academic Press),第396-428页中有描述。可通过多种程序将适当的DNA序列插入到质粒或载体(本文统称为“载体”)中。在一些情况下,使用已知程序将

DNA序列插入合适的限制性核酸内切酶位点中。在其他情况下,可适当地应用不涉及限制性消化和/或连接的载体构建方法。此类程序和相关的亚克隆程序被认为是在本领域的技术人员知识范围内。

[0076] 包含目的蛋白或变体的编码序列的重组丝状真菌可以通过将包含该目的蛋白或变体的编码区的嵌合构建体引入选定的丝状真菌菌株的细胞中而产生。

[0077] 一旦获得所需形式的核酸序列,其就可以多种方式进行修饰。例如,当该序列涉及非编码侧翼区时,可对侧翼区实施切除、诱变等。因而,可对天然存在的序列进行转换、颠换、缺失和插入。

[0078] 可根据已知的重组技术将选定的编码序列插入合适的载体中,并用于转化丝状真菌。由于遗传密码固有的简并性,编码实质上相同或功能上等同的氨基酸序列的其他核酸序列也可用于克隆且表达目的蛋白或变体。因此,编码区中的此类置换属于本发明所涵盖的序列变体。这些序列变体中的任何和全部都可以与本文描述相同的方式进行利用。例如,当目的蛋白或变体是纤维素酶时,可以使用纤维二糖水解酶的序列变体例如CBH2。

[0079] 术语“纤维素酶”或“纤维素分解酶”是指一类能够将纤维素聚合物水解成较短的纤维寡糖低聚物、纤维二糖和/或葡萄糖的酶。纤维素酶的许多实例,诸如外切葡聚糖酶、外切纤维二糖水解酶、内切葡聚糖酶和葡糖苷酶已经从纤维素分解生物中获得,尤其包括真菌、植物和细菌。由这些微生物产生的酶是蛋白质的混合物,其具有三类在纤维素向葡萄糖的转化中有用的作用:内切葡聚糖酶(EG)、纤维二糖水解酶(CBH)和 β -葡糖苷酶。这三种不同类型的纤维素酶协同作用以将纤维素及其衍生物转化成葡萄糖。

[0080] 来自红褐肉座菌的CBH2是糖基水解酶家族6(从而为Ce16)的成员,并且具体来说是该家族在红褐肉座菌(从而为Ce16A)中鉴定的首个成员。糖基水解酶家族6含有内切葡聚糖酶和纤维二糖水解酶/外切葡聚糖酶两者,而CBH2是纤维二糖水解酶/外切葡聚糖酶。因而,词语CBH2、CBH2型蛋白质和Ce16纤维二糖水解酶在本文中通常可互换使用。因而,术语“变体cbh2基因”意指已通过去除、添加和/或操纵编码序列而被改变的红褐肉座菌cbh2基因的核酸序列。

[0081] 本发明还包括包含一个或多个如上所述的编码蛋白质的核酸序列的重组核酸构建体。所述构建体包含其中已正向或反向插入了本发明的序列的载体,例如质粒或病毒载体。

[0082] 嵌合构建体可以包括目的蛋白或变体的编码序列。在一些实施方案中,编码序列可以按以下方式存在:(i)处于分离状态;(ii)与附加编码序列(例如融合蛋白或信号肽编码序列)相组合,其中所述编码序列是主导编码序列;(iii)与对于所述编码序列在合适的宿主中的表达有效的非编码序列相组合,所述非编码序列例如内含子和控制元件(包括例如启动子和终止子元件、或5'和/或3'非翻译区);和/或(iv)处于所述编码序列在其中为天然或非天然基因的载体或宿主环境中。

[0083] 在某些方面,嵌合构建体用于将编码蛋白质的核酸序列在体外转移到细胞中。优选地,编码蛋白质的核酸序列被转移到其中的细胞是已建立的丝状真菌或酵母系。对于目的蛋白或变体的长期生产,稳定表达是优选的。有效生成稳定转化体的各种已知方法可以用于实施本发明。

[0084] 合适的载体通常配有编码可选择标记的核酸序列、插入位点和合适的控制元件,

例如启动子和终止序列。载体可包含调控序列,包括例如非编码序列,如内含子和控制元件,例如,启动子和终止子元件或5'和/或3'非翻译区,其有效连接至所述编码序列,并且对于所述编码序列在宿主细胞中(和/或在其中通常不表达经修饰的可溶性蛋白编码序列的载体或宿主细胞环境中)的表达是有效的。许多合适的载体和启动子是本领域技术人员已知的,并且许多是可商购获得的和/或在Sambrook等人(同上)中描述的。

[0085] 合适启动子的例子包括组成型启动子和诱导型启动子,包括CMV启动子、SV40早期启动子、RSV启动子、EF-1a启动子、如所描述(Clontech和BASF)含有tet-on或tet-off系统中的四环素应答元件(TRE)的启动子、 β 肌动蛋白启动子和可通过加入某些金属盐进行上调的金属硫蛋白启动子。启动子序列是被特定丝状真菌识别以实现表达目的的DNA序列。它被有效连接至编码目的蛋白或变体的DNA序列。这种连接包括将启动子相对于编码目的蛋白的DNA序列的起始密码子进行定位。启动子序列含有介导目的蛋白或变体的表达的转录或翻译控制序列。非限制性例子包括来自以下基因的启动子:黑曲霉、泡盛曲霉或米曲霉葡糖淀粉酶、 α -淀粉酶或 α -葡糖苷酶编码基因;构巢曲霉gpdA或trpC基因;粗糙链孢霉cbh1或trp1基因;黑曲霉或米赫根毛霉(*Rhizomucor miehei*)天冬氨酸蛋白酶编码基因;红褐肉座菌(里氏木霉)cbh1、cbh2、egl1、egl2、或其他纤维素酶编码基因。

[0086] 适当的可选择标记的选择将取决于宿主细胞,并且适合不同宿主的标记是本领域熟知的。合适的可选择标记基因的例子包括来自构巢曲霉或里氏木霉的argB,来自构巢曲霉的amdS,来自粗糙链孢霉或里氏木霉的pyr4,来自黑曲霉或构巢曲霉的pyrG。合适的可选择标记的另外例子包括但不限于trpc、trp1、oliC31、niaD或leu2,其包括在用于转化突变株(如trp-、pyr-、leu-等)的嵌合构建体中。

[0087] 此类可选择标记赋予转化株利用通常不被丝状真菌代谢的代谢物的能力。例如,来自红褐肉座菌的编码乙酰胺酶的amdS基因能让转化株细胞以乙酰胺作为氮源生长。可选择标记(例如pyrG)可恢复营养缺陷型突变株在选择性基本培养基上生长的能力,或者可选择标记(例如olic31)可赋予转化株在抑制性药物或抗生物素的存在下生长的能力。

[0088] 使用本领域一般采用的方法将可选择标记编码序列克隆到任何合适的质粒中。合适质粒的例子包括pUC18、pBR322、pRAX和pUC100。pRAX质粒含有来自构巢曲霉的AMAL序列,这使得它可以在黑曲霉中复制。

[0089] 除非另有说明,否则本发明的实施将采用分子生物学、微生物学、重组DNA和免疫学的常规技术,这些技术在本领域技术范围内。此类技术在文献中充分阐明。参见例如,Sambrook等人,1989年;Freshney, *Animal Cell Culture*(动物细胞培养),1987年;Ausubel,等人,1993年;和Coligan等人, *Current Protocols in Immunology*(免疫学实验手册),1991年。

[0090] B. 用于重组蛋白生产的丝状真菌和培养条件

[0091] 可以为了重组蛋白表达进行处理和/或修饰的亲本丝状真菌菌种的例子包括但不限于木霉属(例如里氏木霉、长梗木霉、绿色木霉、康宁木霉);青霉菌属菌种,腐质霉属菌种(包括特异腐质霉),曲霉属菌种,金孢子菌属菌种,镰刀菌属菌种,肉座菌属菌种和裸孢壳属菌种。

[0092] 在通常用于培养亲本真菌系的条件下培养经转化的细胞。例如,可以在含有生理盐和营养素的的标准培养基中培养细胞,例如如在Pourquie, J.等人, *Biochemistry and*

Genetics of Cellulose Degradation(纤维素降解的生物化学和遗传学),Aubert,J.P.等人编辑,美国学术出版社,第71-86页,1988年和Ilmen,M.等人,Appl.Environ.Microbiol.(应用与环境微生物学)63:1298-1306,1997年中所描述。多种常用培养条件可以是合适的,例如培养物在28℃下在振荡器培养物或发酵罐中温育,直至达到所需水平的重组蛋白表达。

[0093] 适合于给定丝状真菌的培养条件可在科学文献中找到和/或得自真菌的来源,例如美国典型培养物保藏中心(ATCC;www.atcc.org/)。在真菌生长已确立后,使细胞暴露于能有效引起或允许重组蛋白表达的条件。

[0094] 在编码序列处于诱导型启动子的控制下的情况中,向培养基中以能有效诱导重组蛋白表达的浓度加入诱导剂,例如糖、金属盐或抗生素。

[0095] 在一些实施方案中,丝状真菌是黑曲霉,其是可用于获得过表达的目的蛋白的有用菌株。例如,已知黑曲霉变种泡盛曲霉(*A.niger* var *awamori*)dgr246分泌升高的量的纤维素酶(Goedegebuur等人,Curr.Genet(当代遗传学)(2002年)41:89-98)。黑曲霉变种泡盛曲霉的其他菌株例如GCDAP3、GCDAP4和GAP3-4也是已知的。参见例如,Ward等人,Appl.Microbiol.Biotechnol.(应用微生物学和生物技术)39:738-743。

[0096] 在一些实施方案中,丝状真菌是里氏木霉,其是用于获得过表达的目的蛋白的另一种有用菌株。在一些实施方案中,这种丝状真菌宿主细胞可以具有某些与被缺失或减少的有害活性或性状(例如对于表达、稳定性有害的,将使得某些性质的查询或测定变得困难的混杂活性等)相联系的基因(或本文称“有害基因”)。在一些实施方案中,这种真菌宿主细胞可以进行修饰,使得它获得或增强与某些有利的活性或性状相联系的基因(或本文称“有利基因”),所述某些有利的活性或性状例如增加的分泌、增加的稳定性、增加的可溶性等。

[0097] 例如,由Sheir-Neiss,等人,Appl.Microbiol.Biotechnol.(应用微生物学和生物技术)20:46-53(1984年)描述的里氏木霉菌株RL-P37已知能分泌升高的量的纤维素酶。RL-P37的功能等价物包括里氏木霉菌株RUT-C30(ATCC No.56765)和菌株QM9414(ATCC No.26921)。设想到这些菌株还将可用于过表达蛋白质及其变体,包括但不限于某些纤维二糖水解酶,例如CBH1或CBH2,或某些内切葡聚糖酶。

[0098] 举例说,当重组蛋白是CBH2变体时,优选在不存在潜在有害的天然纤维素分解活性的情况下产生变体。因此,获得这样的木霉属宿主菌株是有用的,其在引入含有编码变体CBH2的DNA片段的DNA构建体或质粒前已缺失一个或多个纤维素酶基因。像这样的合适多重缺失菌株可以通过公开于例如美国专利No.5,246,853和PCT公开W0 92/06209的方法进行制备,所述专利的公开内容以引用方式并入本文。通过在缺少一个或多个纤维素酶基因的宿主微生物中表达变体CBH2纤维素酶,简化了鉴定和随后的纯化程序。任何来自木霉属的种的已克隆的基因可以如此进行缺失,例如cbh1、cbh2、eg11和eg12基因以及编码EG III和/或EGV蛋白质的那些基因可以从木霉属宿主菌株中缺失(分别参见例如,美国专利No.5,475,101和PCT公开W094/28117)。

[0099] 可通过将某种形式的待缺失或破坏的所需基因插入质粒中来实现基因缺失。然后在适当的限制性酶位点(在所需的基因编码区的内部)消化该缺失质粒,并且用可选择标记替换基因编码序列或其部分。来自待缺失或破坏的基因的基因座的侧翼DNA序列(优选具有约0.5至约2.0kb之间的大小)可以保留在可选择标记基因的任一侧上。合适的缺失质粒将

通常具有存在于其中的独特的限制性酶位点,以使得含有该缺失基因的片段(包括侧翼DNA序列)和可选择标记基因能够作为单一线性小片被去除。

[0100] 在一些实施方案中,挑选可选择标记以使得经转化的微生物的检测成为可能。任何在选定的微生物中被表达的可选择标记基因都可以是合适的。例如,对于曲霉属菌种,可以挑选可选择标记,以使得可选择标记在转化株中的存在不会显著改变微生物的性质。合适的可选择标记的例子是编码可测定的产物的基因。例如,可使用曲霉属的种的这样的基因的功能拷贝,如果该基因在宿主菌株中缺乏,会导致该宿主菌株显示营养缺陷型表型。对于木霉属的种也存在可选择标记。

[0101] 在一些实施方案中,用功能性pyrG基因转化曲霉属的种的pyrG衍生菌株,该基因提供用于转化的可选择标记。通过选择对氟乳清酸(FOA)有抗性的曲霉属的种菌株可获得pyrG衍生菌株。pyrG基因编码乳清酸核苷-5'-单磷酸脱羧酶,该酶是尿苷生物合成所需要的酶。具有完整pyrG基因的菌株能在缺乏尿苷的培养基中生长但对氟乳清酸敏感。因此,FOA抗性选择可以用于选择缺乏功能性乳清酸核苷单磷酸脱羧酶并且因而需要尿苷进行生长的pyrG-衍生菌株。使用FOA选择技术,还可以获得缺乏功能性乳清酸焦磷酸核糖基转移酶的需要尿苷的菌株。这些细胞可以用编码这个酶的基因的功能拷贝进行转化(Berges&Barreau,Curr.Genet.(当代遗传学)19:359-365(1991年)和van Hartingsveldt等人,(1986年)Mol.Gen.Genet.(分子遗传学与普通遗传学)206:71-75)。使用上文描述的FOA抗性技术来进行衍生菌株的选择。在一些实施方案中,将pyrG基因用作可选择标记。

[0102] 在一些实施方案中,用功能性pyr4基因转化肉座菌属的种(木霉属的种)的pyr4-衍生菌株,所述功能性pyr4基因提供用于转化的可选择标记。通过选择对氟乳清酸(FOA)有抗性的肉座菌属的种(木霉属的种)菌株,可获得pyr4衍生菌株。pyr4基因编码乳清酸核苷-5'-单磷酸脱羧酶,该酶是尿苷生物合成所需要的酶。具有完整pyr4基因的菌株能在缺乏尿苷的培养基中生长但对氟乳清酸敏感。因此,FOA抗性可以用于选择缺乏功能性乳清酸核苷单磷酸脱羧酶并且因而需要尿苷进行生长的pyr4衍生菌株。使用FOA选择技术,还可以获得缺乏功能性乳清酸焦磷酸核糖基转移酶的需要尿苷的菌株。这些细胞可以用编码这个酶的基因的功能拷贝进行转化(Berges&Barreau,1991年)。使用如上文描述的FOA抗性技术来进行衍生菌株的选择。在一些实施方案中,将pyr4基因用作可选择标记。

[0103] 随后从缺失质粒中分离包含被破坏或缺失的有害基因(例如上文例示的一个)的单一DNA片段,并用于转化适当的pyrG-曲霉属或pyr4-木霉属宿主。转化株基于它们分别表达pyrG或pyr4基因产物从而补充宿主菌株的尿苷营养缺陷型的能力进行鉴定和选择。可适当地对所得到的转化株进行DNA印迹分析,以鉴定和证实双交换整合事件,在该事件的过程中该基因的基因组拷贝的部分或全部编码区被缺失且替换为适当的pyr可选择标记。

[0104] 虽然上文描述的特定质粒载体涉及pyr转化株的制备,但本发明并不限于这些载体。使用上文描述的以上技术可在曲霉属的种或肉座菌属的种(木霉属的种)中缺失和替换各种基因。另外,如本文讨论的,有许多可选择标记是合适的。事实上,任何已被鉴定的基因都可使用上述策略从任何宿主的基因组中适当地缺失,所述宿主例如曲霉属的种或肉座菌属的种。

[0105] 在某些实施方案中,使用的宿主菌株可以是肉座菌属的种(木霉属的种)的这样的衍生物,其缺乏或具有与所挑选的可选择标记对应的一个或多个非功能性基因。例如,如果

挑选pyrG可选择标记用于曲霉属的种,则将特定的pyrG衍生菌株用作转化程序中的受体。在另一个实例中,如果挑选pyr4可选择标记用于肉座菌属的种,则将特定的pyr4衍生菌株用作转化程序中的受体。在一些实施方案中,可以使用包含类似于构巢曲霉基因的肉座菌属的种(木霉属的种)基因的可选择标记,包括例如amdS、argB、trpC或niaD。对应的受体菌株相应地为衍生菌株,例如分别为amdS⁻、argB⁻、trpC⁻或niaD⁻菌株。

[0106] 随后可以制备编码目的蛋白或变体的DNA以用于插入适当的微生物中。根据本发明,编码目的蛋白或变体的DNA可以包含编码具有野生型蛋白质的活性的蛋白质或变体的DNA。可将编码目的蛋白或变体的DNA片段功能连接至真菌启动子序列,例如,曲霉属中的glaA基因的启动子,或木霉属中的cbh1或eg11基因的启动子。

[0107] 可通过构建携带编码目的蛋白或目的变体的DNA的表达载体来制备编码该蛋白或变体的DNA。携带编码目的蛋白或变体的插入DNA片段的表达载体可以是例如任何这样的载体,其能够在给定宿主生物中自主复制,或能够整合到宿主的DNA中,通常为质粒的形式。在某些实施方案中,设想到两种类型的表达载体来获得基因的表达。第一类含有这样的DNA序列,其中启动子、基因编码区和终止子序列都源于待表达的基因。如有需要,可通过缺失不需要的DNA序列(例如,编码不希望有的结构域的DNA序列)来获得基因截短,以使待表达的结构域处于其自身转录和翻译调控序列的控制下。可选择标记也可作为载体的一部分被包含,从而允许对多个拷贝的所需基因序列在宿主中的整合情况进行选择。

[0108] 第二类表达载体是预先装配的,含有对于高水平转录有用的序列并含有可选择标记。设想到可将基因或其部分的编码区插入这种通用表达载体中,将其置于表达盒启动子和终止子序列的转录控制下。这种通用表达载体的一个非限制性例子是曲霉属中的pRAX。可将目的基因或变体基因或其部分插入强glaA启动子的下游。这种通用表达载体的一个非限制性例子是肉座菌属中的pTEX。可将目的基因或变体基因或其部分插入强cbh1启动子的下游。

[0109] 在某些实施方案中,在载体中,将编码目的蛋白或变体的DNA序列有效连接至转录和翻译序列,例如合适的启动子序列和信号序列,与结构基因同阅读框。启动子适当地是在任何在特定宿主细胞中显示转录活性的DNA序列,并且可源自编码对宿主细胞而言同源或异源的蛋白质的基因。可以提供任意的信号肽以用于目的蛋白或变体的胞外生产。编码信号序列的DNA优选是与待表达的基因天然结合的DNA。然而,设想到来自任何合适来源的信号序列,例如来自外切纤维二糖水解酶或来自木霉属的内切葡聚糖酶。

[0110] 可以用于将编码目的蛋白或变体的DNA序列连接至启动子和将这种构建体插入合适载体中的方案是本领域已知的。

[0111] 可按照已知的技术如转化、转染、微注射、微穿孔、生物射弹轰击等,将本文描述的DNA载体或构建体引入宿主细胞中。

[0112] 例如,当本文描述的DNA载体或构建体用于转化真菌宿主细胞时,肉座菌属的种(木霉属的种)的细胞壁对于DNA的透性可能很低。因此,所需的DNA序列、基因或基因片段的摄取通常是极微。有许多方法可用来在转化过程前增加衍生菌株(例如,缺乏对应于所使用的可选择标记的功能性基因的衍生菌株)中肉座菌属的种(木霉属的种)细胞壁的透性。

[0113] 在某些实施方案中,为了制备用于转化的曲霉属的种或肉座菌属的种(木霉属的种),制备来自真菌菌丝体的原生质体。参见Campbell等人Curr.Genet.(当代遗传学)16:

53-56;1989年。菌丝体可从萌发的营养孢子获得。用能消化细胞壁的酶处理菌丝体,从而得到原生质体。然后通过悬浮介质中存在的渗透稳定剂来保护原生质体。合适的稳定剂包括例如山梨糖醇、甘露醇、氯化钾、硫酸镁等。通常,稳定剂的浓度可以在0.8M至1.2M之间(例如在0.9M至1.2M之间、在1.0M至1.2M之间、在1.1M至1.2M之间等)变化。在一个具体的实施例中,1.2M山梨糖醇用作悬浮介质中的稳定剂。

[0114] DNA向宿主菌株(例如曲霉属的种或肉座菌属的种(木霉属的种)中的摄取通常可依赖于钙离子浓度。一般来讲,在摄取溶液中使用约10mM CaCl₂至约50mM CaCl₂(例如约15mM至约45mM、约20mM至约40mM、约25mM至约35mM)。除了在摄取溶液中包括钙离子外,通常还包括的其他东西是缓冲系统如TE缓冲液(10mM Tris,pH 7.4;1mM EDTA)或10mM MOPS,pH 6.0缓冲液(吗啉丙磺酸)和聚乙二醇(PEG)。据认为,这个缓冲液中的聚乙二醇的作用是融化细胞膜,从而让介质的内容物被递送至宿主菌株(例如曲霉属的种或肉座菌属的种)的细胞质中,并且质粒DNA被转移至细胞核。在某些实施方案中,这个融化过程使得质粒DNA的多个拷贝整合到宿主染色体内。

[0115] 通常,将含有密度为10⁵至10⁶/mL(优选2×10⁵/mL)、已进行透化处理的曲霉属的种的原生质体或细胞的悬浮液用于转化。类似地,将含有密度为10⁸至10⁹/mL(优选2×10⁸/mL)、已进行透化处理的肉座菌属的种(木霉属的种)的原生质体或细胞的悬浮液用于转化。将这些原生质体或细胞在适当的溶液(例如,1.2M山梨糖醇;50mM CaCl₂)中的100μL体积与所需DNA混合。在一些实施方案中,将大量PEG加入摄取溶液。例如,可将约0.1至约1体积的25%PEG 4000加入原生质体悬浮液。在一个具体的实例中,将约0.25体积的25%PEG 4000加入原生质体悬浮液。也可向摄取溶液加入添加剂例如二甲基亚砷、肝素、亚精胺、氯化钾等,帮助转化。

[0116] 在某些实施方案中,将混合物在约0℃下温育约10至约30分钟的时间。然后向混合物加入另外的PEG以进一步增强所需的基因或DNA序列的摄取。在某些实施方案中,可以以转化混合物体积的5至15倍的体积加入25%PEG 4000;然而,更大和更小的体积也可以是合适的。例如,在一些实施方案中,以转化混合物体积的10倍体积加入25%PEG 4000。加入PEG后,接着使转化混合物在室温下或冰上温育,然后加入山梨糖醇和CaCl₂溶液。然后将原生质体悬浮液进一步加入到生长培养基的融化等分试样。这个生长培养基能使转化株生长。有许多生长培养基可以适当地用于生长本发明的所需转化株。在某些实施方案中,例如,如果选择Pyr⁺转化株,那么优选使用不含尿苷的生长培养基。例如,菌落在耗尽尿苷的生长培养基上进行转移并纯化。

[0117] 在这个阶段,稳定转化株可以通过其更快的生长速率区别于不稳定转化株。此外,对于许多丝状真菌宿主,例如木霉属,在缺乏尿苷的固体培养基上具有平滑轮廓而不是凹凸不平轮廓的圆形菌落的形成可以用作区别特征。在一些实施方案中,可以通过以下方式对稳定性进行进一步测试和选择:在固体非选择性培养基(例如含有尿苷的培养基)上生长转化株,从这个培养基收获孢子,并测定这些孢子的百分比。所选孢子可在缺乏尿苷的选择性培养基上萌发并生长。

[0118] C.将重组蛋白编码核酸序列引入宿主细胞中。

[0119] 本发明进一步提供经遗传修饰以包含重组蛋白编码核酸序列的细胞和细胞组合物。可用克隆载体或表达载体对亲本细胞或细胞系进行遗传修饰(即,转导、转化或转染)。

如上进一步描述,载体可以是例如质粒、病毒粒子、噬菌体等的形式。

[0120] 本发明的转化方法可导致转化载体的全部或部分稳定整合到丝状真菌的基因组中。还设想到导致自主复制的染色体外转化载体得以维持的转化。

[0121] 有许多标准的转染方法可用于生产能表达大量非天然或天然蛋白质的丝状真菌(例如里氏木霉)细胞系。例如,有许多已公布的方法可用于将DNA构建体引入木霉属的产酶菌株中,包括Lorito等人,1993年,Curr.Genet.(当代遗传学)24:349-356;Goldman等人,1990年,Curr.Genet.(当代遗传学)17:169-174;Penttila等人,1987年,Gene(基因)6:155-164;用于将DNA构建体引入曲霉属的产酶菌株中,包括Yelton等人,1984年,Proc.Natl.Acad.Sci.USA(美国国家科学院院刊)81:1470-1474;用于将DNA构建体引入镰刀菌属的产酶菌株中,包括Bajar等人,1991年,Proc.Natl.Acad.Sci.USA(美国国家科学院院刊)88:8202-8212;和用于将DNA构建体引入链霉菌属的产酶菌株中,包括Hopwood等人,1985年,英国诺里奇的约翰恩斯基金会(The John Innes Foundation,Norwich,UK),和用于芽孢杆菌属,包括Brigidi等人,1990年,FEMS Microbiol.Lett.(欧洲微生物学会联合会微生物学快报)55:135-138)。

[0122] 用于将嵌合构建体(表达载体)引入丝状真菌(例如红褐肉座菌)中的其他方法包括但不限于使用粒子枪或基因枪、在转化过程之前透化丝状真菌细胞壁(例如,通过使用高浓度碱,例如,0.05M至0.4M CaCl₂或乙酸锂)、原生质体融合或农杆菌介导的转化。这种通过用聚乙二醇和CaCl₂处理原生质体或原生质球来转化丝状真菌的方法的一个例子在Campbell,E.I.等人,Curr.Genet.(当代遗传学)16:53-56,1989年;和Penttila,M.等人,Gene(基因),63:11-22,1988年中描述。

[0123] 可使用任何已知的用于将外来核苷酸序列引入宿主细胞中的程序。这些程序包括例如使用磷酸钙转染、聚凝胺(polybrene)、原生质体融合、电穿孔、生物射弹(biolistics)、脂质体、微注射、细胞质载体(plasma vector)、病毒载体和任何其他已知的用于将克隆的基因组DNA、cDNA、合成DNA或其他外来遗传材料引入宿主细胞中的方法(参见,例如,Sambrook等人,同上)。同样有用的是在美国专利No.6,255,115中描述的农杆菌介导的转染方法。重要的是具体使用的遗传工程程序能够将至少一种基因成功引入能够表达非天然或内源(但为非天然形式)的基因的宿主细胞中。

[0124] 在一些实施方案中,可以在体外转录包含重组蛋白编码核酸序列的嵌合构建体,并且可通过已知方法例如注射将所得到的RNA引入宿主细胞中。

[0125] 本发明进一步包括用于生产真菌酶、其变体和包含这些分子的组合物的丝状真菌(例如红褐肉座菌和黑曲霉)的新型和有用的转化株。本发明包括丝状真菌尤其是包含某些重组蛋白编码序列或缺失某些内源编码序列的真菌的转化株。

[0126] 在引入包含目的蛋白或其变体的编码序列的嵌合构建体后,可将经遗传修饰的细胞在经过适当改良的常规营养培养基中培养,所述改良是为了激活启动子、选择转化株或扩增重组蛋白编码核酸序列的表达。培养条件(例如温度、pH等)是先前用于被选择进行表达的宿主细胞的那些条件,并且对于本领域的技术人员将显而易见。

[0127] 这种嵌合构建体已引入其中的细胞的后代一般视为包含在嵌合构建体中存在的蛋白编码核酸序列。

[0128] 本发明进一步包括用于生产真菌酶、其变体或包含这些分子的组合物中的丝状真

CGAAGAGACCCTAAAAAAGCTCGTGGATGGCTGTGGCGACGACTCAAGGTTGCCTCCATGGTCGAGGCCATTGACG
 ACTTGAATGAGCCACGAGCAAAGTCGGTCAAGCCTTACAAAACGTACGAAGGTCTCTTGACCTTGGGAGATCCGAAA
 AACGCTCCCGCAGTGGTGGAAATCCGCGTCGAGAGATACTTCAAGACCCATCTAGCCAGGCCACCTGCCGCCAGCAC
 CGTGGTGGTCAAGGAGGAGCAAGCTGGGCCGTCTCAGGCAGACGAGGACGAACAGATGGACGGAGCGGAACCTTACAG
 CTGTGAGGCAGGCCAGGACATAACAAGTCAATGATCCAGATGCCCTGGCGGTAAGCGTGACGTTGAGTTTGTGAGTCT
 CTGGCCAAAGGGTACGAGTACGGCAGGACGGCAGTCCACATCAGCGAGTCTGATCAAAAACGTCACCAAGCTCGCGAC
 AGAAAAGAGCTTCAAGATCATCGGCTTCGTCCAGAAAAGAAAGGTATTGGCTTGGCTCTCAGCATTGACCCGTTGC
 TCTTGGCTAACCCCTTGTTTAGTATGAAATGCTCCTTAATCTTGGCGAAACCTGCGTTACCGTTGCATCCAAGTACGA
 TGAAGTCTGAGCTGGCTTTTGTAGTCTCTGGTGTGGGCGCTCTCGGAGCTCGACGCCTACGCCGTGGCCCGCTAG
 TAACTAAGGACCAAAAGGACCCCATGCTGGTGTACTGATGCCGTATATGGAGCCTGATTATGTTTGTCTCTATGAT
 GTGCCTCTGCCTTTCGACAGGACATCAGGACGTACCAGTTTCCCTCCCTTGGACAGAGTCGTTACCGTCAGTGGCCA
 AACGCTCACCAACCATCGCCTATTGCCATCCGACGAGCTCAACCAAGCGATGAGCGACTACGTAGATGCCATGGACA
 TTTCAAGTTATGGTATCGATGAAGATGGGTGAGTATAGAAGATGATTGTTCAAATCTTCACTTCTAAGCATTGCTT
 CTGATCTAGGCAACCGGCTGAATATGCCACCATCGATGAGTTATAACAACCTGCGATACATCGCATAGGCCATGCGA
 TCAAACAACGAGCGATCCACCCAGAGAAAACCGTGCCGAGATCCCCCAGTCTTGCTTAGATTGCGAGCACCCCG
 ACAGAACTCGTCGAGACTGTGCAGCCTCATATCGATGCACTGATTCACGCTGCAGACGTGAAGAAAGTACTGATTC
 CATTACATATGCTTCTCTGCACACTGATGTTGATTTGTGCTAACGCCCCCTTAGTGCCGCCAAGGCCAAGGGCA
 AGCGCCAAAGAGAAACAGTTAAACCCATCTCGGGACTGGATGTGGATGCCCTTCTGGGAGAAGAGCAGAAAGTTCC
 ATTAGTCCGGAGAATGCCATTCCGGACTTCAAACGAGCCCTCAACTCGTCCGAAGAAGTCGAGCAGATTGCCGACGC
 CAAAAACAAATGGGGGCCATTGTGCGGTCTCTCATTACGGACAGCTTCGGGGATAGCAAATATGCCCAGGCAATGG
 AAGGCATTGGTGCATGCGTGAGGAGCTGATCAACCTGGAAGAGCCTGGCCTGTACAACGACTTTGTGCGCGACTTG
 AAGAAAAGTTTGTATCTGGAGCCTTGGGTGGTGACAGGCGAGATTTCTGGTTCAAGATGAGGTGGGCGAAGCTGGG
 CCTGATTGACAAGAAACAGTCGGAGGTGCTTCCGTCCTTCTGAGGAGGCGGACGAGGTGAGTGGTGCAGCATGCT
 GTCGGATTATACGGACGTTGTTTGTAACTTGTGGGATAGTTTACAAGTCGAGGTGAGGTATCTACGTTGACCAAG
 AATGGGACCATGTATATGAGCGGTGTAACAACAGAATCCTGTGCTTTGAGCATTGTATGA

[0138] 为从quad缺失衍生菌株中缺失里氏木霉ku80基因,将如在例如PCT公开W0 2005/
 001036中一般描述的标准方法为了这个目的进行改编。简言之,利用ku80缺失盒,其采用在
 1.3kb的5'ku80序列和2.3kb的3'ku80序列之间侧接的可选择标记,如图2中示意性显示。赋
 予对于除草剂氯嘧磺隆的抗性的变体里氏木霉a1s用作可选择标记。参见例如,PCT公开
 W02008/039370。ku80敲除盒的核苷酸序列的长度为7685个碱基对:碱基1-1271对应于5'
 ku80同源区;碱基1280-7685对应于a1s-氯嘧磺隆抗性变体(A190D);碱基5381-7685对应于
 3'ku80同源区。ku80敲除盒的核苷酸序列作为SEQ ID NO:1提供:

[0139]

GGCCGCTCAACACCCACACTCGAGGCACACGAGTTCATCGGCGGCTTCCCCACAAGCTCTCGGCCAACCTGCTAC
 CGGCTCTCTCGCGAGACTTCCCAAAGCCTACAAACGAGGTCGACGTCAAGGAGGCCCTCGAGCGCCAGCCCGGCGA
 TGGAGCCTCCAGGGCCAGATCAAGGCCAACACATGAGAGCCAGAGCGCCGCACTCCGGCTCGACGACAAGGAGGG
 CAAGGCGAGAGCCTTTGAGGAGGCCAAGCGCGAGCTACTGGCGTATACCACAGCGCCCTGCGGAAGCCTTCCGGCG
 CAAGATAATGAGCTTGATCGCAATGACGAGTTCACGTACGCTTGGCCATATTGTTGTTGCTTTTTGTTTGGTCTAC
 ATGTACGGCGCATTGGTTGGGAGGATATACCCACGGAGAGTGTCCGAGTGGCTTCTGGGATTTAGAGCGTCATTAGC

AGGATAGAGATGGTTGGCCAGGGGAATGGAATTGACTTTTCACTACAAGGAACCTGTTCACTCTGGTGTGATTCCC
 ATTGCGTGACTGGTAGTAGGGAGGAATGCTTTTACTTTGTGCCACTAGACCGCAGAGAAGGGTTGGTTGCAAGCGGG
 GTCCGTGTATAACCGACCAAGAGTGATGGGCATACAGCAACGTTTCTGAACGACTTCATTTTGTCCGAGTCTACTGGA
 TGCGAGATGCCAGCGTGAAGCCGTACGCCACCAGGGCGACGAACCTCGACAAGGTTGACGAGGGAGGAGATGCCGTGC
 AGCATGCCAAACTTCTTGTGAGGGCACGCATCTCATCCGACTGTGCATCCTTGTACATACCCTCTTTCCGTCTCG
 CTTGGCTGGTGGGAGGGTTCAACAAATCCATCGTCAGCCATCCGGGGTCTCAAATCAATGGCGTGCATGCGGAGTGC
 GGCTTGAGGCTAACCTTGTCCATGGCGGTCTTTCATGGTCTTGACAGTGGCGGAAGCAGCACGGCGAGGTTGACGA
 GGCCGCTGACGAACATGGTTGCGATGGGCACCAAGGAGCTCCACTTGTGGGAGCGTCGACGAGGCCGCCGATGCCG
 CCCTTGATGCCAAGAGGGCGTTTCCGGGGAACGTGAGGGCGAGCAGCGCGGGGATGGCCGTCTGCATGCCAAAGTA
 GATGGGGAACAGCTTGTCTGGATGGCGGAGAAGGAGGGCCGGCTGACGGTGCAGAACATGACGATGCCGTTGACGA
 AGGACTGCAGTAGCGTAGTGTGATGGTAAGCAGCTGGCCGGCGCCTGAGACAATGGCCGGCAATGGTAAAAAGGA
 CCAAGATGTACTAGGTAGTTGCAATGTGGCTTATTACCTACCTACTACCTGGTAGGCACCTACTAGGTACTTGGGTA
 GACGGACAATGAAATTTGAAGTCGGGGTTCAGGAAAAGCAGGGCGCTGGACACATTTGTGCTTCAGGCGGTACCCGTC
 GTCATCGTCAGCCAATGTCGAGGCCCGGCAGCCCGAGGAGCGAGACAACCTTGGCCGGAGGAGCCCGCAGGTACCTG
 CCAAAGCGCGGCTGGTACCTCTCAACCCTCTCAGGCCGTGTTGGATGCCCTATGACATGCCCTGGGGGATGCAGCTGT
 TGCCCCGGCCCCGCACTTTCGGGTGACCGCGAGGCTGCTGATTGGCTGGTTGCCACGGGCTGGGCGGTCCCTGAAGT
 TGTTGCCATCTGAACTCTGTGCGCGCTGGCGTCGGCTGCGCCCAATGGGAGGCGAGACAACCTCAGGGTACTAGAATC
 ACTGACAGAAGAAGAGAATCGAAAGTAGGTAGACAGCCAATTCGTTGCATGGCAGGCAACCGCACAGGAGAAAAATT
 GACTACCCACAATCAGGCACAGTAAGTAGGGCACAGTACGTATGTACAGACAAGGCGCAAGCGATACTGCGCGACC
 CGGTACCTCGCCGGCTTGACACGTGCGACAGGCTACTTTACTAGTATTCGACGCGGCGGGTTCGCGCATTATTACATG
 TACTGTGCCGCCATTTGATGACTGGGCTGCTGCAGTATTAGTAGATCTGCCCGCATCGCCCTTCCATGGGCGCGAC
 CCGGGACTGGACCCTCTGACTCTACCTACATGTACCTAGGCCGGGCCGGGCTTGGTGACTTTTGTCCGATCAGGTGC
 TTCGCCTGGCTACCTATTATTTCTCTTTCTTCTTCTCCATCCTGCTTCTGGCCTTGCAATTTCTTCTCGCCACTCCT
 CCCTCTTCCCCCGCGATACCCTTGAATTCGTCAGAGAGGAAAAGACGAGAAAAAAAAGGGCAGCAGAGACGTCCGT
 CTGGCTCACGTGCTGCATCTCTGCGCACTCTCATTTTTTTTTTATTTGTCGGACCCCTCCCTCAACCTTCTCCTTCGTTG
 ACAGGCTAAGCCTTGCTTCGACGCTCTCTCTTTGAAATTTTTTCTACTTCTACCTTCTTTTCTTGGCGTTACCCACCA
 TAGCTCGATTACGATGCTCCGAAGTCGCAAGTCACAGCCAGGGCCGTCCGGGCTCTGGGCCAGGCGCGCGCCTTT
 ACCTCGACGACCAAGCCTGTCATGATCCAGAGCAGCCAGAGGAAACAGGCCAACGCCAGCGCTGCTCCGTAAGTCGC
 CCATTGCCATTGCATCTTCTGTTTGATATATACTTCTGCTGCTTGGCTGGCGTCGTCTCTCGGTTATGCGTGTCAA
 GGACCAGGTGTGTTGCGATCGTGGTTTTCCAGCGCGGATTACCGGGGACGAATTTTTGGCTGCTCAACTCGCGCGC
 GCGCATTCTGATTCTTCTGTTTTCAATCTTGAGCGACAACCTGGCTAACATAATGGCCATTGGCAATTGCTTCACACAG
 ACAAGTCCGCCCTGTACCGAGCCCTGCTTTCAACGCTGAAGACAAAGACCGCAGCCATGTGCAGCCTCTGGTCAACC
 CGTCGAAGCCCGACATGGATGAATCGTATGTCCACGTCCCCTCGTCCCGCCCTACAAAATGAACACGATTACACCAG
 AATTTTTGCAACAATCGACACTTCTATAACAGACCAATTGAGCTTTGTTCTGACCAATCATGTTGCTCTAGATTTCAT
 TGGCAAAACCGGAGGCGAAATCTTCCAGAGATGATGCTGCGACAGGGTGTCAAGCACATTTGTAGGTTCCGATGCC
 GGCCGCCACACGGGCTCCATCCTTGTCCATCTCTCCAGCTAGGCAAATCTCGCTAACCTTGTGTCACCATCCAGT
 CGGATACCCTGGCGGCGCTATCCTGCCGTCTTTCGACGCCATCTACAACCTAAAACACTTCGACTTCATCCTGCCCC
 GTCATGAGCAGGGAGCTGGCCATATGGCCGAGGGCTATGCCCGTGCCCTCGGGCAAACCCGGTGTGCTCCTGGTGACT
 TCCGGCCCCGGTGCTACCAATGTCATCACGCCATGCAGGATGCCCTGTGCGACGGAACGCCCTTGGTGTCTTCTG

CGGCCAGGTCCCACCACGGCCATCGGCAGCGATGACTTCCAAGAGGCCGACGTCGTGGGCATCTCGCGGGCCTGCA
 CCAAGTGGAAACGTCATGGTCAAGAGCGTTGCTGAGCTGCCGCGAGAATCAACGAGGCCTTTGAGATTGCCACCAGC
 GGCCGCCCTGGCCCCGTCTCGTCGACCTGCCAAGGATGTCACGGCTGGTATCCTGAGGAGAGCCATCCCTACGGA
 GACTGCTCTGCCGTCTTGCCAGTGCCGCCCTCCCGCGCCGCATGGAGCTGAGCTCCAAGCAGCTCAACGCCTCCA
 TCAAGCGTGCCGCCGACCTCATCAACATCGCCAAGAAGCCCGTCATCTACGCCGTCAGGGTGTATCCAGTCCGAG
 GGCGGCGTTGAGCTCCTGAAGCAGCTGGCGGACAAGGCCCTCCATCCCCGTCACCACCACCCTCCATGGCCTGGGTGC
 CTTTGATGAGCTGGACGAGAAGTCGCTGCACATGCTGGGCATGCACGGCTCGGCGTATGCCAACATGGCCATGCAGC
 AGGCCGACCTCATCATCGCCCTCGGCAGCCGATTCGACGACCGTGTACTCTGAATGTCTCCAAATTTGCGCCTGCA
 GCCAGGCAAGCTGCTGCCGAGGGCCGCGGCGGCATCATTCACTTTGAGATCATGCCAAGAACATCAACAAGGTCAT
 CCAGGCGACCGAGGCCGTCGAGGGCGACGTCGCCACCAACCTGAAGCACCTCATTCCCCAGATTGCCGAAAAGTCCA
 TGGCGGACCGAGGAGAGTGGTTTCGGCCTCATCAATGAGTGGAGAAGAAGTGGCCCCTGTCAAACCTACCAGCGCGCG
 GAGCGGGCTGGCCTCATCAAGCCGACGACGGTCATGGAGGAGATTAGCAACCTGACGGCCAACCGAAAGGACAAGAC
 GTACATTGCCACGGGTGTGCGCCAGCACCAGATGTGGTTGCCAGCACTTCCGCTGGAGGCACCCTCGATCCATGA
 TTACCTCTGGTGGTCTGGGCACCATGGGCTACGGTCTCCCCGCGCCATTGGCGCCAAGGTGGCCCAGCCCGACGCT
 CTCGTAATTGACGTTGATGGCGATGCCTCGTTAACATGACGCTGACGGAGCTGTCGACTGCTGCACAGTTCAACAT
 TGGCGTCAAGGTGGTTGTGCTCAACAACGAGGAGCAGGGCATGGTGACGCAGTGGCAGAACCTCTTTTACGAGGACC
 GATATGCCACACGCACCAGAAGAACCCCGACTTCATGAAGCTGGCCGACCCATGGGCGTTCAGCACCAGCGCGTG
 ACGGAGCCGGAGAAGCTGGTCGATGCCCTGACGTGGCTGATCAACACCGATGGCCCGGCCCTGTTGGAGGTTGTAC
 GGACAAGAAGGTGCCTGTCTGCCCATGGTGCCCGCCGATCGGCCCTGCACGAGTTCCTCGTCTTTGAACCTGGTG
 AGTCTACTTCAGACATATTGCTTGCGCATTGCAGATACTAACACTCTCACAGAAAAGGATAAGCAGCGCCGTGAGCT
 GATGAAGGAGAGAACAAAGGGTGTGCACTCCTAAAGCGATGATGTCTGCGAGGGGTTCTTCGTTGAACCCTAGTTCA
 GGCACCATCTTACCCTCTTATTTTTTCCCGTGGGCTTTCATTTGTGTCATCCGAGCATGACGTTGTAGGGTTGGAG
 TTTCTTCCTTTTTATCTTGTCACTTACTGGTACCCATAGGCGCGAGACTAGGCTTCCATGTTTTGTGTTTGGGACTTT
 CAAAAAGTACTTTTAGTGGTTTGGGGCACGACGAGGGGGGCAACCTCTTCTGTCGAAAAAGGTGGCTGGATGGATG
 AGATGAGATGAGATGAGGGTGAAGATAGATACCTGCAGTGTTTTTGACGCGACGGGATGGCGATCGCAGCACCCCG
 ACAGAACTCGTCGAGACTGTGCAGCCTCATATCGATGCACTGATTCACGCTGCAGACGTGAAGAAAGGTACTGATTC
 CATTACATATGCTTCTCTGCACACTGATGTTGATTTGTGCTAACGCCCCCTTAGTGCCGCCAAGGCCAAGGGCA
 AGCGCCAAAGAGAAAACAGTTAAACCCATCTCGGGACTGGATGTGGATGCCCTTCTGGGAGAAGAGCAGAAAGGTTCC
 ATTAGTCCGGAGAATGCCATTCCGGACTTCAAACGAGCCCTCAACTCGTCCGAAGAAGTCGAGCAGATTGCCGACGC
 CAAAAACAAATGGGGGCCATTGTGCGGTCTCTCATTACGGACAGCTTCGGGGATAGCAAATATGCCCAGGCAATGG
 AAGGCATTGGTGCATGCGTGAGGAGCTGATCAACCTGGAAGAGCCTGGCCTGTACAACGACTTTGTGCGCGACTTG
 AAGAAAAGTTTTGCTATCTGGAGCCTTGGGTGGTGACAGGCGAGATTTCTGGTTCAAGATGAGGTGGGCGAAGCTGGG
 CCTGATTGACAAGAAACAGTCGGAGGTGCTTCCGGTCACTCTTGAGGAGGCGACGAGGTGAGTGGTGACGATGCT
 GTCGGATTATACGGACGTTGTTTGTAACTGTGGGATAGTTTTACAAGTCGAGGTGAGGTATCTACGTTGACCAAG
 AATGGGACCATGTATATGAGCGGTGTAACAACAGAAATCCTGTGCTTTGAGCATTGTATGATATGATTATTGATGAAC
 CGGACAAAAGGGGGTAGGGGATTGATGCCATCACGACCGATTGACCAGACCTGGATTCTCGCACAGCATGGCTGCTG
 ATTTTGTGACCTTGGCAGGTAACATCCCTGAAGAACAACCTACTATTAACCTATCATTTAGCAGAAGCTCTGTAAC
 CTTCTTGATTCTTGATTACAGCTTCTGAGTCTGTCAAATGTAATCATTTCGAGGTTGTGTAATTCGGGCCAAGCAGG
 CGGCCGTCTGCCAGCGCCTGCCTAGGCTGCACCGCAATCTGCCAATCAGCTGCCCTTCAGTTTCGTTTGACCTTGC

AGCTGCCCTTCATCCTTTATCTGCACACAATTCTTTTTTCCTCTGCTCTGCGCATTCTTCTCTCTCTCGTCTCCCTTC
TCAAGCTCAACTTCACCTCATCCGCTCCACTACAAGCCCTCCCGTCGTCTCGCATCCTCATCTCGACTGCGGCC
AGCAAAACAAGCAAAGCCGTGATCGATCCTCAGCATGGCTACCTTCAACCTCACCGTCCGCTGGAGATGCTCAAAG
AAATTGGAATCACCGTCCAATACGGCGAGCATGTAGCGAAAAGAAGCAGCCAGCAACGAAGCAGCGATGGCATTTCGAA
GAAGAAGAAGAGTTCCCGCCGTTGTGCCGCCAAGGCAGAACAGCACGCCTCTGAACACGACGCTGGCCACGATGC
TTGGGACGCGGCTGCCACATCTCGACTTCGGCGCAAGAACAGCAGAAGCCCCAGGAGATGGACGACTCGTCTATCG
TGATGCCGCTGGACTACTCCAAGTTTGTCTGGAGAGCCTGCGGACGAATCCATCAGCTTTTGTCTCGTGGAAGGTC
GTCGAGGCTTATCCTGACCAGTTTATCGGCAAGGCAAACAGGCCTCGTGTATGTAGCGATTGCTTTCTCTGCATTAT
GGGAATCTCAAGAGAGTATGGTAGAAGATAACTGACAACCTGCAGGCCAAGCCGTACTTTGACAAGATTTTGGAAGA
CAGAGTCTGGGATTTGTGAGGATCTTGATTGATGTGCATATGGCGACATGCCTGCTAATATCATTGTAGCTTCTATC
TCTACAACCCCGAGAAGCCTTCAGAGAAGCCTCGCGTGCTGGTGCCACTGTTTCAGCTCGAAGGCTTTCTCAAAGC
ATCAACAGAGCGCTCGGTACTTCTCTCACCATTCCAGGAGGGGCAAACCAGGACCGTTTTTATCTGAGGTTCCGCCA
GGGAGACACCCCAAGGCCTCGATATCTACAGAGGTCGAGAGACCAGAAATCCCTAAAGATTGAAACGTTCCCGATT
TTCAACAGGCGGACTACGACAGCTTTAGGAACGCGCATGGCGCCATCCAGGAGGACTGGTTGAAGAACTGGCAGATG
CTGGTACCTCGGCCGAGTTTCGACAAGAAGAAAAATGCAGACAAAAGAGCAGCCAAGAGAAGGCTCGAGCGAGAGCG
AATGCTTACAATACGCAGGAATTTCTTCATTTGGCAGGTAAGGGCAAAGGGGCTGACGTGG。

[0140] 由里氏木霉 Δ ku80 quad缺失衍生菌株产生Archy2菌株

[0141] 首先,从ku80敲除菌株中缺失pyr2基因。图3中示意性描述的pyr2缺失盒含有里氏木霉cbh1启动子、潮霉素抗性基因和侧翼为5'和3'pyr2序列的部分amdS可选择标记。这使得可以在pyr2敲除转化株中筛选对于潮霉素和氟乳清酸的抗性。部分amdS基因含有该基因的3'部分,但缺乏编码区的启动子和氨基末端部分,因而是无功能的。pyr2敲除盒的核苷酸序列的长度为9259碱基对:碱基1-1994对应于pyr2 3'同源区;碱基2002-3497对应于里氏木霉cbh1启动子;碱基3563-5449对应于潮霉素抗性可选择标记;碱基5450-7440对应于构巢曲霉amdS 3'部分标记;碱基7441-9259对应于pyr25'同源区。pyr2敲除盒的核苷酸序列作为SEQ ID NO:2提供:

