



(12) 发明专利

(10) 授权公告号 CN 101506360 B

(45) 授权公告日 2014. 06. 11

(21) 申请号 200780031317. X

C12N 15/12 (2006. 01)

(22) 申请日 2007. 08. 21

C12N 5/10 (2006. 01)

(30) 优先权数据

60/839, 735 2006. 08. 23 US

60/903, 247 2007. 02. 23 US

(56) 对比文件

WO 01/51620 A2, 2001. 07. 19, 摘要, 序列 3.

WO 01/51620 A2, 2001. 07. 19, 摘要, 序列 3.

Jorgensen L et al. Accession: AAA72993.

《NCBI database》. 1993,

(85) PCT国际申请进入国家阶段日

2009. 02. 23

审查员 王慧

(86) PCT国际申请的申请数据

PCT/US2007/018523 2007. 08. 21

(87) PCT国际申请的公布数据

W02008/024372 EN 2008. 02. 28

(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 G·英格兰德 M·孔克曼

B·S·米勒 C·弗勒门

(74) 专利代理机构 北京市中咨律师事务所

11247

代理人 黄革生 刘金辉

(51) Int. Cl.

C12N 9/44 (2006. 01)

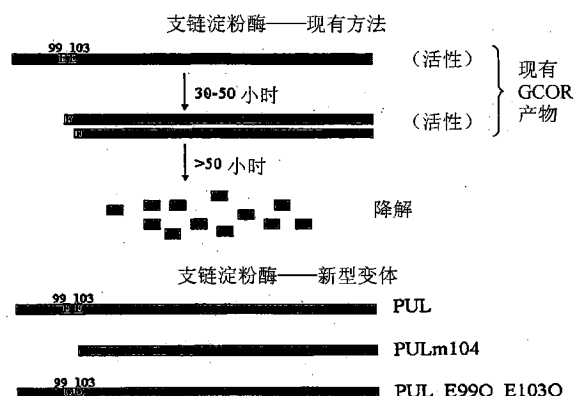
权利要求书2页 说明书20页 附图11页

(54) 发明名称

具有增加产率的支链淀粉酶变体

(57) 摘要

本发明涉及酶促肽支链淀粉酶的新型变体, 编码所述新型肽的基因序列, 包括那些基因序列的表达载体以及表达新型支链淀粉酶变体的生物。本发明的新型支链淀粉酶变体通过使用密码子优化方法, 选择性截短“野生型”分子以及通过置换所选择的氨基酸残基而被经验性地制造。而且, 本发明涉及这些新型支链淀粉酶肽在纺织、发酵、食品和其它工业中的应用。



1. 一种组合物,其包括分离的肽分子,所述分离的肽分子为 SEQ ID NO:6,并包括相对于 SEQ ID NO:2 在位置 99 和 103 处从谷氨酸至谷氨酰胺的替代,其中所述分离的肽分子具有支链淀粉酶活性。

2. 一种组合物,其包括分离的核酸分子,所述分离的核酸分子编码权利要求 1 所定义的具有支链淀粉酶活性的肽分子。

3. 权利要求 2 的组合物,其中所述分离的核酸分子由 SEQ ID NO:5 的核苷酸序列组成。

4. 一种组合物,其包括分离的 DNA,所述分离的 DNA 编码 SEQ ID NO:6 的氨基酸序列的蛋白质。

5. 一种组合物,其包括由分离的 DNA 组成的表达构建物,所述分离的 DNA 编码权利要求 1 所定义的具有支链淀粉酶活性的肽分子。

6. 表达构建物,其包括可操作地连接至控制序列的编码权利要求 1 所定义的具有支链淀粉酶活性的肽分子的核酸分子,所述控制序列可被由所述表达构建物转化的宿主细胞识别。

7. 根据权利要求 6 的表达构建物,其中所述核酸分子为 SEQ ID NO:5 的核苷酸序列。

8. 宿主生物,其转染了根据权利要求 6 或 7 所述的表达