[0142] ATCACGCCCTCGCATAAAAGACCCTCAAGAGTCCATGTGCCCTATCTGCCTGATCTTCCTAACCTTAT
TTAACATTGGCCCTATCACAACTAGTTCTTCTTCAGCCTGCTTTGTCAACACTGTACGGTTCAACTCAACGTAA
TCAGCAGGTAGCAGGACGAGGATAGGGAGAGAAACGAAGAGAAGAAGAGGAGAGAGGAAGAAAAAAGAAAAGA
AAGAAAAAGGGAAAAGGAAAGAAGGAGGAAAAGAGAAGAAAGTCAGATGAAGAAGCAAGAAGACGCCATGGTAGCCA
CCGCTTGTACAGGGCTCCTTAGCAACGAACAACTCTAGCTTGGGACTTGTGATGTGTCGTTTCTTCCCTACCCATC
AGCACCAACGATGAGTTCGATATAGACGAGGACCTCATGGAAGTAGAGACCATTGGGTTTCGACAGGATCTCTCAGTT
TCACTTCTATGAGGTCTGTGCTCGGATGACTTTTTGAGGAGCTTCCCCTTCTGCTTCAACCCCAAACCTCTCTTTCC
TGAAACCGCAGCACGTTGGCACGGCCGTGTTGCTGGAGCAGTTTGCTTTCGAGCACTCTCAGCGTGGTTTCAGCAGC
CCACTGGTGAGTGGCCTCCTTTGACGTCCACACCTTGCTCCTGTGCGATGCGTATCTGGTGGGAACGACTGCTCCAA
GGAGGATTGCTAACGAGGTTGTAGGCCGAATATCGCATCAGATTCTCCGGTAACCTTAGCTACGGCCTCTTCAACAT
CTGTGACATGACGGAGCGCAAGTACTGGTGGTTGGCGACCAAGATGCGCGGCTGGAACATCGACGGCTGCCCCGAAG
ACGTCAGGAGACTCATGTTTGTTCACATCATCGCCACCCTGGGATGCAGCCCCGTCGTGACGGATGAAGACATGGAC
TACCCCAAGAAGTGGGCGCAATTCTCCACGGTAGAGACAGATATCCGAGTGAACCTGTGGGCCACCGGCCTCATGG
GCGCACCATCTGCCTCCACTCGGTGGCCGTCTGCCCTCGTCTCCAGGGCTTGGTCTCGGTACTGCGACTCTGAAGT

CGTATGTGCAGCGCATGAACAGCCTCGGCGCCGCGGACCGTGTGTGCTCTCGTTTGCCGCAAGCCCGAGACGAGATTT
TTTGAAAGATGCGGCTTCAGGAACAGCGGCCGAGTAGTATCAAGACTCTGGTCGGCGAATACTACAACATGGTGTG
TGCTTCCACATCGACTTGCCAGACTCTATACGATTTTCAAACCTCGCTATACGTCATATTGACTTGTTCCTTAGG
TCTTCGATTTGCCCGGGCCAAAGACTTTATCGACTGGAATAGCATTGCCGACGCTGCCAAGAAGATGTGAACCATT
TGACTGATACGATGTGTGCTACGCATGTGCACCTTCTTTGTTTGTTCCTTTGGCGGCTCTTTGTATACCTTGGGACA
CGGCAGACGCATGTCTATGTGAAGAAAACGTTACGCGGCTGTTTGCATCAGGAATATGATCATTAAACATGGAGCG
TAATGGTATTAATGATCAACTAGAAAAATGGTATGGAAGGGCGAGAGGGCGATCAACAAAGCAGCCCGGGCATAGT
CTGGAAGCAGCAGGAATTGGAAGGGAAAAGGAGCTGCACAATGAAGGGATATCGTGAGCGGAGTGGCTCACGAGAG
TATCAACAGACTGGCGAAAGCAAGCAATTGCCAACCGCGGCTATTAGGCCATAAGATGGCCTGTTGTGAGTCCCAGT
TGCACGTATCCCCATATGACTGCTCTGTGCTGACTTGAAAAAAATAGGGAGGATAAAGGAGAAAAGAAAGTGAGAC
AACCCGTGAGGGACTTGGGGTAGTAGGAGAACACATGGGCAACCGGGCAATACACGCGATGTGAGACGAGTTCAACG
GCGAATGGAAAATCTTGAAAAACAAAATAAAATAACTGCCCTCCATACGGGTATCAAATTCAGCAGTTGTACGGAG
GCTAGCTAGAGTTGTGAAGTCGGTAATCCCGCTGTATAGTAATACGAGTCGCATCTAAATACTCCGAAGCTGCTGCG
AACCCGGAGAATCGAGATGTGCTGGAAAGCTTCTAGCGAGCGGCTAAATTAGCATGAAAGGCTATGAGAAAATCTGG
AGACGGCTTGTGAATCATGGCGTTCATTCTTCGACAAGCAAAGCGTTCGTCGAGTAGCAGGCACTCATTCCC
AAAAACTCGGAGATTCCTAAGTAGCGATGGAACCGGAATAATATAATAGGCAATACATTGAGTTGCCTCGACGGTT
GCAATGCAGGGTACTGAGCTTGACATAACTGTTCCGTACCCACCTCTTCTCAACCTTTGGCGTTTCCCTGATTC
AGCGTACCCGTACAAGTCGTAATCACTATTAACCCAGACTGACCGGACGTGTTTTGCCCTTCATTTGGAGAAATAAT
GTCATTGCGATGTGTAATTTGCTGCTTGACCGACTGGGGCTGTTGCAAGCCGAATGTAGGATTGTTATCCGAAC
CTGCTCGTAGAGGCATGTTGTGAATCTGTGTCGGGCAGGACACGCCTCGAAGGTTACGGCAAGGGAAACCACCGAT
AGCAGTGTCTAGTAGCAACCTGTAAGCCGCAATGCAGCATCACTGAAAATACAAACCAATGGCTAAAAGTACATA
AGTTAATGCCTAAAGAAGTCATATAACCAGCGGCTAATAATTGTACAATCAAGTGGCTAAACGTACCGTAATTTGCCA
ACGGCTTGTGGGGTTGCAGAAGCAACGGCAAAGCCCCACTTCCCCACGTTTGTTCCTCACTCAGTCCAATCTCAGC
TGGTGATCCCCAATTGGGTCGCTTGTTCGTTCCGGTGAAGTGAAAGAAGACAGAGGTAAGAATGTCTGACTCGGAG
CGTTTTGCATACAACCAAGGGCAGTGATGGAAGACAGTGAAATGTTGACATTCAGGAGTATTTAGCCAGGGATGCT
TGAGTGTATCGTGAAGGAGGTTGTCTGCCGATACGACGAATACTGTATAGTCACTTCTGATGAAGTGGTCCATAT
TGAAATGTAAAGTCGGCACTGAACAGGCAAAAGATTGAGTTGAAACTGCCTAAGATCTCGGGCCCTCGGGCCTTCGG
CCTTTGGGTGTACATGTTTGTGCTCCGGGCAAAATGCAAAGTGTGGTAGGATCGAACACACTGCTGCCTTTACCAAGC
AGCTGAGGGTATGTGATAGGCAAATGTTACAGGGCCACTGCATGGTTTCGAATAGAAAGAGAAGCTTAGCCAAGAAC
AATAGCCGATAAAGATAGCCTCATTAACCGAATGAGCTAGTAGGCAAAGTCAGCGAATGTGTATATATAAAGGTTCC
GAGGTCCGTGCCTCCCTCATGCTCTCCCCATCTACTCATCAACTCAGATCCTCCAGGAGACTTGTACACCATCTTTT
GAGGCACAGAAACCAATAGTCAACCGGACTGCGCATCATGTATCGGAAGTTGGCCGTATCTCGGCCTTCTTGG
CCACACCTCGTGCTAGACTAGGCGGCCAGGAAGCCCGAAGGTAAGTGGATTCTTCGCCGTGGCTGGAGCAACCGG
TGGATTCCAGCGTCTCCGACTTGACTGAGCAATTCAGCGTCACGGATTCACGATAGACAGCTCAGACCGCTCCACG
GCTGGCGCATTATTGGTTAACCCGAAAACCTCAGTCTCCTTGGCCCCGTCCCGAAGGGACCCGACTTACCAGGCTGG
GAAAGCCAGGGATAGAATACACTGTACGGCTTCGTACGGGAGGTTCCGGCGTAGGGTTGTTCCCAAGTTTTACACAC
CCCCAAGACAGCTAGCGCACGAAAGACGCGGAGGGTTGGTGAAAAAAGGGCGAAAATTAAGCGGGAGACGTATTT
AGGTGCTAGGGCCGGTTTCTCCCCATTTTCTTCGGTTCCTTTCTCTCCTGGAAGACTTCTCTCTCTCTCTCTCT
TCTCTTCTCCATCCTCAGTCCATCTTCTTCCCATCATCCATCTCCTCACCTCCATCTCAACTCCATCACATCAC

AATCGATATGAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAAAGTTCGACAGCGTCTCCG
ACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTCGATGTAGGAGGGCGTGGATATGTCCTGCGG
GTAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGCACCTTGCATCGGCCGCGCTCCCGATTCC
GGAAGTGCTTGACATTGGGGAATTCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTACGTTGC
AAGACCTGCCTGAAACCGAACTGCCCGCTGTTCTGCAGCCGGTTCGCGGAGGCCATGGATGCGATCGCTGCGGCCGAT
CTTAGCCAGACGAGCGGGTTTCGGCCATTTCGGACCCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTTCATATG
CGCGATTGCTGATCCCCATGTGTATCACTGGCAAACCTGTGATGGACGACACCGTCAGTGCCTCCGTGCGCAGGCTC
TCGATGAGCTGATGCTTTGGGCCGAGGACTGCCCGAAGTCCGGCACCTCGTGCACGCGGATTTTCGGCTCCAACAAT
GTCCTGACGGACAATGGCCGATAACAGCGGTCAATGACTGGAGCGAGGCGATGTTCCGGGATTCCCAATACGAGGT
CGCCAACATCTTCTTCTGGAGGCCGTGGTTGGCTTGATGGAGCAGCAGACGCGCTACTTCGAGCGGAGGCATCCGG
AGCTTGACAGGATCGCCGCGGCTCCGGGCGTATATGCTCCGCATTTGGTCTTGACCAACTCTATCAGAGCTTGGTTGAC
GGCAATTTTCGATGATGCAGCTTGGGCGCAGGGTCGATGCGACGCAATCGTCCGATCCGGAGCCGGGACTGTCCGGCG
TACACAAATCGCCCGAGAAGCGCGGCCGTTCGGACCGATGGCTGTGTAGAAGTACTCGCCGATAGTGAAACCGAC
GCCCCAGCACTCGTCCGAGGGCAAAGGAATAGAGTAGATGCCGACCGGGATCCACTAACGTTACTGAAATCATCAA
ACAGCTTGACGAATCTGGATATAAGATCGTTGGTGTGATGTCAGCTCCGGAGTTGAGACAAATGGTGTTCAGGATC
TCGATAAGATACGTTCAATTTGTCCAAGCAGCAAAGAGTGCCTTCTAGTGATTTAATAGCTCCATGTCAACAAGAATA
AAACGCGTTTTCCGGTTTTACCTCTTCCAGATACAGCTCATCTGCAATGCATTAATGCATTGGACCTCGCAACCCTAGT
ACGCCCTTCAGGCTCCGGCGAAGCAGAAGAATAGCTTAGCAGAGTCTATTTTCATTTTCGGGAGACTAGCATTCTGT
AAACGGGCAGCAATCGCCCAGCAGTTAGTAGGGTCCCCTCTACCTCTCAGGGAGATGTAACAACGCCACCTTATGGG
ACTATCAAGCTGACGCTGGCTTCTGTGCAGACAAACTGCGCCACGAGTCTTCCCTGACGCCGCTCTCGCGCAGGC
AAGGGAACCTCGATGAATACTACGCAAAGCACAAGAGACCCGTTGGTCCACTCCATGGCCTCCCCATCTCTCTCAAAG
ACCAGCTTCGAGTCAAGGTACACCGTTGCCCTAAGTCGTTAGATGTCCCTTTTTGTCAGCTAACATATGCCACCAG
GGCTACGAAACATCAATGGGCTACATCTCATGGCTAAACAAGTACGACGAAGGGGACTCGGTTCTGACAACCATGCT
CCGCAAAGCCGGTGCCGTCTTCTACGTCAAGACCTCTGTCCCGCAGACCCTGATGGTCTGCGAGACAGTCAACAACA
TCATCGGGCGCACCGTCAACCCACGCAACAAGAAGTGGTTCGTGCGGCGGCAGTCTGGTGGTGAGGGTGCATCGTT
GGGATTCGTGGTGGCGTCATCGGTGTAGGAACGGATATCGGTGGCTCGATTTCGAGTGCCGGCCGCGTTCAACTTCT
GTACGGTCTAAGGCCGAGTCATGGGCGGCTGCCGTATGCAAAGATGGCGAACAGCATGGAGGGTTCAGGAGACGGTGC
ACAGCGTTGTGGGCCGATTACGCACTCTGTGAGGGTGAGTCCCTTCGCCTCTTCCCTTCTTTTCTGCTCTATACCA
GGCCTCCACTGTCTCTCTTTCTGCTTTTTATACTATATAACGAGACCGGCAGTCACTGATGAAGTATGTTAGACCTC
CGCCTCTTACCAAATCCGTCTCGGTGAGGACCATGGAAATACGACTCCAAGGTATCCCCATGCCCTGGCGCCA
GTCCGAGTCGGACATTATTGCCTCCAAGATCAAGAACGGCGGGCTCAATATCGGCTACTACAACTTCGACGGCAATG
TCCTTCCACACCCTCCTATCTGCGCGGCGTGGAAACCACCGTCGCCGACTCGCCAAAGCCGGTACACCCGTGACC
CCGTGGACGCCATAACAAGCAGATTTTCGGCCACGATCTCATCTCCATATCTACGCGGCTGACGGCAGCGCCGACGT
AATGCGCGATATCAGTGCATCCGGCGAGCCGGCGATTCCAAATATCAAAGACCTACTGAACCCGAACATCAAAGCTG
TTAACATGAACGAGCTCTGGGACACGCATCTCCAGAAGTGGAAATTACCAGATGGAGTACCTTGAGAAATGGCGGGAG
GCTGAAGAAAAGGCCGGGAAGGAAGTGGACGCCATCATCGCGCCGATTACGCCTACCGCTGCGGTACGGCATGACCA
GTTCCGGTACTATGGGTATGCCTCTGTGATCAACCTGCTGGATTTACGAGCGTGGTTGTTCCGGTTACCTTTGCGG
ATAAGAACATCGATAAGAAGAATGAGAGTTTCAAGGCGTTAGTGAGCTTGTGCCCTCGTGCAGGAAGAGTATGAT
CCGGAGGCGTACCATGGGGCACCGGTTGCAGTGCAGGTTATCGGACGGAGACTCAGTGAAGAGAGGACGTTGGCGAT

TGCAGAGGAAGTGGGGAAGTTGCTGGGAAATGTGGTGACTCCATAGCTAATAAGTGTGATAGCAATTTGCACAAG
AAATCAATACCAGCAACTGTAAATAAGCGCTGAAGTGACCATGCCATGCTACGAAAGAGCAGAAAAAACCTGCCGT
AGAACCGAAGAGATATGACACGCTTCCATCTCTCAAAGGAAGAATCCCTTCAGGGTTGCGTTTCCAGTAGTATT
ACCGCTGATGAAATGACTGGACTCCCTCCTGCTTATACGAAAAATTGCCTGACTCTGCAAAGGTTGTTTGTCT
TTGGAAGATGATGTGCCCCCATCGCTCTTATCTCATACCCCGCCATCTTTCTAGATTCTCATCTTCAACAAGAGG
GGCAATCCATGATCTGCGATCCAGATGTGCTTCTGGCCTCATACTCTGCCTTCAGGTTGATGTTCACTTAATTGGTG
ACGAATTCAGCTGATTTGCTGCGATATGCTTTGTGTTGGTTCTTTCCAGGCTTGCGCCAGCCATGAGCGCTTTGAGA
GCATGTTGTCACTTATAAACTCGAGTAACGGCCACATATTGTTCACTACTTGAATCACATACCTAATTTTGATAGAA
TTGACATGTTTAAAGAGCTGAGGTAGCTTTAATGCCTCTGAAGTATTGTGACACAGCTTCTCACAGAGTGAGAATGA
AAAGTTGGACTCCCCCTAATGAAGTAAAAGTTTCGTCTCTGAACGGTGAAGAGCATAGATCCGGCATCAACTACCTG
GCTAGACTACGACGTCAATTCTGCGGCCTTTTGACCTTTATATATGTCCATTAATGCAATAGATTCTTTTTTTTTTT
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGCCCAATTCGCAGATCAAAGTGGACGTTATAGCATCATAACTAAGC
TCAGTTGCTGAGGGAAGCCGTCTACTACCTTAGCCCATCCATCCAGCTCCATACCTTGATACTTTAGACGTGAAGCA
ATTCACACTGTACGTCTCGCAGCTCTCCTTCCCGCTTGTCTCCCACTGGGGTCCATGGTGCCTGTATCGTCCCC
TCCACAATTCTATGCCATGGTACCTCCAGCTTATCAATGCCCGCTAACAAGTCGCCTCTTGCCTTGATAGCTTAT
CGATAAACTTTTTTTCCGCCAGAAAGGCTCCGCCACAGACAAGAAAAAAATTCACCGCTAGCCTTTGGCCCCG
GCATTTGGCTAAACCTCGAGCCTCTCTCCCGTCTTGGGGTATCAGGAAGAAAAGAAAAAAATCCATCGCCAAGGGCT
GTTTTGGCATCACCACCCGAAAACAGCACTTCTCGATCAAAAGTTGCCCGCCATGAAGACCAGTGAAGGACATC
CCTCCGGTGCCTACGCACCAGGAGTTTCTGGACATTGTGCTGAGCAGGACCCAGCGCAAACCTGCCACTCAGATCCG
TGCCGGCTTCAAGATTAGCAGAATTCGAGGTACGTGCGATTGCCATCGCAGGATGTCTCATTATCGGGGTCCTGG
AGAACGATCATGATTGCATGGCGATGCTAACACATAGACAGCCTTCTACACTCGAAAGTCAAGTTCACCCAGGAGA
CGTTTTCCGAAAAGTTCGCCTCCATCCTCGACAGCTTCCCTCGCCTCCAGGACATCCACCCCTTCCACAAGGACCTT
CTCAACACCCTCTACGATGCCGACCACTTCAAGATTGCCCTTGGCCAGATGTCCACTGCCAAGCACCTGGTTCGAGAC
CATCTCGCGGACTACGTCCGTCTTGAATAACGCCAGTCGCTCTACCAGTGCAAGCAGCTCAAGCGGGCCGCTC
TCGGTTCGATGGCCACGCTGGTCAAGCGCCTCAAGGACCCCTGCTGTACCTGGACCAGGTCCGCCAGCATCTCGGC
CGTCTTCCCTCCATCGACCCCAACACCAGGACCCCTGCTCATCTGCGGTTACCCCAATGTTGGCAAGTCCAGCTTCT
GCGAAGTATCACCCGCGCCGATGTGGACGTCCAGCCCTATGCTTTCACCACCAAGAGTCTGTTTGTGCGCCACTTTG
ACTACAAGTACCTGCGATTCCAGGCCATTGATACCCCGGTATTCTGGACCACCCTCTTGAGGAGATGAACACTATC
GAAATGCAGAGGTATGTGGCGCGGCTA

[0143] 由Archy2里氏木霉菌株产生Archy3菌株。

[0144] 用潮霉素缺失盒转化Archy 2菌株,以在相同pyr2基因座整合且用pyr2基因的编
码区替换潮霉素抗性基因。潮霉素缺失盒在图4中示出。这种将pyr2基因再引入回到pyr2基
因座的做法使pyr2基因被置于里氏木霉cbh1启动子和部分amdS可选择标记之间。通过尿苷
原养型和对潮霉素的敏感性来选择这个菌株。hygR敲除盒的核苷酸序列的长度为9088碱基
对:碱基1-1994对应于pyr2 3'同源区;碱基1995-3497对应于里氏木霉cbh1启动子;碱基
3564-5137对应于pyr2可选择标记;碱基5280-7270对应于构巢曲霉amdS3'部分标记;碱基
7271-9088对应于pyr2 5'同源区。hygR敲除盒的核苷酸序列作为SEQ ID NO:3提供:

[0145] ATCACGCCCTCGATAAAAGACCCTCAAGAGTCCATGTGCCCTATCTGCCTGATCTTCCCTAACCTTAT
TTAACATTGGCCCTATCACAACTAGTTCTTCTCAGCCTGCTTTGTCAACACTGTCACGGTTCACCTCAACGTAA

TCAGCAGGTAGCAGGACGAGGATAGGGAGAGAAACGAAGAGAAGAAGAGGAGAGAGGAAGAAAAAAAAAAGAAAAGA
AAGAAAAAGGGAAAAGGAAAAGAAGGAGGAAAAGAGAAAAGTCAGATGAAGAAGCAAGAAGACGCCATGGTAGCCA
CCGCTTGTACAGGGCTCCTTAGCAACGAACAACCTCTAGCTTGGGGACTTGTGATGTGTCGTTTCCTTCTACCCATC
AGCACCAACGATGAGTTCGATATAGACGAGGACCTCATGGAAGTAGAGACCATTGGGTTTCGACAGGATCTCTCAGTT
TCACTTCTATGAGGTCTGTGCTCGGATGACTTTTTGAGGAGCTTCCCCTTCTGCTTCAACCCCAAACCTCTCTTTCC
TGAAACCGCAGCACGTTGGCACGGCCGTGTGCTGGAGCAGTTTGCTTTCGAGCACTCTCAGCGTGGTTTCAGCAGC
CCACTGGTGAGTGGCCTCCTTTGACGTCCACACCTTGCTCCTGTGCGATGCGTATCTGGTGGGAACGACTGCTCCAA
GGAGGATTGCTAACGAGGTTGTAGGCCGAATATCGCATCAGATTCTCCGGTAACCTTAGCTACGGCCTCTTCAACAT
CTGTGACATGACGGAGCGCAAGTACTGGTGGTTGGCGACCAAGATGCGCGGCTGGAACATCGACGGCTGCCCCGAAG
ACGTCAGGAGACTCATGTTTGTTCACATCATCGCCACCCTGGGATGCAGCCCCGTCGTGACGGATGAAGACATGGAC
TACCCCAAGAACTGGGCGGCAATTCTCCACGGTAGAGACAGATATCCGAGTGAACCTGTGGGCCACCGGCCTCATGG
GCGCACCATCTGCCTCCACTCGGTGGCCGTCTGCCCTCGTCTCCAGGGCTTGGGTCTCGGTACTGCGACTCTGAAGT
CGTATGTGCAGCGCATGAACAGCCTCGGCGCGGACCGTGTGCTCTCGTTTGCCGCAAGCCCGAGACGAGATTT
TTTGAAAGATGCGGCTTCAGGAACAGCGGCCGAGTAGTATCAAGACTCTGGTTCGGCGAATACTACAACATGGTGTG
TGCTTCCACATCGACTTGGCCAGACTCTATACGATTTTCAAACCTCGCTATACGTCATATTGACTTGTTCCTTTAGG
TCTTCGATTTGCCCGGGCCAAAGACTTTATCGACTGGAATAGCATTGCCGACGCTGCCAAGAAGATGTGAACCATT
TGACTGATACGATGTGTGCTACGCATGTGCACCTTCTTTGTTTGTTCCTTGGCGGCTCTTTGTATACCTTGGGACA
CGGCAGACGCATGTCTATGTGAAGAAAACGTTACGGCGCTGTTGCATCAGGAATATGATCATTAAACATGGAGCG
TAATGGTATTAATGATCAACTAGAAAAATGGTATGGAAGGGCGAGAGGGCGATCAACAAAGCAGCCCGGGGCATAGT
CTGGAAGCAGCAGGAATTGGAAGGGAAAAGGAACTGCACAATGAAGGGATATCGTGAGCGGAGTGGCTCACGAGAG
TATCAACAGACTGGCGAAAAGCAAGCAATTGCCAACCGCGGCTATTAGGCCATAAGATGGCCTGTTGTGAGTCCCAGT
TGCACGTATCCCCATATGACTGCTCTGTGCTGACTTGAAAAAAAAATAGGGAGGATAAAGGAGAAAAGAAAGTGAGAC
AACCCGTGAGGGACTTGGGGTAGTAGGAGAACACATGGGCAACCGGGCAATACACGCGATGTGAGACGAGTTCAACG
GCGAATGGAAAATCTTGAAAAACAAAATAAAATAACTGCCCTCCATACGGGTATCAAATTCAGCAGTTGTACGGAG
GCTAGATAGAGTTGTGAAGTCGGTAATCCCGCTGTATAGTAATACGAGTGCATCTAAATACTCCGAAGCTGCTGCG
AACCCGGAGAATCGAGATGTGCTGGAAAGCTTCTAGCGAGCGGCTAAATTAGCATGAAAGGCTATGAGAAATTCTGG
AGACGGCTTGTGTAATCATGGCGTTCCATTCTTCGACAAGCAAAGCGTTCCGTGCGAGTAGCAGGCACTCATTCCCG
AAAAACTCGGAGATTCTAAGTAGCGATGGAACCGGAATAATATAATAGGCAATACATTGAGTTGCCTCGACGGTT
GCAATGCAGGGTACTGAGCTTGGACATAACTGTTCCGTACCCACCTCTTCTCAACCTTTGGCGTTTCCCTGATTC
AGCGTACCCGTACAAGTCGTAATCACTATTAACCCAGACTGACCGGACGTGTTTTGCCCTTCATTTGGAGAAATAAT
GTCATTGCGATGTGTAATTTGCCTGCTTGACCGACTGGGGCTGTTTGAAGCCCGAATGTAGGATTGTTATCCGAAC
CTGCTCGTAGAGGCATGTTGTGAATCTGTGTCGGGCAGGACACGCCTCGAAGGTTACGGCAAGGGAAACCACCGAT
AGCAGTGTCTAGTAGCAACCTGTAAAGCCGCAATGCAGCATCACTGGAAAATACAAACCAATGGCTAAAAGTACATA
AGTTAATGCCTAAAGAAGTCATATAACCAGCGCTAATAAATTGTACAATCAAGTGGCTAAACGTACCGTAATTTGCCA
ACGGCTTGTGGGGTTGCAGAAGCAACGGCAAAGCCCCACTTCCCACGTTTGTTCCTCACTCAGTCCAATCTCAGC
TGGTGATCCCCAATTGGGTCGCTTGTGTTTCCGGTGAAGTAAAGAAGACAGAGGTAAGAATGTCTGACTCGGAG
CGTTTTGCATAACAACCAAGGGCAGTGATGGAAGACAGTAAAATGTTGACATTCAAGGAGTATTTAGCCAGGGATGCT
TGAGTGTATCGTGAAGGAGGTTTGTCTGCCGATACGACGAATACTGTATAGTCACTTCTGATGAAGTGGTCCATAT
TGAAATGTAAAGTCGGCACTGAACAGGCAAAAAGATTGAGTTGAAACTGCCTAAGATCTCGGGCCCTCGGGCCTTCGG

CCTTTGGGTGTACATGTTTGTGCTCCGGGCAAATGCAAAGTGTGGTAGGATCGAACACACTGCTGCCTTTACCAAGC
 AGCTGAGGGTATGTGATAGGCAAATGTTTCAGGGGCCACTGCATGGTTTCGAATAGAAAGAGAAGCTTAGCCAAGAAC
 AATAGCCGATAAAAGATAGCCTCATTAAACGGAATGAGCTAGTAGGCAAAGTCAGCGAATGTGTATATATAAAGGTTC
 GAGGTCCGTGCCTCCCTCATGCTCTCCCATCTACTCATCAACTCAGATCCTCCAGGAGACTTGTACACCATCTTTT
 GAGGCACAGAAACCCAAATAGTCAACCGCGACTGCGCATCATGTATCGGAAGTTGGCCGTCTCTCGGCCCTTCTGG
 CCACACCTCGTGCTAGACTAGGCGCGTCAATATGTGGCCGTTACTCGAGTTTATAAGTGACAACATGCTCTCAAAGC
 GCTCATGGCTGGCACAAGCCTGGAAAGAACCAACACAAAGCATACTGCAGCAAATCAGCTGAATTTCGTACCAATTA
 AGTGAACATCAACCTGAAGGCAGAGTATGAGGCCAGAAGCACATCTGGATCGCAGATCATGGATTGCCCTCTTGTT
 GAAGATGAGAATCTAGAAAGATGGCGGGTATGAGATAAGAGCGATGGGGGGCACATCATCTTCCAAGACAAACAA
 CCTTTGCAGAGTCAGGCAATTTTTCGTATAAGAGCAGGAGGAGGGAGTCCAGTCATTTTCATCAGCGTAAAATCACT
 CTAGACAATCTTCAAGATGAGTTCTGCCTTGGGTGACTTATAGCCATCATCATACCTAGACAGAAGCTTGTGGGATA
 CTAAGACCAACGTACAAGCTCGCACTGTACGCTTTGACTTCCATGTGAAAACTCGATACGGCGCGCCTCTAAATTTT
 ATAGCTCAACCACTCCAATCCAACCTCTGCATCCCTCTCACTCGTCCTGATCTACTGTTCAAATCAGAGAATAAGGA
 CACTATCCAAATCCAACAGAATGGCTACCACCTCCCAGCTGCCTGCCTACAAGCAGGACTTCCTCAAATCCGCCATC
 GACGGCGCGTCTCAAGTTTGGCAGCTTCGAGCTCAAGTCCAAGCGGATATCCCCCTACTTCTTCAACGCGGGCGA
 ATTCCACACGGCGCGCCTCGCCGGCGCCATCGCTCCGCTTTGCAAAGACCATCATCGAGGCCAGGAGAAGGCCG
 GCCTAGAGTTTGACATCGTCTTCGGCCCGCCTACAAGGCATCCCGCTGTGCTCCGCCATCACCATCAAGCTCGGC
 GAGCTGGCGCCCCAGAACCTGGACCGCGTCTCCTACTCGTTTGACCGCAAGGAGGCCAAGGACCACGGCGAGGGCGG
 CAACATCGTCGGCGCTTCGCTCAAGGGCAAGAGGGTCTGATTGTGACGACGTCATCACCGCCGGCACCGCCAAGA
 GGGACGCCATTGAGAAGATCACCAAGGAGGGCGGCATCGTCGCCGCATCGTCGTGGCCCTGGACCGCATGGAGAAG
 CTCCCCGCTGCGGATGGCGACGACTCCAAGCCTGGACCGAGTGCCATTGGCGAGCTGAGGAAGGAGTACGGCATCCC
 CATCTTTGCCATCCTCACTCTGGATGACATTATCGATGGCATGAAGGGCTTTGCTACCCCTGAGGATATCAAGAACA
 CGGAGGATTACCGTGCCAAGTACAAGGCGACTGACTGATTGAGGCGTTCAATGTCAGAAGGGAGAGAAAGACTGAAA
 AGGTGAAAGAAGAGGCAAATTGTTGTTATTATTATTCTATCTCGAATCTTCTAGATCTTGTGTAATAAACA
 AGCGTAACTAGCTAGCCTCCGTACAACCTGCTTGAATTTGATACCCGTATGGAGGGCAGTTATTTTATTTTGTTTTC
 AAGATTTTCCATTCGCCGTTGAACTCGTCTCACATCGCGTGTATTGCCCGTTGCCCATGTGTACGCGTTTCGGGTT
 TACCTCTTCCAGATACAGCTCATCTGCAATGCATTAATGCATTGGACCTCGCAACCCTAGTACGCCCTTCAGGCTCC
 GGCGAAGCAGAAGAATAGCTTAGCAGAGTCTATTTTCATTTTCGGGAGACTAGCATTCTGTAAACGGGCAGCAATCG
 CCCAGCAGTTAGTAGGGTCCCCTCTACCTCTCAGGGAGATGTAACAACGCCACCTTATGGGACTATCAAGCTGACGC
 TGGCTTCTGTGCAGACAAACTGCGCCACGAGTTCTCCCTGACGCCGCTCTCGCGCAGGCAAGGGAACCTCGATGAA
 TACTACGCAAAGCACAAGAGACCCGTTGGTCCACTCCATGGCCTCCCCATCTCTCTCAAAGACCAGCTTCGAGTCAA
 GGTACACCGTTGCCCTAAGTCGTTAGATGTCCCTTTTTGTCAGCTAACATATGCCACCAGGGCTACGAAACATCAA
 TGGGCTACATCTCATGGCTAAACAAGTACGACGAAGGGGACTCGGTTCTGACAACCATGCTCCGCAAAGCCGGTGCC
 GTCTTCTACGTCAAGACCTCTGTCCCGCAGACCCTGATGGTCTGCGAGACAGTCAACAACATCATCGGGCGCACCGT
 CAACCCACGCAACAAGAAGTGGTCGTGCGGCGCAGTTCTGGTGGTGAGGGTGCATCGTTGGGATTCGTGGTGGCG
 TCATCGGTGTAGGAACGGATATCGGTGGCTCGATTGAGTGCCGGCCGCTTCAACTTCTGTACGGTCTAAGGCCG
 AGTCATGGGCGGCTGCCGTATGCAAAGATGGCGAACAGCATGGAGGGTCAGGAGACGGTGCACAGCGTTGTGGGCC
 GATTACGCACTCTGTTGAGGGTGAGTCCCTTCGCTCTTCTTCTTTTCTGCTCTATACCAGGCCTCCACTGTCCTC
 CTTTCTTGCTTTTTATACTATATACGAGACCGGCAGTCACTGATGAAGTATGTTAGACCTCCGCCTCTTACCAAAT

CCGTCCTCGGT CAGGAGCCATGGAAATACGACTCCAAGGTCATCCCCATGCCCTGGCGCCAGTCCGAGTCGGACATT
ATTGCCTCCAAGATCAAGAACGGCGGGCTCAATATCGGCTACTACAACCTCGACGGCAATGTCTTCCACACCCTCC
TATCCTGCGCGGCGTGGAAACCACCGTCGCCGCACTCGCCAAAGCCGGT CACACCGTGACCCCGTGGACGCCATA
AGCACGATTTCCGCCACGATCTCATCTCCCATATCTACGCGGCTGACGGCAGCGCCGACGTAATGCGCGATATCAGT
GCATCCGGCGAGCCGGCGATTCCAAATATCAAAGACCTACTGAACCCGAACATCAAAGCTGTTAACATGAACGAGCT
CTGGGACACGCATCTCCAGAAGTGGAAATTACCAGATGGAGTACCTTGAGAAATGGCGGGAGGCTGAAGAAAAGGCCG
GGAAGGAACTGGACGCCATCATCGCGCGATTACGCCCTACCGCTGCGGTACGGCATGACCAGTTCGGTACTATGGG
TATGCCTCTGTGATCAACCTGCTGGATTTACGAGCGTGGTTGTTCCGGTTACCTTTGCGGATAAGAACATCGATAA
GAAGAATGAGAGTTTCAAGGCGGTTAGTGAGCTTGATGCCCTCGTGCAGGAAGAGTATGATCCGGAGGCGTACCATG
GGGCACCGGTTGCAGTGCAGGTTATCGGACGGAGACTCAGTGAAGAGAGGACGTTGGCGATTGCAGAGGAAGTGGGG
AAGTTGCTGGGAAATGTGGTGACTCCATAGCTAATAAGTGT CAGATAGCAATTTGCACAAGAAATCAATACCAGCAA
CTGTAAATAAGCGCTGAAGTGACCATGCCATGCTACGAAAGAGCAGAAAAAACCTGCCGTAGAACCGAAGAGATAT
GACACGCTTCCATCTCTCAAAGGAAGAATCCCTCAGGGTTGCGTTCCAGTAGTGATTTTACCCTGATGAAATGA
CTGGACTCCCTCCTCCTGCTCTTATACGAAAAATGCCTGACTCTGCAAAGGTTGTTTGTCTTGGAAGATGATGTGC
CCCCCATCGCTCTTATCTCATACCCCGCCATCTTTCTAGATTCTCATCTTCAACAAGAGGGGCAATCCATGATCTG
CGATCCAGATGTGCTTCTGGCCTCATACTCTGCCTCAGGTTGATGTTCACTTAATTGGTGACGAATT CAGCTGATT
TGCTGCAGTATGCTTTGTGTTGGTTCTTTCCAGGCTTG GCCAGCCATGAGCGCTTTGAGAGCATGTTGTCACTTAT
AAACTCGAGTAACGGCCACATATTGTTCACTACTTGAATCACATACCTAATTTTGATAGAATTGACATGTTTAAAGA
GCTGAGGTAGCTTTAATGCCTCTGAAGTATTGTGACACAGCTTCTCACAGAGTGAGAATGAAAAGTTGGACTCCCC
TAATGAAGTAAAAGTTTCGTCTCTGAACGGTGAAGAGCATAGATCCGGCATCAACTACCTGGCTAGACTACGACGTC
AATTCTGCGGCCTTTTGACCTTTATATATGTCCATTAATGCAATAGATTCTTTTTTTTTTTTTTTTTTTTTTTTT
TTTTTTTTTTTTTTTGCCCAATTTGCGAGATCAAAGTGGACGTTATAGCATCATAACTAAGCTCAGTTGCTGAGGGAA
GCCGTCTACTACCTTAGCCCATCCATCCAGCTCCATACCTTGATACTTTAGACGTGAAGCAATTCACACTGTACGTC
TCGCAGCTCTCCTTCCCCTCTTGCTTCCCCTGTTGGGTTCCATGGTGCGTGTATCGTCCCCTCCACAATTCTATGCC
ATGGTACCTCCAGCTTATCAATGCCCGCTAACAAAGTCGCCCTTTGCCTTGATAGCTTATCGATAAACTTTTTTT
CCGCCAGAAAGGCTCCGCCACAGACAAGAAAAAAATTCACCGCTAGCCTTTGGCCCCGGCATTGTTGGCTAAACCT
CGAGCCTCTCTCCCGTCTTGGGGTATCAGGAAGAAAAAGAAAAAATCCATCGCCAAGGGCTGTTTTGGCATCACCAC
CCGAAAACAGCACTTCTCGATCAAAAAGTTGCCCGCCATGAAGACCACGTGGAAGGACATCCCTCCGGTGCCTACGC
ACCAGGAGTTTCTGGACATTGTGCTGAGCAGGACCCAGCGCAAACCTGCCACTCAGATCCGTGCCGGCTTCAAGATT
AGCAGAATTCGAGGTACGTCGATTGCCATCGCAGGATGTCTCATTATCGGGGTCCTTGAGAACGATCATGATTG
CATGGCGATGCTAACACATAGACAGCCTTCTACACTCGAAAAGTCAAGTTCACCCAGGAGACGTTTTCCGAAAAGTT
CGCCTCCATCCTCGACAGCTTCCCTCGCTCCAGGACATCCACCCCTTCCACAAGGACCTTCTCAACACCCTCTACG
ATGCCGACCACTTCAAGATTGCCCTTGCCAGATGTCCACTGCCAAGCACCTGGTCGAGACCATCTCGCGGACTAC
GTCCGTCTCTTGAAATACGCCAGTCGCTTACCAGTGAAGCAGCTCAAGCGGGCCGCTCTCGGTGCGATGGCCAC
GCTGGTCAAGCGCCTCAAGGACCCCTGCTGTACCTGGACCAGGTCGCCAGCATCTCGGCCGTCTTCCCTCCATCG
ACCCCAACACCAGGACCCTGCTCATCTGCGGTTACCCCAATGTTGGCAAGTCCAGCTTCTGCGAAGTATCACCCGC
GCCGATGTGGACGTCCAGCCCTATGCTTTACCACCAAGAGTCTGTTTGTGCGCCACTTTGACTACAAGTACCTGCG
ATTCCAGGCCATTGATACCCCCGGATTCTGGACCACCCTCTGAGGAGATGAACACTATCGAAATGCAGAGGTATG
TGGCGCGGCT

[0146] 由Archy3里氏木霉菌株A5D产生菌株。

[0147] 使用本领域已知的双交换重组基因替换载体从Archy 3菌株中缺失天然里氏木霉bg11,例如M.Ward,等人(1990年).Gene(基因)86(2):153-62。潮霉素抗性用作bg11缺失的可选择标记。此外,潮霉素抗性标记侧接着loxP位点。bg11缺失盒在图5中示出。分析随后的潮霉素抗性转化株的bg11缺失情况。随后用编码cre重组酶的端粒载体和用于选择转化株的功能性amdS转化已证实缺失bg11的菌株,以促进通过cre重组酶表达进行的潮霉素去除和促进loxP位点的环出。编码cre重组酶的端粒载体在图7中示意性显示。转化株首先在乙酰胺培养基上获得,随后转移至马铃薯右旋糖琼脂并复制铺板到潮霉素培养基上,以筛选潮霉素敏感性。将对潮霉素敏感的菌株再次转移至马铃薯右旋糖培养基并复制铺板到乙酰胺培养基。选择已丧失其在乙酰胺上生长的能力(这指示端粒载体的丧失)的菌株。bg11敲除盒的核苷酸序列作为SEQ ID NO:4提供:

[0148] AATGGTAGGAATGCTGGGATATAGGCTCTGTGCTGGCAAGTTGATGGATCCTCGAATGAGGCCGCCCTG
CAAGGGGAACATCAGAGATCTACCATTGCCTCCTTGGCCCAATCCACTATCATACCTACCTCATGATCATTCTCGG
AAGGTCTACCAGTAAATATTTCTCGTCCCGTGTTCATCATGTCCAGAACCTCATCTCGCCAAATTGACTTTGCCA
CAGTGTCTGGAGCTGGGTAAGCAGCGTGCCAAGGAATTGTTGTCGAGTCTGTGCCAGGCATTGTGCCCGACATTGTG
AACTTCAGCCAGGAGAACTTTTCGATCGCACCTATGCTGAGCACCGTGGGCATGCGATGGCTTCAATAATGCAGTTC
GAGAGGGAGTGTGTCATGCCCTAAAGCTCATTGGCCACCTCCACAGGCTAGCTCTACCTGCATCTGTAGATGGACTT
TCCTTGTCTCCTCCTTCAGAAAACCTCTTGGTCGCTCGCAGGTAAGTGTGTTGCCGTCATTGTTTGACAGTGGAT
AGCCAAGGCAAAACCGTCTGCTTTCAACGGAAGCATTCCGCGGTTGTTTGTGTCATCGTGTATCGATCGACCAGGAGA
ACCCAGACGAGTGTGTTGTCGAGAGAATCATCGACGATGTGAAGAGGCGACGACTAGTATCTAGAAGATTATAATCGA
ACAAATCAGCGTTTGTCTGTGCGGGCGTTTGAGGGCGCAGTTGCCCGCAAAGCAGCGTCGCAATATATAGGCAGCGA
GAGACTGTCAACAGCCAGCCGCCATGTGATCGATCGTAGCCGCTTCCCGATCTTCCCTAAACCCCTTTCTTTGGGG
GGCGGGGGCAGCGGCTTCTAATATTTGCTGGCTGTCTGGATAACGTGAATGGTAGACATGGTAATGTTCCGGTCTGC
GAAACATTTGTACAATTGGAGTTTACGATCGAGATGGAAGGAAACGCTCCACAACTCGGTGACTGGGTTGCCATCA
GGTGTCTAGGGCATAGCGTTCTCTGCAAAATAGAGGAAAGAGAATAGCACTAGTGAAGTGTGAATCACAATGAAGAG
GAGGTTGTTGCCGGAATGCTTTGAGCAGCGTCAAAGTTGAACTTGAAGCTATCACAATTGCAGGGTAAAGTACATG
TTGGTGCCAGTTTGACAGCACAGTGCAGCGGAGCGGAGGATGTCGCGGAAGAGGCGGACGCTAACCCGGGCTTCTT
CTCAGTGAGCAGAACTCCTGCTGCAAGAGTTCTTCTCTCTGCGAGATGACGTGAGGCCCAATTTGCAGCTTCCCTC
GAACAAGGTGATTGAACATCTCTTCCCTCACATTTTCATCATCACTACCTCCTCAATTCACTTCTGCTTCGGCCGT
CTTCATCATTGTTACTGCTCTGATGCCTATCCTGAAGATTGTATTCCTGCAGTATTCACGCCATCCCACCTTCG
GTCCTCACTCACAGTCACAGGTCAACCGCCTTACCCTCCTCGGATGATGTCGGCAATCTGGTGGATCAATGTGCG
GTTGAGGGCCGCCGTAGTGAGGATGGGCATGGGGAACGAGGTCGCCCATTCGCCACAGATAACTTCGTATAGCATA
CATTATACGAAGTTATCCTGGGCTTGTGACTGGTCGCGAGCTGCCACTAAGTGGGGCAGTACCATTTTATCGGACCC
ATCCAGCTATGGGACCCACTCGCAAATTTTACATCATTTTCTTTTTGCTCAGTAACGGCCACCTTTTGTAAAGCGT
AACCAGCAAACAAATTGCAATTGGCCCGTAGCAAGGTAGTCAGGGCTTATCGTGATGGAGGAGAAGGCTATATCAGC
CTCAAAAATATGTTGCCAGCTGGCGGAAGCCCGGAAGGTAAGTGGATTCTTCGCCGTGGCTGGAGCAACCGGTGGAT
TCCAGCGTCTCCGACTTGGACTGAGCAATTCAGCGTCACGGATTCACGATAGACAGCTCAGACCGCTCCACGGCTGG
CGGCATTATTGGTTAACCCGAAACTCAGTCTCCTTGGCCCGTCCCGAAGGGACCCGACTTACCAGGCTGGGAAAG
CCAGGGATAGAATACTGTACGGGCTTCGTACGGGAGGTTCCGGCTAGGGTTGTTCCCAAGTTTACACACCCCCC

AAGACAGCTAGCGCACGAAAGACGCGGAGGGTTTGGTGAAAAAAGGGCGAAAATTAAGCGGGAGACGTATTTAGGTG
 CTAGGGCCGGTTTCTCCCATTTTTCTTCGGTTCCCTTCTCTCCTGGAAGACTTCTCTCTCTCTTCTTCTCT
 TCTTCCATCCTCAGTCCATCTTCCCTTCCCATCATCCATCTCCTCACCTCCATCTCAACTCCATCACATCACAATCG
 ATATGAAAAAGCCTGAACTCACCGCGACGTCTGTGCGAGAAGTTTCTGATCGAAAAGTTCGACAGCGTCTCCGACCTG
 ATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTCGATGTAGGAGGGCGTGGATATGTCCTGCGGGTAAA
 TAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCGTCCCATTCCGGAAG
 TGCTTGACATTGGGGAATTCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTACGTTGCAAGAC
 CTGCCTGAAACCGAACTGCCCCTGTTCTGCAGCCGGTCGCGGAGGCCATGGATGCGATCGCTGCGGCCGATCTTAG
 CCAGACGAGCGGGTTCGGCCCATTCGGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTTCATATGCGCGA
 TTGCTGATCCCATGTGTATCACTGGCAAACTGTGATGGACGACACCGTCAGTGCCTCCGTGCGCAGGCTCTCGAT
 GAGCTGATGCTTTGGGCCGAGGACTGCCCCGAAGTCCGGCACCTCGTGCACGCGGATTTCCGGCTCCAACAATGTCT
 GACGGACAATGGCCGCATAACAGCGGTCAATTGACTGGAGCGAGGCGATGTTCCGGGATTCCCAATACGAGGTGCGCA
 ACATCTTCTTCTGGAGGCCGTGGTTGGCTTGATGGAGCAGCAGACGCGCTACTTCGAGCGGAGGCATCCGGAGCTT
 GCAGGATCGCCGCGGCTCCGGGCGTATATGCTCCGCATTGGTCTTGACCAACTCTATCAGAGCTTGGTTGACGGCAA
 TTTTCGATGATGCAGCTTGGGCGCAGGGTCGATGCGACGCAATCGTCCGATCCGGAGCCGGGACTGTCCGGCGTACAC
 AAATCGCCCGCAGAAGCGCGGCCGTCTGGACCGATGGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCC
 AGCACTCGTCCGAGGGCAAAGGAATAGAGTAGATGCCGACCGGGATCCACTTAACGTTACTGAAATCATCAAACAGC
 TTGACGAATCTGGATATAAGATCGTTGGTGTGATGTCAGCTCCGGAGTTGAGACAAATGGTGTTCAGGATCTCGAT
 AAGATACGTTTCATTTGTCCAAGCAGCAAAGAGTGCCTTCTAGTGATTTAATAGCTCCATGTCAACAAGAATAAAACG
 CGTTTTCGGGTTTACCTCTTCCAGATACAGCTCATCTGCAATGCATTAATGCATTGGACCTCGCAACCCTAGTACGCC
 CTTTCAGGCTCCGGCGAAGCAGAAGAATAGCTTAGCAGAGTCTATTTTCATTTTCGGGAGACGAGATCAAGCAGATCA
 ACGGTCTCAAGAGACCTACGAGACTGAGGAATCCGCTCTTGGCTCCACGCGACTATATATTTGTCTCTAATTGTAC
 TTTGACATGCTCCTCTTCTTTACTCTGATAGCTTGACTATGAAAAATCCGTCACCAGCCCCCTGGGTTTCGAAAAGATA
 ATTGCACTGTTTCTTCTTGAACCTCTCAAGCCTACAGGACACACATTCATCGTAGGTATAAACCTCGAAAATCATTC
 CTACTAAGATGGGTATAACAATAGTAACCATGGTTGCCAGTGAATGCTCCGTAACACCCAATACGCCGGCCGATAAC
 TTCGTATAGCATAACATTATACGAAGTTATACTTGGCGCGCCTAGTGGAACACGAGCACATAAGCTTTTACCTATGGT
 TATCGCTTGATCTACGCGCCGTTGATGGTGGAGGATGGTGGACGTTCCCGAGACCCCTACGAGCTGTGGCATCGTC
 AAATGTGCCACAGACCTTTGTCTTGCTTTCATAACCTCGAGGAGTGTTCAGACTCATCATCCATACACAAGCA
 GTATTAATCAAAGAACTCGGTGCGAATGGCAAAAAATGGTTTGCAAACAGAAAATATGGCCTCTTCTTATCCATC
 ATTAATACTCTACCCGTTTGTACATAACAACATCATTAACCCTTATGCGTCAGGTGTAGCATCCTTGATCTGTTG
 CTCCTCCAACGGCCAGTTCTCAATCGTTACCTCTTCTCCACCAACTCAAACCTCAAGCTTCACAGACTCGTCGGTGT
 TCAAGGCTAGCTCATACTTGGCGGGTATAACAATCCGGTTTCCGTGAGAATCAACACGGGCGAGAGCACTGACAGGG
 ATGGGGATGCTGAGCTTGGAAAGAGTGACCAGGCTTGATGTCGGCAAGTCGGTCGAATCCGACGAGCCACTTGTTCCG
 GTACGGGGCTGGGCCAGCGTTGCTTGTGCGAACAACAGCATGGCCGTATATGGGGACTCCGTCTTGGCCGAGTTCT
 TGATGTTGGCCTCGAAGGTGAAGACGGGAATCTGCTCGCTGTAAGTGTATCCGGGGTGAGGAGCAGAGAGGATCGAT
 GAGGTGTTGAACTTGAGGCTCTTGGGGTGGCTGGCGAGAGTCTCCTTGAAGGTGGTGTAGAAGAGACCACTGCCAAA
 CTCGTAGACGGGTTTGGCGGTGTACCAGATGTAAGTCTGTCCAGGGTTTACTTTCCATCGGGTCGGAGGTTTCATGT
 CATTCTGGGGGAATTGGTGAACATACTCAGCCGGGTAAGTGTGAGTGGTGACCAGTCGGCCGGCAGGAGCACGTTGCCA
 GAGAGAATGTGAAGAGGGCAACGCCTCCCGACTGGCCGGGATATCCGCCCCAGACGAGGGAGTTGACCTTCTTGT

GCTCTTGAGCGAGGATGAGTCTACCTGACCACCGCCATTTGCAGGACGACAAGGGGTTTGCCGACCTCGCTGAGCT
 GCTTGATGAGATCCAGCTGATTACCGGGCCAAGCAATGTCCGTGCGGTGAGCGCCCTCCTGTTCAATGGTGTGTCA
 ATTCCACCGAGGTAGATGATGGCATCCGACTTCTTGCGGCAGCAATGGCCTTGCAAAGCCAGTGGTGTGTTGCC
 GGGCATCTCTGTGCCGAGTTCAAAGTTGACGTGATAGCCGGCCTTCTTAGCAGCTTCCAGAGGGCTGATGAGGTATG
 GGGCAGGGCCATAGTAGTTGCCTTGCATTTGGGTTGTGGCATTGGCCCATGGTCCGATCAGAGCAATGCTGCGCACC
 TTCTTGACAGAGGGAGAGTGCCATCGTTCTTGAGCAGGACGATGCCCTCAACAGCAGCCTCGTACGAGATGTT

[0149] 端粒质粒pTTT-cre的核苷酸序列作为SEQ ID NO:5提供:

[0150] TTGTACAAAGTGGTGATCGCGCCGCGCCAGCTCCGTGCGAAAGCCTGACGCACCGGTAGATTCTTGG
 TGAGCCCGTATCATGACGGCGGGGAGCTACATGGCCCCGGGTGATTTATTTTTTTTGTATCTACTTCTGACCCTT
 TTCAAATATACGGTCAACTCATCTTTCCTGAGATGCGGCCTGCTTGGTATTGCGATGTTGTCAGCTTGCAAATT
 GTGGCTTTTCAAACACAAAACGATTCCCTAGTAGCCATGCATTTTAAGATAACGGAATAGAAGAAAGAGGAAATTA
 AAAAAAAAAAAAAAAAACAACATCCCGTTCATAACCCGTAGAATCGCCGCTCTTCGTGTATCCCAGTACCAGTTTATT
 TTGAATAGCTCGCCCGTGGAGAGCATCCTGAATGCAAGTAACAACCGTAGAGGCTGACACGGCAGGTGTTGCTAGG
 GAGCGTCGTGTTCTACAAGGCCAGACGTCTTCGCGGTTGATATATATGTATGTTTACTGCAGGCTGCTCAGCGACG
 ACAGTCAAGTTCGCCCTCGCTGCTTGTGCAATAATCGCAGTGGGGAAGCCACACCGTGACTCCCATCTTTCAGTAAA
 GCTCTGTTGGTGTGTTATCAGCAATACACGTAATTTAAACTCGTTAGCATGGGGCTGATAGCTTAATTACCGTTTACC
 AGTGCCATGGTTCTGCAGCTTTCCTTGGCCCGTAAAAATCGGCGAAGCCAGCCAATCACCAGCTAGGCACCAGCTAA
 ACCCTATAATTAGTCTCTTATCAACACCATCCGCTCCCCGGGATCAATGAGGAGAATGAGGGGGATGCGGGGCTAA
 AGAAGCCTACATAACCCCTCATGCCAACTCCCAGTTTACACTCGTCGAGCCAACATCCTGACTATAAGCTAACACAGA
 ATGCCTCAATCCTGGGAAGAACTGGCCGCTGATAAGCGCGCCCGCTCGAAAAACCATCCCTGATGAATGGAAAGT
 CCAGACGCTGCCTGCGGAAGACAGCGTTATTGATTTCCCAAAGAAATCGGGGATCCTTTCAGAGGCCGAACTGAAGA
 TCACAGAGGCCTCCGCTGCAGATCTTGTGTCCAAGCTGGCGGCCGGAGAGTTGACCTCGGTGGAAGTTACGCTAGCA
 TTCTGTAAACGGGCAGCAATCGCCCAGCAGTTAGTAGGGTCCCCTCTACCTCTCAGGGAGATGTAACAACGCCACCT
 TATGGGACTATCAAGCTGACGCTGGCTTCTGTGCAGACAACTGCGCCACGAGTTCTTCCCTGACGCCGCTCTCGC
 GCAGGCAAGGAACTCGATGAATACTACGCAAAGCACAAGAGACCCGTTGGTCCACTCCATGGCCTCCCCATCTCTC
 TCAAAGACCAGCTTCGAGTCAAGGTACACCGTTGCCCTAAGTCGTTAGATGTCCCTTTTTGTCAGCTAACATATGC
 CACCAGGGCTACGAAACATCAATGGGCTACATCTCATGGCTAAACAAGTACGACGAAGGGGACTCGGTTCTGACAAC
 CATGCTCCGCAAAGCCGGTGGCGTCTTCTACGTCAAGACCTCTGTCCCGCAGACCCTGATGGTCTGCGAGACAGTCA
 ACAACATCATCGGGCGCACCGTCAACCCACGCAACAAGAACTGGTCGTGCGGCGGCAGTTCTGGTGGTGAGGGTGCG
 ATCGTTGGGATTCGTGGTGGCGTCATCGGTGTAGGAACGGATATCGGTGGCTCGATTCCAGTGCCGGCCGCGTTCAA
 CTTCTGTACGGTCTAAGGCCGAGTCATGGGCGGCTGCCGTATGCAAAGATGGCGAACAGCATGGAGGGTCAGGAGA
 CGGTGCACAGCGTTGTGCGGCCGATTACGCACTCTGTTGAGGGTGAGTCCTTCGCCTCTTCTTCTTTTCTGCTCT
 ATACCAGGCCTCCACTGTCTCTCTTTCTGCTTTTTATACTATATACGAGACCGGCAGTCACTGATGAAGTATGTTA
 GACCTCCGCCTCTTACCAAATCCGTCTCGGTGAGGAGCCATGGAAATACGACTCCAAGGTATCCCCATGCCCTG
 GCGCCAGTCCGAGTCGGACATTATTGCTCCAAGATCAAGAACGGCGGGCTCAATATCGGCTACTACAACTTCGACG
 GCAATGTCTTCCACACCCCTCTATCTGCGCGGCGTGAAACCACCGTCGCCGACTCGCCAAAGCCGGTACACCC
 GTGACCCCGTGGACGCCATACAAGCAGATTTCGGCCACGATCTCATCTCCATATCTACGCGGCTGACGGCAGCGC
 CGACGTAATGCGCGATATCAGTGCATCCGGCGAGCCGGCGATTCCAAATATCAAAGACCTACTGAACCCGAACATCA
 AAGCTGTTAACATGAACGAGCTCTGGGACACGCATCTCCAGAAGTGAATTACCAGATGGAGTACCTTGAGAAATGG

CGGGAGGCTGAAGAAAAGGCCGGGAAGGAACTGGACGCCATCATCGCGCCGATTACGCCTACCGCTGCGGTACGGCA
TGACCAGTTCGGTACTATGGGTATGCCTCTGTGATCAACCTGCTGGATTCACGAGCGTGGTTGTTCCGGTTACCT
TTGCGGATAAGAACATCGATAAGAAGAATGAGAGTTTCAAGGCGGTTAGTGAGCTTGATGCCCTCGTGAGGAAGAG
TATGATCCGGAGGCGTACCATGGGGCACCGTTGCAGTGCAGGTTATCGGACGGAGACTCAGTGAAGAGAGGACGTT
GGCGATTGCAGAGGAAGTGGGGAAGTTGCTGGGAAATGTGGTGACTCCATAGCTAATAAGTGTGAGATAGCAATTTG
CACAAGAAATCAATACCAGCAACTGTAAATAAGCGCTGAAGTGACCATGCCATGCTACGAAAGAGCAGAAAAAACC
TGCCGTAGAACCGAAGAGATATGACACGCTTCCATCTCTCAAAGGAAGAATCCCTTCAGGGTTGCGTTTCCAGTCTA
GACACGTATAACGGCACAAGTGTCTCTACCAAATGGGTTATATCTCAAATGTGATCTAAGGATGGAAAGCCCAGAA
TATCGATCGCGCAGATCCATATATAGGGCCCGGTTATAATTACCTCAGGAAATAGCTTTAAGTAGCTTATTAAG
TATTAATAATTATATATATTTTTAATAACTATATTTCTTTAATAAATAGGTATTTAAGCTTTATATATAAATATA
ATAATAAAATAATATATTATATAGCTTTTTATTAATAAAATAAATAGCTAAAAATATAAAAAAATAGCTTTAAAT
ACTTATTTTTAATTAGAATTTTATATATTTTTAATATATAAGATCTTTACTTTTTTATAAGCTTCTACCTTAAAT
TAAATTTTTACTTTTTTTACTATTTTACTATATCTTAAATAAAGGCTTTAAAAATATAAAAAAATCTTCTTATAT
ATTATAAGCTATAAGGATTATATATATATTTTTTTTTAATTTTTAAAGTAAGTATTAAGCTAGAATTAAGTTTTTA
ATTTTTTAAGGCTTTATTTAAAAAAGGCAGTAATAGCTTATAAAAGAAATTTCTTTTTCTTTTATACTAAAAGTAC
TTTTTTTTAATAAGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGT
TAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAAACAAAGCCACGTTGTGTCTCAAATCTCTGATGTTAC
ATTGCACAAGATAAAAAATATATCATCATGAACAATAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTAT
GAGCCATATTCAACGGGAAACGTCTTGCTCGAGGCCGCGATTAATTCACATGGATGCTGATTTATATGGGTATA
AATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAATCTATCGATTGTATGGGAAGCCCATGCGCCAGAGTTG
TTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAACTGGCTGACGGAATT
TATGCCTCTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTACTCACCCTGCGATCCCCGGGA
AAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATATGTTGATGCGCTGGCAGTGTTCCTGCGC
CGGTTGCATTGATTCTGTTTGTAAATTGTCCTTTAACAGCGATCGCGTATTTGCTCTGCTCAGGCGCAATCAGG
AATGAATAACGGTTTTGGTTGATGCGAGTGATTTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAG
AAATGCATAAGCTTTTGCCATTCTCACCGGATTCAGTCGTCACCTCATGGTATTTCTCACTTGATAACCTTATTTTT
GACGAGGGGAAATTAATAGGTTGATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCT
ATGGAAGTGCCTCGGTGAGTTTTCTCCTTCATTACAGAAACGGCTTTTTCAAATAATGGTATTGATAATCCTGATA
TGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTCTAATCAGAATTGGTTAATTGGTTGTAACACTGGCAG
AGCATTACGCTGACTTGACGGACGGCGGCTTTGTTGAATAAATCGAACTTTTGCTGAGTTGAAGGATCAGATCAGG
CATCTTCCCAGACAACGCAGACCGTTCCGTGGCAAAGCAAAAGTTCAAAATCACCAACTGGTCCACCTACAACAAAGC
TCTCATCAACCGTGGCTCCCTCACTTTCTGGCTGGATGATGGGGCGATTCAGGCCTGGTATGAGTCAGCAACACCTT
CTTCAGAGGCAGACCTCAGCGGTTAAACCTAACCCTAACCCTAACCCTAACCCTAACCCTAACCCTAACCCTAAC
CCTAACCCTAACCCTAACCCTAACCCTAACCCTAACCCTAACCCTAACCCTAACCCTAATGGGGTCGATCTGAACCGAGGATGA
GGGTTCTATAGACTAATCTACAGGCCGTACATGGTGTGATTGCAGATGCGACGGGCAAGGTGTACAGTGTCCAGAAG
GAGGAGAGCGGCATAGGTATTGTAATAGACCAGCTTACATAATAATCGCCTGTTGCTACTGACTGATGACCTTCTT
CCCTAACCCAGTTTCTAATTACCACTGCAGTGAGGATAACCCTAACTCGCTCTGGGGTTATTATTATACTGATTAGC
AGGTGGCTTATATAGTGCTGAAGTACTATAAGAGTTTCTGCGGGAGGAGGTGAAGGACTATAAACTGGACACAGTT
AGGGATAGAGTGATGACAAGACCTGAATGTTATCCTCCGGTGTGGTATAGCGAATTGGCTGACCTTGCAGATGGTAA

TGGTTTAGGCAGGGTTTTTGCAGAGGGGGACGAGAACGCGTTCCTGCGATTTAACGGCTGCTGCCCAAGCTTTACG
GTTCTCTAATGGGCGGCCCTCAGGTCGACGTCCCATGGCCATTCGAATTCGTAATCATGGTCATAGCTGTTTCCT
GTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTA
ATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCCGCTTCCAGTCGGGAAACCTGTGCTGCCAGCTGC
ATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTTCGCTATTGGGCGCTCTCCGCTTCCCTCGCTCACTGACTCG
CTGCGCTCGGTCGTTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGG
GGATAACGCAGGAAAGAACATGTGAGCAAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGCTGGCGT
TTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGA
CTATAAGATAACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATA
CCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGG
TCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTTCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGT
CTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTA
TGTAGGCGGTGCTACAGAGTTCCTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCG
CTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGT
GGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTTCTACGGG
GTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGA
TCCTTTTTAAATTAATAATGAAGTTTTAAATCAATCTAAAAGTATATAGAGTAAACTTGGTCTGACAGTTACCAATGC
TTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT
AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAG
ATTTATCAGCAATAAACCCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCGCAACTTTATCCGCTCCATCCAG
TCTATTAATTGTTGCCGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTAC
AGGCATCGTGGTGTACGCTCGTCTTTGGTATGGCTTCATTCAGCTCCGGTCCCAACGATCAAGGCGAGTTACAT
GATCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCCTCCGATCGTTGTGAGAAGTAAGTTGGCCGAGTG
TTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTGATGCCATCCGTAAGATGCTTTTCTGTGACTGG
TGAGTACTCAACCAAGTCATTCTGAGAATAGTGATGCGGGCACCAGTTGCTCTTGCCCGGCGTCAATACGGGATA
ATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACCTCTCAAGGATC
TTACCGCTGTTGAGATCCAGTTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTTACCAG
CGTTTTCTGGGTGAGCAAAAAACAGGAAGGCAAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAATGTTGAATAC
TCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGT
ATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACAATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTAT
TATCATGACATTAACCTATAAAAAATAGGCGTATCACGAGGCCCTTTCGCTCTCGCGCTTTCGGTGATGACGGTGAAA
ACCTCTGACACATGCAGCTCCCGGAGACGGTTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCCTCAG
GGCGCGTCAGCGGGTGTGGCGGGTGTGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTAAGTACTGAGAGTGCA
CCATAAAATTGTAAACGTTAATATTTTTGTTAAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAAT
AGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAAGAAATAGCCCGAGATAGGGTTGAGTGTGTTCCAGTTTGAAC
AAGAGTCCACTATTAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACG
TGAACCATCACCCAAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAGCACTAAATCGGAACCCTAAAGGGAGCCCC
GATTTAGAGCTTGACGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGAAGAAAGCGAAAGGAGCGGGCGCTAGG
GCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCTTAATGCGCCGCTACAGGGCGCGTA

CTATGGTTGCTTTGACGTATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCCATTC
 GCCATTACAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGG
 GATGTGCTGCAAGGCGATTAAGTTGGGTAAACGCCAGGGTTTTCCAGTCACGACGTTGTAACGACGGCCAGTGCC
 CAAGCTTACTAGATGCATGCTCGAGCGGCCCGCCAGTGTGATGGATATCTGCAGAATTCGCCCTTGACTAGTGCTCTC
 TATCCTGGTGGCAGGCGTCAAGTACCCAGAGGCAGCAGCGGGCTTAGGAGCGGCCTGGGTTGTTCTCCGCACCCTCT
 ACATGCTGGGCTATATTTATAGCGACAAGCCGAACGGCACCGGCAGGTACAATGGTTCGCTGTACTTGCTTGCGCAA
 GCGGGTCTTTGGGGATTGAGCGCATTTGGTGTGCAAAGGATTTGATGTAAATGTAGTCGACATCTTAGCACAGAGG
 GGAGAGTTGATAAAATGTGGTCTGTTTGAATGATAGTCGGGTTCTGTGACCTATATTCGTGATAGTGGAGATAGGTCT
 GCGCCTATCTTATCGGGCCGAGCAAAAATTCACCGCAGCGGGTGAGTTTTCGTTATACAGCCATCCCACCTTCCA
 GCTTCAAATTTGTCAGTTAATCCAGCCCAATCAATCATTGGAGAACCGCCATCATGTCTTCGAAGTCCCACCTCCC
 CTACGCAATTCGCGCAACCAACCATCCCAACCTTTAACATCTAAACTCTTCCATCGCCGAGGAGAAGAAAACCA
 ACGTCACCGTCTCCGCAGACGTTACTACTTCCGCCGAGCTCCTCGATCTTGCTGACCGTACATCCTGCACCAATGCC
 CCTCCAGGATAACAAATAGCTGATGCGTAGTGAGTACAGGCCTAGGCCCTATATCGCAGTCTGAAAACCCACATC
 GACATCCTCACCGATCTCACCCCGTCGACCCTTTCCTCGCTCCAATCCCTCGCGACAAAGCACAACCTTCTCATCTT
 TGAGGACCGCAAGTTCATCGACATCGGCAACACCGTGCAAAAGCAGTACCACGGTGGCGCTCTCCGCATCTCCGAAT
 GGGCACACATCATCAACTGCGCCATCCTGCCGGCGAAGGGATCGTCGAGGCCCTCGCACAGACAACCAAGTCTCCT
 GACTTTAAAGACGCGAATCAACGAGGTCTCCTGATTCTTGCCGAGATGACGAGTAAGGGATCTCTTGCGACAGGGGA
 GTCACAGGCACGCTCGGTTGAGTACGCGCGAAGTATAAGGGGTTTGATGGGATTCGTGAGTACAAGGGCGTTGA
 GTGAGGTGCTGCCCGAACAGAAAGAGGAGAGCGAGGATTTTGTGCTTTTACGACTGGGGTGAATCTGTGCGATAAG
 GGGGATAAGCTGGGGCAGCAGTATCAGACACCTGGGTCCGCCGTTGGGCGAGGTGCGGACTTTATCATTGCGGGTAG
 GGGCATCTATAAGGCGGACGATCCAGTCGAGGCGGTTTCAGAGTACCAGGAGGAAGGCTGAAAAGCTTACGAGAAAA
 GAGTTGGACTTTGAGTGTGAGTGGAATGTGTAACGGTATTGACTAAAAGGGATCCATATGTTTATTGCAGCCAGCA
 TAGTATTACCAGAAAGAGCCTCACTGACGGCTCTAGTAGTATTGAAACAGATATTATTGTGACCAGCTCTGAACGAT
 ATGCTCCCTAATCTGGTAGACAAGCACTGATCTACCCCTTGGAACGCAGCATCTAGGCTCTGGCTGTGCTCTAACCC
 TAACTAGACGATTGATCGCAGACCATCCAATACTGAAAAGTCTCTATCAGAGGAAATCCCCAACATTGTAGTAGTCA
 GGTTCCCTTTGTGGCTGGGAGAGAATTGGTTCGCTCCACTGATTCCAGTTGAGAAAGTGGGCTAGAAAAAGTCTTGA
 AGATTGGAGTTGGGCTGTGGTTATCTAGTACTTCTCGAGCTCTGTACATGTCCGGTCGCGACGTACGCGTATCGATG
 GCGCCAGCTGCAGGCGGCCGCTGCAGCCACTTGCAGTCCCGTGGAAATTCACGGTGAATGTAGGCCTTTTGTAGG
 GTAGGAATTGTCACTCAAGCACCCCAACCTCCATTACGCCTCCCCATAGAGTTCCCAATCAGTGAGTCATGGCAC
 TGTTCTCAAATAGATTGGGGAGAAGTTGACTTCGCCAGAGCTGAAGTTCGCACAACCGCATGATATAGGGTCGGC
 AACGGCAAAAAAGCACGTGGCTCACCGAAAAGCAAGATGTTTGGATCTAACATCCAGGAACCTGGATACATCCATC
 ATCACGCACGACCACTTTGATCTGCTGGTAAACTCGTATTCGCCCTAAACCGAAGTGCCTGGTAAATCTACACGTGG
 GCCCCTTTCGGTATACTGCGTGTGTCTTCTCTAGGTGCCATTCTTTCCCTTCTCTAGTGTGAATTGTTTGTGTT
 GGAGTCCGAGCTGTAACCTCTGAATCTCTGGAGAATGGTGGACTAACGACTACCGTGCACCTGCATCATGTATA
 TAATAGTGATCCTGAGAAGGGGGTTTTGGAGCAATGTGGGACTTTGATGGTCATCAAACAAAGAACGAAGACGCCTC
 TTTTGCAAAAGTTTTGTTTCGGCTACGGTGAAGAACTGGATACTTGTGTGCTTCTGTGTATTTTTGTGGCAACAAG
 AGGCCAGAGACAATCTATTCAAACACCAAGCTTGCTCTTTGAGCTACAAGAACCTGTGGGGTATATATCTAGAGTT
 GTGAAGTCGGTAATCCCGCTGTATAGTAATACGAGTCGCATCTAAATACTCCGAAGCTGCTGCGAACCCGGAGAATC
 GAGATGTGCTGGAAAGCTTCTAGCGAGCGGCTAAATTAGCATGAAAGGCTATGAGAAATCTGGAGACGGCTTGTG

AATCATGGCGTTCCATTCTTCGACAAGCAAAGCGTTCCGTCGCAGTAGCAGGCACTCATTCCCGAAAAAACTCGGAG
ATTCTAAGTAGCGATGGAACCGGAATAATAATAGGCAATACATTGAGTTGCCTCGACGGTTGCAATGCAGGGGT
ACTGAGCTTGGACATAACTGTTCCGTACCCACCTCTTCTCAACCTTTGGCGTTTCCCTGATTGAGCGTACCCGTAC
AAGTCGTAATCACTATTAACCCAGACTGACCCGACGTGTTTTGCCCTTCATTTGGAGAAATAATGTCATTGCGATGT
GTAATTTGCCTGCTTGACCGACTGGGGCTGTTGGAAGCCCGAATGTAGGATTGTTATCCGAACTCTGCTCGTAGAGG
CATGTTGTGAATCTGTGTCGGCAGGACACGCCTCGAAGGTTACGGCAAGGGAAACCACCGATAGCAGTGTCTAGT
AGCAACCTGTAAAGCCGCAATGCAGCATCACTGGAAAATACAAACCAATGGCTAAAAGTACATAAGTTAATGCCTAA
AGAAGTCATATAACCAGCGGCTAATAATTGTACAATCAAGTGGCTAAACGTACCGTAATTTGCCAACGGCTTGTGGGG
TTGCGAAGCAACGGCAAAGCCCCACTTCCCCACGTTTGTCTTCACTCAGTCCAATCTCAGCTGGTGATCCCCCA
ATTGGGTCGCTTGTGTTCCGGTGAAGTGAAAAGAAGACAGAGGTAAGAATGTCTGACTCGGAGCGTTTTGCATACA
ACCAAGGGCAGTGATGGAAGACAGTGAAATGTTGACATTCAAGGAGTATTTAGCCAGGGATGCTTGAGTGTATCGTG
TAAGGAGGTTTTGTCTGCCGATACGACGAATACTGTATAGTCACTTCTGATGAAGTGGTCCATATTGAAATGTAAAGT
CGGCACTGAACAGGCAAAAAGATTGAGTTGAAACTGCCTAAGATCTCGGGCCCTCGGGCCTTCGGCCTTTGGGTGTAC
ATGTTTGTGCTCCGGGCAAATGCAAAGTGTGGTAGGATCGAACACACTGCTGCCTTTACCAAGCAGCTGAGGGTATG
TGATAGGCAAATGTTGAGGGGCCACTGCATGGTTTCGAATAGAAAGAGAAGCTTAGCCAAGAACAATAGCCGATAAA
GATAGCCTCATTAAACGGAATGAGCTAGTAGGCAAAGTCAGCGAATGTGTATATATAAAGGTTGAGGTCCGTGCCT
CCCTCATGCTCTCCCATCTACTCATCAACTCAGATCCTCCAGGAGACTTGTACACCATCTTTTGAGGCACAGAAAC
CCAATAGTCAACCATCACAAGTTTGTACAAAAAAGCAGGCTCACCATGAGCAACCTGCTCACCGTCCACCAGAACCT
CCCTGCCCTCCCTGTGACGCCACCTCTGACGAGGTCGCAAGAACCTCATGGACATGTTCCGCGACCGCCAGGCCT
TTAGCGAGCACACCTGGAAGATGCTCCTCAGCGTCTGCCGATCTTGGGCCCTGGTGCAAGCTCAACAACCGCAAG
TGGTTCCCCGCGAGCCGGAGGACGTCCGCGACTACCTCCTTACCTGCAGGCCGAGGCCTGGCCGTCAAGACCAT
CCAGCAGCACCTCGGCCAGCTCAACATGCTCCACCGACGCTCTGGCCTGCCTCGCCCTAGCGACTCTAACGCCGTCA
GCCTGGTCATGCGCCGCATCCGCAAGGAGAACGTCGACGCTGGCGAGCGAGCCAAGCAGGCCCTCGCCTTCGAGCGC
ACCGACTTCGACCAGGTCCGACGCTCATGGAGAACAGCGACCGCTGCCAGGATATCCGCAACCTCGCCTTTCTCGG
CATTGCCTACAACACCCTGCTCCGCATTGCCGAGATCGCCGCATCCGCGTCAAGGACATCTCTCGCACCGACGGCG
GCCGCATGCTCATTACATCGGCCGCACCAAGACCCTCGTGTCTACCGCCGGCGTCGAGAAGGCCCTCAGCCTCGGC
GTCACCAAGCTCGTCGAGCGCTGGATTTCTGTCTCCGGCGTCGCTGACGACCCCAACAACCTACCTCTTCTGCCGCGT
CCGAAAGAACGGCGTCGCCGCCCTTCTGCCACCTCTCAGCTCAGCACCCGAGCCCTGGAGGGCATCTTTGAGGCCA
CCCACCGCCTCATCTACGGCGCCAAGGACGACTCTGGCCAGCGCTACCTCGCCTGGTCTGGCCACTCTGCCCGAGTC
GGCGCTGCCCGAGACATGGCCCGAGCCGGCTCAGCATCCCCGAGATTATGCAGGCCGGCGGCTGGACCAACGTCAA
CATCGTCATGAACTACATCCGCACCCTCGACTCTGAGACCGGCGCCATGGTCCGACTCCTCGAGGACGGCGACTAAA
CCCAGCTTTC

[0151] 由A5D里氏木霉菌株产生MAD6菌株。

[0152] 使用与先前针对bg11缺失描述的方法相似的方法,从A5D菌株(上文)中缺失天然eg13(参见上文)。eg13缺失盒的示意图在图6中示出。潮霉素抗性用作eg13缺失的可选择标记。潮霉素抗性标记侧接着loxP位点。转化株被证实缺失了eg13。如针对A5D菌株的产生所描述的,从这个菌株中去除潮霉素标记。如针对A5D菌株的产生所描述的,从这个菌株中去除编码cre重组酶的端粒载体。eg13敲除盒的核苷酸序列作为SEQ ID NO:6提供:

[0153] GGGAGGTAGGCGCAGATACGGTGCATGGGACCCGAACCCGTAACCGAACACGACCTTATCAGCCCTCC

AACTCACACCCTCTCGCCTATCACTATCCTAGATAGTTCATCGGCCAACTCATGTAACCTAGCTACCTACCTACCTG
GTAAGAATGCGGGCTATCATGTCTCACGGCGGGTACATGTCCGGTATCTCGCTGCTTCCCCGAGGTTGACGTCCGA
ATCCATGCAAGTACTCCCTGAAATCGAGACGACAGAGAGAACAACCAACGCGCTTAAACGCTTCATGTTTATCTAAG
AGGCACATTCGAAGAAGTACTGCTTAAACACTAGACCTGGCTTTTCGACCCCCTCCGAGAAAGCCGTTTTCTCCTCA
ATCCTCCCGGGCTTGGCTTTTGTGAGTCCGTACTTGTGCGCTAACAGAGTCTTGGACGCAGCGTTTGGCATCAGT
CTTGCAGGCGGTTACGGGACTAGGACAACAGGGGATGTGACAGGCCGGATAGTAATTATGGGTTATCCGGGGTAAG
CAGGGAATTTACGAGGCCGCTTTACGTGGGGGAACAGCCACTGCGGGGGGAAGAGGAGTAGTAGGCGACTCGGTCCG
ATGAGCTCGAGGTGTCTGGTTGACTTGGACTGCAGAGCGTAGGTAATTGAGATCGGGCAACATTATCGGTGTTCCGGC
TCGGTATGGCCGAGTTGCGACTGCTTGGTCATTCGGCGAAGCTGATGTCGTGGTATCCTGAAGCATCGATATCGGAA
ACCATGATGGTCAGTCTATCTGACGTGTGCGGTGACAAGCGAGTCCGGATTTTGTGACATGACGTTCAACTTCAGTC
AATGCCTTAGGGCTCGATAAGATTAAGATTGGGTTCTGGCAGCGGTCTAGAACACCGCCACAAATTCTGTCCATTGA
GGAGCGTGATGTCTAGGCGCATCACTAACACGGAGCTGTATGACCGGCAGCTCAACGGACTTCTCTTCGTTCAACGG
CAGTCTATTTGCGGTACACGAATGGATCTTTCTTCTGGTCTTGAAGTGCCGAGTGGCGTGCGAATGTATAGATGT
CTCGCTACCTAGAAAAGCTGGCTTTTCTGACAGGGTCCCTTCCACCTCTCCTACCAACGACAAACTGAACAAGTATC
TGGCGTTTTCCCAACGCCGAATAGGCCAGTCGCCAATACTCCCTCCAGCCCTGATTGGGCCCTCGAAGTATCGCCA
TGTCTGTGTGTTGAGATTATTCGATGGACGTCCTCCCCAACCTACAGGAAGAGCAAAATGGGAGCAGTGTCTGC
AATGAGCTATATAATAGATCGCTCGATCTCATACAAATTGTATGCTCAGTCAATACAACGAGCGGTTCCAAGATCCC
TTCTCCAACGACCCTCGAAACATTGCAACCCGGTGCAGCCTGAACTGTTTCGTATAGCCTAGAAAGCGACGCCATCT
TCATCTTTTACGCGATTAGCCTCATGGCTATTTGTGCCGAAGTGGGAGTTGTATGGTAGCAGTGAGGAGATTGTGGC
TACGACACAGGCGGGTTCTCTTGAGCGGCTTACATCTCCGCATTAGGCCTGCGTACGATCCAGATCATGGGAAACTT
TACAATGGCTTACTCGTTTTATCTCAACACTGAGCTTCCAATTCACTCTATGCATTGATTAACACGTTTGGTCATGT
GGTTCTTCAGCTGTAAATCTTCAGCTTCCCAAGAATTGCAACCTCGCTGATTGCTAATAGTGTTCATGCGTTGCAT
CCTGGTGCAGCAGTGCAAAGGAGAGTCAAAGTAGCCGGCAGATTAATTTAAGCTTATATCACTCAGGGGTAAACAGC
CGTAAAGGACCTTTTGATCTAACATGCCGATGTGTATGTAGATCACGCAATGCCACCATATCTTGGCAGTCAGATT
TGTCCGTGGCGCGCCAAGTATAAATTTCGTATAATGTATGCTATACGAAGTTATCGGCCGGCGTATTGGGTGTTACGG
AGCATTCACTAGGCAACCATGGTTACTATTGTATACCCATCTTAGTAGGAATGATTTTCGAGGTTTATACCTACGAT
GAATGTGTGCTCCTGTAGGCTTGAGAGTTCAAGGAAGAAACAGTGCAATTATCTTTGCGAACCCAGGGGCTGGTGACG
GAATTTTCATAGTCAAGCTATCAGAGTAAAGAAGAGGAGCATGTCAAAGTACAATTAGAGACAAATATATAGTCGG
TGGAGCCAAGAGCGGATTCCCTCAGTCTCGTAGGTCTCTTGACGACCGTTGATCTGCTTGTCTCGTCTCCCGAAAAT
GAAAATAGACTCTGCTAAGCTATTCTTCTGCTTCGCCGGAGCCTGAAGGGCGTACTAGGGTTGCGAGGTCCAATGCA
TTAATGCATTGCAGATGAGCTGTATCTGGAAGAGGTAACCCGAAACGCGTTTTATTCTTGTGACATGGAGCTATT
AAATCACTAGAAGGCACTCTTTGCTGCTTGGACAAATGAACGTATCTTATCGAGATCCTGAACACCATTTGTCTCAA
CTCCGGAGCTGACATCGACACCAACGATCTTATATCCAGATTCGTCAAGCTGTTTGTATGATTTTTCAGTAACGTTAAGT
GGATCCCGGTGGCATCTACTCTATTCTTTGCCCTCGGACGAGTGCTGGGGCGTCCGTTTCCACTATCGGCGAGTA
CTTCTACACAGCCATCGGTCCAGACGGCCGCGCTTCTGCGGGCGATTTGTGTACGCCCAGAGTCCCGGCTCCGGAT
CGGACGATTGCGTGCATCGACCCTGCGCCAAAGTGCATCATCGAAATTGCCGTCAACCAAGCTCTGATAGAGTTG
GTCAAGACCAATGCGGAGCATATACGCCGGAGCCGCGGATCCTGCAAGCTCCGGATGCCCTCCGCTCGAAGTAGC
GCGTCTGCTGCTCCATACAAGCCAACCGCCTCCAGAAGAAGATGTTGGCGACCTCGTATTGGGAATCCCCGAAC
ATCGCCTCGCTCCAGTCAATGACCGCTGTTATGCGGCCATTGTCCGTCAGGACATTGTTGGAGCCGAAATCCGCGTG

CACGAGGTGCCGACTTCGGGGCAGTCCTCGGCCAAAGCATCAGCTCATCGAGAGCCTGCGCGACGGACGCACTGA
CGGTGTCGTCCATCACAGTTTGCCAGTGATACACATGGGGATCAGCAATCGCGCATATGAAATCACGCCATGTAGTG
TATTGACCGATTCTTTCGGTCCGAATGGGCCGAACCCGCTCGTCTGGCTAAGATCGGCCGAGCGATCGCATCCAT
GGCCTCCGCGACCGGCTGCAGAACAGCGGGCAGTTCGGTTTCAGGCAGGTCTTGCAACGTGACACCCTGTGCACGGC
GGGAGATGCAATAGGTCAGGCTCTCGCTGAATTCCCCAATGTCAAGCACTTCGGGAATCGGGAGCGCGGCCGATGCA
AAGTGCCGATAAACATAACGATCTTTGTAGAAACCATCGGGCAGCTATTTACCCGACGGACATATCCACGCCCTCC
TACATCGAAGCTGAAAGCACGAGATTCTTCGCCCTCCGAGAGCTGCATCAGGTCCGAGACGCTGTGCAACTTTTCGA
TCAGAACTTCTCGACAGACGTCGCGGTGAGTTCAGGCTTTTCATATCGATTGTGATGTGATGGAGTTGAGATGGA
GGTGAGGAGATGGATGATGGGAAAGGAAGATGGACTGAGGATGGAAGAAGAGAAGAAGAGAGAGAGAGAAAGTCTTC
CAGGAGAGAAAGGGAACCGAAGAAAAATGGGGAGGAAACCGGCCCTAGCACCTAAATACGTCTCCCGCTTAATTTTC
GCCCTTTTTTACCAAACCTCCGCGTCTTTCGTGCGCTAGCTGTCTTGGGGGTGTGTA AAACTTGGGAACAACCC
TACGCCGAACCTCCCGTACGAAGCCCGTACAGTGATTCTATCCCTGGCTTTCCAGCCTGGTAAGTCGGGTCCCTT
CGGGACGGGGCCAAGGAGACTGAGTTTCCGGTTAACCAATAATGCCGCCAGCCGTGGAGCGGTCTGAGCTGTCTAT
CGTGAATCCGTGACGCTGAATTGCTCAGTCCAAGTCGGAGACGCTGGAATCCACCGTTGCTCCAGCCACGGCGAAG
AATCCACTTACCTTCCGGGCTTCCGCCAGCTGGCAACATATTTTTGAGGCTGATATAGCCTTCTCCTCCATCACGAT
AAGCCCTGACTACCTTGCTACGGGCCAATTGCAATTTGTTTGTGTTACGCTTTACAAAAGGTGGCCGTTACTGAG
CAAAAAGAAAATGATGTA AAAATTTGCGAGTGGGTCCCATAGCTGGATGGTCCGATAAAAATGGTACTGCCCCACTT
AGTGGCAGCTCGCGACCAGTCACAAGCCCAGGATAACTTCGTATAATGTATGCTATACGAAGTTATCTGTGGGCGTT
ATGAATAATAGACTGGAACCGGGCCCTTTGATTGACGACTCCATATTTTGTAGATGTAGCAACTCGGCAAGAGCATT
ATGTGCAATACATTTGTTACCATACAAAGGCAGCTGCCAGACGACTTGATTGCGTACAATTCTCACGGCAAGCTTT
CCAGGTGTTATGCATTATGCGCAAATGCTTGATGCTTACCGCAGGATTAATCTCGGAAGAAGCGCTGCAAGCTATAT
GGGTGTAGTAGATATGTAGATGTACCAACCAATGAAGAACATTTATGGTCTAGAACGTAGTGATGAAGTTTTGAGT
AATTTGTATCAAGTAAGACGATATTATTGATATAATACCAAGCATATATTCATGATAAATTACTTGGAAACCACCCTT
GCGTCCGGCCTCACGAGCCTTCTCACTGCCGGGCTCGAAGGAGCCACTGGAGGCCTGTCCACCCTTGGATGCGATTT
CCTGCACCTTTTCTTGGGCTGCACGTGATTAGACATGATTCAAATCGAGATCTTGGAAATATCTTACATGCTGGC
GAAGCCACCGGTGTGGCTGGACTGTCCGCCCTTCTGCGCAATGCTTTGAACCTCCTCCTTGGGGCTGTGTAGAAAGG
TTTGTTAGCAACATTACTACAACCTCAGGACTCGGTGGTTCGTACCGGTTGGCGAAGTTTCCGGGGTTATCGTTGCC
AGACATTGTGTGATTATTTGGTGTGCAAAATGTGTGCTATGTGTGTTGTTGCTGTTGGTGATGATGCTGAAGCTGTTG
AAAGCAGGCTGGTTCTGTGGGAGAGACTTGGGATATTTATATCCAAAGTTCGGTCGTGTTCCCTTCTGGAAGCTCTTC
TCTACTCCATACAATCATCCAAAGTTGTCGTATTGAGCGTTGATCAGTAGTAGCCTCTGAGGTCATCACCATGATC
CTTCCGGCCAACAGTCGGCACTCATCAACAGCAACAATCAGCCGCCACAAACATAGGTACAGTAAGGAGTTAGATAT
CATGTAGTCGTGAGTACTCGACATCATGACGTACAAGCTTTGCCAGTGTGGTAGGTGCAAGTATGATGATCGTAT
CCGCCGTTGTTGATCGAACAGAGTCCGGTCAGATTCACGGTTTCTCTCACCTTGAACATTGGATGCAATTGGATTG
ATCCACAATCCTGGAGAATGGCTTCAAGCTCACTGCTCCAGTCGCAAGCTTCAGAGCCTATTACTAAGGGTAGAGCT
ACCTATGTCAAGAGTTTTCAAGGTACCTAAGCTACATGTGATAGTCGGCAAGCCATTTTGAACGCAGACCGTGAACG
GTGATGTA AATCCGGGATAGACGCCAAGCGTGCCGTGCAATGACGCTAGATACACCTCGATTTACGTAGAGTGAA
TGCCAGCCAATGGAGTCATGCACATAACCCGCTTAGACTCTGCTCGGGCGATACCCGATCGCAGAGGCAGAGCCGC
TTAAACGCGATCGCGTAACCTGTAATCAGAGCCAGCGCTCGATGAATTGCATCATGGAAGCCATTGATGTGGAATG
TTGAGCGTATAACAACACGAATTGAAGACGACATTGACTTGCTTCAAGTGAGTGGAGAATTGCCGGCAGACAAGAT

AGGTAGGCTCTTGGTGCCTGTACATCAATCCATTCCTTTTCCTCTGTTCAATCTCTATGTTGACATTCTGATAGG
 GATCATTGGATGCCAATGCAAAGAACATGAGAGTGTGGTCTGCATTCAAGTATCCTGGTTCGTAAGCTGTGGCCATGG
 GCGCTGCGGTCAAGGTCAATCGCGATGACTAATCAGTCTCGGTGACTCTGGGGCGGTAGAGGCAGTGTCTGTAACCA
 AAGCTTGAGCCGAGGGCAAAAAACAACGGCGCATCAAACAATCAACGAAAGCATCGTCAACAGTGTCTCTTCCCAGTC
 AATTACTTCGAAAAACCTTCTCGATAGAACCCTTCAGACGATGAACAGGCCACGCAACCGTCAGCCGCGCCCCCAG
 GACAGACTCAGCGCCCGGGAGGCAGATCGTCACACCTTGGTCGACGAGCTC

[0154] 实施例2

[0155] 酶表达比较:Qnad缺失菌株与MAD6菌株比较

[0156] 1.载体的制备

[0157] 在里氏木霉菌株、Archy3和quad缺失菌株中表达两种糖基水解酶家族43蛋白质。使用Gateway®克隆系统(英杰公司(Invitrogen)),由拟轮生镰刀菌(*Fusarium verticillioides*)基因组DNA克隆基因fv43B和fv43C,并装配成表达载体pTrex3gM。如图9中所示,最初将这两个基因克隆到pENTR/D-TOPO载体(英杰公司)中。

[0158] 随后将基因重组到载体pTrex3gM中,在其中cbhI启动子处于目的基因的编码序列的上游,并且cbhI终止子处于目的基因的终止密码子的下游。该载体另外含有位于cbhI终止子3'的构巢曲霉酰胺酶(amdS)可选择标记。该载体在图10中示出。

[0159] 所得到的Fv43B表达载体pTrex3gM-Fv43B,示意性显示于图11中。

[0160] 表达载体pTrex3gM中的Fv43B糖基水解酶的完全核苷酸序列作为SEQID NO:14提供如下。

[0161] TTGTACAAAGTGGTGATCGCGCCGCGCCAGCTCCGTGCGAAAGCCTGACGCACCGGTAGATTCTTGG
 TGAGCCCGTATCATGACGGCGGCGGAGCTACATGGCCCCGGTGATTTATTTTTTTTGTATCTACTTCTGACCCTT
 TTCAAATATACGGTCAACTCATCTTTCCTGAGATGCGGCCCTGCTTGGTATTGCGATGTTGTCAGCTTGCCAAATT
 GTGGCTTTCGAAAACAAAAACGATTCCCTTAGTAGCCATGCATTTTAAGATAACGGAATAGAAGAAAGAGGAAATTA
 AAAAAAAAAAAAAAAAACAAAACATCCCGTTCATAACCCGTAGAATCGCCGCTCTTCGTGTATCCCAGTACCAGTTTATT
 TTGAATAGCTCGCCCGCTGGAGAGCATCCTGAATGCAAGTAACAACCGTAGAGGCTGACACGGCAGGTGTTGCTAGG
 GAGCGTCGTGTTCTACAAGGCCAGACGTCTTCGCGGTTGATATATATGTATGTTTACTGCAGGCTGCTCAGCGACG
 ACAGTCAAGTTCGCCCTCGCTGCTTGTGCAATAATCGCAGTGGGGAAGCCACACCGTACTCCCATCTTTCAGTAAA
 GCTCTGTTGGTGTATCAGCAATACACGTAATTTAAACTCGTTAGCATGGGGCTGATAGCTTAATTACCGTTTACC
 AGTGCCATGGTTCTGCAGCTTTCCTTGGCCGTAATAATCGGCGAAGCCAGCCAATCACCAGCTAGGCACCAGCTAA
 ACCCTATAATTAGTCTCTTATCAACACCATCCGCTCCCCGGGATCAATGAGGAGAATGAGGGGGATGCGGGGCTAA
 AGAAGCCTACATAACCCTCATGCCAACTCCCAGTTTACACTCGTCGAGCCAACATCCTGACTATAAGCTAACACAGA
 ATGCCTCAATCCTGGGAAGAACTGGCCGCTGATAAGCGCGCCCGCTCGAAAAACCATCCCTGATGAATGGAAAGT
 CCAGACGCTGCCTGCGGAAGACAGCGTTATTGATTTCCCAAAGAAATCGGGGATCCTTTCAGAGGCCGAACTGAAGA
 TCACAGAGGCCTCCGCTGCAGATCTTGTGTCCAAGCTGGCGGCCGAGAGTTGACCTCGGTGGAAGTTACGCTAGCA
 TTCTGTAAACGGGCAGCAATCGCCCAGCAGTTAGTAGGGTCCCCTCTACCTCTCAGGGAGATGTAACAACGCCACCT
 TATGGGACTATCAAGCTGACGCTGGCTTCTGTGCAGACAACTGCGCCCACGAGTTCTTCCCTGACGCCGCTCTCGC
 GCAGGCAAGGGAACCTCGATGAATACTACGCAAAGCACAAGAGACCCGTTGGTCCACTCCATGGCCTCCCCATCTCTC
 TCAAAGACCAGCTTCGAGTCAAGGTACACCGTTGCCCTAAGTCGTTAGATGTCCCTTTTTGTGAGCTAACATATGC
 CACCAGGGCTACGAAACATCAATGGGCTACATCTCATGGCTAAACAAGTACGACGAAGGGGACTCGGTTCTGACAAC

CATGCTCCGCAAAGCCGGTGCCGTCTTCTACGTCAAGACCTCTGTCCCGCAGACCCTGATGGTCTGCGAGACAGTCA
 ACAACATCATCGGGCGCACCGTCAACCCACGCAACAAGAACTGGTCGTGCGGCGGCAGTTCCTGGTGGTGAGGGTGCG
 ATCGTTGGGATTCGTGGTGGCGTCATCGGTGTAGGAACGGATATCGGTGGCTCGATTTCAGTGC CGGCCGCGTTCAA
 CTTCTGTACGGTCTAAGGCCGAGTCATGGGCGGCTGCCGTATGCAAAGATGGCGAACAGCATGGAGGGTCAGGAGA
 CGGTGCACAGCGTTGTGCGGCCGATTACGCACTCTGTTGAGGGTGAGTCCCTTCGCCTCTTCTTTCTTTCTGCTCT
 ATACCAGGCCTCCACTGTCTCTCTTTCTTGTCTTTTATACTATATAACGAGACCGGCAGTCACTGATGAAGTATGTTA
 GACCTCCGCCTCTTACCAAAATCCGTCTCTCGGTCAGGAGCCATGGAAATACGACTCCAAGGTCATCCCCATGCCCTG
 GCGCCAGTCCGAGTCGGACATTATTGCCTCCAAGATCAAGAACGGCGGGCTCAATATCGGCTACTACAACTTCGACG
 GCAATGTCTTCCACACCCCTCTATCTGCGCGGCGTGGAAACCACCGTCGCCGCACTCGCCAAAGCCGGTCACACC
 GTGACCCCGTGGACGCCATAACAAGCACGATTTCCGCCAGGATCTCATCTCCCATATCTACGCGGCTGACGGCAGCGC
 CGACGTAATGCGCGATATCAGTGCATCCGGCGAGCCGGCGATTCCAAATATCAAAGACCTACTGAACCCGAACATCA
 AAGCTGTTAACATGAACGAGCTCTGGGACACGCATCTCCAGAAGTGAATTACCAGATGGAGTACCTTGAGAAATGG
 CGGGAGGCTGAAGAAAAGGCCGGGAAGGAACTGGACGCCATCATCGCGCCGATTACGCCTACCGCTGCGGTACGGCA
 TGACCAGTTCGGTACTATGGGTATGCCTCTGTGATCAACCTGCTGGATTTACGAGCGTGGTTGTTCCGGTTACCT
 TTGCGGATAAGAACATCGATAAGAAGAATGAGAGTTTCAAGGCGTTAGTGAGCTTGATGCCCTCGTG CAGGAAGAG
 TATGATCCGGAGGCGTACCATGGGGCACCGTTGCAGTGCAGGTTATCGGACGGAGACTCAGTGAAGAGAGGACGTT
 GGGCATTGCAGAGGAAGTGGGGAAGTTGCTGGGAAATGTGGTGACTCCATAGCTAATAAGTGT CAGATAGCAATTTG
 CACAAGAAATCAATACCAGCAACTGTAAATAAGCGCTGAAGTGACCATGCCATGCTACGAAAGAGCAGAAAAAACC
 TGCCGTAGAACCGAAGAGATATGACACGCTTCCATCTCTCAAAGGAAGAATCCCTTCAGGGTTGCGTTTCCAGTCTA
 GAGGCCATTTAGGCCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAA
 GTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCT
 GTTCCGACCCTGCCGTTACCGGATACCTGTCCGCCTTCTCCCTTCGGGAAGCGTGGCGTTTTCTCATAGCTCAGG
 CTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTT CAGCCCGACC
 GCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACT
 GGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACAC
 TAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCG
 GCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAA
 GAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAAACGAAAACCTCACGTTAAGGCCTGCAGGGCCGATT
 TTGGTCATGAGATTATCAAAAAGGATCTTACCTAGATCCTTTTAAATTA AAAATGAAGTTTTAAATCAATCTAAAG
 TATATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCT
 GTTCATCCATAGTTGCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCT
 GCAATGATAACCGCAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCG
 CAGAAGTGGTCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGAAGCTAGAGTAAGTAGTTCGC
 CAGTTAATAGTTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCTTGGTATGGCTTCA
 TTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGG
 TCCTCCGATCGTTGTGAGAAGTAAGTTGGCCGAGTGTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTA
 CTGTATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGG
 CGACCGAGTTGCTCTTGCCCGGCTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCAT
 TGAAAAACGTTCTTCGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTG

CACCCAACTGATCTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCA
 AAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTATCA
 GGGTTATTGTCTCATGGCCATTTAGGCCTCTAGAGTTGTGAAGTCGGTAATCCCGCTGTATAGTAATACGAGTCGCA
 TCTAAATACTCCGAAGCTGCTGCGAACCCGGAGAATCGAGATGTGCTGGAAAGCTTCTAGCGAGCGGCTAAATTAGC
 ATGAAAGGCTATGAGAAATTCTGGAGACGGCTTGTGAATCATGGCGTTCCATTCTCGACAAGCAAAGCGTTCCGT
 CGCAGTAGCAGGCACTCATTCCCAGAAAACTCGGAGATTCCCTAAGTAGCGATGGAACCGGAATAATATAATAGGCA
 ATACATTGAGTTGCCTCGACGGTTGCAATGCAGGGGTACTGAGCTTGGACATAACTGTTCCGTACCCACCTCTTCT
 CAACCTTTGGCGTTTCCCTGATTCAGCGTACCCGTACAAGTCGTAATCACTATTAACCCAGACTGACCGGACGTGTT
 TTGCCCTTCATTTGGAGAAATAATGTCATTGCGATGTGTAATTTGCCTGCTTGACCGACTGGGGCTGTTGCAAGCCC
 GAATGTAGGATTGTTATCCGAACCTGCTCGTAGAGGCATGTTGTGAATCTGTGTCGGGCAGGACACGCCTCGAAGG
 TTCACGGCAAGGGAAACCACCGATAGCAGTGTCTAGTAGCAACCTGTAAAGCCGCAATGCAGCATCACTGGAAAATA
 CAAACCAATGGCTAAAAGTACATAAGTTAATGCCTAAAGAAGTCATATACCAGCGGCTAATAATTGTACAATCAAGT
 GGCTAAACGTACCGTAATTTGCCAACGGCTGTGGGGTTGCAGAAGCAACGGCAAAGCCCCACTTCCCACGTTTGT
 TTCTTCACTCAGTCCAATCTCAGCTGGTGATCCCCAATTGGGTCGCTTGTGTTCCGGTGAAGTGAAAGAAGACA
 GAGGTAAGAATGTCTGACTCGGAGCGTTTTGCATACAACCAAGGGCAGTGATGGAAGACAGTGAAATGTTGACATTC
 AAGGAGTATTTAGCCAGGGATGCTTGAGTGATCGTGTAAGGAGTTTGTCTGCCGATACGACGAATACTGTATAGT
 CACTTCTGATGAAGTGGTCCATATTGAAATGTAAAGTCGGCACTGAACAGGCAAAGATTGAGTTGAAACTGCCTAA
 GATCTCGGGCCCTCGGGCCTTCGGCCTTTGGGTGTACATGTTGTGCTCCGGGCAAATGCAAAGTGTGGTAGGATCG
 AACACACTGCTGCCTTTACCAAGCAGCTGAGGGTATGTGATAGGCAAATGTTCCAGGGGCCACTGCATGGTTTTGCAAT
 AGAAAGAGAAGCTTAGCCAAGAACAATAGCCGATAAAGATAGCCTCATTAAACGGAATGAGCTAGTAGGCAAAGTCA
 GCGAATGTGTATATATAAAGGTTTCGAGGTCCGTGCCTCCCTCATGCTCTCCCCATCTACTCATCAACTCAGATCCTC
 CAGGAGACTTGTACACCATCTTTTGAGGCACAGAAACCCAATAGTCAACCATACAAGTTTGTACAAAAAAGCAGGC
 TCCGCGGCCGCCCTTACCATGCGCTTCTCTTGGCTATTGTGCCCTTCTAGCGATGGGAAGTGTCTTCTCTGA
 AACGAAGACGGATGTTTCGACATACACCAACCTGTCCTTCCAGGATGGCACTCGGATCCATCGTGTATCCAGAAAG
 ATGGCCTCTTCTCTGCGTCACTTCAACATTCATCTCCTTCCCAGGTCTTCCCGTCTATGCCTCAAGGGATCTAGTC
 AACTGGCGTCTCATCAGCCATGTCTGGAACCGCGAGAAACAGTTGCCTGGCATTAGCTGGAAGACGGCAGGACAGCA
 ACAGGGAATGTATGCACCAACCATTCGATACCACAAGGGAACATACTACGTCATCTGCGAATACCTGGGCGTTGGAG
 ATATTATTGGTGTCTCTTCAAGACCACCAATCCGTGGGACGAGAGTAGCTGGAGTGACCCTGTTACCTTCAAGCCA
 AATCACATCGACCCCGATCTGTTCTGGGATGATGACGGAAAGGTTTATTGTGCTACCCATGGCATCACTCTGCAGGA
 GATTGATTTGGAAACTGGAGAGCTTAGCCCGGAGCTTAATATCTGGAACGGCACAGGAGGTGTATGGCCTGAGGGTC
 CCCATATCTACAAGCGCGACGGTTACTACTATCTCATGATTGCCGAGGGTGGAAGTCCGAAGACCACGCTATCACA
 ATCGCTCGGGCCCAGATACCCGGCCCCTATGAAGCCTACAATAACAACCCAATCTTGACCAACCGCGGGACATC
 TGAGTACTTCCAGACTGTCCGTACGGTGATCTGTTCCAAGATACCAAGGGCAACTGGTGGGGTCTTTGTCTTGCTA
 CTCGCATCACAGCACAGGGAGTTTACCCATGGGCCGTGAAGCTGTTTTGTTCAATGGCACATGGAACAAGGGCGAA
 TGGCCCAAGTTGCAACCAGTACGAGGTCGCATGCCTGGAAACCTCCTCCCAAAGCCGACGCGAAACGTTCCCGGAGA
 TGGGCCCTTCAACGCTGACCCAGACAACACTACAACCTGAAGAAGACTAAGAAGATCCCTCCTCACTTTGTGCACCATA
 GAGTCCCAAGAGACGGTGCCTTCTCTTTGTCTTCCAAGGGTCTGCACATCGTGCCTAGTCGAAACAACGTTACCGGT
 AGTGTGTTGCCAGGAGATGAGATTGAGCTATCAGGACAGCGAGGTCTAGCTTTCATCGGACGCCGCAAACCTCACAC
 TCTGTTCAAATATAAGTGTGATATCGACTTCAAGCCCAAGTCCGATGATCAGGAAGCTGGAATCACCGTTTTCCGCA

CGCAGTTCGACCATATCGATCTTGGCATTGTTTCGTCTTCCTACAAACCAAGGCAGCAACAAGAAATCTAAGCTTGCC
 TTCCGATTCCGGGCCACAGGAGCTCAGAATGTTCCCTGCACCGAAGGTAGTACCGGTCCCGGATGGCTGGGAGAAGGG
 CGTAATCAGTCTACATATCGAGGCAGCCAACGCGACGCACTACAACCTTGGAGCTTCGAGCCACAGAGGCAAGACTC
 TCGACATCGCGACAGCATCAGCAAGTCTTGTGAGTGGAGGCACGGGTTCAATTTGTTGGTAGTTTGCTTGGACCTTAT
 GCTACCTGCAACGGCAAAGGATCTGGAGTGGAAATGCCCAAGGGAGGTGATGTCTATGTGACCCAATGGACTTATAA
 GCCCGTGGCACAAGAGATTGATCATGGTGTTTTTGTGAAATCAGAATTGTAGAAGGGTGGGCGCGCCGACCCAGCTT
 TC

[0162] 所得到的Fv43C表达载体pTrex3gM-Fv43C,示意性显示于图12中。

[0163] 表达载体pTrex3gM中的Fv43C糖基水解酶的完全核苷酸序列作为SEQID NO:15提供如下。

[0164] TTGTACAAAGTGGTGATCGCGCCGCGGCCAGCTCCGTGCGAAAGCCTGACGCACCGGTAGATTCTTGG
 TGAGCCCGTATCATGACGGCGGGGAGCTACATGGCCCCGGTGATTTATTTTTTTTGTATCTACTTCTGACCCTT
 TTCAAATATACGGTCAACTCATCTTTCCTGAGATGCGGCCTGCTTGGTATTGCGATGTTGTCAGCTTGCAAATT
 GTGGCTTTCGAAAAACAAAAACGATTCCCTTAGTAGCCATGCATTTTAAGATAACGGAATAGAAGAAAGAGGAAATTA
 AAAAAAAAAAAAAAAAAACAAACATCCCGTTCATAACCCGTAGAATCGCCGCTCTTCGTGTATCCAGTACCAGTTTATT
 TTGAATAGCTCGCCCGCTGGAGAGCATCCTGAATGCAAGTAACAACCGTAGAGGCTGACACGGCAGGTGTTGCTAGG
 GAGCGTCGTGTTCTACAAGGCCAGACGTCTTCGCGGTTGATATATATGTATGTTTACTGACGGCTGCTCAGCGACG
 ACAGTCAAGTTCGCCCTCGCTGCTTGTGCAATAATCGCAGTGGGGAAGCCACACCGTACTCCCATCTTTCAGTAAA
 GCTCTGTTGGTGTATATCAGCAATACACGTAATTTAAACTCGTTAGCATGGGGCTGATAGCTTAATTACCGTTTACC
 AGTGCCATGGTTCTGCAGCTTTCCTTGGCCCGTAAAAATTCGGCGAAGCCAGCCAATCACCAGCTAGGCACCAGCTAA
 ACCCTATAATTAGTCTCTTATCAACACCATCCGCTCCCCGGGATCAATGAGGAGAATGAGGGGGATGCGGGGCTAA
 AGAAGCCTACATAACCCCTCATGCCAACTCCCAGTTTACACTCGTCGAGCCAACATCCTGACTATAAGCTAACACAGA
 ATGCCTCAATCCTGGGAAGAAGTGGCCGCTGATAAGCGCGCCCGCTCGAAAAACCATCCCTGATGAATGGAAAGT
 CCAGACGCTGCCTGCGGAAGACAGCGTTATTGATTTCCCAAAGAAATCGGGGATCCTTTCAGAGGCCGAACTGAAGA
 TCACAGAGGCCTCCGCTGCAGATCTTGTGTCCAAGCTGGCGGCGGAGAGTTGACCTCGGTGGAAGTTACGCTAGCA
 TTCTGTAAACGGGCAGCAATCGCCCAGCAGTTAGTAGGGTCCCCTCTACCTCTCAGGGAGATGTAACAACGCCACCT
 TATGGGACTATCAAGCTGACGCTGGCTTCTGTGCAGACAACTGCGCCACGAGTTCTTCCCTGACGCCGCTCTCGC
 GCAGGCAAGGGAAGTTCGATGAATACTACGCAAAAGCACAAGAGACCCGTTGGTCCACTCCATGGCCTCCCCATCTCTC
 TCAAAGACCAGCTTCGAGTCAAGGTACACCGTTGCCCTAAGTCGTTAGATGTCCCTTTTTGTCAGCTAACATATGC
 CACCAGGGCTACGAAACATCAATGGGCTACATCTCATGGCTAAACAAGTACGACGAAGGGGACTCGGTTCTGACAAC
 CATGCTCCGCAAAGCCGGTGGCGTCTTCTACGTCAAGACCTCTGTCCCGCAGACCCTGATGGTCTGCGAGACAGTCA
 ACAACATCATCGGGCGCACCGTCAACCCACGCAACAAGAACTGGTCGTGCGGCGGCAGTTCTGGTGGTGAGGGTGGC
 ATCGTTGGGATTTCGTGGTGGCGTCATCGGTGTAGGAACGGATATCGGTGGCTCGATTTCGAGTGCCGGCCGCGTTCAA
 CTTCTGTACGGTCTAAGGCCGAGTCATGGGCGGCTGCCGTATGCAAAGATGGCGAACAGCATGGAGGGTTCAGGAGA
 CGGTGCACAGCGTTGTGGGCGGATTACGCACTCTGTTGAGGGTGAGTCCTTCGCTCTTCTTCTTTTCTGCTCT
 ATACCAGGCCTCCACTGTCTCTCTTCTTGTCTTTTATACTATATACGAGACCGGCAGTCACTGATGAAGTATGTTA
 GACCTCCGCCTCTTACCAAATCCGTCCCTCGGTCAGGAGCCATGGAAATACGACTCCAAGGTCATCCCCATGCCCTG
 GCGCCAGTCCGAGTCGGACATTATTGCCTCCAAGATCAAGAACGGCGGGCTCAATATCGGCTACTACAACCTTCGACG
 GCAATGTCCTTCCACACCCTCCTATCCTGCGCGGCGTGGAAACCACCGTCCCGCACTCGCCAAAGCCGGTACACCC

GTGACCCCGTGGACGCCATACAAGCACGATTTTCGGCCACGATCTCATCTCCCATATCTACGCGGCTGACGGCAGCGC
CGACGTAATGCGCGATATCAGTGCATCCGGCGAGCCGGCGATTCCAAATATCAAAGACCTACTGAACCCGAACATCA
AAGCTGTAAACATGAACGAGCTCTGGGACACGCATCTCCAGAAGTGAATTACCAGATGGAGTACCTTGAGAAATGG
CGGGAGGCTGAAGAAAAGGCCGGGAAGGAACTGGACGCCATCATCGCGCCGATTACGCCTACCGCTGCGGTACGGCA
TGACCAGTTCGGTACTATGGGTATGCCTCTGTGATCAACCTGCTGGATTTACAGAGCGTGGTTGTTCCGGTTACCT
TTGCGGATAAGAACATCGATAAGAAGAATGAGAGTTTCAAGGCGGTTAGTGAGCTTGATGCCCTCGTGAGGAAGAG
TATGATCCGGAGGCGTACCATGGGGCACCGTTGCAGTGCAGGTTATCGGACGGAGACTCAGTGAAGAGAGGACGTT
GGCGATTGCAGAGGAAGTGGGGAAGTTGCTGGGAAAATGTGGTGACTCCATAGCTAATAAGTGTGAGATAGCAATTTG
CACAAGAAATCAATACCAGCAACTGTAAATAAGCGCTGAAGTGACCATGCCATGCTACGAAAGAGCAGAAAAAACC
TGCCGTAGAACCGAAGAGATATGACACGCTTCCATCTCTCAAAGGAAGAATCCCTTCAGGGTTGCGTTTCCAGTCTA
GAGGCCATTTAGGCCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCAGAAAAATCGACGCTCAA
GTCAGAGGTGGCGAAAACCCGACAGGACTATAAAGATACCAGGCGTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCT
GTTCCGACCCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGTTTTCTCATAGCTCAGG
CTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTACGCCCCGACC
GCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACT
GGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACAC
TAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCG
GCAAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAA
GAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGCCTGCAGGGCCGATT
TTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAATAATGAAGTTTTAAATCAATCTAAAG
TATATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCT
GTTTATCCATAGTTGCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCT
GCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCG
CAGAAGTGGTCCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTGTTGCCGGAAGCTAGAGTAAGTAGTTCGC
CAGTTAATAGTTTGCACAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCTGTTGGTATGGCTTCA
TTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAGCGGTTAGCTCCTTCGG
TCCTCCGATCGTTGTGAGAAGTAAGTTGGCCGAGTGTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTA
CTGTGATGCCATCCGTAAGATGCTTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGG
CGACCGAGTTGCTCTTGCCCGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCAT
TGGAACACGTTCTTCGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTG
CACCCAACTGATCTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCA
AAAAAGGGAATAAGGGCGACACGAAAATGTTGAATACTCATACTCTTCTTTTCAATATTATTGAAGCATTTATCA
GGGTTATTGTCTCATGGCCATTTAGGCCTCTAGAGTTGTGAAGTCGGTAATCCCGCTGTATAGTAATACGAGTCGCA
TCTAAATACTCCGAAGCTGCTGCGAACCCGGAGAATCGAGATGTGCTGGAAAGCTTCTAGCGAGCGGCTAAATTAGC
ATGAAAGGCTATGAGAAATTCTGGAGACGGCTGTTGAATCATGGCGTTCCATTCTTCGACAAGCAAAGCGTTCCGT
CGCAGTAGCAGGCACTCATTCCCGAAAAAATCGGAGATTCCTAAGTAGCGATGGAACCGGAATAATATAATAGGCA
ATACATTGAGTTGCTCGACGTTGCAATGCAGGGTACTGAGCTTGACATAACTGTTCCGTACCCACCTCTTCT
CAACCTTTGGCGTTTTCCCTGATTCAGCGTACCCGTACAAGTCGTAATCACTATTAACCCAGACTGACCGGACGTGTT
TTGCCCTTCATTTGGAGAAATAATGTCATTGCGATGTGTAATTTGCCTGCTTGACCGACTGGGGCTGTTGGAAGCCC

GAATGTAGGATTGTTATCCGAACTCTGCTCGTAGAGGCATGTTGTGAATCTGTGTGGGCAGGACACGCCTCGAAGG
TTCACGGCAAGGGAAAACCCGATAGCAGTGTCTAGTAGCAACCTGTAAAGCCGCAATGCAGCATCACTGGAAAATA
CAAACCAATGGCTAAAAGTACATAAGTTAATGCCTAAAGAAGTCATATACCAGCGGCTAATAATTGTACAATCAAGT
GGCTAAACGTACCGTAATTTGCCAACGGCTTGTGGGGTTGCAGAAGCAACGGCAAAGCCCCACTTCCCCACGTTTGT
TTCTTCACTCAGTCCAATCTCAGCTGGTGATCCCCCAATTGGGTGCTTGTTTGTTCCGGTGAAGTGAAAAGAAGACA
GAGGTAAGAATGTCTGACTCGGAGCGTTTTGCATACAACCAAGGGCAGTGATGGAAGACAGTGAAATGTTGACATTC
AAGGAGTATTTAGCCAGGGATGCTTGAGTGATCGTGTAAGGAGGTTTGTCTGCCGATACGACGAATACTGTATAGT
CACTTCTGATGAAGTGGTCCATATTGAAATGTAAAGTCGGCACTGAACAGGCAAAGATTGAGTTGAAACTGCCTAA
GATCTCGGGCCCTCGGGCCTTCGGCCTTTGGGTGTACATGTTTGTGCTCCGGGCAAATGCAAAGTGTGGTAGGATCG
AACACACTGCTGCCTTTACCAAGCAGCTGAGGGTATGTGATAGGCAAATGTTCCAGGGGCCACTGCATGGTTTTCGAAT
AGAAAGAGAAGCTTAGCCAAGAACAATAGCCGATAAAGATAGCCTCATTAACCGAATGAGCTAGTAGGCAAAGTCA
GCGAATGTGTATATATAAAGGTTGAGGTCCGTGCCTCCCTCATGCTCTCCCCATCTACTCATCAACTCAGATCCTC
CAGGAGACTTGTACACCATCTTTTGGGACACAGAAACCAATAGTCAACCATCACAAGTTTGTACAAAAAAGCAGGC
TCCGCGGCCGCCCTTACCATGCGTCTTCTATCGTTTCCAGCCATCTCCTCGTGGCCTTCTAACCCTCAAAGA
GGCTTCATCCCTCGCCCTCAGCAAACGGGATAGCCCTGTCCCTCCCGGCCCTCTGGGCGGACCCCAACATCGCCATCG
TCGACAAGACATACTACATCTTCCCTACCACCGACGGTTTTCGAAGGCTGGGGCGGCAACGTCTTCTACTGGTGGAAA
TCAAAGATCTCGTATCATGGACAAAGAGCGACAAGCCATTCCCTACTCTCAATGGTACGAATGGCAACGTTCCCTG
GGCTACAGGTAATGCCTGGGCTCCTGCTTTCGCTGCTCGCGGAGGCAAGTATTACTTCTACCATAGTGGGAATAATC
CCTCTGTGAGTGATGGGCATAAGAGTATTGGTGCGGCGTGGCTGATCCTGAGGGGCCGTGGAAGGCACAGGAT
AAGCCGATGATCAAGGGAACCTTCTGATGAGGAGATTGTCAGCAACCAGGCTATCGATCCCGCTGCCTTGAAGACCC
TGAGACTGGAAAGTGGTATATCTACTGGGGAAACGGTGTCCTCCATGTCGAGAGCTCAACGACGACATGGTCTCTC
TCAAAGCAGGCTGGCACAAAATCACAGGTCTTTCAGAAATTTCCGCGAGGGTCTTTTCGTCAACTATCGCGATGGAACA
TATCATCTGACATACTCTATCGACGATACGGGCTCAGAGAACTATCGCGTTGGGTACGCTACGGCGGATAACCCCAT
TGGACCTTGGACATATCGTGGTGTCTTCTGAGAGAAGGACGAATCGAAGGGCATTCTTGCTACGGGACATAACTCCA
TCATCAACATTCCTGGAACGGATGAGTGGTATATCGCGTATCATCGCTTCCATATTCGGATGGAAATGGGTATAAT
AGGGAGACTACGATTGATAGGGTACCCATCGACAAGGATACGGGTTTGTGTTGGAAAGGTTACGCCGACTTTGCAGAG
TGTTGATCCTAGGCCTTTGTAGAAGGGTGGGCGCGCCGACCCAGCTTTC

[0165] Fv43B, (来自拟轮生镰刀菌的GH43家族酶)的核苷酸序列作为SEQ IDNO:16提供如下:

[0166] ATGCGCTTCTTTGGCTATTGTGCCCTTCTAGCGATGGGAAGTGCTCTTCCCTGAAACGAAGACGGAT
GTTTCGACATACACCAACCCTGTCTTCCAGGATGGCACTCGGATCCATCGTGTATCCAGAAAGATGGCCTCTTTCT
CTGCGTCACTTCAACATTCATCTCCTTCCAGGTCTTCCCGTCTATGCCTCAAGGGATCTAGTCAACTGGCGTCTCA
TCAGCCATGTCTGGAACCGCGAGAAACAGTTGCCTGGCATTAGCTGGAAGACGGCAGGACAGCAACAGGGAATGTAT
GCACCAACCATTGATAACCACAAGGGAACATACTACGTCATCTGCGAATACCTGGGCGTTGGAGATATTATTGGTGT
CATCTTCAAGACCACCAATCCGTGGGACGAGAGTAGCTGGAGTGACCCTGTTACCTTCAAGCCAAATCACATCGACC
CCGATCTGTTCTGGGATGATGACGGAAGGTTTATTGTGCTACCCATGGCATCACTCTGCAGGAGATTGATTTGGAA
ACTGGAGAGCTTAGCCCGGAGCTTAATATCTGGAACGGCACAGGAGGTGATGGCCTGAGGGTCCCCATATCTACAA
GCGCGACGGTTACTACTATCTCATGATTGCCGAGGTGGAAGTCCGAAGACCACGCTATCACAATCGCTCGGGCCC
GCAAGATCACCGGCCCTATGAAGCCTACAATAACAACCAATCTTGACCAACCGGGACATCTGAGTACTTCCAG

ACTGTCGGTCACGGTGATCTGTTCCAAGATACCAAGGGCAACTGGTGGGGTCTTTGTCTTGCTACTCGCATCACAGC
ACAGGGAGTTTCACCCATGGGCCGTGAAGCTGTTTTGTTCAATGGCACATGGAACAAGGGCGAATGGCCCAAGTTGC
AACCAGTACGAGGTGCATGCCTGGAAACCTCCTCCCAAAGCCGACGCGAAACGTTCCCGGAGATGGGCCCTTCAAC
GCTGACCCAGACAACACTACAACCTGAAGAAGACTAAGAAGATCCCTCCTCACTTTGTGCACCATAGAGTCCCAAGAGA
CGGTGCCTTCTCTTTGTCTTCCAAGGGTCTGCACATCGTGCCTAGTCGAAACAACGTTACCGGTAGTGTGTTGCCAG
GAGATGAGATTGAGCTATCAGGACAGCGAGGTCTAGCTTTCATCGGACGCCGCAAACACTCACACTCTGTTCAAATAT
AGTGTGATATCGACTTCAAGCCCAAGTCCGATGATCAGGAAGCTGGAATCACCGTTTTCCGCACGCAGTTCCGACCA
TATCGATCTTGGCATTGTTTCGTCTTCTACAAACCAAGGCAGCAACAAGAAATCTAAGCTTGCCCTCCGATTCCGGG
CCACAGGAGCTCAGAATGTTCTGCACCGAAGGTAGTACCGGTCCCCGATGGCTGGGAGAAGGGCGTAATCAGTCTA
CATATCGAGGCAGCCAACGCGACGCACTACAACCTTGGAGCTTCGAGCCACAGAGGCAAGACTCTCGACATCGCGAC
AGCATCAGCAAGTCTTGTGAGTGGAGGCACGGGTTCACTTGTGTTAGTTGCTTGGACCTTATGCTACCTGCAACG
GCAAAGGATCTGGAGTGAATGTCCCAAGGGAGGTGATGTCTATGTGACCCAATGGACTTATAAGCCCGTGGCACAA
GAGATTGATCATGGTGTGTTTTGTGAAATCAGAATTGTAG

[0167] Fv43B的蛋白质序列作为SEQ ID NO:17提供如下:

[0168]

MRFSWLLCPLLAMGSALPETKTDVSTYTNPVLPGWHS DPSCIQKDGLFLCVTSTFI SFPGLPVYASRDLVNWRLISH
VWNREKQLPGISWKTAGQQQGMYP TIRYHKGYTYV ICEYLG VGD IIGVIFKTTNPWDES
SWSDPVTFKPNHIDPDLFWDDDGKVVY CATHGITLQEIDLETGEL SPELNIWNGTGGVWPEGPHIYKRDGYYYLMI AE
GGTAEDHAIT IARARKITGPEAYNNNP ILTNRGTSEYFQTVGHGDLFQDTKGNWWGLCLATRIT AQGVSPMGREAV
LFNGTWNKGEWPKLQPVRRMPGNLLPKPTRNVPDGPFNADPDNYNLKTKKI PPHFVHHRVPRDGAFSLSSKGLH
IVPSRNNVTGSVLPGDEIELSGQRGLAF IGRRQTHLTKYSVDIDFKPKSDDQEAGITVFRTQFDHIDLGI VRLPTN
QGSNKKSKLAFRFRATGAQNVPAPKVVPD GW EKGVISLHIEAANATHY NLGASSHRGKTLDIATASASLVSGGTG
SFVGSLLGPYATCNGKGSVVECPKGGDVYVTQW TYKPV AQEIDHGVFVKSEL

[0169] Fv43C,(来自拟轮生镰刀菌的GH43家族酶)的核苷酸序列作为SEQ IDNO:18提供如下:

[0170] ATGCGTCTTCTATCGTTTCCCAGCCATCTCCTCGTGGCCTTCCTAACCCCTCAAAGAGGCTTCATCCCTC
GCCCTCAGCAAACGGGATAGCCCTGTCTCCCCGGCTCTGGGCGGACCCCAACATCGCCATCGTCGACAAGACATA
CTACATCTTCCCTACCACCGACGGTTTTCGAAGGCTGGGGCGGCAACGCTTCTACTGGTGGAAATCAAAGATCTCG
TATCATGGACAAAGAGCGACAAGCCATTCCCTACTCTCAATGGTACGAATGGCAACGTTCCCTGGGCTACAGGTAAT
GCCTGGGCTCCTGCTTTTCGCTGCTCGCGAGGCAAGTATTACTTCTACCATAGTGGGAATAATCCCTCTGTGAGTGA
TGGGCATAAGAGTATTGGTGC GGCGGTGGCTGATCATCCTGAGGGGCCGTGGAAGGCACAGGATAAGCCGATGATCA
AGGGAACCTTCTGATGAGGAGATTGTCAGCAACCAGGCTATCGATCCCGCTGCCTTTGAAGACCCTGAGACTGGAAAG
TGGTATATCTACTGGGGAACGGTGTCCCCATTGTCGCAGAGCTCAACGACGACATGGTCTCTCTCAAAGCAGGCTG
GCACAAAATCACAGGTCTTCAGAATTTCCGCGAGGGTCTTTTCGTCAACTATCGCGATGGAACATATCATCTGACAT
ACTCTATCGACGATACGGGCTCAGAGAACTATCGCGTTGGGTACGCTACGGCGGATAACCCCATTTGGACCTTGGACA
TATCGTGGTGTCTTCTGGAGAAGGACGAATCGAAGGGCATTCTTGCTACGGGACATAACTCCATCATCAACATTCC
TGGAACGGATGAGTGGTATATCGCGTATCATCGCTCCATATTTCCCGATGGAAATGGGTATAATAGGGAGACTACGA
TTGATAGGGTACCCATCGACAAGGATACGGGTTTGTGTTGGAAAGTTACGCCGACTTTCAGAGTGTGATCCTAGG
CCTTTGTAG

[0171] Fv43C的蛋白质序列作为SEQ ID NO:19提供如下:

[0172] MRLLSFPSHLLVAFLLTLKEASSLALSKRDSPLPLGLWADPNIAIVDKTYIIFPTTDGFEGWGGNVFYWW
KSKDLVSWTKSDKPFLLTLNGTNGNVPWATGNAWAPAFARGGGKYFYHSGNNSVSDGHKSIGAAVADHPEGPWKAQ
DKPMIKGTSDEEIVSNQAIDPAAFEDPETGKWIYWGNGVPIVAELNDDMVSLKAGWHKITGLQNFREGLFVNYRDG
TYHLTYSIDDTGSENYRVGYATADNP IGPWTYRGVLLKDESKGILATGHNSI INIPGTDEWYIAYHRFHIPDGNGY
NRETTIDRVPIDKDTGLFGKVTPTLQSVDPRL

[0173] 将pTrex3gM-Fv43B和pTrex3gM-Fv43C载体各自独立地通过PEG介导的原生质体融合转化到MAD6菌株中,并通过粒子轰击转化到quad缺失的菌株中。

[0174] 2. quad缺失的里氏木霉菌株的转化。

[0175] 根据制造商的说明书且如美国专利申请公开US 2006/0003408的实施例2中所述,通过PDS-1000/氩系统(位于加利福尼亚州Hercules的伯乐公司(Biorad, Hercules, CA)),使用生物射弹粒子轰击,将载体pTrex3gM-Fv43B和载体pTrex3gM-Fv43C独立地转化到里氏木霉quad缺失的菌株中。

[0176] 3. 用fv43B和fv43C转化的里氏木霉quad缺失克隆的SDS-PAGE

[0177] 如PCT公开WO 2011/038019中所述,将稳定转化株在96孔微量滴定板中生长。将培养上清液在SDS-PAGE上跑胶,然后用Simply Blue染剂(英杰公司)进行考马斯蓝染色。通过密度测定法扫描并分析凝胶。使用ImageJ(来自美国国立卫生研究院(National Institutes of Health))并通过采用Analyze Gel子菜单功能(如用户指南的小节27.13中所述),完成图像处理 and 条带亮度定量。将对应于Fv43B蛋白的条带进行定量并以占总蛋白质的百分比报告。图13提供了从用fv43B转化的里氏木霉quad缺失克隆表达的蛋白质的SDS-PAGE照片。

[0178] 将对应于Fv43C蛋白的条带进行定量并以占总蛋白质的百分比报告。Fv43C在凝胶上跑出代表不同糖形的两个条带,将这两个条带汇总在密度测定法分析中。从用Fv43C转化的里氏木霉quad缺失克隆表达的蛋白质的SDS-PAGE显示于图14中。

[0179] 4. PEG介导的MAD6里氏木霉菌株的原生质体融合转化

[0180] 使用引物1061F和1085R,通过PCR各自扩增载体pTrex3gM-Fv43B和pTrex3gM-Fv43C的表达盒部分,以分别生成5.1kb和4.4kb的线性DNA片段,将它们用于PEG介导的MAD6菌株的原生质体融合转化(参见例如,Pentilla, M., 等人(1987)Gene(基因)61(2):155-164)。

[0181] 1061F:5'-GACCGGACGTGTTTTGCCCTTCAT-3'(SEQ ID NO:20)

[0182] 1085R:5'-GTGTGACCGGCTTTGGCGAGTG-3'(SEQ ID NO:21)

[0183] 为了制备原生质体,使用10mg/mL的来自哈茨木霉的裂解酶(西格玛公司(Sigma)目录号L1412)。在与转化DNA和PEG一起温育后,将原生质体加到冷却的具有0.5%尿苷的熔化山梨糖醇/乙酰胺琼脂。将板在30°C下温育。24小时后,将相等体积的补充有0.5%尿苷和1.2g/L 5-氟乳清酸(FOA)的相同培养基以覆盖层的形式加入板中。将板在30°C下温育一周。熔化的山梨糖醇/乙酰胺琼脂使用下述配方制备:

[0184] 山梨糖醇/乙酰胺琼脂

[0185] 第I部分

	乙酰胺	0.6g
	CsCl	1.68g
	葡萄糖	20g
[0186]	KH ₂ PO ₄	6g
	MgSO ₄ ·7H ₂ O	0.6g
	CaCl ₂ ·2H ₂ O	0.6g
	1000X 盐	1mL

[0187] 用milliQ H₂O定容到300mL

[0188] 第II部分

[0189] 山梨糖醇218g

[0190] 低熔点琼脂糖20g

[0191] 定容到700mL

[0192] 将第I部分和第II部分分别高压灭菌,随后合并。

[0193] 1000x盐(每L)

	FeSO ₄ ·7H ₂ O	5g
[0194]	MnSO ₄ ·H ₂ O	1.6g
	ZnSO ₄ ·7H ₂ O	1.4g
	CoCl ₂ ·6H ₂ O	1.0g

[0195] 过滤灭菌(0.22微米)

[0196] 5. 用fv43B或fv43C转化的里氏木霉MAD6克隆的SDS-PAGE

[0197] 如PCT公开WO 2011/038019中所述,使Fv43B的三个转化株和Fv43C的两个转化株在96孔微量滴定板中生长。将培养上清液在SDS-PAGE上跑胶,然后用Simply Blue染剂(英杰公司)进行考马斯蓝染色。通过密度测定法扫描并分析凝胶。使用ImageJ(来自美国国立卫生研究院)且通过采用Analyze Gel子菜单功能(如用户指南的小节27.13中所述),完成图像处理和条带亮度定量。图15显示了用fv43B和fv43C转化的里氏木霉MAD6克隆的SDS-PAGE。将对应于Fv43C蛋白质的条带定量并以占总蛋白质的百分比报告。Fv43C蛋白质在凝胶上跑出代表不同糖形的两个条带,将这两个条带汇总在密度测定法分析中。

[0198] 6. 所表达蛋白质的量的定量测量

[0199] 将quad缺失菌株所表达的蛋白质的量与通过MAD6菌株达到的量进行比较。如上所述,通过密度测定法扫描并分析相关凝胶,图13、14和15。使用ImageJ(来自美国国立卫生研究院)且通过采用Analyze Gel子菜单功能(如用户指南的小节27.13中所述),完成图像处理和条带亮度定量。将对应于每种目的蛋白的条带进行定量并以占总蛋白质的百分比报告。Fv43C在凝胶上跑出代表不同糖形的两个条带,将这两个条带汇总在密度测定法分析中。这个分析的结果汇总于下表2-1中:

[0200]

泳道	Fv43B/Quad缺失	泳道	Fv43C/Quad缺失	Fv43B/MAD6	Fv43C/MAD6
1	3%	1	21%	38%	
2	11%	2	30%	38%	
3	23%	3	27%	36%	
4	17%	4	2%		26%
5	7%	5	2%		26%
6	9%	6	21%		

7	3%	7	4%		
8	17%	8	40%		
M	(宿主)	9	28%		
M	(宿主)	10	19%		
		M	(宿主)		

[0201] 这个比较明确表明,使用MAD6菌株进行的表达导致可靠得多的表达,表达水平的可变性很低(例如小于20%可变性)。相比之下,使用quad缺失菌株,大部分的转化株未能表达目的蛋白,并且表达的可变性很大(例如大于50%可变性)。

[0202] 实施例3

[0203] 红褐肉座菌CBH2DNA文库的生成

[0204] 含有红褐肉座菌CBH2蛋白编码序列的pTTTpyrG-cbh2质粒(参见例如,PCT公开WO 2010/141779)用作参考序列,以用于产生编码CBH2变体酶的DNA文库。

[0205] SEQ ID NO:7示出参考红褐肉座菌CBH2编码DNA序列:

[0206]

ATGATTGTCGGCATTCTCACCACGCTGGCTACGCTGGCCACACTCGCAGCTAGTGTGCCTCTAGAGGAGCGGCAAGC
 TTGCTCAAGCGTCTGGGGCCAATGTGGTGGCCAGAAATGGTTCGGGTCGACTTGCTGTGCTTCCGGAAGCACATGCG
 TCTACTCCAACGACTATTACTCCCAGTGTCTCCCGGCGCTGCAAGCTCAAGCTCGTCCACGCGCGCCGCGTGCAGC
 ACTTCTCGAGTATCCCCACAACATCCCGGTCGAGCTCCGCGACGCCTCCACCTGGTTCTACTACTACCAGAGTACC
 TCCAGTCGGATCGGGAACCGCTACGTATTCAGGCAACCCCTTTGTGGGGTCACTCCTTGGGCCAATGCATATTACG
 CCTCTGAAGTTAGCAGCCTCGCTATTCCTAGCTTGACTGGAGCCATGGCCACTGCTGCAGCAGCTGTCGCAAAGGTT
 CCCTCTTTTATGTGGCTAGATACTCTTGACAAGACCCCTCTCATGGAGCAAACCTTGGCCGACATCCGCACCGCCAA
 CAAGAATGGCGGTAAGTATGCCGGACAGTTTGTGGTGTATGACTTGCCGGATCGCGATTGCGCTGCCCTTGCCTCGA
 ATGGCGAATACTCTATTGCCGATGGTGGCGTCGCCAAATATAAGAAGTATATCGACACCATTTCGTCAAATTGTCGTG
 GAATATTCCGATATCCGGACCCTCCTGGTTATTGAGCCTGACTCTCTTGCCAACCTGGTGACCAACCTCGGTACTCC
 AAAGTGTGCCAATGCTCAGTCAGCCTACCTTGAGTGCATCAACTACGCCGTCACACAGCTGAACCTTCAAATGTTG
 CGATGTATTTGGACGCTGGCCATGCAGGATGGCTTGGCTGGCCGGCAAACCAAGACCCGGCCGCTCAGCTATTTGCA
 AATGTTTACAAGAATGCATCGTCTCCGAGAGCTCTTCGCGGATTGGCAACCAATGTCGCCAACTACAACGGGTGGAA
 CATTACCAGCCCCCATCGTACACGCAAGGCAACGCTGTCTACAACGAGAAGCTGTACATCCACGCTATTGGACCTC
 TTCTTGCCAATCACGGCTGGTCCAACGCCTTCTTCATCACTGATCAAGGTCGATCGGAAAGCAGCCTACCGGACAG
 CAACAGTGGGGAGACTGGTGCAATGTGATCGGCACCGGATTTGGTATTCGCCATCCGCAAACACTGGGGACTCGTT
 GCTGGATTGTTTTGTCTGGGTCAAGCCAGGCGGAGTGTGACGGCACCAGCGACAGCAGTGCGCCACGATTTGACT
 CCCACTGTGCGCTCCCAGATGCCCTTGCAACCGGCGCCTCAAGCTGGTGCTTGGTTCCAAGCCTACTTTGTGCAGCTT
 CTCACAAACGCAAACCCATCGTTCCTGTAA.

[0207] SEQ ID NO:8是红褐肉座菌CBH2全长蛋白序列:

[0208] MIVGILTTLATLATLAASVPLEERQACSSVWQCGGQNWSGPTCCASGSTCVYSNDYYSQCLPGAASSS
 SSTRAASTTSRVSPPTSRSSSATPPPGSTTRVPPVSGTATYSGNPFVGVTPWANAYYASEVSSLAIPSLTGAMAT
 AAAAVAKVPSFMWLDLTKTPLMEQTLADIRTANKNGGNYAGQFVVYDLPDRDCAALASNGEYSIADGGVAKYKNYI
 DTIRQIVVEYSDIRTLVIEPDSLNLVLTNLGTPKCANAQSAYLECIYAVTQLNLPNVAMYLDAGHAGWLGWPANQ
 DPAAQLFANVYKNASSPRALRGLATNVANYNGWNI TSPPSYTGQNAVYNEKLYIHAIGPLLANHGWSNAFFITDQGR
 SGKQPTGQQQWGDWCNVI GTGFGIRPSANTGDSLLDSFVWVKPGGECGDTSDSSAPRFDSHCALPDALQPAPQAGAW

TQAYFVQLLTNANPSFL

[0209] SEQ ID NO:9是红褐肉座菌CBH2成熟蛋白序列:

[0210] QACSSVWGQCGGQNWGPTCCASGSTCVYSNDYYSQCLPGAASSSSSTRAASTTSRVSPPTSRSSSATP
PPGSTTRVPPVSGTATYSGNPFVGVTPWANAYYASEVSSLAIPSLTGAMATAAAVAKVPSFMWLDLTKTPLME
QTLADIRTKNGGNYAGQFVVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDTIRQIVVEYSDIRTLVIEPDSL
ANLVTNLGTPKCANAQSAYLECIYAVTQLNLPNVAMYLDAGHAGWLGWPANQDPAAQLFANVYKNASSPRALRGLA
TNVANYNGWNITSPPSYTGQNAVYNEKLYIHAIGPLLANHGWSNAFFITDQGRSGKQPTGQQQWGDWCNVI GTGFGI
RPSANTGDSLLDSFVWVKPGGECGTSDSSAPRFDSDHICALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

[0211] 合成的CBH2组合文库由GeneOracle(加利福尼亚州山景城(Mountain View,CA))制备。用多个其他氨基酸残基置换CBH2的多个氨基酸残基。表3-1列出了CBH2组合文库的成员的置换(编号根据CBH2成熟氨基酸序列)。

靶向位置	野生型残基	置换
98	P	P,Q,L
111	L	L,S
182	N	N,W
291	S	S,E
316	S	S,P
362	Q	Q,I,L
400	C	C,S

[0212]

[0213] 文库作为纯化的PCR产物提供,其中引物

[0214] GACCGGACGTGTTTTGCCCTTCAT(SEQ ID NO:10)和

[0215] GTGTGACCGGCTTTGGCGAGTG(SEQ ID NO:11)

[0216] 用于扩增cbh2基因,该基因上游侧接着约1.1kb的cbh1启动子且下游侧接着约1.85kb的amdS标记,以在红褐肉座菌宿主菌株的pyr2基因座中强迫整合。该表达盒同源重组到筛选菌株中的示意图在图8中示出。使用以上引物从pTTTpyrG-CBH2扩增的PCR片段(部分cbh1启动子、cbh2基因和部分amdS基因)的核苷酸序列作为SEQ ID NO:12提供如下:

[0217]

GACCGGACGTGTTTTGCCCTTCATTTGGAGAAAATAATGTCATTGCGATGTGTAATTTGCCTGCTTGACCGACTGGGG
CTGTTCTGAAGCCCGAATGTAGGATTGTTATCCGAACTCTGCTCGTAGAGGCATGTTGTGAATCTGTGTCGGGCAGGA
CACGCCTCGAAGGTTACGGCAAGGAAACCACCGATAGCAGTGCTAGTAGCAACCTGTAAAGCCGCAATGCAGCA
TCACTGGAAAATACAAACCAATGGCTAAAAAGTACATAAGTTAATGCCTAAAGAAGTCATATACCAGCGGCTAATAAT
TGTACAATCAAGTGGCTAAACGTACCGTAATTTGCCAACGGCTTGTGGGGTTCAGAAGCAACGGCAAAGCCCCACT
TCCCCACGTTTGTCTTCACTCAGTCCAATCTCAGCTGGTGATCCCCAATTGGGTCGCTGTTTGTTCGGTGAA
GTGAAAGAAGACAGAGGTAAGAATGTCTGACTCGGAGCGTTTGCATACAACCAAGGGCAGTGATGGAAGACAGTGA
AATGTTGACATTCAAGGAGTATTTAGCCAGGGATGCTTGAGTGTATCGTGTAAAGAGGTTTGTCTGCCGATACGACG
AATACTGTATAGTCACTTCTGATGAAGTGGTCCATATTGAAATGTAAAGTCGGCACTGAACAGGCAAAGATTGAGT
TGAAACTGCCTAAGATCTCGGGCCCTCGGGCCTTCGGCCTTTGGGTGTACATGTTTGTGCTCCGGGCAAATGCAAAG
TGTGGTAGGATCGAACACACTGCTGCCTTTACCAAGCAGCTGAGGATATGTGATAGGCAAATGTTTCAGGGGCCACTG
CATGGTTTCTGAATAGAAAGAGAAGCTTAGCCAAGAACAATAGCCGATAAAGATAGCCTCATTAAACGGAATGAGCTA

GTAGGCAAAGTCAGCGAATGTGTATATATAAAGGTTTCGAGGTCGGTGCCTCCCTCATGCTCTCCCATCTACTCATC
AACTCAGATCCTCCAGGAGACTTGTACACCATCTTTTGAGGCACAGAAACCAATAGTCAACCATCACAAGTTTGTA
CAAAAAAGCAGGCTCCGCGGCCGCCCTTCACCCACCATGATTGTCGGCATTCTCACCACGCTGGCTACGCTGGCC
ACACTCGCAGCTAGTGTGCCTCTAGAGGAGCGGCAAGCTTGCTCAAGCGTCTGGTAATTATGTGAACCCTCTCAAGA
GACCCAAATACTGAGATATGTCAAGGGGCAATGTGGTGGCCAGAATTGGTTCGGGTCCGACTTGCTGTGCTTCCGGA
AGCACATGCGTCTACTCCAACGACTATTACTCCCAGTGTCTTCCCGGCGCTGCAAGCTCAAGCTCGTCCACGCGCGC
CGCGTCGACGACTTCTCGAGTATCCCCACAACATCCCGGTCGAGCTCCGCGACGCCTCCACCTGGTTCTACTACTA
CCAGAGTACCTCCAGTCGGATCGGGAACCGCTACGTATTCAGGCAACCCTTTTGTTGGGGTCACTCCTTGGGCCAAT
GCATATTACGCTCTGAAGTTAGCAGCCTCGCTATTCCTAGCTTGACTGGAGCCATGGCCACTGCTGCAGCAGCTGT
CGCAAAGGTTCCCTCTTTTATGTGGCTGTAGGTCCTCCCGGAACCAAGGCAATCTGTTACTGAAGGCTCATCATTCA
CTGCAGAGATACTCTTGACAAGACCCCTCTCATGGAGCAAACCTTGGCCGACATCCGCACCGCCAACAAGAATGGCG
GTAACATATGCCGGACAGTTTGTGGTGTATGACTTGCCGGATCGCGATTGCGCTGCCCTTGCCCTCGAATGGCGAATAC
TCTATTGCCGATGGTGGCGTCGCCAAATATAAGAACTATATCGACACCATTTCGTCAAATTGTCGTGGAATATTCCGA
TATCCGGACCCCTCCTGGTTATTGGTATGAGTTAAACACCTGCCTCCCCCCCCCTTCCCTTCTTTCCCGCCGGCA
TCTTGTGCTTGTGCTAACTATTGTTCCCTCTCCAGAGCCTGACTCTCTTGCCAACCTGGTGACCAACCTCGGTACT
CCAAAGTGTGCCAATGCTCAGTCAGCCTACCTTGAGTGCATCAACTACGCCGTCACACAGCTGAACCTTCCAAATGT
TGCGATGTATTTGGACGCTGGCCATGCAGGATGGCTTGGCTGGCCGGCAAACCAAGACCCGGCCGCTCAGCTATTTG
CAAATGTTTACAAGAATGCATCGTCTCCGAGAGCTCTTCGCGGATTGGCAACCAATGTCGCCAACTACAACGGGTGG
AACATTACCAGCCCCCATCGTACACGCAAGGCAACGCTGTCTACAACGAGAAGCTGTACATCCACGCTATTGGACC
TCTTCTTGCCAATCACGGCTGGTCCAACGCCTTCTTCATCACTGATCAAGGTCGATCGGGAAAGCAGCCTACCGGAC
AGCAACAGTGGGGAGACTGGTGAATGTGATCGGCACCGGATTTGGTATTCGCCCATCCGCAAACACTGGGGACTCG
TTGCTGGATTGTTTTGTCTGGGTCAAGCCAGGCGGCGAGTGTGACGGCACCAGCGACAGCAGTGCGCCACGATTTGA
CTCCCCTGTGCGCTCCCAGATGCCCTTGAACCCGGCGCCTCAAGCTGGTGTGGTTCCAAGCCTACTTTGTGCAGC
TTCTCACAAACGCAAACCCATCGTTCCCTGTAAGGGTGGGCGCGCCGACCCAGCTTCTTGTACAAAGTGGTGATC
GCGCCGCGGCCAGCTCCGTGCGAAAAGCCTGACGCACCGGTAGATTCTTGGTGAGCCCGTATCATGACGGCGGGGG
AGCTACATGGCCCCGGGTGATTTATTTTTTTTGTATCTACTTCTGACCCTTTTCAAATATACGGTCAACTCATCTTT
CACTGGAGATGCGGCCTGCTTGGTATTGCGATGTTGTCAGCTTGGCAAATTGTGGCTTTCGAAAACACAAAACGATT
CCTTAGTAGCCATGCATTTTAAAGATAACGGAATAGAAGAAAGAGGAAATTAACAAAAACAAAACATCCCG
TTCATAACCCGTAGAATCGCCGCTCTTCGTGTATCCCAGTACCAGTTTATTTTGAATAGCTCGCCCGCTGGAGAGCA
TCCTGAATGCAAGTAACAACCGTAGAGGCTGACACGGCAGGTGTGCTAGGGAGCGTCGTGTTCTACAAGGCCAGAC
GTCTTCGCGGTTGATATATATGTATGTTTACTGCAGGCTGCTCAGCGACGACAGTCAAGTTCGCCCTCGCTGCTTG
TGCAATAATCGCAGTGGGGAAGCCACACCGTACTCCCATCTTTCAGTAAAGCTCTGTTGGTGTATATCAGCAATAC
ACGTAATTTAAACTCGTTAGCATGGGGCTGATAGCTTAATTACCGTTTACCAGTGCCATGGTTCTGCAGCTTTCCTT
GGCCCGTAAAATTCGGCGAAGCCAGCCAATCACCAGCTAGGCACCAGCTAAACCCCTATAATTAGTCTCTTATCAACA
CCATCCGCTCCCCGGGATCAATGAGGAGAATGAGGGGGATGCGGGGCTAAAGAAGCCTACATAACCCTCATGCCAA
CTCCCAGTTTACACTCGTCGAGCCAACATCCTGACTATAAGCTAACACAGAATGCCCTCAATCCTGGGAAGAAGTGGC
CGCTGATAAGCGCGCCCGCCTCGCAAAAACCATCCCTGATGAATGAAAAGTCCAGACGCTGCCTGCGGAAGACAGCG
TTATTGATTTCCCAAAGAAATCGGGGATCCTTTCAGAGGCCGAAGTCAAGATCACAGAGGCCCTCCGCTGCAGATCTT
GTGTCCAAGCTGGCGGCCGGAGAGTTGACCTCGGTGGAAGTTACGCTAGCATTCTGTAACGGGCAGCAATCGCCCA

GCAGTTAGTAGGGTCCCCTCTACCTCTCAGGGAGATGTAACAACGCCACCTTATGGGACTATCAAGCTGACGCTGGC
TTCTGTGCAGACAAACTGCGCCCACGAGTTCTTCCCTGACGCCGCTCTCGCGCAGGCAAGGGAACCTCGATGAATACT
ACGCAAAGCACAAAGAGACCCGTTGGTCCACTCCATGGCCTCCCCATCTCTCTCAAAGACCAGCTTCGAGTCAAGGTA
CACCGTTGCCCTAAGTCGTTAGATGTCCCTTTTTGTGTCAGCTAACATATGCCACCAGGGCTACGAAACATCAATGGG
CTACATCTCATGGCTAAACAAGTACGACGAAGGGGACTCGGTTCTGACAACCATGCTCCGCAAAGCCGGTGCCGTCT
TCTACGTCAAGACCTCTGTCCCGCAGACCCTGATGGTCTGCGAGACAGTCAACAACATCATCGGGCGCACCGTCAAC
CCACGCAACAAGAAGTGGTCGTGCGGCGGAGTTCTGGTGGTGAGGGTGCGATCGTTGGGATTCGTGGTGCGTCAT
CGGTGTAGGAACGGATATCGGTGGCTCGATTCGAGTGCCGGCCGCTTCAACTTCCTGTACGGTCTAAGGCCGAGTC
ATGGGCGGCTGCCGTATGCAAAGATGGCGAACAGCATGGAGGGTCAGGAGACGGTGCACAGCGTTGTGGGCGGATT
ACGCACTCTGTTGAGGGTGAGTCCTTCGCCTCTTCTTCTTTTCTGCTCTATACCAGGCCTCCACTGTCCTCCTTT
CTTGCTTTTTTATACTATATACGAGACCCGCGAGTCACTGATGAAGTATGTTAGACCTCCGCCTCTTCACCAAATCCGT
CCTCGGTCAGGAGCCATGGAAATACGACTCCAAGGTCATCCCCATGCCCTGGCGCCAGTCCGAGTCGGACATTATTG
CCTCCAAGATCAAGAACGGCGGGCTCAATATCGGCTACTACAACCTTCGACGGCAATGTCCTTCCACACCCTCCTATC
CTGCGCGGCGTGAAACCACCGTCGCCGCACTCGCCAAAGCCGGTCAACACC。

[0218] 如美国专利申请公开US 2006/0094080所描述,用线性DNA文库转化实施例1中所述的AD5红褐肉座菌菌株($\Delta eg11$ 、 $\Delta eg12$ 、 $\Delta cbh1$ 、 $\Delta cbh2$ 、 $\Delta bgl1$)的原生质体,并且在含有乙酰胺的选择性琼脂上在28℃下生长7天(0.6g/L乙酰胺、1.68g/L CsCl、20g/L葡萄糖、6g/L KH_2PO_4 、0.6g/L $MgSO_4 \cdot 7H_2O$ 、0.6g/L $CaCl_2 \cdot 2H_2O$ 、0.5g/L尿苷、微量元素盐、10g/L低熔点琼脂糖)。24小时后,将琼脂用补充有1.2g/L氟乳清酸(FOA)的选择性琼脂覆盖。将总共380个菌落转移至含有1.2g/L FOA的马铃薯右旋糖琼脂板,并在28℃下温育4-5天。将孢子转移至新鲜的马铃薯右旋糖琼脂板,将琼脂板在28℃下温育3天。

[0219] 作为另一种选择,可以采用实施例1中所述的MAD6菌株的原生质体代替AD5以用于变体文库成员的表达。同样地,其中另外的纤维素酶已被灭活的MAD6菌株衍生生物的原生质体也可以用于这个目的。这种衍生物将显示更少的本底纤维素酶活性。

[0220] 对于CBH2变体蛋白生产,使用96针复制器将孢子转移至在PVDF滤板中的补充有2%葡萄糖/槐糖混合物的200 μ L甘氨酸基本培养基:6.0g/L甘氨酸、4.7g/L $(NH_4)_2SO_4$;5.0g/L KH_2PO_4 ;1.0g/L $MgSO_4 \cdot 7H_2O$;33.0g/L PIPPS;pH 5.5;并无菌添加作为碳源的2%葡萄糖/槐糖混合物、10ml/L的100g/L的 $CaCl_2$ 、2.5ml/L里氏木霉微量元素(400X):175g/L无水柠檬酸;200g/L $FeSO_4 \cdot 7H_2O$;16g/L $ZnSO_4 \cdot 7H_2O$;3.2g/L $CuSO_4 \cdot 5H_2O$;1.4g/L $MnSO_4 \cdot H_2O$;0.8g/L H_3BO_3 。使每种CBH2变体一式四份地生长。在用透氧膜将板密封后,将板在28℃下温育6天,同时以200rpm振荡。通过在低压下将培养基转移至微量滴定板收获上清液。

[0221] 针对目的性质测试CBH2变体。使用SDS PAGE检查各个变体的表达。图16A是四个SDS-PAGE的照片,显示多个变体的表达;图16B示出了这些变体的平均生产水平,其中误差条指示表达的可变性。测定了对于洗涤的稀酸预处理的玉米秸秆(PCS 50℃)、对于在50℃下的玉米穗轴(CC 50℃)和对于在57℃下的玉米穗轴(CC 57℃)的比活性。分离了对57℃玉米穗轴显示改善的活性的总共十个变体。分离这些菌株的基因组DNA并测定其cbh2基因序列。组合文库变体的置换以及对玉米穗轴和玉米秸秆的比活性的性能指数在表3-2中显示。基于变体的归一化蛋白质表达水平测定比活性的性能指数。

[0222]

变体	CC 57℃	CC 50℃	PCS 50℃
CBH2.S291E/Q362I	1.2	0.97	1.13
CBH2.P98Q/S316P/Q362L/L439P	1.52	0.97	0.1
CBH2.P98Q/N182W/S291E/S316P/C400S	1.26	0.97	0.75
CBH2.P98Q/N182W/S291E/S316P/C400S	1.24	0.97	1.02
CBH2.P98L/N182W/S291E/Q362I/C400S	1.26	0.97	1.05
CBH2.P98L/N182W/S291E/Q362I/C400S	1.28	0.97	1.14
CBH2.T74S/P98L/N182W/S291E	1.31	0.97	1.13
CBH2.P98L/N182W/S291E/S316P/Q362I	1.26	0.97	1.22
CBH2.N182W/S291E/Q362L/C400S	1.3	0.97	1.08
CBH2.S291E/Q362L	1.29	0.97	1.18

[0223] PCS 50℃。如所描述(Schell等人, J Appl Biochem Biotechnol(应用微生物学与生物技术期刊), 105:69-86, 2003年), 将玉米秸秆用2%w/w H₂SO₄预处理, 然后用去离子水多次洗涤, 以获得具有pH 4.5的糊。然后加入乙酸钠缓冲液(pH 5.0)至50mM乙酸钠的最终浓度, 并且如果适当的话, 使用1N NaOH将这个混合物滴定至pH5.0。反应混合物中的纤维素浓度为大约7%。向96孔微量滴定板(Nunc Flat Bottom PS)中每孔加入65μL的这个纤维素悬浮液。向每个孔加入10μL酶样品, 每个样品含有在来自quad缺失菌株(Δegl1、Δegl2、Δcbh1、Δcbh2)的上清液中的49μg蛋白质。

[0224] 加入最多20μL来自表达野生型CBH2或CBH2变体的红褐肉座菌细胞的培养上清液。加入补偿体积的乙酸盐缓冲液以补足在总体积方面的差异。将板密封后, 将其在50℃下温育, 同时在200rpm震荡。2天后, 将板置于冰上5分钟并加入100μL 100mM甘氨酸pH 10.0。在混合后, 将板以3000rpm离心5分钟。将体积10μL上清液在190μL水中稀释。将10μL稀释溶液转移至新的96孔微量滴定板(Costar Flat Bottom PS), 在每个孔中含有100μLABTS葡萄糖测定混合物(2.74mg/mL 2,2'连氨基-双-(3-乙基苯并-噻唑啉-6-磺酸)、1U/mL辣根过氧化物酶VI型、1U/mL葡萄糖氧化酶), 并在微量滴定板分光光度计(Spectramax Plus 384, 分子仪器公司(Molecular Devices))中记录A420的增加。在每块板上包括一系列葡萄糖浓度作为标准品(0; 0.008; 0.016; 0.031; 0.063; 0.125; 0.25; 0.5; 1mg/mL)。测定一式两份地进行。通过使数据与Temkin等温方程($y = a + b(\ln(1 + c \cdot x))$)拟合生成野生型CBH2的剂量响应曲线, 并将CBH2变体的活性除以相同板的野生型CBH2的计算活性, 以获得性能指数。

[0225] 玉米穗轴50℃。将玉米穗轴磨碎以通过0.9mm筛, 然后依照PCT公开W0 2006/110901的实施例4中所述方法进行预处理。将经预处理的玉米穗轴作为在50mM乙酸钠pH 5.0中的7%纤维素悬浮液使用。向96孔微量滴定板(Nunc Flat Bottom PS)中每孔加入70μL的悬浮液。向每个孔中加入10μL溶液, 其含有来自quad缺失菌株(Δegl1、Δegl2、Δcbh1、Δcbh2)的上清液的46.55μg蛋白质, 补充有4.90μg里氏木霉CBH1、6.84μg里氏木霉Xyn2Y5(Xiong等人, Extremophiles(极端微生物)8:393-400, 2004年)、2.28μg拟轮生镰刀菌(Fv)51A、5.32μg Fv3A、0.76μg Fv43D, 和2.45μg里氏木霉BGL1。拟轮生镰刀菌的酶已在PCT公开W0/2011/038019中描述。

[0226] 将最多20μL来自表达野生型CBH2或CBH2变体的红褐肉座菌细胞的上清液加入每个孔中。加入补偿体积的乙酸盐缓冲液以补足在总体积方面的差异。将板在50℃下温育, 同

时在200rpm震荡。2天后,将板置于冰上5分钟并加入100 μ L 100mM甘氨酸pH 10.0。在混合后,将板以3000rpm离心5分钟。将体积10 μ L上清液在190 μ L水中稀释。将10 μ L稀释溶液转移至含有100 μ L ABTS葡萄糖测定混合物的新的96孔微量滴定板(Costar Flat Bottom PS),并如上所述进行测定和分析。

[0227] 玉米穗轴57 $^{\circ}$ C。将玉米穗轴磨碎以通过0.9mm筛,然后依照PCT公开W0 2006/110901的实施例4中所述方法进行预处理。将经预处理的玉米穗轴作为在50mM乙酸钠pH 5.0中的7%纤维素悬浮液使用。向96孔微量滴定板(Nunc Flat Bottom PS)中每孔加入70 μ L的悬浮液。向每个孔中加入10 μ L溶液,其含有来自quad缺失菌株(Δ egl1、 Δ egl2、 Δ cbh1、 Δ cbh2)的上清液的46.55 μ g蛋白质、4.90 μ g CBH1变体(S8P/T41I/N89D/S92T/S113N/S196T/P227L/D249K/T255P/S278P/E295K/T296P/T332Y/V403D/S411F/T462I)、6.84 μ g里氏木霉Xyn2Y5(Xiong等人,Extremophiles(极端微生物)8:393-400,2004年)、2.28 μ g Fv51A、5.32 μ g Fv3A、0.76 μ g Fv43D、2.45 μ g埃默森篮状菌(*Talaromyces emersonii*) β -葡萄糖苷酶。加入最多20 μ L来自表达野生型CBH2或CBH2变体的红褐肉座菌细胞的上清液。加入补偿体积的乙酸盐缓冲液以补足在总体积方面的差异。将板在57 $^{\circ}$ C下温育,同时在200rpm震荡。在2天后,将板置于冰上5分钟并加入100 μ L 100mM甘氨酸pH 10.0。在混合后,将板以3000rpm离心5分钟。将体积10 μ L上清液在190 μ L水中稀释。将10 μ L稀释溶液转移至含有100 μ L ABTS葡萄糖测定混合物的新的96孔微量滴定板(Costar Flat Bottom PS),并如上所述进行测定和分析。

序列表

- <110> 美国丹尼斯克有限公司
B·S·鲍尔
T·卡佩尔
B·R·凯莱门
- <120> 丝状真菌宿主菌株和 DNA 构建体以及它们的使用方法
- <130> 31551W0-2
- <140> PCT/US2011/039092
- <141> 2011-06-03
- [0001]
- <150> US 61/351,286
- <151> 2010-06-03
- <160> 21
- <170> PatentIn version 3.5
- <210> 1
- <211> 7685
- <212> DNA
- <213> 人工序列
- <220>

<223> 合成的Ku80敲除基因

<400> I

```

ggccgcctca acaccacac tcgaggcaca cgagttcate ggcggettec cccacaagct      60
ctcggccaac ctgctaccgg ctctctcgcg agacttccca aagcctacaa acgaggtcga      120
cgtcaaggag gccctcgagc gccagcccgg cagatggagc ciccagggcc agatcaaggc      180
caacaacatg agagcccaga ggcgcgact ccggctcgac gacaaggagg gcaaggcgag      240
agcctttgag gaggccaagc gcgagctact ggcgtatcac cacagcgccc tgcggaagcc      300
ttccggcgca agataatgag cttgatcgca atgacgagtt cacgtacget ttgccatatt      360
gttgtttgctt ttigtttggt cctacatgta cggcgcaattg gttgggagga tatacccacg      420
gagagtgtcc gagtggcttc tgggatttag agcgtcaatta gcaggataga gatggttggc      480
caggggaatg gaattgactt tteactacaa ggaacttgtt cactctggtg ttgattccca      540
ttgcgtgact ggtagtaggg aggaatgett ttactttgtg ccactagacc gcagagaagg      600
[0002] gttggttgca agcggggctc gtgtataccg accaagagtg atgggcatac agcaacgttt      660
ctgaacgaact tcattttgtc cgagtctact ggatgcgaga tgcacagctg aagccgtacg      720
ccaccagggc gacgaactcg acaaggttga cgagggagga gatgccgtgc agcatgccaa      780
actttttggt gagggcacgc atctcatccg actgtgcate cttgtcatac cactccttc      840
cgtctcgctt ggctggtggg agggttcaac aaatecatcg tcagccatcc ggggtctcaa      900
atcaatggcg tgcattcgga gtcgggcttg aggctaacct tgiccatggc ggtccttcat      960
ggtcttgaca gtggcgggaa gcagcacggc gaggttgacg aggccgctga cgaacatggt      1020
tgcgatgggc accaaggagc tccacttggt gggagcgtcg acgaggccgc cgatgccgcc      1080
cttgatgccc aagagggcgt ttccggggaa cgtgagggcg agcagcgcgg ggatggccgt      1140
ctgcatgcca aagtagatgg ggaacagctt gctctggatg gcggagaagg agggccggct      1200
gacggtgcgg aacatgacga tgccgttgac gaaggactgc agtagcgtag tgtgatggta      1260
agcagctggc cggcgcgct gagacaatgg ccggcaatgg taaaaggac caagatgtac      1320
taggtagttg caatgtggtt tattacctac ctactacctg gtaggcacct actaggtact      1380
tgggtagacg gacaatgaaa ttggaagtcg gggttgacag aaagcagggc gctggacaca      1440

```

	ttgtgcttca ggcggtacct gtcgtcatcg tcagccaatg tcgaggcccg gcagcccag	1500
	gagcgagaca accttggccg gaggagcccg caggtaacct ccaaagcgcg gctggtaacct	1560
	etcaaccctc tcaggcctgt tggatgccct atgacatgcc ctgggggatg cagctgttgc	1620
	cccgceccc cacttiteggg tgaccgcgag gctgctgatt ggctggttgc cacgggctgg	1680
	gcggtccttg aagtgtttgc catctgaact ctgtcggcgc tggcgtcggc tgcgcccatt	1740
	gggaggcgag acaactcagg gtactagaat cactgacaga agaagagaat cgaaagiagg	1800
	tagacagcca attcgttgca tggcaggcaa ccgcacagga gaaaaattga ctaccccaca	1860
	atcaggcaca gtaagtaggg cacagtacgt atgtacagac aaggcgcag cgatactgcg	1920
	cgacccggtc cctcgcggc ttgacaactg cgacaggcta ctttactagt attcgcagcg	1980
	gcgggtcggc cattattaca tgtactgtgc cgccatttga tgactgggct gctgcagtat	2040
	tagtagatct gcccggeatc gcccttccat gggcgcgacc cgggactgga ccctctgact	2100
	ctaectacat gtacctaggc cgggcggggc ttggtgactt ttgtccgac aggtcgttcg	2160
	cctggctacc tattatttct ctttcttctt ctccatcctg cttctggcct tgcaattctt	2220
[0003]	cttcgccact cctccctctt cccccgcga tacctttgaa ttctgcagag aggaaaagac	2280
	gagaaaaaaaa agggcagcag agacgtcggc ctggctcagc tgcctcatct ctgcgcactc	2340
	tcattttttt tattgtccga cccctccctc aacctctctc ttctgtgaca ggctaagcct	2400
	tgtctgcagc ctctctcttt gaattttctt acttctacct tcttttcttg cgtgttacc	2460
	accatagctc gattcagcat gctccgaagt cgccaagtca cagccagggc cgtccgggct	2520
	ctgggccagg cgcgcgcctt tacctcagc accaagcctg tcatgateca ggcagccag	2580
	aggaaacagg ccaacgccag cgtctcctcg taagtcgcc attgccattg catctctctg	2640
	ttgatataata ctctctctg cttgcgtggc gtcgtctctc ggttatgcgt gtcaaggacc	2700
	agggtgtgct gcategtggt tttccagcgc cgattaccgg gggacgaatt ttggctgct	2760
	caactcgcgc gcgcgcattc tgattctctg ttttcaatct tgagcgacaa ctggctaaca	2820
	taatggccat tggcaattgc ttacacacaga caagtcgcc ctgtaccgag ccctgctttc	2880
	aacgtgaag acaaagaccg cagccatgtg cagcctctgg tcaaccctc gaagcccagc	2940
	atggatgaat cgtatgtcca cgtccctctg tcccgccta caaatgaac acgattacac	3000
	cagaattttt gcaacaatcg acacttctat aacagaccaa ttgactttg ttctgaccaa	3060

	tcattgttgc	ctagattcat	tggcaaaacc	ggaggcgaaa	tcttccacga	gatgatgctg	3120
	cgacaggggtg	tcaagcacat	ttgtagggtc	cgatgccggc	cgcccacacg	ggetccatcc	3180
	ttgctccatc	tctccagcta	ggcaaatctc	gctaaccttg	agtcaccate	cagtcggata	3240
	ccctggcggc	gctatcctgc	ccgtcttcga	cgccatctac	aactcaaac	acttcgactt	3300
	catectgccc	cgctcatgagc	agggagctgg	ccatatggcc	gagggetatg	cccgtgcctc	3360
	gggcaaaccc	ggtgtcgtcc	tggtgacttc	cgccccgggt	gctaccaatg	tcatacagcc	3420
	catgcaggat	gccctgtcgg	acggaacgce	cttggctcgtc	ttctgcggcc	aggtocccac	3480
	cacggccatc	ggcagcgaatg	acttccaaga	ggccgacgtc	gtgggcatct	cgcgggcctg	3540
	caccaagtgg	aacgctcatgg	tcaagagcgt	tgctgagctg	ccgcggagaa	tcaacgagge	3600
	ctttgagatt	gccaccagcg	gccgccctgg	ccccgtcctc	gtcgacctgc	ccaaggatgt	3660
	cacggctggt	atcctgagga	gagccatccc	tacggagact	gctctgccgt	ctctgcccag	3720
	tgccgcctcc	cgcgccgcca	tggagctgag	ctccaagcag	ctcaacgcct	ccataacgcg	3780
	tgccgccgac	ctcatcaaca	tcgccaagaa	gcccgtcacc	tacgccggtc	agggtgctcat	3840
[0004]	ccagtcagag	ggcggcgctt	agctcctgaa	gcagctggcg	gacaaggcct	ccatccccgt	3900
	caccaccacc	ctccatggcc	tgggtgcctt	tgatgagctg	gacgagaagt	cgtctgcat	3960
	gctgggcatg	cacggctcgg	cgtaigccaa	catggccatg	cagcagggcg	acctcatcat	4020
	cgccctcggc	agccgattcg	acgaccgtgt	tactctgaat	gtctccaaat	ttgcgcctgc	4080
	agccaggcaa	gctgctgcgg	agggcgcggg	cggcatcatt	cactttgaga	tcatagcccc	4140
	gaaatcaaac	aaggctcatcc	aggcgaccga	ggccgtcgag	ggcgacgtcg	ccaccaacct	4200
	gaagcacctc	attccccaga	ttgccgaaaa	gtccatggcg	gaccgaggag	agtggttcgg	4260
	cctcatcaat	gagtggaaga	agaagtggcc	cctgtcaaac	taccagcgcg	cggagcgggc	4320
	tggcctcatc	aagccgcaga	cggctcatgga	ggagattagc	aacctgacgg	ccaaccgaaa	4380
	ggacaagacg	tacattgcca	cgggtgtcgg	ccagcaccag	atgtggggtg	cccagcactt	4440
	ccgctggagg	caccctcgat	ccatgattac	ctctgggtgt	ctgggcacca	tgggetacgg	4500
	tctccccgcg	gccattggcg	ccaaggtggc	ccagcccagc	gtctctgtaa	ttgacgttga	4560
	tggcgatgcc	tcgtttaaca	tgacgtgac	ggagctgtcg	actgctgcac	agttcaacat	4620
	tggcgtcaag	gtggttgtgc	tcaacaacga	ggagcagggc	atggtgacgc	agtggcagaa	4680

	cctcttttac gaggaccgat atgcccacac gcaccagaag aaccccgaact tcatgaagct	4740
	ggccgacgcc atgggcgttc agcaccageg cgtgacggag cgggagaagc tggtegatgc	4800
	cctgacgtgg ctgatcaaca ccgatggccc ggcccigtig gaggttgta cggacaagaa	4860
	ggtgcctgic ctgcccattg tgcccgcgg atcggccctg cagcagttcc tcgtctttga	4920
	acctggtgag tctacttcag acatattget tgcgcattgc agatactaac actctcacag	4980
	aaaaggataa gcagcgcctg gagctgatga aggagagaac aaagggtgtg cactcctaaa	5040
	gcgatgatgt ctgaggggg ttcttcgttg aacctagtt caggeacccat cttacctct	5100
	tattttttcc cgtgggcttt cattttgtgt catccgagca tgacgttga gsgttggagt	5160
	ttcttcttt ttatctgtc attactggt acccataggc gcgagactag gettccatgt	5220
	ttgttttgc gactttcaaa aagtactttt agtggtttg ggacgacga gggggggcaa	5280
	cctcttctgt cgaaaaaggt gctggatgg atgagatgag atgagatgag ggtgaagata	5340
	gatacctgea gtgtttttga cgcgacggga tggcgatgc agcacccecg acagaactcg	5400
	tcgagactgt gcagcctcat atcgatgcac tgattcacgc tgcagacgtg aagaaagfta	5460
[0005]	ctgattccat tacatatget tcctgcaca ctgatgtttg atttgtgcta acgccccct	5520
	tagtgccgcc caaggccaag ggcaagcgc aaagagaaac agttaaccc atctcgggac	5580
	tggatgtgga tgcccctctg ggagaagagc agaaagttc cattagtccg gagaatgcca	5640
	ttccggactt caaacgagcc ctcaactcgt ccgaagaagt cgagcagatt gccgacgcca	5700
	caaaacaaat gggggccatt gtgcggtctc tcattacgga cagcttcggg gatagcaaat	5760
	atgcccaggc aatggaaggc attggtgcca tgcgtgagga gctgatcaac ctggaagagc	5820
	ctggcctgia caacgacttt gtgcgcgact tgaagaaaag tttgctatct ggagccttg	5880
	gtggtgacag gcgagatttc tggttcaaga tgaggtgggc gaagctgggc ctgattgaca	5940
	agaaacagtc ggagggtctc tcggtcactc ttgaggagc ggacgaggtg agtgggtcag	6000
	catgctgtcg gattatacgg acgttgtttg ctaactttg ggatagtttt acaagtcgag	6060
	gtgaggtatc tacgttgacc aagaatggga ccatttatat gagcgggtgta acaacagaat	6120
	cctgtgcitt gagcattgta tgatatgatt attgatgaac cggacaaaag ggggtagggg	6180
	attgatgcca tcacgaccga ttgaccagac ctggattctc gcacagcatg gctgctgatt	6240
	ttgttgacct tgcgacgtaa catcccigaa gaacaacctt ctattaacct atcatttagc	6300

	agaagctctg taaccttctt gattettgta ttcagettct gagtcigtca aatgtaatca	6360
	tttcgagggt gtgtaattcc ggccaagcag ggggccgtct gccagcgcct gcctaggetg	6420
	caccgcaatc tgeccaatca getgcccctc agtttctgtt gaccttgacg ctgcccctca	6480
	tcctttatct gcacacaatt cttttctctc tgctctgcgc attcttctct ctctcgtctc	6540
	ccttctcaag ctcaacttca cctcatccgc tccaactaaa gccctcccgt cgtcgtctcg	6600
	catcctcacc tcgactgagg ccagcaaac aagcaaagcc gtgactgac ctcagcatgg	6660
	ctaccttcaa cctcaccgtc cgcctggaga tgctcaaaga aattggaatc accgtccaat	6720
	acggcgagca tgtagcgaac gaagcagcca gcaacgaagc agcagatggca ttcgaagaag	6780
	aagaagagtt ccccgccgtt gtcccgcca aggcagaaca gcacgcctct gaacacgacg	6840
	ctggccacga tgcttgggac gcggctgccc acatctcgac ttggcgcaa gaacagcaga	6900
	agccccagga gatggacgac tegtctatcg tgatgccgct ggactactcc aagtttctcg	6960
	ttggagagcc tgcggacgaa tccatcagct tttgctcgtg gaaggctcgc gaggettatc	7020
	ctgaccagtt tatcggaag gcaaacagcc ctctgtatg tagcagattgc tttctctgca	7080
[0006]	ttatgggaat ctcaagagag tatggtagaa gataactgac aacttgcagg ccaagccgta	7140
	ctttgacaag attttgggaag acagagtctg ggatttctga ggatcttgat tgatgtgeat	7200
	atggcgacat gcctgctaata acattgttag cttctatctc tacaaccctg agaagcctc	7260
	agagaagcct cgcgtgctgg tgcccactgt tcagctcgaa ggettictca aaagcatcaa	7320
	cagagcgtc ggtacttctc tcaccattcc aggaggggca aaccaggacc gtttttatct	7380
	gaggttcggc cagggagaca cccaaggcc tcgatatcta cagaggtcga gagaccagaa	7440
	atcccfaaag attgaaacgt tccccgattt tcaacaggcg gactacgaca gcttiaggaa	7500
	cgcgcacggc gccatccagg aggactgggt gaagaactgg cagatgctgg tacctcggcc	7560
	gagtttcgac aagaagaaa atgcagacaa aagagcagcc aagagaaggc tcgagcgaga	7620
	gogaatgctt cacaatacgc aggaatttct tcaittggca ggtaagggca aaggggctga	7680
	cgtgg	7685

<210> 2

<211> 9259

<212> DNA

<213> 人工序列

<220>

<223> 合成的敲除盒

<400> 2

atcacgcect cgataaaaag accctcaaga gtccatgtgc cctatctgcc tgatcttct 60
 aacccttatt taacattggc cctatcacia cctagttctt cttcagcctg ctttgtaac 120
 acttgtaacg gttcaactca acgtaatcag caggtagcag gacgaggata gggagagaaa 180
 cgaagagaag aagaggagag aggaagaaaa aaaaaagaaa agaagaaaa agggaaaagg 240
 aaagaaggag gaaaagagaa gaaagtcaga tgaagaagca agaagacgcc atggtagcca 300
 ccgcttgta gggctcctta gcaacgaaca actctagctt ggggacttgt cgatgtgtcg 360
 [0007] ttctcttct acccatcagc accaacgatg agttcgatat agacgaggac ctcatggaag 420
 tagagacat tgggttcgac aggatctctc agtttcaett ctatgaggtc tgcgctcgg 480
 atgacttttt gaggagcttc ccttctgct tcaaccctaa actctcttct ctgaaaccgc 540
 agcacgtttg cacggcctg ttgctggagc agtttgettt cgagcactct cagcgtggtt 600
 tcagcagccc actggtagt ggctccttt gacgtccaca ccttctctct gtcgcatgag 660
 tatctggtgg gaacgactgc tccaaggagg attgctaacg aggtttagg ccgaatatag 720
 catcagattc tceggtaacc ttagctacgg cctcttcaac atctgtgaca tgacggagag 780
 caagtactgg tggttggcga ccaagatgag cggctggaac ategacggct gccccgaaga 840
 cgtcaggaga ctcatgtttg ttacatcat cgccacctg ggatgcagcc ccgtcgtgac 900
 ggatgaagac atggactacc ccaagaactg ggcggcaatt ctccacgta gagacagata 960
 tccgagtga cctgtgggccc accggcctca tgggcgcacc atctgcctcc actcgggtggc 1020
 cgtctgccct cgtctccagg gcttgggtct cggtagctgc actctgaagt cgtatgtgca 1080
 gcgcatgaac agcctcggcg ccgaggaccg tgttgcctct gtttgcgca agcccagac 1140
 gagatttttt gaaagatgag gcttcaggaa cagcggcccg agtagtatca agactctggt 1200

	eggcgaatac tacaacatgg tgtgtgcttc cacatcgact tggccagact ctatacgatt	1260
	ttcaaaocctc gctatacgtc atattgactt gtttcttttag gtcttegatt tgcccgggce	1320
	caaagacttt atcgactgga atagcattgc cgacgetgcc aagaagatgt gaaccatttg	1380
	actgatacga tgtgtgctac gcatgtcgac cttctttggtt tgtttctttg gcggetcttt	1440
	gtataccttg ggacacggca gacgcatgic taigtgaaga aaacgittac ggcgctgttt	1500
	gcatcaggaa taigalcatt aaacatggag cgtaatggta ttaatgatca actagaaaaa	1560
	tggataggaa gggcgagagg gcgatcaaca aagcagcccg gggcatagtc tggaagcagc	1620
	aggaattgga agggaaaagg aagctgcaca atgaagggat atcgtgagcg gagtggctca	1680
	cgagagtatc aacagactgg cgaagcaag caattccaa cgccggctat taggccataa	1740
	gatggcctgt tgtgagtccc agttgcacgt atcccatac gactgctctg tcgctgactt	1800
	gaaaaaaaaat agggaggata aaggagaaag aaagtgagac aaccctgtag ggacttgggg	1860
	tagtaggaga acacatgggc aaccgggcaa tacacgcgat gtgagacgag ttcaacggcg	1920
	aatggaaaat cttagaaaac aaaataaaat aactgcctc catacgggta tcaaattcaa	1980
[0008]	gcagttgtac ggagctagc tagagttgtg aagtcggtaa tcccgtgta tagtaatac	2040
	agtcgcactc aaatactccg aagctgctgc gaaccgggag aatcgagatg tgetgaaag	2100
	cttctagcga gcggciaaat tagcatgaaa ggctatgaga aattctggag acggcttgtt	2160
	gaatcatggc gttccattct tcgacaagca aagcgttccg tcgcagtage aggcactcat	2220
	tcccgaaaaa actcggagat tctaagtag cgatggaacc ggaataatat aataggcaat	2280
	acattgagtt gcctcgacgg ttgcaatgca ggggtactga gcttggacal aactgttccg	2340
	taccccacct cttctcaacc ttggcggtt ccctgattca gcttaccctt acaagtegta	2400
	atcactatta acccagactg accggacgtg ttttgcctt catttgaga aataatgta	2460
	ttgcgatgtg taatttgctt gcttgaccga ctggggctgt tcgaagcccg aatgtaggat	2520
	tgttatccga actctgctcg tagaggcatg ttgtgaatct gtgtcgggca ggacacgect	2580
	cgaaggttca cggcaaggga aaccaccgat agcagtgtct agtagcaacc tgtaaagccg	2640
	caatgcagea tcaactgaaa atacaaacca atggctaaaa gtacataagt taatgcctaa	2700
	agaagtcata taccageggc taataattgt acaatcaagt ggctaaacgt accgtaattt	2760
	gccaacggct tgtggggttg cagaagcaac ggcaaagccc cactteccca cgtttgtttc	2820

	ttcactcagt ccaatctcag ctggtgatcc cccaattggg tcgcttgitt gttccgggta	2880
	agtgaagaa gacagaggta agaatgtctg actcggagcg ttttgatac aaccaagggc	2940
	agtgatggaa gacagtgaaa tgttgacatt caaggagtat ttagccaggg atgctigagt	3000
	gtatcgtgia aggaggtttg tctgccgata cgacgaatac tgtatagtea cttctgatga	3060
	agtgggccat attgaaatgt aaagtggca ctgaacaggc aaaagattga gttgaaactg	3120
	cctaagatct cgggccctcg ggccttcggc ctttgsgtgt acatgtttgt gctccgggca	3180
	aatgcaaagt gtggtaggat cgaacacact gctgccttta ccaagcaget gagggtatgt	3240
	gataggcaaa tgttcagggg ccaactgcatg gtttcgaata gaaagagaag cttagccaag	3300
	aacaatagcc gataaagata gcctcattaa acggaatgag ctagtaggca aagtcagcga	3360
	atgtgtatat ataaaggttc gaggtcctg cctccctcat gctctccca tctactcatc	3420
	aactcagatc ctccaggaga cttgtacac atcttttgag gcacagaaac ccaatagtca	3480
	accgcggaact gcgcacatg tatcgggaagt tggccgtcat ctggccttc ttggccaac	3540
	ctcgtgctag actagcggcg ccaggaagcc cggaagtaa gtggattctt cggcgtggct	3600
[0009]	ggagcaaccg gtggattcca gcgtctcga cttggactga geaattcagc gtcacggatt	3660
	cacgatagac agctcagacc gctccacgge tgccggcatt attggttaac ccgaaactc	3720
	agtctcttg gccccgtccc gaagggacce gacttaccag gctgggaaag ccagggatag	3780
	aatacactgt accggcttcg tacgggaggt tcggcgtagg gttgttccca agttttacac	3840
	accccccaag acagctagcg cacgaaagac gcggagggtt tggtgaaaaa agggcgaaaa	3900
	ttaagcggga gacgiattta ggtgctaggg ccggttctct cccatltttt cttcggttcc	3960
	ctttctctcc tggaagactt tctctctctc tcttctctc ttcttccatc ctcagttccat	4020
	cttctttcc catcatccat ctctcactt ccctctcaac tccatcacat cacaatcgat	4080
	atgaaaaagc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga aaagttcgac	4140
	agcgtctcgg accgatgca gctctcggag ggcgaagaat ctctgctttt cagcttcgat	4200
	gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt ctacaaagat	4260
	cgttatgttt atcggcaactt tgcacggcc gcgctccga ttcgggaagt gcttgacatt	4320
	ggggaattca gcgagagcct gacctattgc atctccgcc gtgcacaggg tgtcacgttg	4380
	caagacctgc ctgaaaccga actgcccgtt gttctgcagc cggctcggga ggccatggat	4440

	gcgatcgetg eggecgatct tagccagaac agcgggttcg gcccattcgg accgcaagga	4500
	atcgggcaat acactacatg gcgigatttc atatgcgcga ttgctgatcc ccatgtgtat	4560
	cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgcagge tctcgatgag	4620
	ctgatgcctt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc ggatttcggc	4680
	tccaacaatg tctgacgga caatggccgc ataacagcgg tcattgactg gagegaggcg	4740
	atgttcgggg atlcccaata cgaggctgcc aacatcttct tetggaggcc gtggttgct	4800
	igtatggagc agcagacgcg ctacttcgag cggaggcacc cggagcttc aggatcggc	4860
	cggctccggg cgtatatgct ccgcatttgt cttgaccaac tctatcagag cttggttgac	4920
	ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt ccgatccgga	4980
	gccgggactg tcgggcgtac acaaatcgc cgcagaagcg cggecgtctg gaccgatggc	5040
	tgtgtagaag tactcgccga tagtggaac cgacgcccc gactcgtcc gagggcaag	5100
	gaatagagta gatgcgacc gggatccact taacgttact gaaatcatca aacagcttga	5160
	cgaatctgga tataagatcg ttgggtcga tctcagctcc ggagtgaga caaatgtgt	5220
[0010]	tcaggatctc gataagatac gtccatttgt ccaagcagca aagagtgcct tctagtatt	5280
	taatagctcc atgtcaacaa gaataaaacg cgtttcgggt ttacctcttc cagatacagc	5340
	tcatctgcaa tgeattaatg cattggacct cgeaaccta gtacgcectt caggctccgg	5400
	cgaagcagaa gaatagctta gcagagtcta ttttcatttt cgggagacta gcattctgta	5460
	aacgggcagc aatcgcccag cagttagtag ggccccctct acctctcagg gagatgtaac	5520
	aacgccacct tatgggacta tcaagctgac gctggcttct gtgcagaaa actgcgcccc	5580
	cgagttcttc cctgacgccc ctctcgcgca ggcaaggaa ctcgatgaat actacgcaaa	5640
	gcacaagaga cccgttggtc cactccatgg cctccccatc tetctcaaag accagcttcg	5700
	agtcaaggta caccgttccc cetaagtcgt tagatgtccc tttttgtcag ctaacatag	5760
	ccaccagggc tacgaacat caatgggcta catctcatgg ctaaacaagt acgacgaagg	5820
	ggactcgggt ctgacaacca tgctccgcaa agccgggtcc gtctctctac tcaagacctc	5880
	tgteccgcag acctgatgg tetgcgagac agtcaacaac atcatcgggc gcaccgtcaa	5940
	cccacgcaac aagaactggt cgtgcggcgg cagttctggt ggtgagggtg cgatcgttgg	6000
	gattcgtggt ggcgtcatcg gtgtaggaac ggatatcggg ggtcgtatc gattgcggc	6060

	cgcggtcaac ttectgtaag gtctaaggcc gagtcatggg cggetgccgt atgcaaagat	6120
	ggcgaacagc atggagggtc aggagacggt gcacagcgtt gtcgggcccga ttacgeactc	6180
	tgttgagggt gagtcctteg cctcttccit ctttctctgc tctataaccag gcctccactg	6240
	tcctccttcc ttgcttttta tactatafac gagaccggca gtcactgatg aagtaigtta	6300
	gacctccgcc tcttcaccaa atccgtcctc ggtcaggagc catggaaata cgaactccaag	6360
	gtcatcccca tgccttggcg ccagtccgag tcggacatta ttgectccaa gatcaagaac	6420
	ggcgggctca atafcggeta ctacaacttc gacggcaatg tccittccaca cctctctatc	6480
	ctgcgcggcg tggaaaccac cgtcgcggca ctgcctaaag ccggtcacac cgtgaccccg	6540
	tggacgccat acaagcacga ttteggcacc gatctcatct cccatctcta cgcggctgac	6600
	ggcagcggcg acgtaatgcg cgatatcagt gcatccggcg agccggcgat tccaaatctc	6660
	aaagacctac tgaaccgaa catcaaagct gttaacatga acgagctctg ggacacgcct	6720
	ctccagaagt ggaattacca gatggagtag cttagaaaat ggcgggaggc tgaagaaaag	6780
	gccgggaagg aactggagcc catcatcgcg ccgattaccg ctaccgctgc ggtacggcat	6840
[0011]	gaccagttcc ggtactatgg gtatgcctct gtgatcaacc tgctggattt cacgagcgtg	6900
	gttgttccgg ttaccttgcg ggataagaac atcgataaga agaattgagag ttccaaggcg	6960
	gttagtgagc ttgatgccct cgtgcaggaa gagtatgata cggaggcgta ccatggggca	7020
	ccggttgtag tgcaggttat cggacggaga ctcatgtag agaggacgtt ggcgattgca	7080
	gaggaagtgg ggaagtgtct gggaaatgig gtgactccat agctaataag tctcagatag	7140
	caatttgcaac aagaaatcaa taccagcaac tgtaaataag cgtgaaagt accatgccat	7200
	gctacgaaag agcagaaaaa aacctgccgt agaaccgaag agatatgaca cgetteccatc	7260
	tctcaaagga agaattccctt cagggttgcg tttccagtag tgattttacc gctgatgaaa	7320
	tgactggact cctctctctc gctcttatac gaaaaattgc ctgactctgc aaaggttgtt	7380
	tgtcttggaa gatgatgtgc cccccatcg ctcttatctc atacccegcc atctttctag	7440
	attctcatct tcaacaagag gggcaatcca tgatctgaga tccagatgtg cttctggcct	7500
	catactctgc cticaggttg atgttcactt aattggtgac gaattcagct gatttgcctc	7560
	agtatgcctt ggttgggttc ttccagget tgtgccagcc atgagcgcct tgagagcatg	7620
	ttgtcactta taaactcgag taacggccac atattgttca ctacttgaat cacataccta	7680

	atTTTgatag aattgacatg tttaaagagc tgaggtagct ttaatgcctc tgaagtattg	7740
	tgacacaget tctcacagag tgagaatgaa aagtTggact cccctaatag aagtaaaagt	7800
	ttcgtctctg aacggTgaag agcaatagatc cggcaTcaac tacctggeta gactaegacg	7860
	tcaattctgc ggcctTTTga cttttatata Tgtccattaa Tgcaatagat tctTTTTTTT	7920
	TTTTTTTTT TTTTTTTTT TTTTTTTTT TTTGCCaAT ttcgcagatc aaagtggacg	7980
	ttatagcacc ataactaagc teagTtctg agggaagccg tctactacct tagcccatcc	8040
	atccagctcc atacctTgat actTtagacg tgaagcaatt cacactgtac gtctcgcage	8100
	ictccttccc gctcttgett ccccaactggg gtccatggTg cgtgtatcgt cccctccaca	8160
	attctatgcc atggTaccctc cagcttatca atgccccgt aacaagTcgc ctctttgcct	8220
	Tgatagctta Tcgataaaac TTTTTTccg ccagaaagcc Tccgcccaca gacaagaaaa	8280
	aaaattcacc gcctagcctt Tggccccggc attTggetaa acctcgagcc tctctcccgt	8340
	ctTggggtat caggaagaaa agaaaaaaat ccctcgccaa gggctgtTtt ggcaTaccA	8400
	cccgaaaaca gcactTcctc gatcaaaagt Tgcccccat gaagaccacg Tggaaggaca	8460
[0012]	tccctccggt gcctacgcaC caggagtTtc Tggacattgt gctgagcagg acccagcgcA	8520
	aactgcccac Tcagatecgt gccgcttca agattagcag aattcgaggt acgtcgcatt	8580
	gcccatcgca ggatgtctca ttatcggggt cctTggagaa cgatcatgat Tgcatggcga	8640
	Tgtaaacaca tagacagcct tctacaetcg aaaggTcaag Ttcaccagcag agacgtTtC	8700
	cgaaaagtTc gcctccatcc Tcgacagctt cctcgcctc caggacatcc accccttcca	8760
	caaggacctt ctcaacacc tctacgatgc cgaccacttc aagattgccc Ttggccagat	8820
	gtccactgcc aagcacctgg Tcgagaccat ctcgccgac tacgtccgTc tctTgaaata	8880
	cgcccagTcg ctctaccagt gcaagcagct caagcgggcc gctctcggTc gcatggccac	8940
	gctggTcaag cgcctcaagg acccctgct gtacctggac caggtccgcc agcatctcgg	9000
	ccgtcttccc tccatcgacc ccaacaccag gacctgctc atctcgggtt accccaatgt	9060
	Tggcaagtcc agcttctctg gaagTatcac ccgcgccgat Ttggacgtcc agccctatgc	9120
	tttaccacc aagagtctgt Ttgtcggeca cttTgaetac aagtacctgc gattccagcc	9180
	cattgatacc cccggtattc Tggaccacc cctTgaggag atgaacacia Tcgaaatgca	9240
	gaggTatgtg gcgcggcta	9259

<210> 3
 <211> 9088
 <212> DNA
 <213> 人工序列

<220>

<223> 合成的敲除盒

<400> 3

[0013]

```

atcacgcect cgcataaaag accctcaaga gtccatgtgc cctatctgcc tgatcttctt      60
aacccttatt taacattgga cctatcacia cctagttctt ctteagectg ctttgtcaac     120
acttgtcacg gttcaactca acgtaatcag caggtagcag gacgaggata gggagagaaa     180
cgaagagaag aagaggagag aggaagaaaa aaaaaagaaa agaaagaaaa agggaaaagg     240
aaagaaggag gaaaagagaa gaaagtcaga tgaagaagca agaagacgcc atggtagcca     300
ccgcttgica gggctcetta gcaacgaaca actctagctt ggggacttgt cgatgigtcg     360
tttcttctct acccatcage accaacgatg agttcgatat agacgaggac ctcatggaag     420
tagagaccat tgggttcgac aggatctctc agtttcaact ctatgaggtc tgtcgtcgg      480
atgacttttt gaggagcttc cccttctgct tcaaccccaa actctcttct ctgaaaccgc     540
agcacgttgg cacggccgtg ttgctggagc agtttgcttt cgagcactct cagcgtggtt     600
tcagcagccc actggtgagt ggccctcctt gacgtccaca ccttgctcct gtcgcatgcg     660
tatctggtgg gaacgactgc tccaaggagg attgctaacg aggttgtagg ccgaatatcg     720
catcagattc tccgtaacc ttagctacgg cctcttcaac atctgtgaca tgacggagcg     780
caagtactgg tggttggcga ccaagatgcg cggttggaac atcgacggct gccccgaaga     840
cgtcaggaga ctcatgtttg ttcacatcat cgcacacctg ggatgcagcc ccgtctgac     900
ggatgaagac atggactacc ccaagaactg gggcgcaatt ctccacggta gagacagata     960
tccagtgtaa cctgtgggcc accggcctca tgggcgcacc atctgcctcc actcggttgce    1020

```

	cgctctgccct cgtctccagg gcttgggtct cggctactgcg actctgaagt cgtatgtgca	1080
	gcgcattgaac agcctcggcg ccgcggaccc tgttctcttc gtttgcgca agcccagac	1140
	gagatttttt gaaagatgcg gcttcaggaa cagcggccgg agtagtatca agactctggt	1200
	cggcgaatac tacaacatgg tgtgtgcttc cacatcgact tgcccagaact ctatacgatt	1260
	ttcaaacctc gctatacgtc atattgactt gtttctttag gtcttcgatt tgcccgggccc	1320
	caaagacttt atcgactgga atagcattgc cgacgctgce aagaagatgt gaaccattig	1380
	actgatacga tgtgtgctac gcattgtgac ctctcttgtt tgtttctttg gcggtctttt	1440
	gtataccttg ggacacggca gacgcatgtc tatgtgaaga aaacgttccac ggcgcigtct	1500
	gcattcaggaa tatgatcatt aaacatggag cgtaatggta ttaatgatca actagaaaaa	1560
	tggtatggaa gggcgagagg gcgatcaaca aagcagcccg gggcatagtc tggaagcagc	1620
	aggaattgga agggaaaagg aagctgcaca atgaagggat atcgtgagcg gagtggctca	1680
	cgagagtatc aacagactgg cgaaagcaag caattgcca cgcggctat taggccataa	1740
	gatggectgt tgtgagtccc agttgcacgt atcccatac gactgctctg tcgctgactt	1800
[0014]	gaaaaaaaaat agggaggata aaggagaaag aaagtgagac aaccctgtag ggacttgggg	1860
	tagtaggaga acacatgggc aaccgggcaa tacacggat gtgagacgag ttcaacggcg	1920
	aatggaaaat cttgaaaaac aaaataaaat aactgccctc catacgggta tcaaatcaa	1980
	gcagttgiac ggaggetaga tagagttgtg aagtcggtaa tcccgtgtga tagtaatac	2040
	agtcgcatct aaatactccg aagctgctgc gaaccggag aatcgagatg tgctggaaag	2100
	cttctagcga gcggctaaat tagcatgaaa ggctatgaga aattctggag acggcttgtt	2160
	gaatcatggc gttccattct tcgacaagca aagcgttccg tcgcagtage aggcactcat	2220
	tcccgaaaaa actcggagat tectaagtag cgatggaacc ggaataatat aataggcaat	2280
	acattgagtt gcctcgacgg ttgcaatgca gggglactga gcttggacat aactgttccg	2340
	taccccacct ctctcaacc ttggcgctt ccctgatcca gcgtaccctg acaagtcgta	2400
	atcactalta acccagactg accggacgtg ttttgcctt catttgaga aataatgta	2460
	ttgcgatgtg taatttgcct gcttgaccga ctggggctgt tcgaagcccg aatgtaggat	2520
	tgttatccga actctgctcg tagaggcatg ttgtgaatct gtgtcgggca ggacacgct	2580
	cgaaggttca cggcaaggga aaccaccgat agcagtgtct agtagcaacc tgtaaagccc	2640

	caatgcagca tcaactggaaa atacaaacca atggctaaaa gtacataagt taatgcctaa	2700
	agaagtcata taccagcgge taataatigt acaatcaagt ggctaaacgt accgtaattt	2760
	gccaacggct tgtggggttg cagaagcaac ggcaaaagccc cacttcecca cgtttgtttc	2820
	ttcaactcagt ccaatctcag ctggatgatc cccaattggg tcgcttgttt gttccggatga	2880
	agtgaagaa gacagaggta agaatgtctg actcggagcg ttttcatac aaccaagggc	2940
	agtgatggaa gacagtgaaa tgttgacatt caaggagtat ttagccaggg atgcttgagt	3000
	gtatcgtgta aggaggtttg tctgccgata cgacgaatac tgtatagtea cttctgatga	3060
	agtggccat attgaaatgt aaagtcggca ctgaacaggc aaaagattga gttgaaactg	3120
	cctaagatct cgggcctcgc gcccttcgce ctttgggtgt acatgtttgt gctccgggca	3180
	aatgcaaagt gtggtaggat cgaacacact gctgccttta ccaagcagct gagggtatgt	3240
	gataggcaaa tgttcagggg ccaactgcatg gtttcgaata gaaagagaag cttagccaag	3300
	aacaatagcc gataaagata gctcattaa acggaatgag ctagtaggea aagtcagcga	3360
	atgtgtatat ataaaggttc gaggctcgtg cctccctcat gctctccca tctactcact	3420
[0015]	aactcagatc ctccaggaga ctgtacacc atcttttgag gcacagaaac ccaatagtea	3480
	accgcgact gcgcatcatg tateggaagt tggcgcctat ctccgcttc ttggccacac	3540
	ctcgtgctag actaggcgcg tcaatatgig gccgttactc gagtttataa gtgacaacat	3600
	gctctcaaag cgctcatgce tggcacaagc ctggaaagaa ccaacacaaa gcatactgca	3660
	gcaaatcagc tgaattcgtc accaattaag tgaacatcaa cctgaaggca gagtatgagg	3720
	ccagaagcac atctggatcg cagatcatgg attgccccctc ttgttgaaga tgagaatcta	3780
	gaaagatgce ggggtatgag ataagagcga tgggggggca catcatcttc caagacaac	3840
	aaccittgca gagtcaggca atttttcgta taagagcagg aggagggagt ccagtcattt	3900
	catcagcggg aaaatcactc tagacaatct tcaagatgag ttctgccttg ggtgacttat	3960
	agccatcact atacctagac agaagcttgt gggatactaa gaccaacgta caagctcgca	4020
	ctgtacgctt tgacttccat gtgaaaactc gatacggcgc gcctctaaat tttatagctc	4080
	aaccactcca atccaacctc tgcatectc tcaactcgtc tgaactactg ttcaaatcag	4140
	agaataagga cactatccaa atccaacaga atggctacca cctcccagct gcctgcctac	4200
	aagcaggact tctctaaatc cgcctcgcac ggccggctcc tcaagtttgg cagcttcgag	4260

	ctcaagtcca agcggatate cccctacttc ttcaacgcgg gcgaattcca cacggcgcgc	4320
	ctcgcgggeg ccatagccte cgcctttgca aagaccatca tcgaggecca ggagaaggcc	4380
	ggcctagagt tcgacatcgt ctteggeceg gcctacaagg gcatcccgt gtgtccgcc	4440
	atcacatca agctcggcga gctggcgccc cagaacctgg accgcgtctc ctactcgttt	4500
	gaccgcaagg aggecaagga ccacggcgag ggcggcaaca tcgtcggcgc ttcgtcaag	4560
	ggcaagaggg tcttgattgt cgacgacgtc atcaccgccc gcaccgcaa gagggacgcc	4620
	attgagaaga tcaccaagga gggcggcacc gtgcceggca tcgtcgtggc cctggaccgc	4680
	atggagaagc tccccgtgc ggatggcgac gactccaage ctggaccgag tgccattggc	4740
	gagctgagga aggagtacgg catcccate ttgccatcc tcaactctgga tgacattate	4800
	gatggcatga agggctttgc taccctgag gatataaga acacggagga ttaccgtgcc	4860
	aagtacaagg cgactgactg attgaggcgt tcaatgtcag aaggagaga aagactgaaa	4920
	aggtggaag aagaggcaaa ttgtgttat tattattatt ctatctgaa tcttctagat	4980
	cttgctgtaa ataaacaagc gtaactagct agcctccgta caactgctg aatttgatac	5040
[0016]	ccgtatggag ggcagttatt ttattttgtt ttcaagatt ttcattcgc cgttgaactc	5100
	gtctcacatc gcgtgtattg cccggtgcc catgtgtacg cgtttcgggt ttacctctc	5160
	cagatacagc tcatctgcaa tgcattaatg cattggacct cgcaacctc gtacgccctt	5220
	caggctccgg cgaagcagaa gaatagctta gcagagtcta ttttcatttt cgggagacta	5280
	gcattctgta aacgggcagc aatcggccag cagttagtag ggteccctct acctctcagg	5340
	gagatgtaac aacgccacct tatgggacta tcaagctgac gctggettct gtgcagacaa	5400
	actgcgceca cgagttctc cctgacgccg ctctcgcga ggcaaggga ctcgatgaat	5460
	actacgcaaa gcacaagaga cccgttggtc cactccatgg cctcccate tctctcaaag	5520
	accagcttcg agtcaaggta caccgttccc cetaagtcgt tagatgtccc tttttgtcag	5580
	ctaacatatg ccaccagggc tacgaaacat caatgggcta catctcatgg ctaaacaagt	5640
	acgacgaagg ggactcggt ctgacaacca tgetccgcaa agccggtgcc gtcttctacg	5700
	tcaagacctc tglcccag accctgatgg tetgcgagac agtcaacaac atcatcgggc	5760
	gcaccgcaa cccacgcaac aagaactggt cgtgcggcgg cagttctggt ggtgagggtg	5820
	cgatcgttgg gattcgtggt ggcgtcatcg gtgtaggaac ggatcctggt ggctcgattc	5880

	gagtgccggc cgcgttcaac ttctgtacg gtctaaggcc gagtcatggg cggcigccgt	5940
	atgcaaagat ggcgaacage atggagggtc aggagacggt gcacagcgtt gtcgggccga	6000
	ttacgcactc tgtigagggt gagtccttcg cctcttcott cttttccigc tctataccag	6060
	gccitceactg tctctcttcc ttgcttttta tactatatac gagaccggca gtcactgatg	6120
	aagtatgta gacctccgcc tcttcaccaa atccgtcctc ggtcaggagc catggaaata	6180
	cgactccaag gtcaccccc a gccctggcg ccagtcggag tcggacatia ttgcttccaa	6240
	gatcaagaac ggcgggctca atatcggcta ctacaacttc gacggcaatg tctttccaca	6300
	ccctcctatc ctgcccggcg tggaaaccac cgtcgcggca ctgcctaaag ccggtcacac	6360
	cgtgaccocg tggacgccat acaagcacga tttcggccac gatctcatct cccatateta	6420
	cgccgctgac ggcagcggcg acgtaatgcg cgatctcagt gcatccggcg agccggcgat	6480
	tccaaatata aaagacctac tgaacccgaa catcaaagct gtaacatga acgagctctg	6540
	ggacacgcat ctccagaagt ggaattacca gatggagtac ctgagaaat ggcgggaggc	6600
	tgaagaaaag gccgggaagg aactggacgc catcatcgcg ccgattacgc ctaccgctgc	6660
[0017]	ggtaaggcat gaccagttcc ggtactatgg gtatgcctct gtgatcaacc tgctggattt	6720
	caagagcglg gttgttccgg ttacctttgc ggataagaac atcgataaga agaattgagag	6780
	tttcaaggcg gttagtgage ttgatgccct cgigcaggaa gagtatgac cggaggcgta	6840
	ccatggggca ccggttgcag tgcaggttat cggacggaga ctccagtgaag agaggacgtt	6900
	ggcgattgca gaggaagtgg ggaagttgct gggaaatgtg gtgactccat agctaataag	6960
	tgtcagatag caatttgcac aagaaatcaa taccagcaac tgtaaataag cgctgaagtg	7020
	accatgccat gctacgaaag agcagaaaaa aacctgcctt agaaccgaag agatatgaca	7080
	cgcttccatc tctcaaagga agaattccctt cagggttgcg tttccagtag tgattttacc	7140
	gctgatgaaa tgactggact cctctctcct gctcttatac gaaaaattgc ctgactctgc	7200
	aaaggttggt tgtcttggaa gatgatgtgc cccccatcg ctcttatctc ataccocgcc	7260
	atctttctag attctcatct tcaacaagag gggcaatcca tgatctgcga tccagatgtg	7320
	cttctggcct calactctgc cttcagggtg atggttcaact aattgggtgac gaattcagct	7380
	gatttgctgc agtatgettt gigtggttc tttccaggtc tgtgccagcc atgagcgtt	7440
	tgagagcatg ttgtcactta taaactcgag taacggccac atattgttca ctacttgaat	7500

	cacataccta attttgatag aattgacatg tttaaagagc tgaggtagct ttaatgcctc	7560
	tgaagtattg tgacacagct telcacagag tgagaatgaa aagtiggaet ccccctaag	7620
	aagtaaaagt ttctctctg aacgggtgaag agcatagatc eggcatcaac tacctggeta	7680
	gactacgacg tcaattctgc gcccttttga cctttatata tglccattaa tgcaatagat	7740
	tctttttttt tttttttttt tttttttttt tttttttttt ttgcccatt ttctgcagatc	7800
	aaagtggacg ttatagcacc ataaactaagc tcagttgctg agggaagccg tctactacct	7860
	tagcccatcc atccagctcc atacctgat accttagacg tgaagcaatt cacactgtac	7920
	gtctcgcagc tctctctccc gctcttgctt ccccaactggg gtccatgggt cgtgtatcgt	7980
	cccctcaca atctctatgcc atggtacctc cagcttatca atgccccgct aacaagtcgc	8040
	ctctttgect tgatagctta tcgataaaac tttttttccg ccagaaagge tccgccaca	8100
	gacaagaaaa aaaattcacc gcctagcctt tggccccgga atttggctaa acctcgagcc	8160
	tctctcccgt cttgggggat caggaagaaa agaaaaaat ccctcgccaa gggctgtttt	8220
	ggcatcaeca cccgaaaaca gcacttctc gatcaaaagt tgcccgcct gaagaccacg	8280
[0018]	tggaaggaca tccctccggt gcctacgcac caggagtctc tggacattgt gctgagcagg	8340
	accagcgca aactgccac tcagatcctt gccggetta agattagcag aattcgaggt	8400
	acgtcgcaat gcccatcgca ggaatgtctc ttatcgggt ccttgagaa cgaicagat	8460
	tgcattggca tgctaacaca tagacagcct tctacactcg aaaggtcaag ttcaccagc	8520
	agacgttttc cgaaaagttc gcctccatcc tcgacagctt cctcgcctc caggacatcc	8580
	acccttcca caaggacctt ctcaacacc tctacgatgc cgaccacttc aagattgccc	8640
	ttggccagat gtccactgcc aagcactgg tcgagaccat ctgcgcgac tacgtccgtc	8700
	tcttgaaata cgcccagteg ctctaccagt gcaagcagct caagcgggc gctctcggtc	8760
	gcatggccac gctggtaag cgcctcaagg acccctgct gtacctggac caggctccgc	8820
	agcatctcgg ccgtcttccc tccatcgacc ccaacaccag gacctgctc atctgcggtt	8880
	accccaatgt tggcaagtc agcttctgc gaagtatcac ccgcgccat gtggacgtcc	8940
	agccctatgc tttaccacc aagagtctgt ttgtcgcca ctttgactac aagtacctgc	9000
	gattccagc cattgatacc cccgtattc tggaccacc tcttgaggag atgaacacta	9060
	tcgaaatgca gaggtatgt gcgcggct	9088

<210> 4

<211> 5610

<212> DNA

<213> 人工序列

<220>

<223> 合成的敲除盒

<400> 4

[0019]

```

aatggtagga atgctgggat ataggcictg tgctggcaag ttgatggatc ctgaaatgag      60
gccgccctgc aaggggaaca tcagagatct accattgcct ccttggceca atccactatc      120
atacctacct catgatcatt cctgcgaagg tetaccagta aatatttcct cgtcccgtgt      180
ttcatcatgt ccagaacctc atctcgccaa attgaacttg ccacagtgtc tggagctggg      240
taagcagcgt gccaaaggaat tgttgtcgag tctgtgccag gcattgtgcc cgacattgtg      300
aacttcagcc aggagaactt ttcgatcgca cctatgciga gcaccgtggg catgcgatgg      360
cttcaataat gcagttcgag agggagtgtg tcatgcecta aagctcattg gccacctcca      420
caggctagct ctacctgcat ctgtagatgg actttctctg tctctctctc tcagaaaacc      480
tcttggtcgc tcgcaggtaa ctgttgttgc cgtcattgtt tgacagtgga tagccaaggc      540
aaaaccgtct gctttcaacg gaagcattcg ggggttgtt gtcategtgt tatcgatcga      600
ccaggagaac ccagacgagt gttgttcgag agaatcatcg acgatgtgaa gaggcgacga      660
ctagtatcta gaagattata atcgaacaaa tcagcgittg tctgtcgggc gtttgagggc      720
gcagttgccc gccaaagcag cgtcgcaata tataggcagc gagagactgt caacagccag      780
ccgccatgtg atcgatcgta gccgtcttcc egatcttccc taaaccctt tctttggggg      840
gcgggggcag cggcgttcia atatttgctg gctgtcigga taactgtaat ggtagacatg      900
gtaatgttcg gtctgcgaaa catttgtaaa attggagttt acgatcgaga tggaaggaaa      960
cgctccacaa actcgggtgac tgggttgcca tcaggtgctc agggcatagc gtctctgca     1020

```

	aatagaggaa agagaatage actagtghaaa gtgtgaatca caatgaagag gaggttgttg	1080
	ccggaatgct ttgagcageg tcaaagtiga acttgaaget atcacaatt gcagggtaaa	1140
	gtacatgttg gtgccagttt gacagcacag tgcgcggagc ggaggatgte gcggaagagg	1200
	cgcgacgcta acccggcct tctctcagt gacgagaact cctgcigcaa gatttcttc	1260
	tctctgcgag atgacgtgag gcccaatttg cagcttccct cgaacaagg gattgaacat	1320
	ctctcttccc tcacattca tcateactac ctctcaatt cacttcigt tggccgtct	1380
	tcacattca tttactgct ctgatgcta tctgaagat tttattctg cagtattca	1440
	gccatcccac ctccggtct cactcacagt cacagggtcaa ccgcttcac cctctcgcg	1500
	atgatgtcgg caatcigggt gatcaatgtg cggttgaggg ccgccgtagt gaggatgggc	1560
	atggggaaac aggtcgcaca ttgcgccaca gataacttcg tatagcatac attatacgaa	1620
	gttatcttgg gcttggact ggtcgcgagc tgccactaag tggggcagta ccattttate	1680
	ggacceatcc agctatggga cccactcga aatttttaca tcattttct tttgtcagt	1740
	aacggccacc tttgtaaag cgtaccagc aaacaaattg caattggccc gtagcaagg	1800
[0020]	agtcagggt tctctgatg gaggagaagg ctatatcagc ctcaaaaata ttttccagc	1860
	tggcggaaag ccggaaggta agtggattct tgcgcgtggc tggagcaacc gttgattcc	1920
	agcgtctcgc acttggactg agcaattcag cgtcacggat tcacgataga cagctcagac	1980
	cgctccacgg ctggcggcat tatggttaa cccggaaact cagtctcctt ggccccgtcc	2040
	cgaagggacc cgacttaca ggctgggaaa gccagggata gaatacactg tacggcttc	2100
	gtacgggagg ttgcgcgtag gttgttccc aagttttaca cccccccaa gacagctagc	2160
	gcacgaaaga cgcggagggt ttgtgaaaa aaggcgaaa attaagcggg agacgtattt	2220
	aggtgctagg gccggttcc tcccatttt tcttcggtc cttttctctc ctggaagact	2280
	ttctctctct ctctctctct etctctccat cctcagtea tcttcttct ccatcatcca	2340
	tctctcacc tccatctcaa ctccatcaca tcacaatga tatgaaaaag cctgaactca	2400
	ccgcgacgtc tgtcgagaag tttctgatc aaaagttega cagcgtctcc gacctgatgc	2460
	agctctcggg gggcgaagaa tctctgctt tcagcttca tttaggagg cgtggatag	2520
	tcttgcgggt aaatagctgc gccgatggtt ttacaaaaga tctttatgtt tctcggcact	2580
	ttgcatcggc cgcgctccc attccggaag tcttgacat tggggaattc agcgagagcc	2640

	tgacctattg catctcccgc cgtgcacagg gtgtcacgtt gcaagacctg cctgaaaccg	2700
	aactgcccgc tgttctgcag ccggtcgcgg aggccatgga tgcgatcgct gcggccgate	2760
	ttagccagac gagcgggttc ggcccattcg gaccgcaagg aatcgggtcaa tacaactacat	2820
	ggcgtgattt catatgcgcg attgctgata cccatgtgta tcaactggcaa actgtgatgg	2880
	acgacaccgt cagtgcgtcc gtcgcgcagg ctctcgatga gctgatgctt tgggcccagg	2940
	actgcccoga agtcggcac ctctgtcacg cggatttcgg ctccaacaat gtctgacgg	3000
	acaatggccg cataacagcg gtcattgact ggagcagggc gatgttcggg gattcccaat	3060
	acgaggtcgc caacatcttc ttctggaggc cgtggttggc ttgtatggag cagcagacgc	3120
	gctacttcga gcggaggcat ccggagcttg caggatcgcc gggctccgg gcgtataatgc	3180
	tccgcatigg tcttgaccaa ctctatcaga gcttggttga cggcaatttc gatgatgcag	3240
	cttggggcca gggtcgatgc gacgcaatcg tccgatccgg agccgggact gtcggggcta	3300
	cacaaatcgc ccgcagaagc gggccctct ggaccgatgg ctgtgtagaa gtactcgcgc	3360
	atagtggaaa ccgaccccc agcactcgtc cgagggcaaa ggaatagagt agatgcgcac	3420
[0021]	cgggatccac ttaacgttac tgaatcacc aaacagcttg acgaatctgg atataagatc	3480
	gttgggtgct atgtcagctc cggagttgag acaaatggtg ttcaggatct cgataagata	3540
	cgttcatttg tccaagcagc aaagagtgcc ttctagtat ttaatagctc catgtcaaca	3600
	agaataaac gcgttcggg tttacctctt ccagatacag ctcatctgca atgcattaat	3660
	gcattggacc tcgcaacct agtacgcct tcaggctccg gcgaagcaga agaatagctt	3720
	agcagagtct attttcattt tcgggagacg agatcaagca gatcaacggt cgtcaagaga	3780
	cctacgagac tgaggaatcc gctcttggct ccacgcgact atatatttgt ctetaattgt	3840
	actttgacat gctctcttc ttaactctga tagcttgact atgaaaattc cgtcaccage	3900
	cccigggttc gcaaagataa ttgcactgtt tcttcttga actctcaagc ctacaggaca	3960
	cacattcacc gtaggtataa acctcgaaaa tcattctac taagatgggt atacaatagt	4020
	aaccatggtt gcctagttaa tgcctcgtaa cacccaatac gccggccgat aacttcgtat	4080
	agcatacatt ataagaagtt ataactggcg gcctagtgg aacacgagca cataagcttt	4140
	tacctatggt taccgcttgc atctacgcgc cgttgatggt ggaggatggt ggaegtccc	4200
	gagacccta cgagctgtgg cactgtcaaa ctgtgcccac agaccttgt ctgctttca	4260

	taacctegag gagggtttcc agacatca tccatacaca agcagtatta atcaaagaaa	4320
	ctcggtcgca atggcaaaaa tggtttgeaa acagaaaact atgacctett cctattccat	4380
	cattaactac tetaccggtt tgcataaca acatcattaa aaccttatg cgteaggtgt	4440
	agcatccttg atctgttgc cctccaacgg ccagttctca atcgttacct cttctccac	4500
	caactcaaac tcaagcttca cagactcgtc ggtgttcaag gctagctcat acttgccggg	4560
	gtatacaatc cggtttccgt gagaatcaac acgggcgaga gcactgacag ggatggggat	4620
	gctgagcttg gaagagtgac caggcttgat gtcggcaagt cggtcgaatc cgacgagcca	4680
	cttgittcggg tacggggctg gccagcgtt gcttgtgcga acaaacagca tggccgtata	4740
	tggggactcc gtcttcccc agttcttgat gttggcctcg aaggatgaaga cgggaatctg	4800
	ctcgcgtgaa gtgtatccgg ggtgaggagc agagaggatc gatgaggtgt tgaacttgag	4860
	gctcttgggg tggctggcga gagtctcctt gaaggtggtg tagaagagac cactgccaaa	4920
	ctcgtagacg ggtttgcggg tgtaccagat gtaagtctgt ccagggttg actttccatc	4980
	gggtcggagg ttcattgcat tctgggggaa ttggtgaaca tactcagccg ggtactgagt	5040
[0022]	ggtgaccagt cggeccgcag gacacgctt gccagagaga atgicgaaga gggcaacgcc	5100
	tcccgaactgg ccgggatata cgcgccagac gagggagttg acctcttctt tgcctttgag	5160
	cgaggatgag tctacctgac caccgccat ttgcaggacg acaagggtt tgcgacctc	5220
	gctgagctgc ttgatgagat ccagctgatt accgggcca acaatgtccg tgcggtcagc	5280
	gccctctgt tcaatggtgt tgcattcc accgaggtag atgatggcat ccgaactctt	5340
	ggcggcagca atggccttg caaagccagt ggtgctgtt cggcgatct ctgtgccgag	5400
	ttcaaagtig acgtgatagc cggecttctt agcagcttc agaggctga tgaggtatgg	5460
	ggcaggcca tagtagtgc cttgcatttg gttgtggca ttggccatg gtccgatcag	5520
	agcaatgetg cgcacctct tggacagagg gagagtcca tcgttcttga gcaggacgat	5580
	gccctcaaca gcagcctct acgagatgtt	5610

<210> 5

<211> 13862

<212> DNA

<213> 人工序列

<220>

<223> 合成的载体

<400> 5

	ttgtacaaag tgggatcgc gccgcgcgcc agctccgtgc gaaagcctga cgcaccggta	60
	gattcttggg gageccgtat catgacggcg gcgggagcta catggccccg ggtgatttat	120
	tttttttcta tctacttctg acccttttca aatatacggg caactcatct ttcactggag	180
	atgeggcctg cttggtattg cgaigtgtgc agcttggcaa atgttggett tcgaaaacac	240
	aaaacgattc cttagtagcc atgcatttta agataacgga atagaagaaa gaggaaatta	300
	aaaaaaaaa aaaaacaaac atcccgttca taaccgtag aatcgccctc cttcgtgtat	360
	cccagtagca gtttatittg aatagctcgc ccgctggaga gcactcctgaa tgcaagtaac	420
[0023]	aaccgtagag gctgacacgg caggtgttgc tagggagcgt cgtgttctac aaggccagac	480
	gtcttcgcgg ttgatataata tgtaigtgtg actgcaggct gctcagcagc gacagtcagc	540
	ttcgccctcg ctgcttgtgc aataatcgcg gtggggaage cacaccgtga ctcccatctt	600
	tcagtaaage tctgttggtg ttatcagca atacacgtaa tttaaactcg ttageatggg	660
	gctgatagct taattaccgt ttaccagtcg catggttctg cagctttcct tggcccgtaa	720
	aattcggcga agccagccaa tcaccagcta ggcaccagct aaacctata attagtctct	780
	tatcaacacc atccgctccc ccgggatcaa tgaggagaat gaggggatg cggggctaaa	840
	gaagcctaca taacctcat gccaaactcc agttlacact cgtcgagcca acatcctgac	900
	tataagctaa cacagaatgc ctcaatcctg ggaagaactg gccgctgata agcgcgcccg	960
	cctcgcaaaa accatccctg atgaatggaa agtcagacg ctgcctgcgg aagacagcgt	1020
	tattgatttc ccaagaaat cggggatcct ttcagaggcc gaactgaaga tcacagagcc	1080
	ctccgctgca gatcttgtgt ccaagctggc ggcccagagc ttgacctcgg tggaagttac	1140
	gctagcattc tgtaaaccgg cagcaatcgc ccagcagtta gtagggctcc ctctacctct	1200
	cagggagatg taacaacgcc accttatggg actatcaagc tgacgctggc ttctgtgcag	1260

	acaaactgeg cccacgagtt cttccctgac gccgctctcg cgcaggcaag ggaactcgat	1320
	gaatactaeg caaagcacia gagaccggtt ggtccactec atggcctccc catctctctc	1380
	aaagaccage ttcgagicaa ggtacaccgt tgccectaag tegttagatg tccctttttg	1440
	tcagctaaca tatgccacca gggctaogaa acatcaatgg gctacatctc atggctaaac	1500
	aagtacgacg aaggggactc ggttctgaca accatgctcc gcaaagccgg tgccgtcttc	1560
	tacgtcaaga cctctgtccc gcagaccctg atggctctcg agacagtoaa caacatcate	1620
	gggcgcaccg tcaacccacg caacaagaac tggctctgag gcggcagttc tgggtgtgag	1680
	ggtgcgatcg ttgggattcg tggtagctc atcgggtgag gaacggatat cggtagctcg	1740
	attcgagtgc cggcccggtt caacttctg tacggctctaa ggccgagtca tgggaggctg	1800
	ccgtatgcaa agatggcgaa cagcatggag ggtcaggaga cggtagcacag cgttctcggg	1860
	ccgattacgc actctgttga gggtagtcc ttccctctt ccttcttttc ctgctctata	1920
	ccaggcctcc actgtctccc tttcttgctt ttatactat atacgagacc ggcagtcact	1980
	gatgaagtat gttagacctc cgcctctca ccaaaccgt cctcggtcag gagccatgga	2040
[0024]	aatacgactc caaggtcacc cccatgcctt ggcgccagtc cgagtcggac attattgcct	2100
	ccaagatcaa gaacggcggg ctcaatatcg gctactacaa cttcgacggc aatgtctctc	2160
	cacaccctcc tatctctcgc ggcgtggaaa ccaccgtcgc cgcactcgc aaagccggtc	2220
	acaccgtgac cccgtggacg ccatacaagc acgatttcgg ccacgatctc atctccata	2280
	tctacggcgc tgacggcagc gccgacgtaa tgcgcgatat cagtgcctcc ggcgagccgg	2340
	cgattccaaa tatcaagac ctactgaacc cgaacatcaa agctgttaac atgaacgagc	2400
	tctgggacac gcctctccag aagtggaatt accagatgga gtaccttgag aatggcggg	2460
	aggtgaaga aaaggccggg aaggaactgg acgccatcat cgcgccgatt acgcctaccg	2520
	ctgcggtagc gcctgaccag ttccggtact atgggtatgc ctctgtgate aacctgctgg	2580
	atttcacgag cgtggttggt cgggttacct ttgcggataa gaacatcgat aagaagaatg	2640
	agagtttcaa ggcggttagt gagcttgatg cctcctgca ggaagagtat gatccggagg	2700
	cgtaccatgg ggcaccggtt gcagtgcagg ttatcggacg gagactcagt gaagagagga	2760
	cgttggcgat tgcagaggaa gtggggaagt tgcctggaaa tgtggtgact ccatagctaa	2820
	taagtgtcag atagcaattt gcacaagaaa tcaataccag caactgtaaa taagcgtgta	2880

	agtgaccatg ccatgctaag aaagagcaga aaaaaacctg ccgtagaacc gaagagatat	2940
	gacacgettc catctctcaa aggaagaate ccttcagggt tgcgtttcca gtctagacac	3000
	gtataacggc acaagtgtct ctccaccaat gggttatata tcaaatgtga tctaaggatg	3060
	gaaagcccag aatatcgatc gcgcgcagat ccatatatag ggcccgggtt ataattacct	3120
	caggaaatag ctttaagtag cttattaagt attaaaatta tatatatttt taatataact	3180
	atatttcttt aataaatagg tattttaagc tttatatata aatataataa taaaataata	3240
	tattatatag ctttttatta ataaataaaa tagctaaaaa tataaaaaaa atagctttaa	3300
	aatacttatt ttttaattaga attttatata tttttaatat ataagatctt ttactttttt	3360
	ataagcttcc taccttaaat taaattttta ctttttttta ctattttact atatcttaaa	3420
	taaaggcttt aaaaatataa aaaaaatctt cttatatatt ataagctata aggattatat	3480
	atataatttt ttttaatttt taaagtaagt attaaagcta gaattaaagt ttttaatttt	3540
	taaggcttta tttaaaaaaa ggcagtaata gcttataaaa gaaatttctt tttcttttat	3600
	actaaaagta cttttttttt aataaggtta gggttagggt tagggtagg gttagggtta	3660
[0025]	gggttagggt tagggtagg gttagggtta gggttagggt tagggtagg gttagggtta	3720
	gggtttaaac aaagccacgt tgtgtctcaa aatctctgat gttacattgc acaagataaa	3780
	aatatatacat catgaacaat aaaactgtct gcttacataa acagtaatac aaggggtgtt	3840
	atgagccata ttcaacggga aacgtcttgc tcgaggccgc gatlaaatc caacatggat	3900
	gctgatttat atgggtataa atgggctcgc gataatgctg ggcaatcagg tgcgacaatc	3960
	tatcgattgt atgggaagcc cgatgcgcca gagttgittc tgaaacatgg caaaggtagc	4020
	gttgccaatg atgttacaga tgagatggtc agactaaact ggcctgacgga atttatgect	4080
	cttccgacca tcaagcattt tafccgtact cctgatgatg catggttact caccactgcg	4140
	atccccggga aaacagcatt ccaggtatta gaagaatate ctgaitcagg tgaaaatatt	4200
	gttgatgcgc tggcagtgtt cctgcgccgg ttgcattcga ttctgtttg taattgtect	4260
	tttaacagcg atcgcgtatt tcgtctcgtc caggcgcaat cacgaatgaa taacggtttg	4320
	gttgatgcga gtgattitga tgacgagcgt aaiggctggc ctgittgaaca agtctggaaa	4380
	gaaatgcata agcttttgcc attctcaccg gattcagtcg tcaetcatgg tgatttctca	4440
	cttgataacc ttatttttga cgaggggaaa ttaataggit gtattgatgt tggacgagtc	4500

	ggaatgcag accgatacca ggatcttgcc atcctatgga actgcctcgg tgagttttct	4560
	ccttcattac agaaacggct ttttcaaaaa tatggtattg ataatectga tatgaataaa	4620
	ttgcagttc attigatgct cgalgagttt ttctaatacag aattggttaa ttggttgtaa	4680
	cactggcaga gcattacgct gacttgacgg gacggcggtt ttgtigaata aatcgaactt	4740
	ttgctgagtt gaaggatcag atcacgcate tteccgacaa cgcagaccgt tccgtggcaa	4800
	agcaaaagtt caaaatcacc aactgggtcca cctacaacaa agctctcate aaccgtggct	4860
	ccctcactt ctggctggat gatggggcga ttcaggcctg giatgagtea gcaacacctt	4920
	cttcacgagg cagacctcag cggtttaaac ctaaccctaa ccctaaccct aaccttaacc	4980
	ctaaccctaa ccctaaccct aaccttaacc ctaaccctaa ccctaaccct aaccttaacc	5040
	taaccctaata ggggtcgatc tgaaccgagg atgagggttc tatagactaa tctacagccc	5100
	gtacatggtg tgattgcaga tgcgacgggc aaggtgtaca gtgtccagaa ggaggagagc	5160
	ggcataggtg ttgtaataga ccagcittac ataataatcg cctgttgcta ctgactgatg	5220
	accttcttcc ctaaccagtt tctaattac cactgcagtg aggataaccc taactcgtc	5280
[0026]	tggggttatt attatactga ttagcaggtg gcttatatag tgctgaagta ctataagagt	5340
	ttctgcggga ggaggtggaa ggactataaa ctggacacag ttagggatag agtgatgaca	5400
	agacctgaat gttatcctcc gggtgggat agegaattgg ctgacctgc agatggtaat	5460
	ggtttaggca gggtttttgc agagggggac gagaacgcgt tctgcgatti aacggctgct	5520
	gccgccaagc ttacggttc tclaatgggc ggccgccca ggtcgacgtc ccatggccat	5580
	tcaattcgt aatcatggtc atagctgttt cctgtgtgaa attgttatcc gctcacaatt	5640
	ccacacaaca tacgagccgg aagcataaag tgtaaagcct ggggtgccta atgagtgage	5700
	taactcacat taattcggtt gcgctcactg cccgctttcc agtcgggaaa cctgtcgtgc	5760
	cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat tgggcgtct	5820
	tccgcttctt cgtcactga ctgctgcgc tgggtcgttc ggctgcggcg agcgtatca	5880
	gctcactcaa aggcggtaat acggttatcc acagaatcag gggataacgc aggaaagaac	5940
	atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt	6000
	ttccatagge tccgcccc tgacgagcat cacaaaaate gacgtcaag tcagaggtgg	6060
	cgaaacccga caggactata aagataaccag gcgtttcccc ctggaagctc cctcgtgcgc	6120

tctcctgttc	cgacctgccc	gettaccgga	tacctgtccg	cctttctccc	ttcgggaagc	6180	
gtggcgcttt	ctcatagctc	acgctgtagg	taictcagtt	cgggtgtaggt	cgttcgctcc	6240	
aagctgggct	gtgtgcacga	aceccccgtt	cagccccgacc	gctgcgcctt	atccggtaac	6300	
tatcgtcttg	agtccaacce	ggtaagacac	gacttatcgc	cactggcagc	agccactggt	6360	
aacaggatta	gcagagcgag	gtatgtaggc	ggigctacag	agttcttgaa	gtggtgacct	6420	
aactacggct	acactagaag	aacagtattt	ggtatctgcg	ctctgctgaa	gccagttacc	6480	
ttcggaaaaa	gagttggtag	ctcttgatcc	ggcaaaaaaa	ccaccgctgg	tagcggtggt	6540	
ttttttgitt	gcaageagca	gattacgcgc	agaaaaaaag	gatctcaaga	agatcctttg	6600	
atcttttcta	cggggtctga	cgctcagtgg	aacgaaaact	cacgttaagg	gattttggtc	6660	
atgagattat	caaaaaggat	cttcacctag	atccttttaa	attaanaaig	aagttttaaa	6720	
tcaatctaaa	gtatataiga	gtaaacttgg	tctgacagtt	accaatgctt	aatcagtgag	6780	
gcacctatct	cagcgatctg	tctatttcgt	teatccatag	ttgcctgact	ccccgtcgtg	6840	
tagataacta	cgatacggga	gggettacca	tctggccccca	gtgctgcaat	gataccgcga	6900	
[0027]	gaccacgct	caccggetcc	agatttatca	gcaataaacc	agccagccgg	aagggccgag	6960
cgcagaagtg	gtcctgcaac	tttatccgcc	tecatccagt	ctattaattg	ttgccgggaa	7020	
gctagagtaa	gtagttcgcc	agttaatagt	ttgcgcaacg	ttgttgccat	tgctacaggc	7080	
atcgtgggtg	cacgctcgtc	gtttgggatg	gcttcattca	gctccgggtc	ccaacgatca	7140	
aggcgagtta	catgatcccc	catgttgtgc	aaaaaagcgg	ttagctcctt	cggctcctccg	7200	
atcgttgta	gaagtaagtt	ggccgcagtg	ttatactca	tggttatggc	agcactgcat	7260	
aattctctta	ctgicatgcc	atccgtaaga	tgtttttctg	tgactggtga	gtactcaacc	7320	
aagtcattct	gagaatagtg	tatgcggcga	cagagttgct	ctgccccggc	gtcaatacgg	7380	
gataataaccg	cgccacatag	cagaacttta	aaagtgtcca	tcattggaaa	acgttcttcg	7440	
gggcgaaaaac	tctcaaggat	cttaccgctg	ttgagatcca	gttcgatgta	accactcgt	7500	
gcacccaact	gatcttcagc	atcitttiact	ttcaccagcg	ttctgggtg	agcaaaaaca	7560	
ggaaggcaaa	atgccgcaaa	aaagggaata	aggcgcacac	ggaaatggtg	aatactcata	7620	
ctcttccttt	ttcaatatta	ttgaagcatt	tatcagggtt	atigtctcat	gagcggatac	7680	
atatttgaat	gtatttagaa	aaataaacia	ataggggttc	cgcgcacatt	tccccgaaaa	7740	

	gtgccacctg acgtctaaga aaccattatt aicaiacacat taacctataa aaatagcctg	7800
	atcacgagge cctttcgtct cgcgcgttcc ggtgatgacg gtgaaaaect ctgacacatg	7860
	cagctcccgg agacggtcac agcttctctg taageggatg ccgggagcag acaagcccgt	7920
	cagggcgcgt cagcgggtgt tggcgggtgt cggggctggc ttaactatgc ggcatcagag	7980
	cagattgtac tgagagtgca ccataaaatt gtaaactgta atatcttctt aaaattcgcg	8040
	ttaaattttt gttaaatcag ctcatctttt aaccaatagg ccgaaatcgg caaaatccct	8100
	tataaatcaa aagaatagcc cgagataggg ttgagtggtg ttccagtttg gaacaagagt	8160
	ccactattaa agaactgga ctccaacgtc aaaggcgaa aaaccgtcta tcagggcgt	8220
	ggcccactac gtgaaccatc acccaaatca agttttttgg ggtcagagtg ccgtaaagca	8280
	ctaaatcgga accctaaagg gagccccga tttagagctt gacggggaaa gccggcgaac	8340
	gtggcgagaa aggaaggaa gaaagcgaaa ggagcggcgg ctaggcgcgt ggcaagtga	8400
	gcggtcacgc tgcgcgtaac caccacacc gccgcgctta atgcgccgt acagggcgcg	8460
	tactatggtt gctttagctt atgcggtgtg aaataaccga cagatgcgta aggagaaat	8520
[0028]	accgcacag gcgccattcg ccattcagcc tgcgcaactg ttgggaagg cgatcggctc	8580
	gggcctcttc getattacgc cagctggcga aagggggatg tcttcaagg cgattaagtt	8640
	gggtaacgcc aggttttcc cagtcacgac gttgtaaac gacggccagt gcccagctt	8700
	actagatgca tgcctgagcg gccgccagtg tgatggatat ctgcagaatt cgcccttgac	8760
	tagtgctctc tatcctgggt gcaggcgtca agtaccaga ggcagcagcg ggcttaggag	8820
	cggcctgggt tgttctccgc accctctaca tgctgggcta tatttatagc gacaagccga	8880
	acggcaccgg caggtacaat ggttcgctgt acttgcctgc gcaagcgggt ctttggggat	8940
	tgagcgcatt tgggtttgca aaggatttga tgtaaatgta gtcacatct tagcacagag	9000
	gggagagttg ataaaatgtg gtctgtttga atgatagtcg ggttcgtgac ctalatctgt	9060
	gatagtggag ataggtctgc gcctatctta tcgggcgga gcaaaaatc caccgcagcg	9120
	gggtgagttt tcgttataca gccatcccac ttccagcttc aaattgtcag tttaatccag	9180
	cccaattcaa tcattggaga accgccatca tgtcttcgaa gtcccacctc ccctacgcaa	9240
	ttcgcgcaac caaccatccc aaccctttaa catctaaact cttctccatc gcgaggaga	9300
	agaaaaccaa cgtcacctgc tcgcagacg ttactacttc cggcagctc ctcgatcttg	9360

	ctgaccgtac atcctgcacc aatgcccctc caggataaca aatagctgat gcgtagtgag	9420
	tacaggccta ggcccctata tcgcagttct gaaaaccac atcgacatcc tcaccgatct	9480
	caccccgtcg accctttect cgtccaatc cctcggaca aagcacaact tectcatott	9540
	tgaggaccgc aagttcateg acatcggcaa caccgtgcaa aagcagtacc acggtggcgc	9600
	tctccgcata tcggaatggg cacacatcat caactgcgce atcctgccgg gogaagggat	9660
	cgctcaggcc ctgcacaga caaccaagtc tctgacttt aaagacgga atcaacgagg	9720
	tctcctgatt cttgccgaga tgacagtaa gggatctctt gcgacagggg agtcacaggc	9780
	acgctcggtt gactacgccc ggaagtataa ggggtttgtg atgggatteg tgagtacaag	9840
	ggcgttgagt gaggctctgc ccgaacagaa agaggagagc gaggattttg tcgtctttac	9900
	gactggggtg aatctgtcgg ataaggggga taagcgggg cagcagtacc agacacctgg	9960
	gtcggcggtt gggegaggtg cggactttat cattgcccgt aggggcatct ataaggcgga	10020
	cgatccagtc gaggcggttc agaggtaccg ggaggaaaggc tggaaagctt acgagaaaag	10080
	agttggactt tgagigtgag tggaaatgtg taacggatatt gactaaaagg gatccatag	10140
[0029]	tttattgcag ccagcatagt attaccagaa agagcctcac tgacggctct agtagtattc	10200
	gaacagatat tattgtgacc agctctgaac gatatgctcc ctaatctggt agacaagcac	10260
	tgatctacce cttggaacgc agcatctagg ctctgctgt gctetaacc taactagacg	10320
	attgatcgca gaccatccaa tactgaaaag tctctatcag aggaaatccc caacattgia	10380
	gtagtcaggt tcccttgtgg ctgggagaga attggttcgc tccactgatt ccagttgaga	10440
	aagtgggcta gaaaaagtc ttgaagattg gagttggct gtggttatct agtacttctc	10500
	gagctctgta catgtccggt cgcgacgtac gcgtatcgat ggcgccagct gcaggcggcc	10560
	gcctgcagcc acctgcagtc ccgtggaatt ctccaggta atgtaggect tttgtagggt	10620
	aggaattgtc actcaagcac ccccacctc cattacgct ccccataga gtteccaatc	10680
	agtgagtcac ggcactgttc tcaaatagat tggggagaag ttgacttccg ccagagctg	10740
	aaggctgcac aaccgcatga tatagggtcg gcaacggcaa aaaagcacgt ggetcaccga	10800
	aaagcaagat gtttgcgac taacatccag gaacctggat acatccatca tcacgcacga	10860
	ccactttgat ctgctggtaa actcgtatc gccctaaacc gaagtgcgtg gtaaacttac	10920
	acgtgggccc ctttcggtat actgcgtgtg tcttctctag gtgccattct titeccttc	10980

	tctagtgttg aattgtttgt gttggagtcc gagctgtaac tacctctgaa tctctggaga	11040
	atggtggact aacgactacc gtgcacctgc atcatgtata taatagtgat cctgagaagg	11100
	ggggtttggg gcaatgtggg actttgatgg tcatcaaaca aagaacgaag acgcctcttt	11160
	tgcaaagttt tgtttcgget acgggtgaaga actggatact tgttgtgtct tctgtgtatt	11220
	tttgtggcaa caagaggcca gagacaatct attcaaacac caagcttgct cttttgagct	11280
	acaagaacct gtgggtata tatctagagt tgtgaagtcg gtaatcccgc tgtatagtaa	11340
	taagagtcgc atctaaatac tccgaagctg ctgcgaacct ggagaatcga gatgtgctgg	11400
	aaagcttcta gcgagcggct aaattagcat gaaagctat gagaaattct ggagacggct	11460
	tgttgaatca tggcgttcca ttcttcgaca agcaaagcgt tccgtcgcag tagcaggcac	11520
	tcattcccga aaaaactcgg agattcctaa gttagcgtgg aaccggaata atataatagg	11580
	caatacattg agttgcctcg acggttgcaa tgcaggggta ctgagcttgg acataactgt	11640
	tccgtacccc acctcttctc aacctttggc gtttcctga ttcagcgtac cegtacaagt	11700
	cgtaatcact attaaaccag actgaccgga cgtgttttgc ccttcatttg gagaaataat	11760
[0030]	gtcattgcca tgtgtaattt gcctgcttga ccgactgggg ctgttcgaag cccgaatgta	11820
	ggattgttat ccgaactctg ctctagaggg catgttttga atctgtgtcg ggcaggacac	11880
	gcctcgaagg ttcacggcaa gggaaaccac cgatagcagt gtctagtagc aacctgtaa	11940
	gccgcaatgc agcatcactg gaaaatacaa accaatggct aaaagiact aagttaatgc	12000
	ctaaagaagt catataccag ccgctaataa ttgtacaatc aagtggctaa acgtaccgta	12060
	atttgccaac ggcttgtggg gttgcagaag caacggcaaa gcccacttc cccacgtttg	12120
	tttcttcact cagtcgaatc tcagctgggt atcccccaat tgggtcgett gttgttccg	12180
	gtgaagtgaa agaagacaga ggtaagaatg tetgactcgg agcgttttgc atacaaccaa	12240
	gggcagtgat ggaagacagt gaaatgttga caticcaagga gtatttagcc agggatgctt	12300
	gagtgtatcg tgtaaggagg tttgtctgcc gatacgcaga atactgtata gtcacttctg	12360
	atgaagtggc ccatattgaa atgtaaagtc ggcactgaac aggcaaaaaga ttgagttgaa	12420
	actgcctaag atctcgggcc ctccggcctt ccgccttttg gtgtacatgt ttgtctccg	12480
	ggcaaatgca aagtglggta ggatcgaaca cactgctgcc ttaccacaagc agctgagggt	12540
	atgtgatagg caaatgttca gggccactg catggtttcg aatagaaaaga gaagcttagc	12600

	caagaacaat agccgataaa gatagccca ttaaaccgaa tgagctagta ggcaaagtca	12660
	gcgaatgtgt atatataaag gttcgaggtc cgtgcctccc tcatgtctct cccatctact	12720
	catcaactca gatcctccag gagacttgta caccatcttt tgaggcacag aaaccaata	12780
	gtcaaccatc acaagtttgt acaaaaaagc aggetcacca tgagcaacct gctcaccgtc	12840
	caccagaacc tcctgccct cctgtctgac gccacctctg acgaggtecg caagaacctc	12900
	atggacatgt tccgcgaccg ccaggccttt agcgagcaca cctggaagat gctcctcagc	12960
	gtctgccgat cttgggccgc ctggtgcaag ctcaacaacc gcaagtggtt ccccgccgag	13020
	ccggaggacg tccgcgacta cctcctctac ctgcaggccc gaggcctggc cgtcaagacc	13080
	atccagcagc acctcggcca gctcaacatg ctccaccgac gctctggcct gcctcgcctt	13140
	agcgactcta acgccgtcag cctggctcatg cgccgcatcc gcaaggagaa cgtcagcgtt	13200
	ggcgagcgag ccaageaggc cctcgccttc gagegcaccg acttcgacca ggtccgcagc	13260
	ctcatggaga acagcgaccg ctgccaggat atccgcaacc tcgctttctt cggcattgcc	13320
	tacaacaccc tgcctccgat tgcgagatc gcccgcatcc gcgtaagga catctctcgc	13380
[0031]	accgaeggcg gccgatgct catteacatc ggccgcacca agacctctgt gtctaccgcc	13440
	ggcgtegaga aggcctcag cctcggcgtc accaagctcg tcgagcctg gatttctgtc	13500
	tccggcgctg ctgacgaccc caacaactac ctcttctgcc gcgtccgaaa gaacggcgtc	13560
	gccgccccct ctgccacctc tcagctcagc acccgagccc tggagggcat ctttgaggcc	13620
	accaccgcc tcactctacg cgcgaaggac gactctggcc agcgtacct cgcctggctt	13680
	ggccactctg cccgagtcgg cgtctcccga gacatggccc gagccggcgt cagcatcccc	13740
	gagattatgc aggccggcgg ctggaccaac gtcaacatcg tcatgaacta catccgcacc	13800
	ctcgactctg agaccggcgc catggctcca ctctctgagg acggcgacta aaccagctt	13860
	tc	13862
	<210> 6	
	<211> 6511	
	<212> DNA	
	<213> 人工序列	

<220>

<223> 合成的缺失盒

<400> 6

[0032]

```

gggaggtagg cgcagatacg gtgcattgga cccgaaccgg taaccggaac acgaccttat      60
cagccctcca actcacacce tctcgctat cactatccta gatagtteat cggccaactc      120
atgtaaccta gctacctacc tacctggtaa gaatgcgggc tatcatgtct cacggcgcgg      180
tacatgtcgg tatctcgtg ctccccgca ggttgacgtc ggaatccatg caagtactcc      240
ctgaaatcga gacgacagag agaacaacca acgccttaa acgcttcatg ttcattcgaag      300
aggeacatte gaagaactag cttaacacac tagacctggc ttttcgacc cctccgcaga      360
aagccgtttt ctctcaate ctcccgggtt tggttttgt cagtccttac ttgctgcgct      420
aacagagtct tggacgcagc gtttgcgat cagtcttga gccggttac gggactagga      480
caacagggga tgtgacaggc cggatagtaa ttatgggta tccgggtaa gcaggaatt      540
tacgaggcgg ctctacgtgg gggaacagcc acttgcgggg ggaagaggag tagtaggcca      600
ctcggtcgat gagctcagg tgtctggtt acttggactg cagagcgtag gtaattgaga      660
tcgggcaaca ttatcgggtt tcggctcggg atggccgagt tgcgactgct tggteattcg      720
gcgaagctga tgcctggta tectgaagca tcgatatcgg aaacctatg ggtcagtcta      780
tctgacgtgt gcggtgaaa gcgagtcgg atttgtgac atgacgttca acttcagtca      840
atgccttagg gctcgataag attaagattg gttctggca gcggtctaga acaccgccac      900
aaattctgtc cattgaggag cgtgatgtct aggcgatca ctaacacgga gctgtatgac      960
cggcagcica acggacttet ctctgtcaa cggcagtcta tttgcggtac acgaatgat      1020
ctttcttctt ggtcttgaag tgcgcagtg gcgtgcgaat gtatagatgt ctcgctacct      1080
agaaaagctg gctttctga cagggtcctt tccacctctc ctaccaacga caaactgaac      1140
aaglatctgg cggtttcca acgccgaata ggccagtcgc caatactccc tccagccctg      1200
attggcccc tcgaagtatc gccatgtctg tgtgtgaga ttattgatg gacgtcactc      1260
ccccaaccta caggaagage aaaatgggag cagtgttctg caatgagcta tataatagat      1320

```

	cgctcgatct catacaaatt gtatgetcag tcaatacaac gagcggttcc aagatccctt	1380
	ctccaaegac cctcgaaaca ttgcaacecg gtgeagcctg aactigtteg tatagcctag	1440
	aaagcgacgc catcttcate ttttacgcga ttagcctcat ggctatttgt gccgaagtgg	1500
	gagttgtaig gtagcagtga ggagattgtg getacgacac aggegggttc tcttgagcgg	1560
	cttacatcic cgcattagge ctgcgtacga tecagatcat gggaaacttt acaatggctt	1620
	actcgtttta tctcaacact gagcttccaa ttcactctat gcattgatta acacgtttgg	1680
	tcatgtggtt cticagctgt aaatcttcag ctteccaaga attgcaacct cgctgattgc	1740
	taatagtgtt gcatgogttg catcctggtg cggcagtgca aaggagagtc aaagtagccg	1800
	gcagattaat ttaagcttat atcactcagg ggtaaacagc cgtaaaggac cttttgatct	1860
	aacatgccga tgtgtatgta gatcacgcaa tgcccacat atcttggcag tcagatttgt	1920
	ccgtggcgcg ccaagtataa cttcgtataa tgtatgctat acgaagtat cggccggcgt	1980
	attgggtgtt acggagcatt cactaggcaa ccatggttac tattgtatac ccatcttagt	2040
	aggaatgatt ttcgaggitt atacctacga tgaatgtgtg tctgttaggc ttgagagttc	2100
[0033]	aaggaagaaa cagtcaatt atctttgcga acccaggggc tggtagcgga attttcatag	2160
	tcaagctatc agagtaaaga agaggagcat gtcaaagtac aattagagac aaatatatag	2220
	tcgctggag ccaagagcgg attcctcagt ctcgtaggtc tcttgacgac cgttgatctg	2280
	cttgatctcg tctcccgaat atgaaaatag actctgctaa getattcttc tgetteccg	2340
	gagcctgaag ggcglactag ggttgcgagg tccaatgcat taatgcattg cagatgagct	2400
	gtatctggaa gaggtaaacc cgaaacggt tttattcttg ttgacatgga getattaaat	2460
	cactagaagg cactctttgc tgcctggaca aatgaacgta tcttatcgag atcctgaaca	2520
	ccatttgict caactccgga getgacatcg acaccaacga tcttatatcc agattcttca	2580
	agctgtttga tgatttcagt aacgttaagt ggatcccggt cggcatctac tctattcctt	2640
	tgccctcgga cgagtgetgg ggcgtcggtt tccactatcg gcgagtactt ctacacagcc	2700
	atcggctcag acggccgcgc ttcctcgggc gatttgtgta cgcgccagag tcccggctcc	2760
	ggatcggacg attgcgtcgc atcgacctg cgcceaagct gcatcatcga aattgccttc	2820
	aaccaagctc tgatagagtt ggtcaagacc aatgcggagc atatacgccc ggagccgcgg	2880
	cgatcctgca agctccggat gcctccctc gaagtagcgc gtctgtctgt ccatacaagc	2940

	caaccacgge ctccagaaga agatgttgge gacctcgtat tgggaatccc cgaacatcgc	3000
	ctcgctccag tcaatgaccg ctgttatgcg gccattgtcc gtcaggacat tgttgagacc	3060
	gaaatccgcg tgcacgaggt gccggacttc ggggcagtec teggeccaaa gcatcagctc	3120
	atcgagagcc tgcgcgacgg acgcactgac ggtgtcgtcc atcacagttt gccagtgata	3180
	cacatgggga tcagcaatcg cgcataatgaa atcacgccat gtagtgtatt gaccgattee	3240
	ttgcggtcgg aatgggcoga acccgctcgt ctggctaaga tcggccgcag cgatcgcate	3300
	catggcctcc gcgaccggct gcagaacagc gggcagttcg gtttcaggca ggtcttgcga	3360
	cgtgacacce tgtgcacggc gggagatgca ataggtcagg ctctcgtcga attcccgaat	3420
	gtcaagcact tccggaatcg ggagcggcgc cgatgcaaag tggcgataaa cataacgac	3480
	ttttagaaaa ccatcggcgc agctatttac ccgcaggaca tatccacgcc ctctacate	3540
	gaagctgaaa gcacgagatt cttegccctc cgagagctgc atcaggtcgg agacgctgc	3600
	gaacttttcg atcagaaaact tctcgacaga cgtcgcggtg agttcaggt ttttcata	3660
	gatigtgatg tgatggagtt gagatggagg tgaggagatg gatgatggga aaggaagatg	3720
[0034]	gactgaggat ggaagaagag aagaagagag agagagaaag tcttccagga gagaaaggga	3780
	accgaagaaa aatggggagg aaaccggccc tagcaactaa atacgtctcc cgcttaattt	3840
	tegccctttt ttaccaaacc cctccgcgtc tttegtgcgc tagctgtctt ggggggtgtg	3900
	taaaacttgg gaacaacct acgccgaacc tcccgtacga agcccgtaca gtgtattcta	3960
	tccttggett tcccagectg gtaagtcggg tcccttcggg acggggccaa ggagactgag	4020
	ttccggggtt aaccaataat gccgccagcc gtggagcggc ctgagctgct tatcgtgaat	4080
	ccgtgacgct gaattgctca gtccaagtcg gagacgctgg aatccaccgg ttgctccagc	4140
	cacggcgaag aatccactta cttccgggc ttcgccagc tggcaacata tttttgaggc	4200
	tgatatagcc ttctctcca tcacgataag ccttgactac cttgetacgg gccaatgca	4260
	attgtttgct tggttacgct ttacaaaagg tggccgttac tgagcaaaaa gaaaatgatg	4320
	taaaaatttg cgagtgggct ccatagctgg atgggtccga taaaatggta ctgccccact	4380
	tagtggcagc tcgcgaccag tcacaagccc aggataaact cgtataatgt atgctatacg	4440
	aagttatctg tgggcgttat gaataataga ctggaaccgg gccctttgat tgacgactcc	4500
	atattttgta gatgtagcaa ctccgcaaga gcattatgtg caatacattt gttaccatac	4560

	aaaggeagct gccagacgac ttgtattgcg tacaaltctc acggcaagct ttccaggtgt	4620
	taigcattat gcgcaaatge ttgatgctta ccgcaggatt aatctcggaa gaagecgtgc	4680
	aagctatatt ggtgtagtag atatgtagat giaccaacca atgaagaaca tttatggtct	4740
	agaacgiagt gatgaaggtt ttgagtaatt tgtatcaagt aagacgatat taitgatata	4800
	ataccaagca tatattcatg ataaattact tggaaccace ctgcegtcgg gcctcacgag	4860
	ccttctcact gccgggctcg aaggagccac tggaggcctg tccacccttg gatgcgattt	4920
	cctgcacctt ttctttgggc ctgcacgtcg attagacatg attcaaatcg agatcttggg	4980
	atatcttaca tgctggcgaa gccaccggtg tggctggact gtcgcacctt ctgcgcaatg	5040
	ctttgaaact cctccttggg gctgtgtaga aaggtttgtt agcaacatta ctacaactct	5100
	caggactcgg tggctgtacc ggttggcgaa gtttccgggg ttatcgttgc cagacattgt	5160
	gtgattattt ggtgtgcaaa tgtgtgctat gtgtgttgtt gctgttgggt atgatgetga	5220
	agctgttgaa agcaggctgg ttctgtggga gagacttggg atatttata ccaaagttcg	5280
	gtcgtgttcc ttctggaagc tcttctctac tccatacaat catccaaagt tctcgtcatt	5340
[0035]	gagcgttgat cagtagtagc ctctgaggte atcaecatga tcttccggc caacagtcgg	5400
	cactcatcaa cagcaacaat cagccgccac aaacataggt acagtaagga gttagatata	5460
	atgtagtctg cgagtactcg acatcatgac gtacaagett tgccagtgtc ggtaggtgca	5520
	agfatgatga tcgtatccgc cgttgttcca tcgaacagag tgcggtcaga ttcacggtt	5580
	ctctcacctt gaacattgga tgcaattgga ttgatccaca atcctggaga atggettcaa	5640
	gctcactgct ccagtcgcaa gcttcagagc ctattactaa gggtagagct acctatgtca	5700
	agagttttca aggtacctaa gctacatgtg atagtcggca agccattttg aacgcagacc	5760
	gtgaacgggtg atgtaaatcc gggatagaag cccaagcgtg ccgltgcaat gacgctagat	5820
	acacctgat ttacgiagag tgaatgccag ccaatggagt catgcacata acccgcttag	5880
	actctgctcg gggcgatacc cgatcgaga ggcagagccg cttaaaccgg atcgcggtaa	5940
	cctgtaatca gagccagcgc tcgatgaatt gcatcatgga agccattgat gtggaatgtt	6000
	gagcgtataa caacacgaat tgaagacgac attgacttgc ttcaagtggg tggagaattg	6060
	ccgggcagac aagataggta ggcctttggt gcgctgtcac atcaatecat tcttttctct	6120
	ctgttcaate tetatgttga cattctgata gggatcattg gatgccaatg caaagaacat	6180

gagagtgtgg tctgcattca agtaicctgg tcgtaagctg tggccatggg cgctgcggtc 6240
aaggtcaatc gcgatgacta atcagtcctg gtgactctgg ggcggtagag gcagtgctgt 6300
gaaccaaagc ttgagccgag ggcaaaaaca acggcgcata aaacaatcaa cgaaagcacc 6360
gtcaacagtg tctcttccca gtcaattact tcgcaaaacc ttctcgatag aacccttcag 6420
acgatgaaca ggccacgcaa ccgtcagccg cgccecccag gacagactca gcgcccggga 6480
ggcagatcgt cacaccttgg tcgacgagct c 6511

<210> 7

<211> 1416

<212> DNA

<213> 红褐肉座菌(*Hypocrea jecorina*)

<400> 7

[0036] atgattgtcg gcattctcac cacgttggt acgttggcca cactcgcagc tagtgtgcct 60
ctagaggagc ggcaagcttg ctcaagcgtc tggggccaat gtggtggcca gaattggctc 120
ggcccgactt gctgtgcttc cggagcaca tgegtctact ccaacgaata ttactcccag 180
tgtcttcccg gcgctgcaag ctcaagctcg tccacgcgcg ccgctcgcac gacttctega 240
gtatccccca caacatcccg gtcgagctcc gegacgcctc cacctgggtc tactactacc 300
agagtaacctc cagtcggatc gggaaccgct acgtattcag gcaacccttt tgttggggtc 360
actccttggg ccaatgcata ttacgcctct gaagttagca gctcgcctat tcttagcttg 420
actggagcca tggccactgc tgcagcagct gtcgcaaagg ttccctcttt tatgtggcta 480
gatactcttg acaagacccc tctcatggag caaaccttgg ccgacatccg caccgccaac 540
aagaatggcg gtaactatgc eggacagttt gtggtgtatg acttgccgga tcgcgattgc 600
gctgcccttg cctcgaatgg cgaatactct attgccgatg gtggcgtegc caaatataag 660
aactatatcg acaaccctcg tcaaattgtc gtggaatatt ccgatatecg gaccctctcg 720
gttattgagc ctgactctct tgcacaactg gtgaccaacc tcggtactcc aaagtgtgcc 780
aatgetcagt cagcctacct tgagtgcata aactacgccg tcacacagct gaaccttcca 840

aatgttgcga tgtatttga cgctggccat gcaggatggc ttgctggcc ggcaaaccaa 900
 gaccggccg ctcagctatt tgcaaatgt tacaagaatg catcgtccc gagagctctt 960
 cgcgattgg caaccaatgt cgccaactac aacgggtgga acattaccag cccccatcg 1020
 tacacgaag gcaacgctgt ctacaacgag aagctgiaca tccacgctat tggacctctt 1080
 cttgccaatc acggetggtc caacgccttc ttcactactg atcaaggctg atcgggaaag 1140
 cagcctaccg gacagcaaca gtggggagac tgggtcaatg tgatcggcac cggatttgg 1200
 attegccat ccgcaaacac tggggactcg ttgctggatt cgtttgtctg ggtaagcca 1260
 ggcggcgagt gtgacggcac cagcgacagc agtgcgccac gatttgactc ccactgtgcg 1320
 ctcccagatg cettgcaacc ggcgcctcaa getgggtgctt ggttccaagc ctactttgtg 1380
 cagcttctca caaacgcaaa cccatcgttc ctgtaa 1416

<210> 8

<211> 471

[0037] <212> PRT

<213> 红褐肉座菌

<400> 8

Met Ile Val Gly Ile Leu Thr Thr Leu Ala Thr Leu Ala Thr Leu Ala
 1 5 10 15
 Ala Ser Val Pro Leu Glu Glu Arg Gln Ala Cys Ser Ser Val Trp Gly
 20 25 30
 Gln Cys Gly Gly Gln Asn Trp Ser Gly Pro Thr Cys Cys Ala Ser Gly
 35 40 45
 Ser Thr Cys Val Tyr Ser Asn Asp Tyr Tyr Ser Gln Cys Leu Pro Gly
 50 55 60
 Ala Ala Ser Ser Ser Ser Thr Arg Ala Ala Ser Thr Thr Ser Arg
 65 70 75 80

Val Ser Pro Thr Thr Ser Arg Ser Ser Ser Ala Thr Pro Pro Pro Gly
 85 90 95
 Ser Thr Thr Thr Arg Val Pro Pro Val Gly Ser Gly Thr Ala Thr Tyr
 100 105 110
 Ser Gly Asn Pro Phe Val Gly Val Thr Pro Trp Ala Asn Ala Tyr Tyr
 115 120 125
 Ala Ser Glu Val Ser Ser Leu Ala Ile Pro Ser Leu Thr Gly Ala Met
 130 135 140
 Ala Thr Ala Ala Ala Ala Val Ala Lys Val Pro Ser Phe Met Trp Leu
 145 150 155 160
 Asp Thr Leu Asp Lys Thr Pro Leu Met Glu Gln Thr Leu Ala Asp Ile
 165 170 175
 Arg Thr Ala Asn Lys Asn Gly Gly Asn Tyr Ala Gly Gln Phe Val Val
 [0038] 180 185 190
 Tyr Asp Leu Pro Asp Arg Asp Cys Ala Ala Leu Ala Ser Asn Gly Glu
 195 200 205
 Tyr Ser Ile Ala Asp Gly Gly Val Ala Lys Tyr Lys Asn Tyr Ile Asp
 210 215 220
 Thr Ile Arg Gln Ile Val Val Glu Tyr Ser Asp Ile Arg Thr Leu Leu
 225 230 235 240
 Val Ile Glu Pro Asp Ser Leu Ala Asn Leu Val Thr Asn Leu Gly Thr
 245 250 255
 Pro Lys Cys Ala Asn Ala Gln Ser Ala Tyr Leu Glu Cys Ile Asn Tyr
 260 265 270
 Ala Val Thr Gln Leu Asn Leu Pro Asn Val Ala Met Tyr Leu Asp Ala
 275 280 285
 Gly His Ala Gly Trp Leu Gly Trp Pro Ala Asn Gln Asp Pro Ala Ala

290 295 300
 Gln Leu Phe Ala Asn Val Tyr Lys Asn Ala Ser Ser Pro Arg Ala Leu
 305 310 315 320
 Arg Gly Leu Ala Thr Asn Val Ala Asn Tyr Asn Gly Trp Asn Ile Thr
 325 330 335
 Ser Pro Pro Ser Tyr Thr Gln Gly Asn Ala Val Tyr Asn Glu Lys Leu
 340 345 350
 Tyr Ile His Ala Ile Gly Pro Leu Leu Ala Asn His Gly Trp Ser Asn
 355 360 365
 Ala Phe Phe Ile Thr Asp Gln Gly Arg Ser Gly Lys Gln Pro Thr Gly
 370 375 380
 Gln Gln Gln Trp Gly Asp Trp Cys Asn Val Ile Gly Thr Gly Phe Gly
 385 390 395 400
 [0039] Ile Arg Pro Ser Ala Asn Thr Gly Asp Ser Leu Leu Asp Ser Phe Val
 405 410 415
 Trp Val Lys Pro Gly Gly Glu Cys Asp Gly Thr Ser Asp Ser Ser Ala
 420 425 430
 Pro Arg Phe Asp Ser His Cys Ala Leu Pro Asp Ala Leu Gln Pro Ala
 435 440 445
 Pro Gln Ala Gly Ala Trp Phe Gln Ala Tyr Phe Val Gln Leu Leu Thr
 450 455 460
 Asn Ala Asn Pro Ser Phe Leu
 465 470
 <210> 9
 <211> 447
 <212> PRT

<213> 红褐肉座菌

<400> 9

Gln Ala Cys Ser Ser Val Trp Gly Gln Cys Gly Gly Gln Asn Trp Ser

1 5 10 15

Gly Pro Thr Cys Cys Ala Ser Gly Ser Thr Cys Val Tyr Ser Asn Asp

 20 25 30

Tyr Tyr Ser Gln Cys Leu Pro Gly Ala Ala Ser Ser Ser Ser Ser Thr

 35 40 45

Arg Ala Ala Ser Thr Thr Ser Arg Val Ser Pro Thr Thr Ser Arg Ser

 50 55 60

Ser Ser Ala Thr Pro Pro Pro Gly Ser Thr Thr Thr Arg Val Pro Pro

65 70 75 80

[0040] Val Gly Ser Gly Thr Ala Thr Tyr Ser Gly Asn Pro Phe Val Gly Val

 85 90 95

Thr Pro Trp Ala Asn Ala Tyr Tyr Ala Ser Glu Val Ser Ser Leu Ala

 100 105 110

Ile Pro Ser Leu Thr Gly Ala Met Ala Thr Ala Ala Ala Ala Val Ala

 115 120 125

Lys Val Pro Ser Phe Met Trp Leu Asp Thr Leu Asp Lys Thr Pro Leu

 130 135 140

Met Glu Gln Thr Leu Ala Asp Ile Arg Thr Ala Asn Lys Asn Gly Gly

145 150 155 160

Asn Tyr Ala Gly Gln Phe Val Val Tyr Asp Leu Pro Asp Arg Asp Cys

 165 170 175

Ala Ala Leu Ala Ser Asn Gly Glu Tyr Ser Ile Ala Asp Gly Gly Val

 180 185 190

Ala Lys Tyr Lys Asn Tyr Ile Asp Thr Ile Arg Gln Ile Val Val Glu
195 200 205

Tyr Ser Asp Ile Arg Thr Leu Leu Val Ile Glu Pro Asp Ser Leu Ala
210 215 220

Asn Leu Val Thr Asn Leu Gly Thr Pro Lys Cys Ala Asn Ala Gln Ser
225 230 235 240

Ala Tyr Leu Glu Cys Ile Asn Tyr Ala Val Thr Gln Leu Asn Leu Pro
245 250 255

Asn Val Ala Met Tyr Leu Asp Ala Gly His Ala Gly Trp Leu Gly Trp
260 265 270

Pro Ala Asn Gln Asp Pro Ala Ala Gln Leu Phe Ala Asn Val Tyr Lys
275 280 285

Asn Ala Ser Ser Pro Arg Ala Leu Arg Gly Leu Ala Thr Asn Val Ala
[0041] 290 295 300

Asn Tyr Asn Gly Trp Asn Ile Thr Ser Pro Pro Ser Tyr Thr Gln Gly
305 310 315 320

Asn Ala Val Tyr Asn Glu Lys Leu Tyr Ile His Ala Ile Gly Pro Leu
325 330 335

Leu Ala Asn His Gly Trp Ser Asn Ala Phe Phe Ile Thr Asp Gln Gly
340 345 350

Arg Ser Gly Lys Gln Pro Thr Gly Gln Gln Gln Trp Gly Asp Trp Cys
355 360 365

Asn Val Ile Gly Thr Gly Phe Gly Ile Arg Pro Ser Ala Asn Thr Gly
370 375 380

Asp Ser Leu Leu Asp Ser Phe Val Trp Val Lys Pro Gly Gly Glu Cys
385 390 395 400

Asp Gly Thr Ser Asp Ser Ser Ala Pro Arg Phe Asp Ser His Cys Ala

405	410	415
Leu Pro Asp Ala Leu Gln Pro Ala Pro Gln Ala Gly Ala Trp Phe Gln		
420	425	430
Ala Tyr Phe Val Gln Leu Leu Thr Asn Ala Asn Pro Ser Phe Leu		
435	440	445

<210> 10

<211> 24

<212> DNA

<213> 人工序列

<220>

<223> 合成的引物

[0042]

<400> 10

gaccggacgt gttttgccct tcat

24

<210> 11

<211> 22

<212> DNA

<213> 人工序列

<220>

<223> 合成的引物

<400> 11

gtgtgaccgg etttggcgag tg

22

<210> 12
 <211> 4979
 <212> DNA
 <213> 人工序列

<220>

<223> 合成的 PCR 片段

<400> 12

[0043]

```

gaccggacgt gttttgccct tcatttggag aaataatgtc attgcgatgt gtaatttgcc      60
tgcttgaccg actggggctg ttcgaagccc gaatgtagga ttgttatccg aactctgctc      120
gtagaggcat gttgtgaatc tgtgtcgggc aggacacgcc tcgaaggttc acggcaaggg      180
aaaccaccga tagcagtgtc tagtagcaac ctgtaaagcc gcaatgcagc atcactggaa      240
aatacaaaacc aatggctaaa agtacataag ttaatgccta aagaagtcac ataccagcgg      300
ctaataattg tacaatcaag tggctaaacg taccgtaatt tgccaacggc ttgtggggtt      360
gcagaageaa eggcaaagcc ccaactcccc acgtttgttt cticactcag tccaatctca      420
gctggtgatc ccccaatigg gtcgcttggt tgttccgggt aagtgaaga agacagaggt      480
aagaatgtct gactcggagc gttttgcata caaccaaggg cagtgatgga agacagtgaa      540
atgttgacat tcaaggagta tttagccagg gatgcttgag tgtatcgtgt aaggaggttt      600
gtctgccgat acgacgaata ctgtatagtc acttctgatg aagtggteca tattgaaatg      660
taaagtcggc actgaacagg caaaagattg agttgaaact gccaaagatc tcgggcctc      720
gggccttcgg cctttgggtg tacatgtttg tgctccgggc aaatgcaaag tgttgtagga      780
tcgaacacac tgctgccctt accaagcagc tgagggtatg tgataggcaa atgttcaggg      840
gccactgcat ggtttcgaat agaaagagaa gcttagccaa gaacaatagc cgataaagat      900
agcctcatta aacggaatga gctagtaggc aaagtcagcg aatgtgtata tataaaggtt      960
cgagggtccgt gccctccctca tgctctcccc atctactcat caactcagat cctccaggag     1020

```

	actgtiacac catcttttga ggcacagaaa cccaatagtc aaccatcaca agtttgtaca	1080
	aaaaagcagg ctccgcggcc gccccttca cccaccatga ttgtcggcat tctcaccacg	1140
	ctggctacgc tggccacaact cgcagctagt gtgcctctag aggagcggca agcttgcctca	1200
	agegtctggg aattatgta accctctcaa gagaccaaa tactgagata tgcaagggg	1260
	ccaatgtggg gccagaatt ggtcgggtcc gacttgcctg gcttccggaa gcacatgcgt	1320
	ctactccaac gactattact cccagtgtct tcccggcget gcaagctcaa gctcgtccac	1380
	gcgcgccgcg tcgacgaact ctcgagtatc ccccacaaca tcccggctga gctccgcgac	1440
	gcctccacct ggttctaact ctaccagagt acctccagtc ggatcgggaa ccgctacgta	1500
	ttcaggcaac cttttgttg gggtcactcc ttgggccaat gcatattacg cctctgaagt	1560
	tagcagcctc gctattceta gcttgactgg agccatggcc actgctgcag cagctgtcgc	1620
	aaagttccc tcttttatgt ggctgtaggt cctcccggaa ccaaggcaat ctgttactga	1680
	aggctcatca ttactgcag agatactctt gacaagacce ctctcatgga gaaaaccttg	1740
	gccgacatcc gcaccgcaa caagaatgga gtaactatg ccggacagtt tgggtgtat	1800
[0044]	gacttgccgg atcgcgattg cgtgcctctt gctcgaatg gcaataactc tattgccgat	1860
	ggtggcgtcg ccaaatataa gaactatctc gacaccattc gtcaaatgt cgtggaatat	1920
	tccgatatcc ggaccctcct ggttattggt atgagtttaa acaactgcct ccccccccc	1980
	ttcccttctt tcccgcggc catcttcteg ttgtgctaac tattgttccc tctccagag	2040
	cctgactctc ttgccaacct ggtgaccaac ctccgtactc caaagtgtgc caatgctcag	2100
	tcagcctacc ttgagtgcct caactacgcc gtcacacagc tgaaccttc aaatgttggc	2160
	atgtatttgg acgctggcca tgcaggatgg cttggctgga cggcaaacca agaccggcc	2220
	gctcagctat ttgcaaatgt ttacaagaat gcatcgtctc cgagagctct tcgctggattg	2280
	gcaaccaatg tcgccaacta caacgggtgg aacattacca gcccccatc gtacacgcaa	2340
	ggcaacgctg tctacaacga gaagctgtac atccacgcta ttggacctct tcttccaat	2400
	cacggctggg ccaacgcctt ctctcact gatcaaggtc gatcgggaaa gcagcctacc	2460
	ggacagcaac agtggggaga ctggtgcaat gtgatcggca ccggatttgg tattcgcaca	2520
	tccgaaaaca ctggggactc gttgctggat tcgtttgtct gggtaagcc aggcggcgag	2580
	tgtgacggca ccagcgacag cagtgcgcca cgatttgact cccactgtgc gctcccagat	2640

	gccttgcaac egggccctca agctgggtgct tggttccaag cctactttgt gcagcttctc	2700
	acaaacgcaa acccatcggt cctgtaaaag ggtgggcgag cccgaccagc ttctctgtac	2760
	aaagtgggga tcgcgccgag ccccagctcc gtcgaaaage ctgacgcacc ggtagattct	2820
	tggtagccc gtatcatgac ggcggcggga gctacatggc cccgggtgat ttatTTTTT	2880
	tgtatctact tctgacctt ttcaaatata cggccaactc atctttcact ggagatcggg	2940
	cctgcttggg attgcgatgt tctcagcttg gcaaatgtg gctttcgaaa acacaaaacg	3000
	attccttagt agccatgcat tttaagataa cggaaataga gaaagaggaa attaaaaaaa	3060
	aaaaaaaaac aaacatcccg ttcataaccg gtagaatcgc cgtcttctgt gtatcccagt	3120
	accagtttat ttgaaatagc tcgcccgtg gagagcatcc tgaatgcaag taacaaccgt	3180
	agaggctgac accgcagggt ttgctaggga gcgtcgtgtt ctacaaggcc agacgtcttc	3240
	gcggttgata tatatgtaig ttgactgca ggctgctcag cgaacacagt caagttgcc	3300
	ctcgtctctt gtgcaataat cgcagtgagg aagccacacc gtgactcca tctttcagta	3360
	aagctctgtt ggtgtttatc agcaatcac gtaattitaa ctgcttagca tggggctgat	3420
[0045]	agcttaatta ccgttiacca gtgccatggt tctgcagctt tctttggccc gtaaaatcgc	3480
	gcgaagccag ccaatcacca gctaggcacc agctaaacc tataattagt ctcttatcaa	3540
	caccatccgc tccccggga tcaatgagga gaatgagggg gatgcggggc taaagaagcc	3600
	tacataacc tcatgccaac tcccagtta cactcgtcga gccaacatcc tgactataag	3660
	ctaacacaga atgccicaat cctgggaaga actggccgct gataagcgcg cccgcctcgc	3720
	aaaaaccatc cctgatgaat ggaaagtcca gacgctgect gcggaagaca gcgttatgta	3780
	ttcccgaag aaatcgggga tctttcaga ggccgaactg aagatcacag aggcctccgc	3840
	tgcagatctt gtgtccaagc tggcggccgg agagttgacc tcggtggaag ttacgctagc	3900
	attctgtaaa cgggcagcaa tcgcccagca gttagtaggg tccccctac ctctcagga	3960
	gatgtaaaa cgcacctta tggactatc aagctgacgc tgcttctgt gcagacaaac	4020
	tgcgccacg agttcttccc tgacccgct ctgcgcagg caagggaact cgatgaalac	4080
	tacgcaaagc acaagagacc cgttggcca ctccatgccc tcccctctc tctcaaagac	4140
	cagcttcgag tcaaggtaca ccgttcccc taagtcgta gatgtccctt ttgtcagct	4200
	aacatatgcc accagggeta cgaaacatca atggctaca tctcatggct aaacaagtac	4260

gacgaagggg actcggttct gacaaccatg ctccgcaaag ccggtgccgt cttctacgtc 4320
 aagacctctg tcccgcagac cctgatggtc tgcgagacag tcaacaacat catcgggccc 4380
 accgtcaacc cagcgaacaa gaactggctc tgcggcggca gttctgggtg tgagggtgag 4440
 atcgttggga ttctgggtgg cgtcatcggg gtaggaacgg atatcgggtg ctcgattcga 4500
 gtgccggccg cgttcaactt cctgtaaggc ctaaggccga gtcattggcg gctgccgtat 4560
 gcaagatgg cgaacagcat ggagggtcag gagaccggtc acagcgttgt cgggccgatt 4620
 acgeactctg ttgagggtga gtcttccgct tcttcttctt ttctctctc tataaccaggc 4680
 ctccactgtc ctcttttctt gttttttata ctatatacga gaccggcagt cactgatgaa 4740
 gtatgttaga cctcgcctc ttcaccaaact ccgtcctcgg tcaggagcca tggaaatagc 4800
 actccaaggc catcctccatg cctggcggcc agtcgagtc ggacattatt gcctccaaga 4860
 tcaagaacgg cgggctcaat atcggctact acaacttoga cggcaatgtc cttecacacc 4920
 ctctatctct gcgcgccgtg gaaaccaccg tcgcccactt cggcaagcc ggtcacacc 4979

[0046]

<210> 13

<211> 2593

<212> DNA

<213> 里氏木霉(Trichoderma reesei)

<400> 13

atggcggaca aggaagcaac cgtcttcac atcgacctc gcgcgtccat ggcagctgtc 60
 aatgggggtc gagaagaatc cgaccttgat tggagcatga gctacgtctg ggacaagatc 120
 agcaacgtcg tggcctogaa tgcgaagacg ctgtgcgttg gcgtcgtggg gttcagaacc 180
 gacgagacaa accacacgct gagcgaggat gggtagcaga acatctccat attgcagccc 240
 ctggggccga tgagcatgtc cagcctcaag gctcttcagc ccaaggtgaa gccgagcagg 300
 acgggtggaag gcgatgccat ctccggcatt gtcattgccg tcgacatgat tgacaagtac 360
 acgaagaaga acaaatggaa gcggcagatt gttctcatta ccgacggcca aggcgagatt 420
 gatccagatg atattggcga cattgctaga aagatgcgag actcgaatat tgaattgaca 480

gtcttgtag	ttggcgagac	cgtttggcgg	acggtaatgg	tgctgacggt	gatgcaaggg	540
gcgtcgactt	tgatgctccc	gattacggct	tcaaagagga	ggacaaacct	tcagtcgaag	600
tactccatat	gttcacttct	tttcttttct	ttctttatct	tcttttcttt	tgaagcttct	660
attaacctct	tcgttagaag	caaaacgaag	agaccctaaa	aaagctcgtg	gatggctgtg	720
gcgacgactc	aaggitcgcc	tccatggctg	aggccattga	cgacttgaat	gagccacgag	780
caaagtcggt	caagccttac	aaaacgtacg	aaggctctct	gaccttgga	gatccgaaaa	840
acgctcccgc	agtggtggaa	atccgcgtcg	agagatactt	caagacctat	ctagccaggg	900
caectgcccg	cagcaccgtg	gtggcaagg	aggagcaagc	tgggccgtct	caggcagacg	960
aggacgaaca	gatggacgga	gcggaactta	cagctgtgag	gcaggccagg	acatacaagg	1020
tcaatgatcc	agatgcccct	ggcggtaagc	gtgacgttga	gtttgagtct	ctggccaaa	1080
ggtacgagta	cggcaggacg	gcagtcacac	tcagcgagtc	tgatcaaaa	gtcaccgaag	1140
tcgcgacaga	aaagagcttc	aagatcatcg	gcttcgtcca	gaaagaaaag	gtattggctt	1200
ggctctcagc	atctgaccgg	ttgctcttgg	ctaacccttg	tttagtatga	aatgctcctt	1260
[0047]	aatcttggcg	aaacctgcgt	taccgttcca	tccaagtacg	atgaaaagtc	1320
tttagctctc	tggtgtgggc	gctctcggag	ctcgacgctt	acgccgtggc	ccgcctagta	1380
actaaggacc	aaaaggaccc	catgctgggt	ttactgatgc	cgtatatgga	gcctgattat	1440
gtttgtctct	atgatgtgcc	tctgccttcc	gcagaggaca	tcaggacgta	ccagtttctt	1500
cccttggaca	gagtcgttac	cgtcagtggc	caaacgctca	ccaaccatcg	cctattgcca	1560
tccgacgagc	tcaaccaagc	gatgagcgac	tacgtagaig	ccatggacat	ttcaagttat	1620
ggtatcgatg	aagatgggtg	agtatagaag	atgattgttc	aaatctttca	cttctaagca	1680
ttgcttctga	tctaggcaac	cggctgaata	tgccaccatc	gatgagttat	acaaccttgc	1740
gatacatcgc	ataggccaig	cgatcaaaac	acgagcgatc	caccagaga	aacccgtgcc	1800
cgagatcccc	ccagtcttgc	ttagattcgc	agcaccctcg	acagaactcg	tcgagactgt	1860
gcagccicat	atcgatgcac	tgattcacgc	tgcagacgtg	aagaaaggta	ctgattccat	1920
tacatatget	tctctgcaca	ctgatgtttg	atttgtgcta	acgccccctt	tagtgcgcgc	1980
caaggccaag	ggcaagcgcc	aaagagaaac	agttaaacc	atctcgggac	tgatgtgga	2040
tgcccttctg	ggagaagagc	agaaggcttc	cattagtccg	gagaatgcca	ttccggactt	2100

caaacgagcc ctcaactcgt ccgaagaagt cgagcagatt gccgacgcca caaaacaaat 2160
 gggggccatt gtgeggcttc tcattacgga cagcttcggg gatagcaaat atgcccagge 2220
 aatggaaggc attggtgcga tgcgtgagga gctgatcaac ctggaagage ctggcctgta 2280
 caacgacttt gtgcgcgact tgaagaaaag ttigtatct ggagccttgg gtggtgacag 2340
 gcgagatttc tggttcaaga tgaggtgggc gaagctgggc ctgattgaca agaaacagtc 2400
 ggaggtgtct tcggtcactc ttgaggagc ggacgaggtg agtgggtgcag catgctgtcg 2460
 gattatacgg acgttgtttg ctaacttggg ggatagtttt acaagtcgag gtgaggtatc 2520
 tacgttgacc aagaatggga ccatgtatat gagcgggtga acaacagaat cctgtgcttt 2580
 gagcattgta tga 2593

<210> 14

<211> 8002

<212> DNA

[0048] <213> 人工序列

<220>

<223> 合成的表达载体

<400> 14

ttgtacaaag tggatgatgc gccgcgcgcc agctccgtgc gaaagcctga cgcaccggta 60
 gattcttggg gagcccgtat catgacggcg gcgggagcta catggccccg ggtgatttat 120
 tttttttgta tetacttctg acccttttca aatatacggg caactcatct ttcactggag 180
 atgcggcctg ctgggtattg cgatgttgtc agcttggcaa attgtggctt tcgaaaacac 240
 aaaacgattc cttagtagcc atgcatttta agataacgga atagaagaaa gaggaaatta 300
 aaaaaaaaa aaaaacaaac atcccgttca taaccgtag aatcggcgtc cttcgtgtat 360
 cccagtacca gtttatttig aatagctcgc ccgctggaga gcatcctgaa tgcaagtaac 420
 aaccgtagag gctgacacgg cagggtgtgc tagggagcgt cgtgttctac aaggccagac 480

	gtcttcgctg ttgatata tgatgtttg actgcagget gctcagcgac gacagtcaag	540
	ttcgccctcg ctgcttgtgc aataatcgca gtggggaage cacacegtga ctcccatctt	600
	tcagtaaage tctgtttgtg tttatcagca atacacgtaa tttaaactcg ttagcatggg	660
	gctgatagct taattaccgt ttaccagtgc catggttctg cagctttect tggcccgtaa	720
	aattcggcga agccagccaa tcaccagctt ggcaccagct aaacctata attagtctct	780
	tatcaacacc atccgctccc cgggatcaa tgaggagaat gaggggatg cggggctaaa	840
	gaagctaca taacctcat gccaaactcc agtttacct cgtcagacca acatcctgac	900
	tataagctaa cacagaatgc ctcaatctg ggaagaactg gccgctgata agcgcgccc	960
	cctcgcaaaa accatccctg atgaatggaa agtccagacg ctgcttgcgg aagacagcgt	1020
	tattgatttc ccaaagaaat cggggatcct ttcagaggcc gaactgaaga tcacagagc	1080
	ctcgcctgca gatcttgtgt ccaagctggc ggccggagag ttgacctcgg tggaaagttac	1140
	gctagcattc tgtaaaccgg cagcaatcgc ccagcagtta gtagggctcc ctctacctt	1200
	caggagatg taacaacgcc accttatggg actatcaagc tgacgtggc ttctgtgcag	1260
[0049]	acaaactgcg cccacaggtt ctccctgac gccctctcg cgcaggcaag ggaactgat	1320
	gaatactacg caaagcaca gagaccegtt ggtccactcc atggctccc catctctctc	1380
	aaagaccagc ttcgagtcaa ggtacaccgt tgcccctaag tcgtagatg tccctttttg	1440
	tcagetaaca tatgccacca gggctacgaa acatcaatgg gctacatctc atggetaaac	1500
	aagtacgacg aaggggactc ggtctgaca accatgctcc gcaaagccgg tgcctcttc	1560
	tacgtcaaga cctctgtccc gcagaccctg atggctctgc agacagtcaa caacatcatc	1620
	ggcgcaccg tcaaccacg caacaagaac tggctgtgc gcggcagttc tgggtgtgag	1680
	ggtgcgatcg ttgggaitcg tgggtggctc atcggtgtag gaacggatat cgggtgctcg	1740
	attcgagtgc cgcccgctt caacttctg tacggtctaa ggccgagfca tgggcgctg	1800
	ccgtatgcaa agatggcga cagcatggag ggtcaggaga cgggtcacag cgttgtcggg	1860
	ccgattacgc actctgttga gggtagtcc ttcgctctt ccttcttttc ctgctctata	1920
	ccaggcctcc actgtctctc ttcttgttt ttatactat atacgagacc ggcagtcact	1980
	gatgaagtat gttagacctc cgcctcttca ccaaatccgt cctcggctcag gagccatgga	2040
	atacagactc caaggtcatc cccatgccct ggcgccagtc cgagtcggac attattgcct	2100

	ccaagatcaa gaacggggg ctcaatatcg gctactacaa cttegcacgc aatgtccttc	2160
	cacaccctcc tatcctgegc ggegtggaaa ccaccgtcgc cgcactcgc aaagccggtc	2220
	acaccgtgac cccgtggacg ccatacaagc acgatttcgg ccacgatctc atctcccata	2280
	tctacggcgc tgacggcagc gccgacgtaa tgegcgatat cagtgcatec ggcgagccgg	2340
	cgattccaaa tatcaaagac ctactgaacc cgaacatcaa agctgttaac atgaacgagc	2400
	tctgggacac gcctctccag aagtggaaatt accagatgga gtaccttgag aaatggcggg	2460
	aggctgaaga aaaggccggg aaggaactgg acgccatcat cgcgccgatt acgcctaccg	2520
	ctgcggtagc gcctgaccag ttccggtaact atgggtatgc ctctgtgac aacctgctgg	2580
	atctcagag cgtggttgtt ccggttacct ttgcggataa gaacatcgat aagaagaatg	2640
	agagtttcaa ggccggttagt gagcttgatg ccctcgtgca ggaagagtat gatccggagg	2700
	cgtaccatgg ggcaccggtt gcagtgcagg ttatcggacg gagactcagt gaagagagga	2760
	cgttggcgat tgcagaggaa gtggggaagt tgctgggaaa tgggtgact ccatagctaa	2820
	taagtgtcag atagcaatll gcacaagaaa tcaataccag caactgtaa taagcctga	2880
[0050]	agtgaccatg ccctgctacg aaagagcaga aaaaaacctg ccgtagaacc gaagagatat	2940
	gacacgcttc catctctcaa aggaagaatc ccttcagggt tgcgtttcca gtctagaggc	3000
	catttaggac gttgcigggc tttttccata ggctccggcc cctgacgag catcacaaaa	3060
	atcgacgcic aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc	3120
	cccctggaag ctcccctcgtg cgtctctctg ttccgacct gccgcttacc ggataaccgt	3180
	ccgcctttct cccttcggga agegtggcgc tttctcatag ctacgcctgt aggtatcica	3240
	gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc gttcagcccc	3300
	accgctgcgc cttatccggt aactatcgtc ttgagtcaa cccgtaaga cacgacttat	3360
	cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta ggccgtgcta	3420
	cagagttctt gaagtgttg cctaaactacg gctacaciag aaggacagta ttggtatct	3480
	gcgctctgt gaagccagtt accttcggaa aaagagtgg tagctcttga tccggcaaac	3540
	aaaccaccgc tggtagcggg gtttttttg ttgcaagca gcagattacg cgcagaaaaa	3600
	aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag tgaacgaaa	3660
	actcagtta aggcctgcag ggcgatttt ggctatgaga ttatcaaaaa ggatcttcac	3720

	ctagatcctt ttaaattaa aatgaagttt taaatcaatc taaagtatat atgagtaaac	3780
	ttggctctgac agttaccaat gcttaatcag tgaggeacct atctcagega tctgtctatt	3840
	tegttcatec atagtigeci gactccccgt cgtgtagata actacgatac gggagggett	3900
	accatctggc cccagtgtg caatgatacc gcgagacca cgtcaccgg ctccagattt	3960
	atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggctctg caattttatc	4020
	cgectccatc cagtctatta attgtigcgg ggaagctaga gtaagtagtt cgccagttaa	4080
	tagtttggc aacgtgttg ccattgctac aggcctcgtg gtgtcacgt cgtcgtttgg	4140
	tatggcttca ttcagctcgg gtcccaacg atcaaggcga gttacatgat cccccatgtt	4200
	gtgcaaaaaa gcggttagct ccttcggctc tccgatcgtt gtcagaagta agttggccgc	4260
	agtgttatca ctcatggtta tggcagcact gcataattct ctactgtca tgecatccgt	4320
	aagatgcttt tctgtgactg gtgagtactc aaccaagtea tctgagaat agtgtatgcg	4380
	gcgaccgagt tgcctttgcc cggcgtcaat acgggataat accgcgccac atagcagaac	4440
	tttaaaagtg ctcatcattg gaaaacgttc ttccggggcga aaactctcaa ggatcttacc	4500
[0051]	getgttgaga tccagttcga tgtaaccac tegtgcacc aactgatctt cagcatcttt	4560
	tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg	4620
	aataagggcg acacggaaat gttgaatact caactcttc ctitttcaat attattgaag	4680
	catttateag ggttattgtc tcatggccat ttaggectct agagtgtga agtcggtaat	4740
	cccgcigtat agtaatacga gtccatctc aatactccga agctcctgcg aaccgggaga	4800
	atcgagatgt gctggaaaag ttctagcgg cggtcaatt agcatgaaag gctatgagaa	4860
	attctggaga cggcttgttg aatcatggcg ttccattctt cgacaagcaa agcgttccgt	4920
	cgcagtagca ggcactcatt cccgaaaaaa ctccggagatt cctaagtagc gatggaaccg	4980
	gaataatata ataggcaata cattgagttg cctcgacggt tgcaatgcag gggactgag	5040
	cttggacata actgttccgt accccacctc ttctcaacct ttggcgttcc cctgattcag	5100
	cgtaccgta caagtcgtaa tcactattaa cccagactga ccggacgtgt tttgcccttc	5160
	atttgagaaa ataatgtcat tgegatgtgt aatttgctg cttgaccgac tggggctgtt	5220
	cgaagcccga atgtaggatt gttatccgaa ctctgctcgt agaggcatgt tgtgaatctg	5280
	tgtcgggcag gacacgctc gaaggttcac ggcaagggaa accaccgata gcagigtcta	5340

	gtagcaacct gtaaagccgc aatgcagcat caciggaana tacaaccaa tggctaaaag	5400
	tacataagtt aatgcctaaa gaagtcatat accagcgget aataattgta caatcaagtg	5460
	gctaaaagta ccgtaatttg ccaacggett gtggggttgc agaagcaacg gcaaagcccc	5520
	actteccac gtttgtttet tcactcagtc caatctcagc tggatgacc ccaattgggt	5580
	cgcttgtttg ttccggtgaa gtgaaagaag acagaggtaa gaatgtctga ctccgagcgt	5640
	tttgataca accaagggca gtgatggaag acagigaaat gttgacattc aaggagtatt	5700
	tagccaggga tgetttagtg tatcgtgtaa ggaggttgt ctgccgatac gacgaatact	5760
	gtatagtcac ttctgatgaa gtggtccata ttgaaatgta aagtcggcac tgaacaggca	5820
	aaagattgag ttgaaactgc ctaagatctc gggccctcgg gccttcggcc tttgggtgta	5880
	catgtttgtg ctccgggcaa atgcaaagtg tggtaggata gaacacactg ctgccittac	5940
	caagcagctg agggtatgtg ataggcaaat gttcaggggc cactgcatgg tttcgaatag	6000
	aaagagaagc ttagccaaga acaatagccg ataaagatag cctcattaaa cggaatgagc	6060
	tagtaggcaa agtcagcgaa tgtgtatata taaaggttcg aggtccgtgc ctccctcatg	6120
[0052]	ctctceecat ctactcatca actcagatec tccaggagac ttgtacacca tcttttgagg	6180
	cacagaaaacc caatagtcac ccatcacaag ttgtacaaa aaagcaggct ccgcgccgcg	6240
	ccccttcacc atgccttctc cttggctatt gtgccccctt cttagcgttg gaagtgtctc	6300
	tcctgaaacg aagacggatg ttctgacata caccaaccct gtccttcag gatggcactc	6360
	ggatccatcg tgatccaga aagatggcct ctttctctgc gtcacttcaa cattcatctc	6420
	cttcccaggt ctcccgctct atgcctcaag ggatctagtc aactggcgtc tcatcagcca	6480
	tgtctggaac cggagaaaac agttgcctgg cattagctgg aagacggcag gacagcaaca	6540
	gggaatgtat gcaccaacca ttcgatacca caagggaaca tactacgtca tctgcaata	6600
	cctggcggtt ggagatatta ttggtgtcat cttcaagacc accaatccgt gggacgagag	6660
	tagctggagt gacctgtta cttcaagcc aatcacatc gaccccgatc tgttctggga	6720
	tgatgacgga aaggtttatt gtgtacca tggtatcact ctgcaggaga ttgatttggg	6780
	aactggagag cttagcccg agcttaatat ctggaacggc acaggaggtg tatggcctga	6840
	gggtcccat atctacaagc ggcaggtta ctactatctc atgattgccg aggggtggaac	6900
	tgccgaagac caagctatca caatcgctcg ggccccgaag atcaccggcc cctatgaagc	6960

	ctacaataac aacceaatct tgaccaaccg cgggacatct gagtacttcc agactgtcgg	7020
	tcacggtgat ctgttccaag ataaccaaggg caactgggigg ggtetttgtc ttgtactcgg	7080
	catcacagca cagggagttt caeccatggg ccgigaaget gttttgttca atggcacatg	7140
	gaacaagggc gaatggcca agttgcaacc agtacgaggt cgcattgctg gaaacctcct	7200
	cccaaagccg acgcgaaacg ttcccggaga tgggcccttc aacgctgacc cagacaacta	7260
	caactigaag aagactaaga agatccctcc tcactttgtg caccatagag tcccaagaga	7320
	cggtgccttc tctttgtctt ccaaggtctt gcacatcgtg cctagtcgaa acaacgttac	7380
	cggtagtgtg ttgccaggag atgagattga gctatcagga cagcgaggtc tagctttcat	7440
	cggacgccgc caaactcaca ctctgttcaa atatagtgtt gatatcgact tcaagcccaa	7500
	gtccgatgat caggaagctg gaatcacctt ttccgcacg cagttcgacc atatcgatct	7560
	tggcattggt cgtcttcta caaaccaagg cagcaacaag aaatctaagc ttgccttccg	7620
	attccgggcc acaggagctc agaatgttcc tgcaccgaag gtagtaaccg tccccgatgg	7680
[0053]	ctgggagaag ggcgtaatca gtctacatat cgaggcagcc aacgcgacgc actacaacct	7740
	tggagcttcg agccacagag gcaagactct cgacatcgcg acagcatcag caagtcttgt	7800
	gagtgagggc acgggttcat ttgttgtag ttgtcttggc ccttatgcta cctgcaacgg	7860
	caaaggatct ggagtggaat gtcccagggg agtgatgtc tatgtgacc aatggactta	7920
	taagcccgtg gcacaagaga ttgatcatgg tgtttttgtg aaatcagaat tgtagaaggg	7980
	tgggcgcgcc gaccagctt tc	8002

<210> 15

<211> 7279

<212> DNA

<213> 人工序列

<220>

<223> 合成的表达载体

<400> 15

ttgtacaaag tggatgcgc gccgcgcgc agctccgtgc gaaagcctga cgcaccggtg	60
gattcttggg gagcccgtat catgacggcg gcgggagcta catggccccg ggtgatitai	120
lllllltltg tetacttctg acccttttca aatatacggg caactcatct ttcactggag	180
aigcggcctg cttggatttg cgatgttgc agcttggcaa attgtggctt tggaaaacac	240
aaaacgattc cttagtagcc atgcatttta agataacgga atagaagaaa gaggaaatta	300
aaaaaaaaa aaaaacaaac atcccgttca taacccttag aatcggcctt cttcgtgtat	360
cccaglacca gtttattttg aatagctcgc ccgctggaga gcacctgaa tgcaagtaac	420
aaccctagag gcctgacacgg cagggttgc tagggagcgt cgtgttctac aaggccagac	480
gtcttcggg tlgalatata tglatgittg actgcaggct gctcagcagc gacagtcaag	540
ttgcctctg ctgcttgcgc aalaatcgc gtggggaagc cacaccgtga ctcccatctt	600
tcagtaaagc tctgttggg tllatcagca atacacgtaa tttaaactcg ttagcatggg	660
gctgatagct taattaccgt ttaccagtgc catggllctg cagctttcct tggcccgtaa	720
[0054] aattcggcga agccagccaa tcaccagcta ggcaccagct aaacctata attagtctct	780
tatcaacacc atccctccc ccgggatcaa tgaggagaat gagggggaig cggggctaaa	840
gaagcctaca taacctcat gccaaactcc agtttacact cgtcgagcca acatcctgac	900
tataagctaa cacagaatgc ctcaatctg ggaagaactg gcgctgata agcgcgcccg	960
ctctgcaaaa acctccctg atgaatggaa agtccagacg ctgcctcggg aagacagcgt	1020
tattgatttc ccaaagaaat cggggatcct ttccagagcc gaactgaaga tcacagagge	1080
ctccgctgca gatcttgtgt ccaagctggc ggccggagag ttgacctcgg tggaaattac	1140
gctagcattc tgtaaacggg cagcaatcgc ccagcagtta gtagggctcc ctctacctct	1200
caggagatg taacaacgcc accttatggg actatcaage tgacgttggc ttctgtgcag	1260
acaaactgcg cccacagatt ctccctgac gccctctcg cgcaggcaag ggaactcgat	1320
gaatactacg caaagcaca gagaccctt ggtccactcc atgacctccc catctctctc	1380
aaagaccagc ttcgagtcaa ggtacaccgt tgcacctaaag tegttagaig tccctttttg	1440
tcagctaaca tatgccacca gggctacgaa acatcaatgg gctacatctc atggetaaac	1500
aagtacgacg aaggggactc gttctgaca acctatctcc gcaaagccgg tgcctcttc	1560

	tacgtcaaga cctctgtccc gcagacctg atggctgcg agacagtcaa caacatcatc	1620
	ggggcaccg tcaacccaag caacaagaac tggctgtgcg gggcagttc tgggtgtgag	1680
	ggtgcatcg ttggattcg tggggcgtc atcgggtgag gaacggatat cgggtgctcg	1740
	attcgagtgc cggcccggtt caacttctg tacggctaa gcccgagtea tgggcgctg	1800
	ccgtatgcaa agatggcgaa cagcatggag ggtcaggaga cggtgcaag cgttgtcggg	1860
	ccgattacgc actctgttga gggtagtcc ttcgectctt cttcttttc ctgctctata	1920
	ccaggcctcc actgtcctcc tttcttgett ttatactat atacgagacc ggcagtcact	1980
	gatgaagiat gttagacctc ccctcttca ccaaaccgt cctcggtcag gagccatgga	2040
	aatacgactc caaggteac cccatgcctt ggcgccagtc cgagtcggac attattgect	2100
	ccaagatcaa gaacggcggg ctcaatatg gctactaaa cttcgacggc aatgtcctc	2160
	cacacctcc tatctgccc ggcgtgaaa ccacctgccc cgcactgccc aaagccggtc	2220
	acaccgtgac cccgtggacg ccatacaagc acgattcgg ccacgatctc atctccata	2280
	tctacggcgc tgacggcagc gcgacgtaa tgcgcatat cagtgcctcc ggcgagccgg	2340
[0055]	cgattccaaa tatcaagac ctactgaacc cgaacatcaa agctgttaac atgaacgagc	2400
	tctgggacac gcctctccag aagtggaatt accagatgga gtacctgag aaatggcggg	2460
	aggctgaaga aaagccggg aaggaaactgg acgccatcat cgcgccgatt acgctaccg	2520
	ctgcggtacg gcatgaccag ttcgggtact atgggtatgc ctctgtgac aacctgctg	2580
	atttcacgag cgtggttgtt ccggtacct ttgcggataa gaacatgat aagaagaatg	2640
	agagtttcaa ggcggttagt gagcttgatg ccctcgtgca ggaagagtat gatccggagg	2700
	cgtaccatgg ggcaccggtt gcagtcagg ttatcggacg gagactcagt gaagagagga	2760
	cgttggcgat tgcagaggaa gtggggaagt tgetgggaaa tgtggtgact ccatagctaa	2820
	taagtgtcag atagcaattt gcacaagaaa tcaataccag caactglaaa taagcgctga	2880
	agtgaccatg ccatgctacg aaagagcaga aaaaaacctg ccglagaacc gaagagatat	2940
	gacacgcttc catctctcaa aggaagaalc ccttcagggt tgcgtttcca gctagaggc	3000
	catttaggcc gttgetggcg ttttccata ggtccgcc ccctgacgag catcacaaaa	3060
	atcgacctc aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgcttc	3120
	cccctggaag ctcctctgtg cgtctctctg ttcgacctt gccgcttacc ggatacctgt	3180

	cegcctttct cccttcggga agcgtggcgc tttctcatag ctcacgctgt aggtatctca	3240
	gttcgggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc gttcagcccc	3300
	accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtaaga caccacttat	3360
	cgccactgge agcagccaact ggtaacagga ttagcagagc gaggtatgta ggcggigcta	3420
	cagagttctt gaagtgggtg cctaactacg gctacactag aaggacagta ttiggtatct	3480
	gcgctctgct gaagccagtt accttcggaa aaagagtgg tagctcttga tccggcaaac	3540
	aaaccaccgc tggtagcggt ggTTTTTTTg ttgcaagca gcagattacg cgcagaaaaa	3600
	aaggatctca agaagatcct ttgatcttt ctacggggtc tgacgctcag tggaacgaaa	3660
	actcacgta aggcctgcag ggccgatttt ggctatgaga ttatcaaaaa ggatcttcac	3720
	ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac	3780
	ttggctgac agttaccaat gettaatcag tgaggacct atctcagcga tctgtctatt	3840
	tegttcatcc atagttgcct gactccccgt cgtgtagata actacgatac gggagggctt	3900
	accatctggc cccagtctg caatgatacc gcgagacca cgtcaccgg ctccagattt	3960
[0056]	atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggctctg caactttatc	4020
	cgctccate cagtctatta atgtttgcg ggaagctaga gtaagtagtt cgcagttaa	4080
	tagtttgcgc aacgttggtg ccattgctac aggcatcgtg gtgtcaegct cgtcgtttgg	4140
	tatggcttca ttcagctccg gttcccaacg atcaaggcga gtiacatgat cccccatggt	4200
	gtgcaaaaaa gcggttagct ccttcggtec tccgatcgtt gtcagaagta agttggccgc	4260
	agtgttaica ctcatggta tggcagcact gcataattct ctactgtca tgccatccgt	4320
	aagatgettt tctgtgactg gtgagtactc aaccaagtca ttctgagaat agtgtatgag	4380
	gcgaccgagt tgetcttgc cggcgtcaat acgggataat accgcgccac atagcagaac	4440
	ttfaaaagt ctcacattg gaaaacgttc ttccggggcga aaactctcaa ggatcttacc	4500
	gctgttgaga tccagttcga tgtaaccac tcgtgcacc aactgatctt cagcatctt	4560
	tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg caaaaaagg	4620
	aataaggcgc acacggaaat gttgaatact catactcttc cttttcaat attattgaag	4680
	catttatcag ggttattgtc tcatggccat ttaggcctct agagtgtga agtcggtaat	4740
	cccgtgtat agtaatacga gtcgcatcta aatactccga agctgctgcg aaccggaga	4800

	atcgagatgt gctggaaagc ttctagcggag cggctaaatt agcatgaaag gctatgagaa	4860
	attctggaga cggettgttg aatcatggcg ticcattctt cgacaagcaa agcgttccgt	4920
	cgcagtagca ggcactcatt ccgaaaaaa ctccggagatt cctaagtagc gatggaaccg	4980
	gaataatata ataggcaata cattgagttg cctcgcacggg tgcaatgcag gggtagctgag	5040
	cttggacata actgttccgt accccacctc ttctcaacct ttggcgttcc cctgattcag	5100
	cgtaccogta caagtcgtaa tcactattaa cccagactga cggacgtgt tttgcccctc	5160
	atttggagaa ataatgcat tgcgatgtgt aatttgcctg cttgaccgac tgggctgtt	5220
	cgaagcccga atgtaggatt gttatccgaa ctctgctcgt agaggcatgt tgtgaatctg	5280
	tgtcgggcag gacacgcctc gaaggttcac ggcaaggga accaccgata gcagtgctca	5340
	gtagcaacct gtaaagccgc aatgcagcat cactggaaaa tacaaccaa tggetaaaag	5400
	tacataagtt aatgcctaaa gaagtcatat accagcggct aataatigta caatcaagtg	5460
	gctaaacgta ccgtaatttg ccaacggctt gtggggttgc agaagcaacg gcaaaagccc	5520
	acttccccac gtttgittct tcactcagtc caatctcagc tggatgcc ccaattgggt	5580
[0057]	cgcttgtttg ttccggtgaa gtgaaagaag acagaggtaa gaatgtctga ctccggacgt	5640
	tttgcataca accaaggcca gtgatggaag acagtgaat gttgacatc aaggagtatt	5700
	tagccagggg tgcctgagtg tctcgtgtaa ggaggtttgt ctgccgatac gacgaatact	5760
	gtatagtcac ttctgatgaa gtggccata ttgaaatgta aagtcggcac tgaacaggca	5820
	aaagattgag ttgaaactgc ctaagatctc gggccctcgg gccttcggcc tttgggtgta	5880
	catgitttg ctcgggcaa atgcaaagtg tggtaggata gaacacactg ctgcccttac	5940
	caagcagctg agggtaigtg ataggcaaat gttcagggc cactgcatgg tttcgaatag	6000
	aaagagaagc ttagccaaga acaatagccg ataaagatag cctcattaaa cggaatgagc	6060
	tagtaggcaa agtcagcgaa tgtgtatata taaagtttc aggtccgtgc ctcccctatg	6120
	ctctcccct ctactcatca actcagatcc tccaggagac ttgtacacca tcttttgagg	6180
	cacagaaacc caatagtcaa ccateacaag ttgtacaaa aaagcaggt cgcggccgc	6240
	ccccttacc atgcgtcttc tatcgtttcc cagccatctc ctctggcct tectaaccct	6300
	caaagaggct tcatccctcg ccctcagcaa acgggatagc cctgtctctc cggcctctg	6360
	ggcggacccc aacatcgcca tcgtcgaaa gacatactac atcttcccta ccaccgacgg	6420

	tttcgaaggc tggggcggca acgtcttcta ctgggtgaaa tcaaaagate tcgtatcatg	6480
	gacaaagagc gacaagccat tccttactct caatggtagc aatggcaacg ttccttgggc	6540
	tacaggtaat gcctgggctc ctgcttctgc tgctcgcgga ggcaagtatt acttctacca	6600
	tagtgggaat aatccctctg tgagtgatgg gcataagagt attggtgcgg cgggtgctga	6660
	tcctcctgag gggccgtgga aggcacagga taagccgatg atcaaggga cttctgatga	6720
	ggagattgtc agcaaccagg ctatcgatcc cgtctccttt gaagaccctg agactggaaa	6780
	gtgtatatac tactggggaa acggtgtccc cattgtcgca gagctcaacg acgacatggt	6840
	ctctctcaaa gcaggctggc acaaaatcac aggtcttcag aatttccgcg agggctcttt	6900
	cgtaactat cgcgatggaa catatcatct gacatactct atcgacgata cgggctcaga	6960
	gaactatcgc gttgggtacg ctacggcgga taaccctatt ggaccttggg catatcgtgg	7020
	tgctctctg gagaaggacg aatcgaaggc cattcttctg acgggacata actccatcat	7080
	caacattcct ggaacggatg agtggtatat cgcgtatcat cgttccata ttcctgatgg	7140
	aatgggtat aatagggaga ctacgatgga taggttacc ctcgacaagg atacgggttt	7200
[0058]	gtttgaaag gttacgccg ctttgcagag tgttgatcct aggcctttgt agaagggtgg	7260
	gcgcgccgac ccagctttc	7279

<210> 16

<211> 1725

<212> DNA

<213> 拟轮生镰刀菌(*Fusarium verticillioides*)

<400> 16

atgccttct ctggctatt gtccccctt ctacgatgg gaagtctct tctgaaacg	60
aagacggatg tttgacata caccaacct gtccttcag gatggcactc ggatccatg	120
tgtatccaga aagatggcct cttctctgc gtcacttcaa cattcatctc ctcccaggt	180
cttccgctc atgcctcaag ggatctagtc aactggcgtc tcatcagcca tgcttggac	240
cgcgagaaac agttgctgg cattagctgg aagacggcag gacagcaaca gggaatgtat	300

	gcaccaacca ttegatacca caagggaaaca tactacgtca tetgccaata cctgggcgtt	360
	ggagatatta ttgggtcat ctccaagacc accaatccgt gggacgagag tagctggagt	420
	gaccctgta cctccaagcc aaatcacatc gaccccgatc tgttctggga tgatgacgga	480
	aaggtttatt gtgctacca tggcatcact ctgcaggaga ttgatttga aactggagag	540
	cttagcccg agcttaatat ctggaacggc acaggagggt taiggcctga gggctcccat	600
	atctacaagc gcgacggtta ctactatctc atgattgccg aggggtggaac tgccgaagac	660
	cacgctatca caatgctcg gcccgcgaag atcaccggcc cctatgaagc ctacaataac	720
	aacccaatct tgaccaaccg cgggacatct gagtacttcc agactgtcgg tcacgggat	780
	ctgttccaag ataccaagg caactgggtg ggttttgtc ttgctactcg catcacagca	840
	cagggagtt caccatggg cgtgaagct gttttgtca atggcacatg gaacaaggcc	900
	gaatggccca agttgcaacc agtacgaggt cgcattcctg gaaacctcct cccaaagccg	960
	acgggaaacg ttcccggaga tggcccttc aacgctgacc cagacaacta caacttgaag	1020
	aagactaaga agatccctcc tcactttgtg caccatagag tccaagaga cgggtgccctc	1080
[0059]	tcittgtctt ccaagggtct gcacatcgtg cctagtcgaa acaacgttac cggtagtgtg	1140
	ttgccaggag atgagattga gctatcagga cagcgaggtc tagctttcat cggacgccgc	1200
	caaactcaca ctctgttcaa atatagtgtt gatatcgact tcaagcccaa gtccgatgat	1260
	caggaagctg gaatcacctg ttccgcacg cagttcgacc atatcgatct tggcattgtt	1320
	cgctttccta caaaccaagg cagcaacaag aaatctaagc ttgccttccg attccgggcc	1380
	acaggagctc agaatgttcc tgcaccgaag gtagtaccgg tcccgatgg ctgggagaag	1440
	ggcgtaatca gtctacatat cgaggcagcc aacgcgacc actacaacct tggagcttcg	1500
	agccacagag gcaagactct cgacatcgcg acagcatcag caagtcttgt gactggagge	1560
	acgggttcat ttgttggtag ttgcttggga ccttatgcta cctgcaacgg caaaggatct	1620
	ggagtggaat gtcccgaagg aggtgatgtc tatgtgacc aatggactta taagcccgtg	1680
	gcacaagaga ttgatsatgg tgttttgtg aaatcagaat tgtag	1725

<210> 17

<211> 574

<212> PRT

<213> 拟轮生镰刀菌

<400> 17

Met Arg Phe Ser Trp Leu Leu Cys Pro Leu Leu Ala Met Gly Ser Ala

1 5 10 15

Leu Pro Glu Thr Lys Thr Asp Val Ser Thr Tyr Thr Asn Pro Val Leu

20 25 30

Pro Gly Trp His Ser Asp Pro Ser Cys Ile Gln Lys Asp Gly Leu Phe

35 40 45

Leu Cys Val Thr Ser Thr Phe Ile Ser Phe Pro Gly Leu Pro Val Tyr

50 55 60

Ala Ser Arg Asp Leu Val Asn Trp Arg Leu Ile Ser His Val Trp Asn

[0060] 65 70 75 80

Arg Glu Lys Gln Leu Pro Gly Ile Ser Trp Lys Thr Ala Gly Gln Gln

85 90 95

Gln Gly Met Tyr Ala Pro Thr Ile Arg Tyr His Lys Gly Thr Tyr Tyr

100 105 110

Val Ile Cys Glu Tyr Leu Gly Val Gly Asp Ile Ile Gly Val Ile Phe

115 120 125

Lys Thr Thr Asn Pro Trp Asp Glu Ser Ser Trp Ser Asp Pro Val Thr

130 135 140

Phe Lys Pro Asn His Ile Asp Pro Asp Leu Phe Trp Asp Asp Asp Gly

145 150 155 160

Lys Val Tyr Cys Ala Thr His Gly Ile Thr Leu Gln Glu Ile Asp Leu

165 170 175

Glu Thr Gly Glu Leu Ser Pro Glu Leu Asn Ile Trp Asn Gly Thr Gly

	180	185	190	
	Gly Val Trp Pro Glu Gly Pro His Ile Tyr Lys Arg Asp Gly Tyr Tyr			
	195	200	205	
	Tyr Leu Met Ile Ala Glu Gly Gly Thr Ala Glu Asp His Ala Ile Thr			
	210	215	220	
	Ile Ala Arg Ala Arg Lys Ile Thr Gly Pro Tyr Glu Ala Tyr Asn Asn			
	225	230	235	240
	Asn Pro Ile Leu Thr Asn Arg Gly Thr Ser Glu Tyr Phe Gln Thr Val			
	245	250	255	
	Gly His Gly Asp Leu Phe Gln Asp Thr Lys Gly Asn Trp Trp Gly Leu			
	260	265	270	
	Cys Leu Ala Thr Arg Ile Thr Ala Gln Gly Val Ser Pro Met Gly Arg			
	275	280	285	
[0061]	Glu Ala Val Leu Phe Asn Gly Thr Trp Asn Lys Gly Glu Trp Pro Lys			
	290	295	300	
	Leu Gln Pro Val Arg Gly Arg Met Pro Gly Asn Leu Leu Pro Lys Pro			
	305	310	315	320
	Thr Arg Asn Val Pro Gly Asp Gly Pro Phe Asn Ala Asp Pro Asp Asn			
	325	330	335	
	Tyr Asn Leu Lys Lys Thr Lys Lys Ile Pro Pro His Phe Val His His			
	340	345	350	
	Arg Val Pro Arg Asp Gly Ala Phe Ser Leu Ser Ser Lys Gly Leu His			
	355	360	365	
	Ile Val Pro Ser Arg Asn Asn Val Thr Gly Ser Val Leu Pro Gly Asp			
	370	375	380	
	Glu Ile Glu Leu Ser Gly Gln Arg Gly Leu Ala Phe Ile Gly Arg Arg			
	385	390	395	400

<400> 18	
atgCGtette tatcgtttcc cagccatctc etcgtggcct tectaacctt caaagagget	60
tcattccctcg cectcagcaa aegggatage cctgtcctcc ccggcctctg ggeggacccc	120
aacatcgcca tegtgcacaa gacatactac atcttcccta ccaccgacgg tttegaagge	180
tggggcggca acgtcttcta ctggtggaaa tcaaaagatc tcgtatcatg gacaaagage	240
gacaagcoat tecttactct caatggtacg aatggcaacg ttccctgggc tacaggtaat	300
gcttgggctc ctgcttccgc tgcctcggga ggcaagtatt acttctacca tagtgggaat	360
aatecctctg tgagtgatgg gcataagagt attggtgcgg cgggtgctga tcattctgag	420
gggccgtgga aggcacagga taagccgatg atcaaggga cttctgatga ggagattgtc	480
agcaaccagg ctatcgatcc cgtctccttt gaagaccctg agactggaaa gtggtatata	540
tactggggaa acggtgtccc cattgtcga gagctcaacg acgacatggt ctctctcaaa	600
gcaggctggc acaaaatcac aggtcttcag aatttccgcg agggctcttt cgtcaactat	660
cgcgatggaa catatcatct gacatactct atcgacgata cgggctcaga gaactatcgc	720
[0063] gttgggtaac ctacggcgga taacccatt ggaccttga catatcgtgg tgttctctg	780
gagaaggacg aatcgaaggc cattcttget acgggacata actccatcat caacattcct	840
ggaacggatg agtggatat cgcgtateat cgttccata ttcccgatgg aatgggtat	900
aatagggaga ctacgattga taggtaccc atcgacaagg atacgggitt gtttggaaag	960
gttacgccga ctttgcagag tgttgatcct aggcctttgt ag	1002

<210> 19

<211> 333

<212> PRT

<213> 拟轮生镰刀菌

<400> 19

Met Arg Leu Leu Ser Phe Pro Ser His Leu Leu Val Ala Phe Leu Thr

1 5 10 15
 Leu Lys Glu Ala Ser Ser Leu Ala Leu Ser Lys Arg Asp Ser Pro Val
 20 25 30
 Leu Pro Gly Leu Trp Ala Asp Pro Asn Ile Ala Ile Val Asp Lys Thr
 35 40 45
 Tyr Tyr Ile Phe Pro Thr Thr Asp Gly Phe Glu Gly Trp Gly Gly Asn
 50 55 60
 Val Phe Tyr Trp Trp Lys Ser Lys Asp Leu Val Ser Trp Thr Lys Ser
 65 70 75 80
 Asp Lys Pro Phe Leu Thr Leu Asn Gly Thr Asn Gly Asn Val Pro Trp
 85 90 95
 Ala Thr Gly Asn Ala Trp Ala Pro Ala Phe Ala Ala Arg Gly Gly Lys
 100 105 110
 [0064] Tyr Tyr Phe Tyr His Ser Gly Asn Asn Pro Ser Val Ser Asp Gly His
 115 120 125
 Lys Ser Ile Gly Ala Ala Val Ala Asp His Pro Glu Gly Pro Trp Lys
 130 135 140
 Ala Gln Asp Lys Pro Met Ile Lys Gly Thr Ser Asp Glu Glu Ile Val
 145 150 155 160
 Ser Asn Gln Ala Ile Asp Pro Ala Ala Phe Glu Asp Pro Glu Thr Gly
 165 170 175
 Lys Trp Tyr Ile Tyr Trp Gly Asn Gly Val Pro Ile Val Ala Glu Leu
 180 185 190
 Asn Asp Asp Met Val Ser Leu Lys Ala Gly Trp His Lys Ile Thr Gly
 195 200 205
 Leu Gln Asn Phe Arg Glu Gly Leu Phe Val Asn Tyr Arg Asp Gly Thr
 210 215 220

<211> 22

<212> DNA

<213> 人工序列

[0066] <220>

<223> 合成的引物

<400> 21

gtgtgaccgg ctttggcgag tg

22

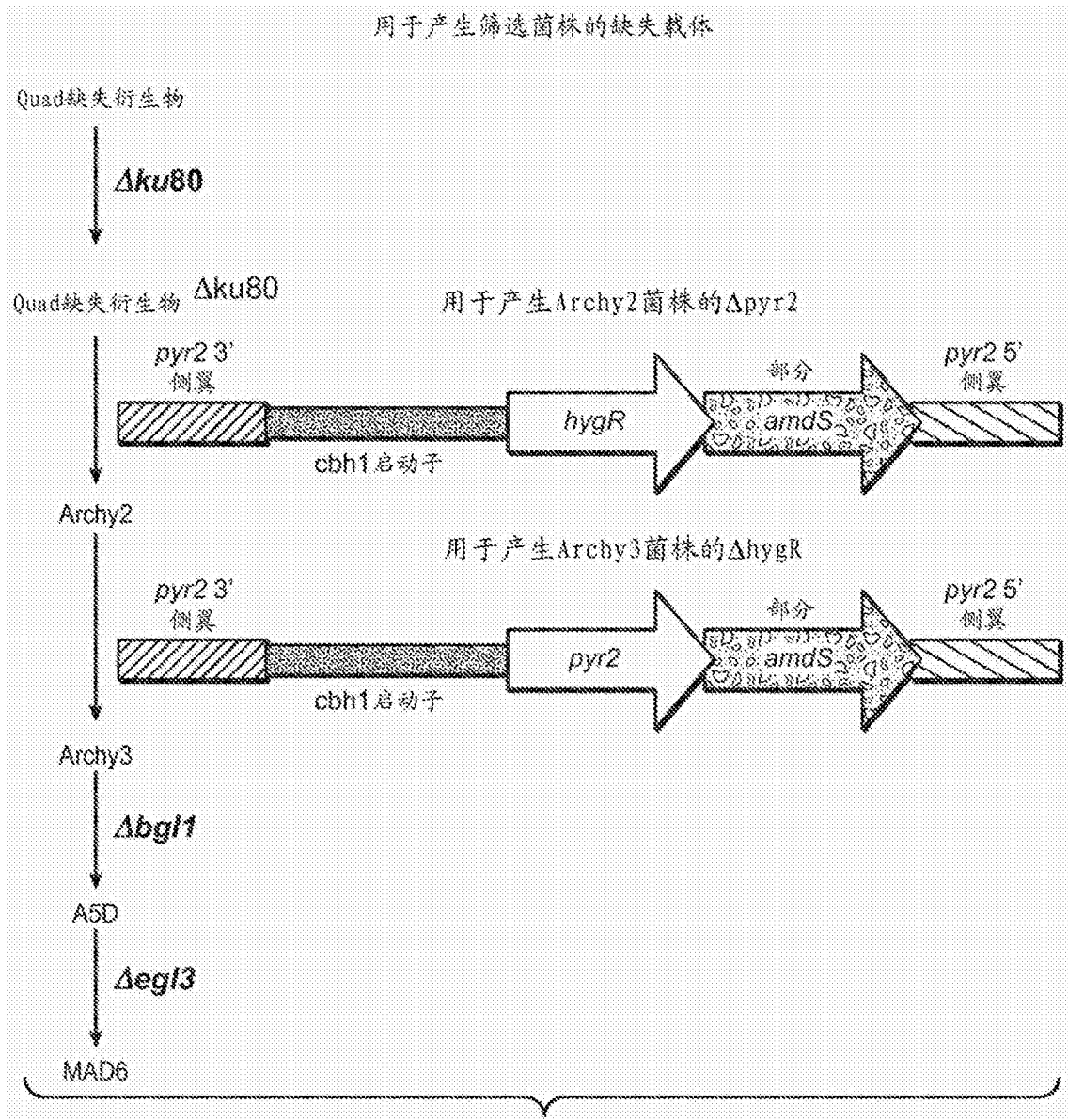


图1

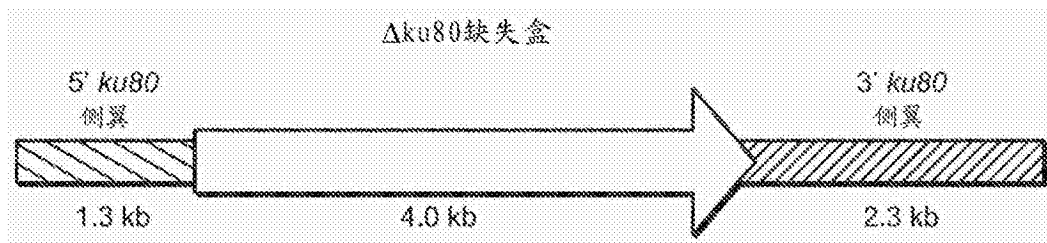


图2

用于产生Archy2菌株的Δpyr2缺失盒

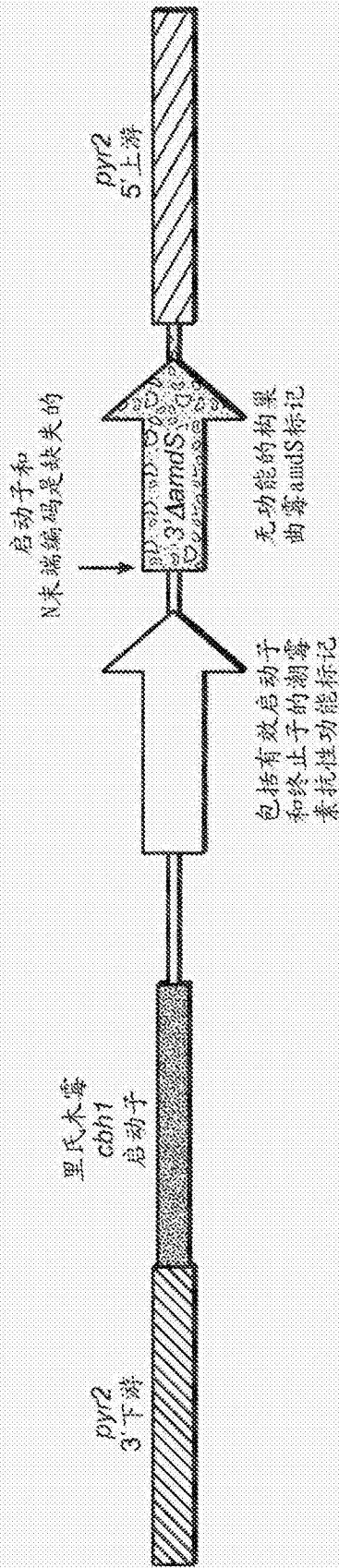


图3

用于产生Archy3菌株的ΔhygR缺失盒

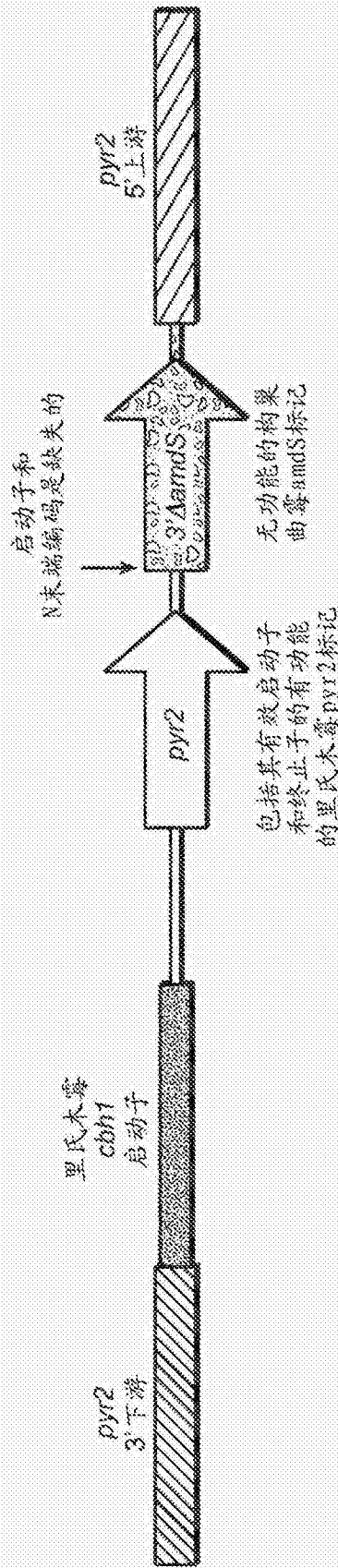


图4

Δbgl1缺失盒 (5610bp)

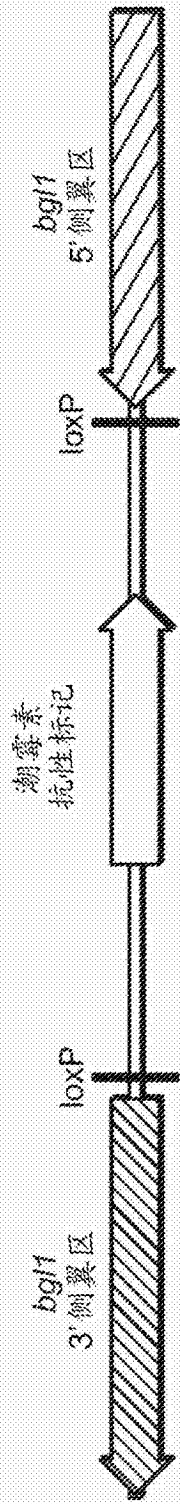


图5

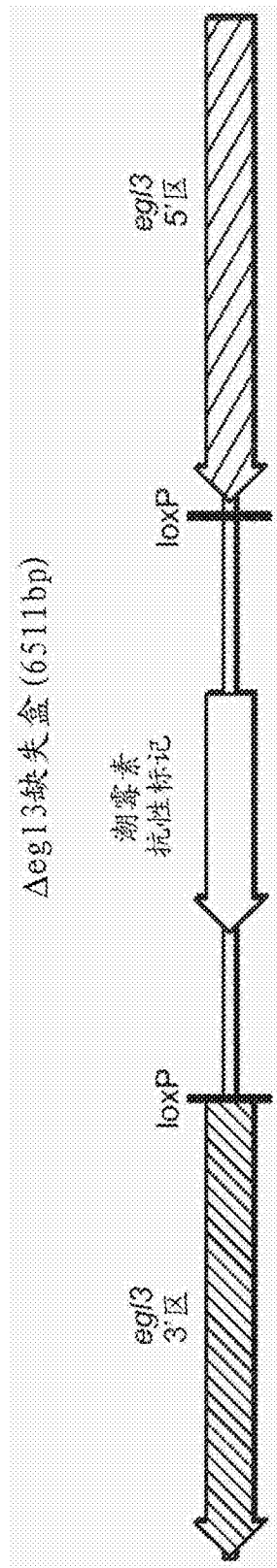


图6

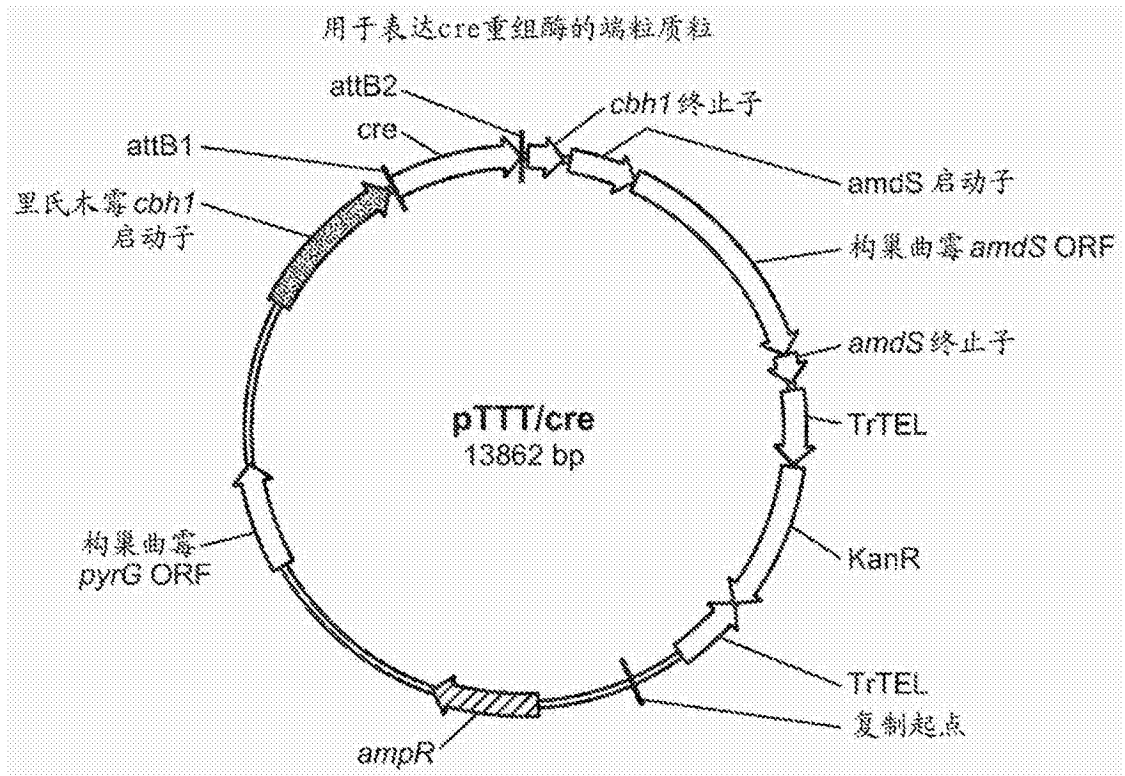


图7

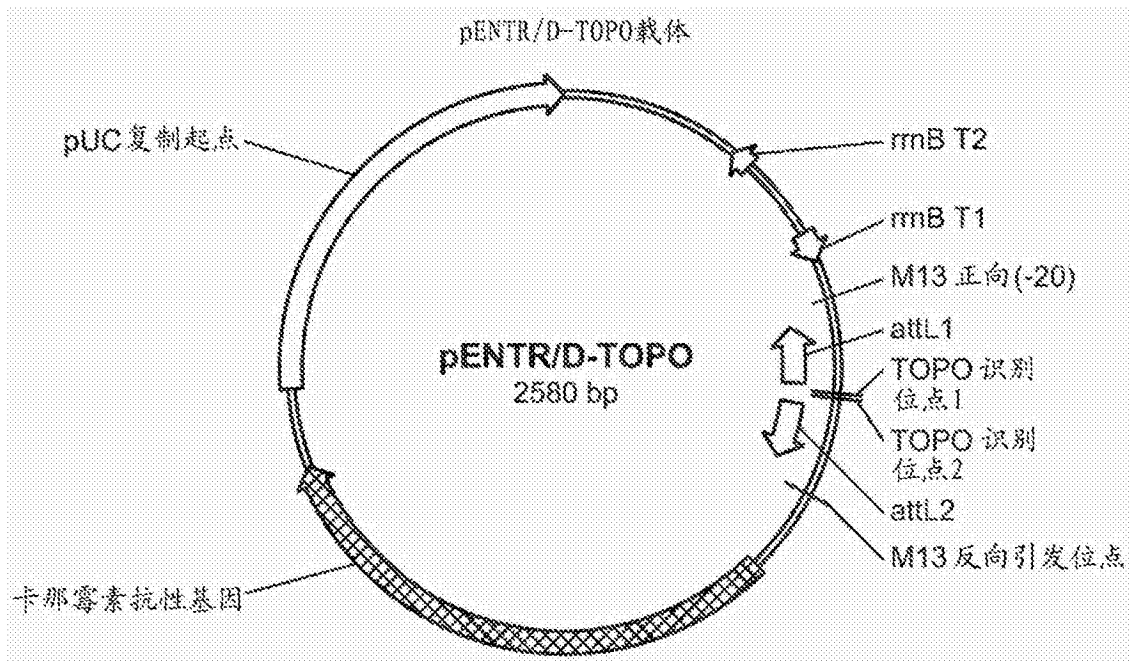


图9

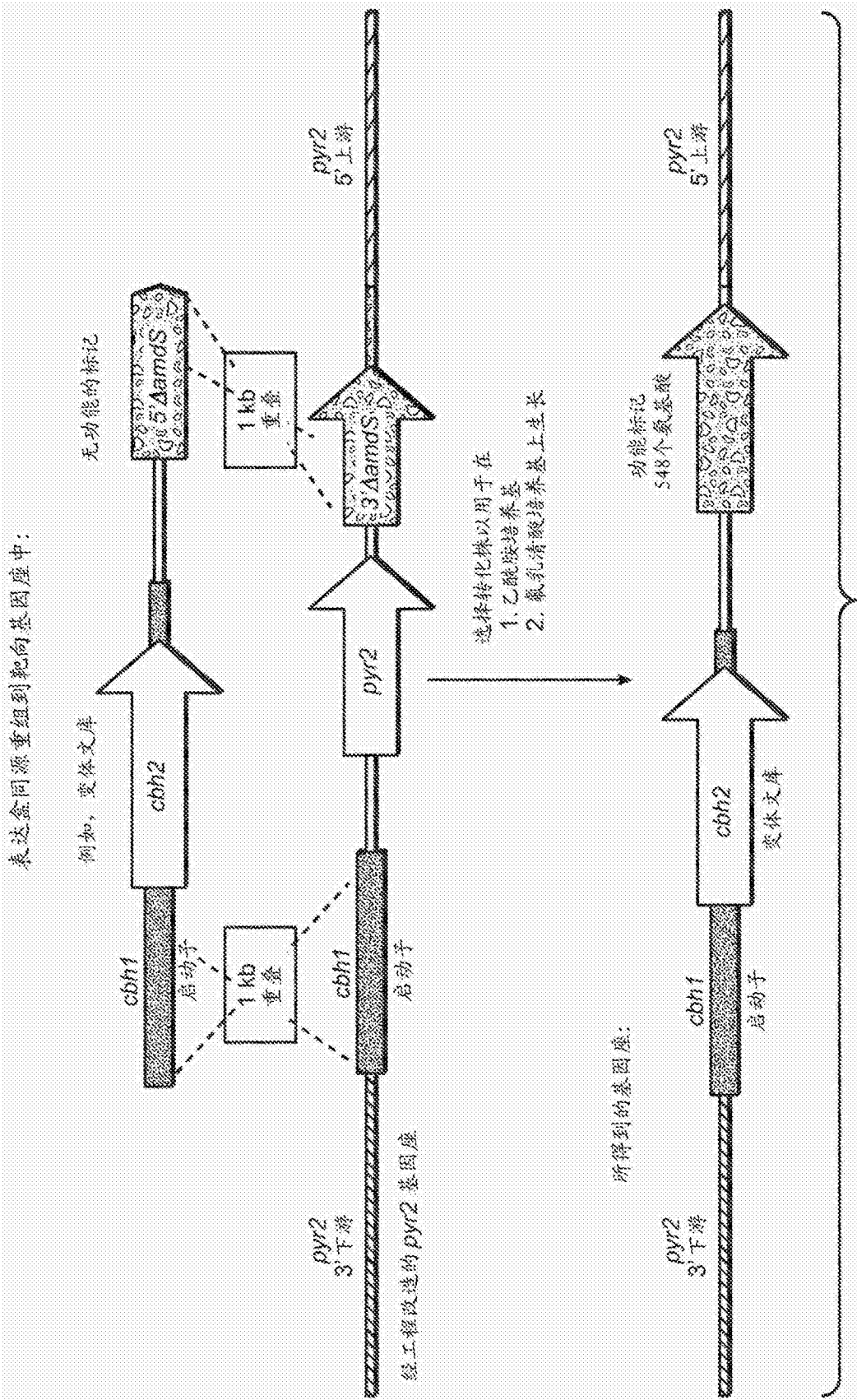


图8

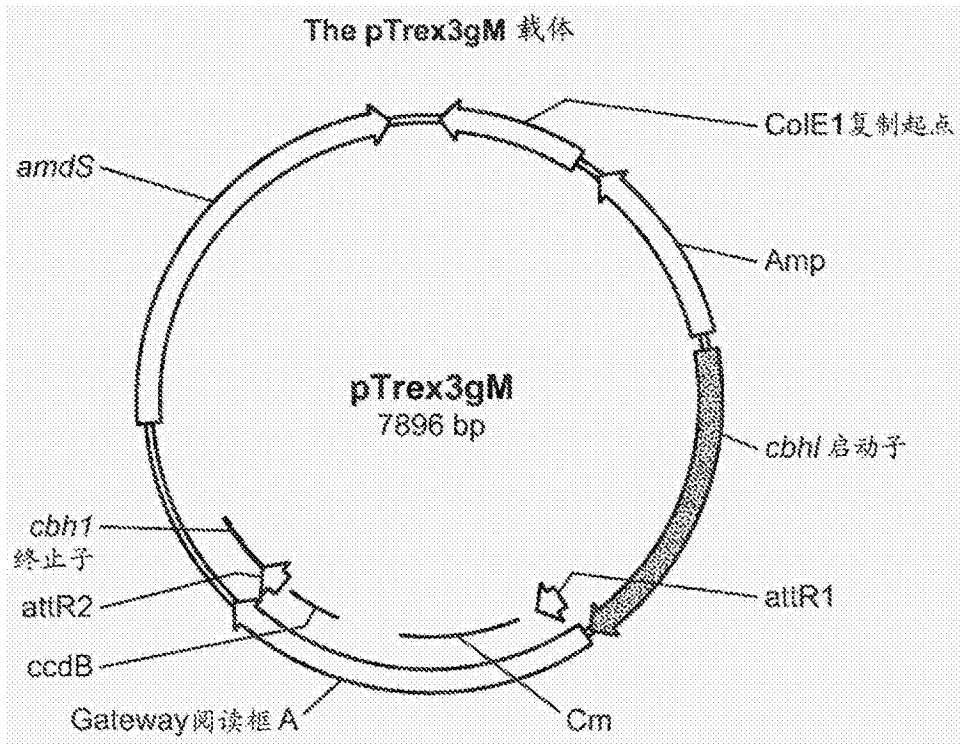


图10

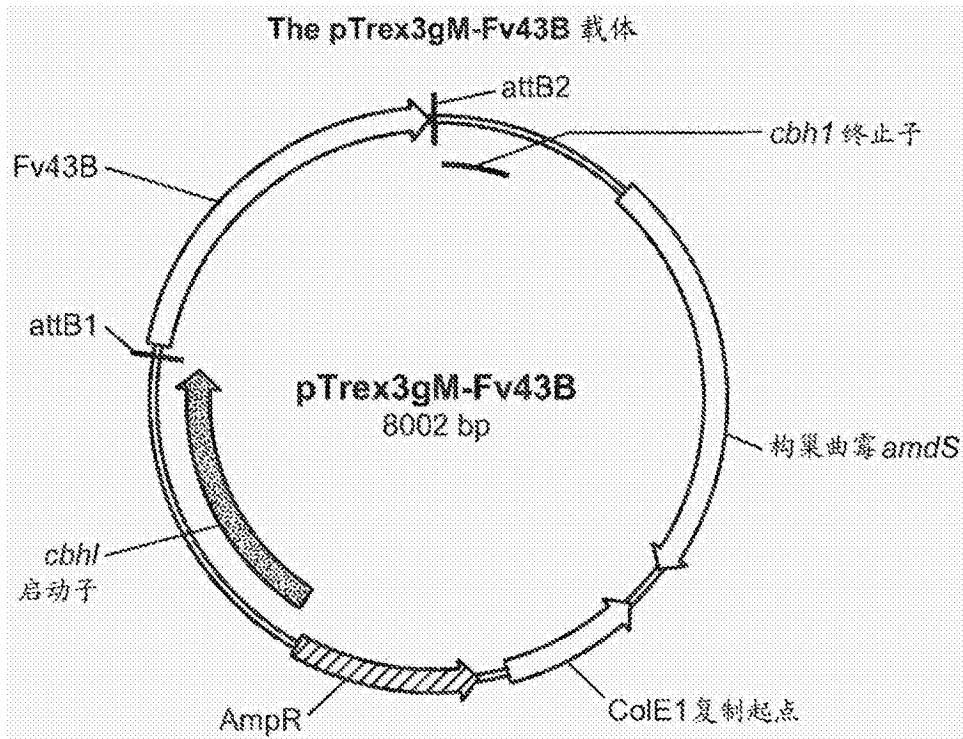


图11

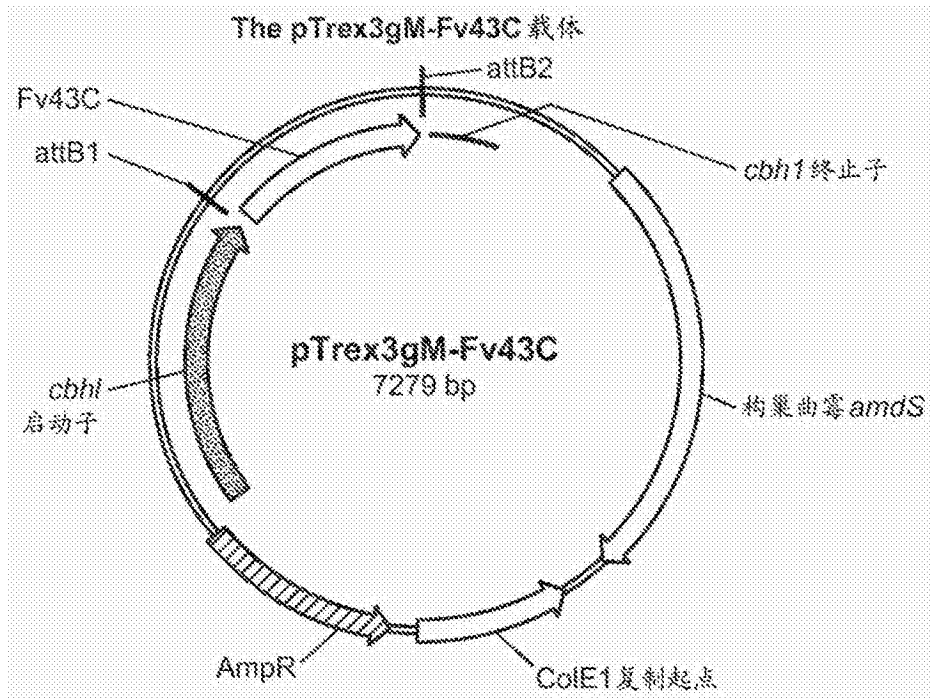


图12

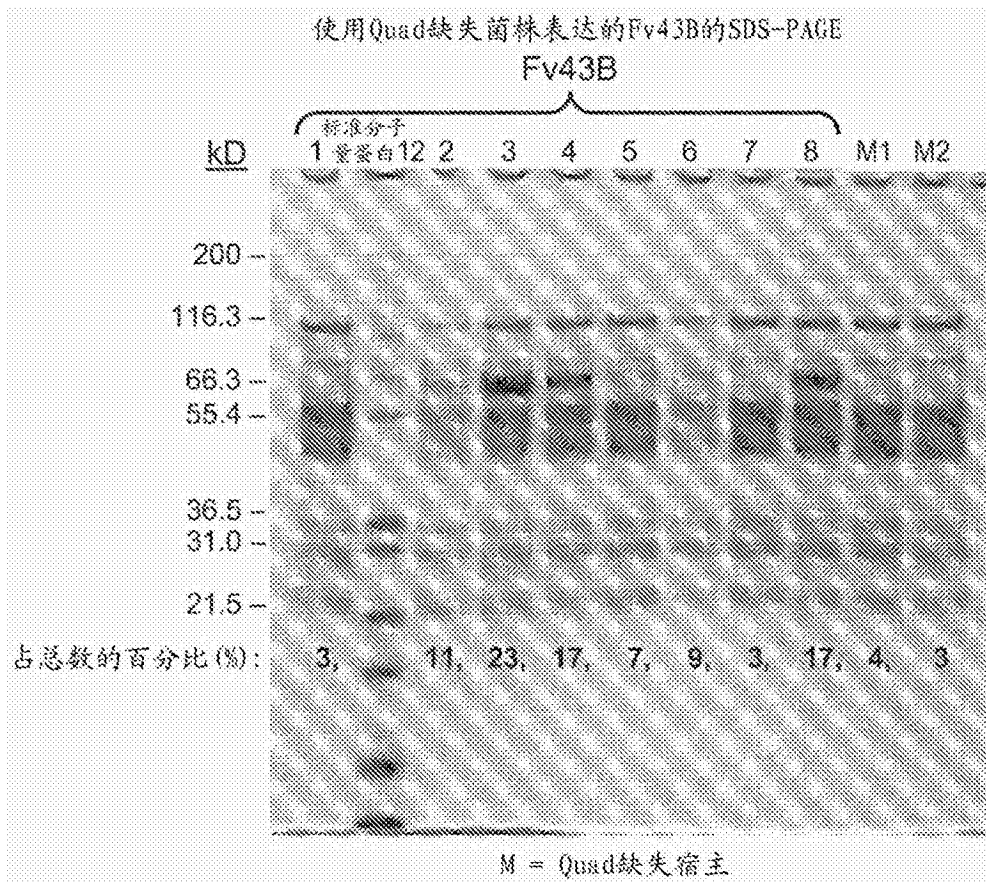


图13

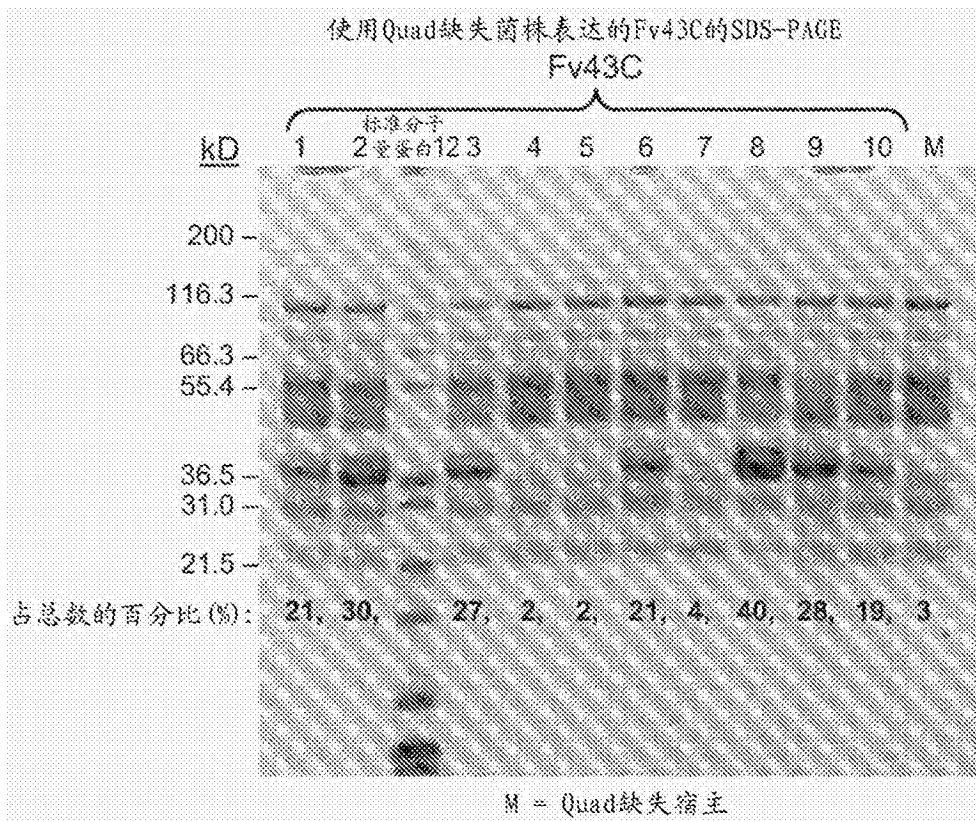


图14

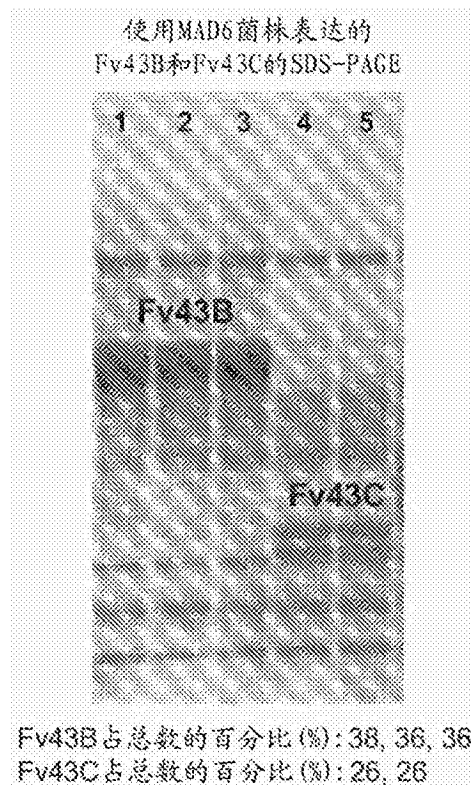


图15

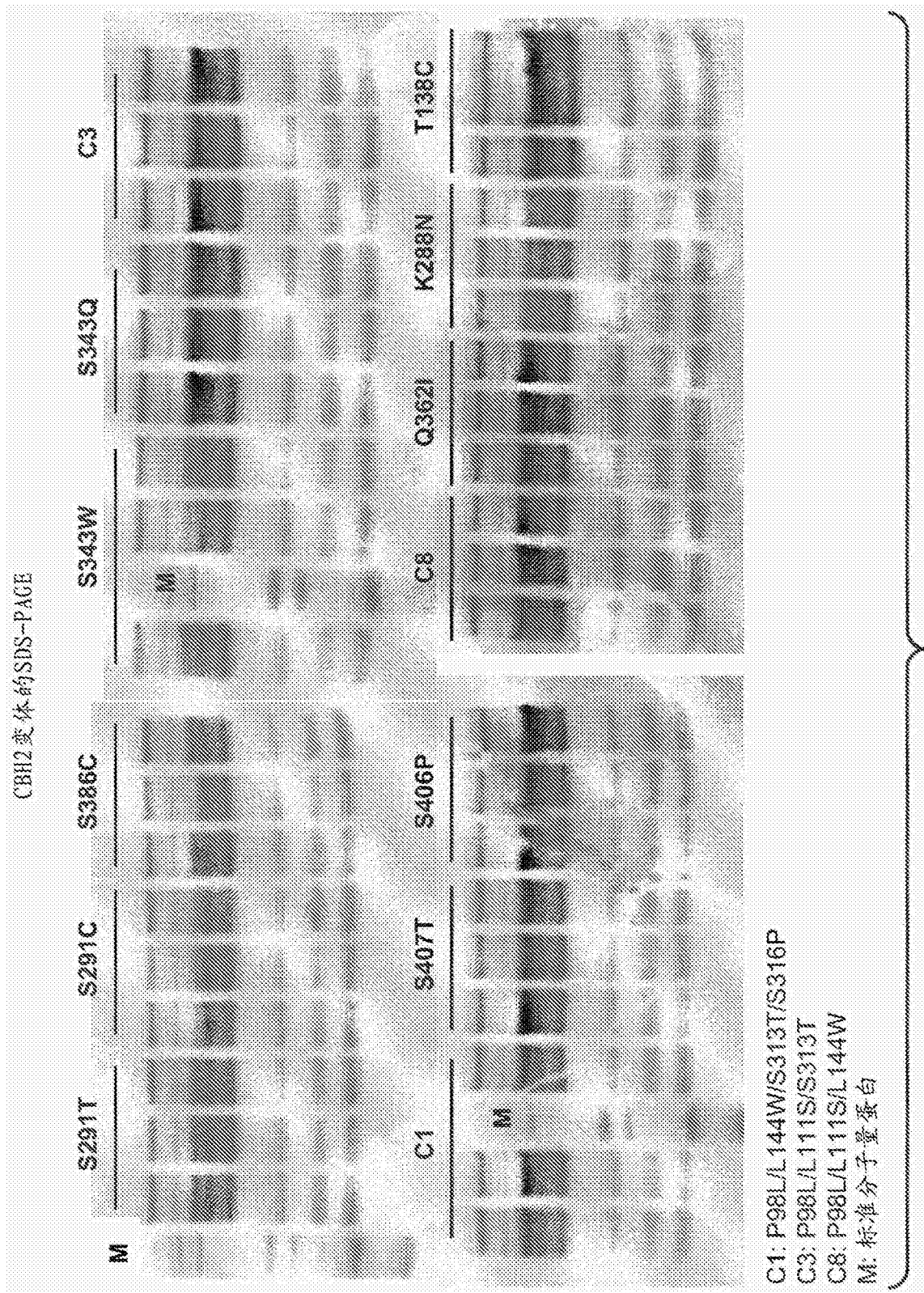


图16A

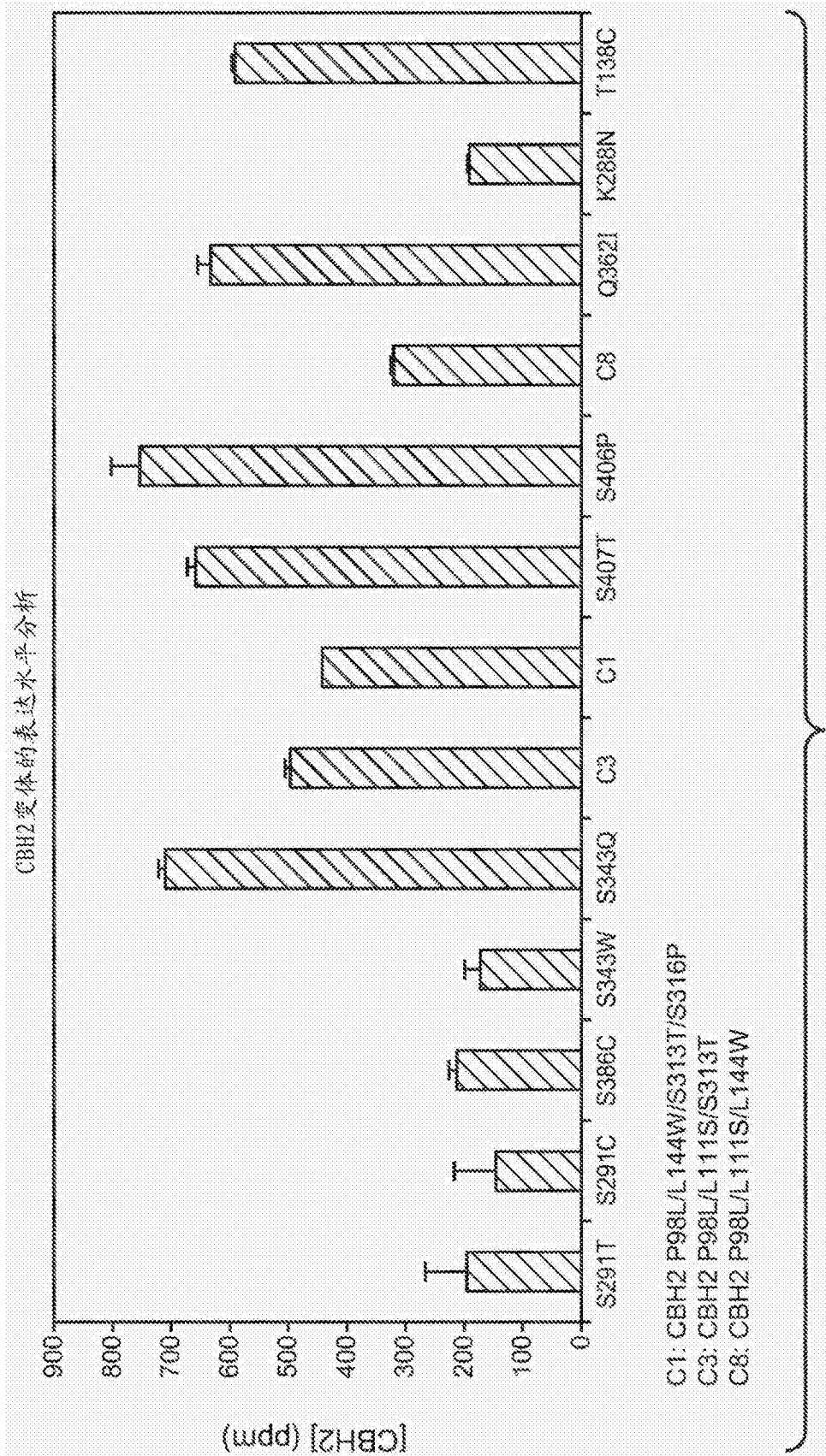


图16B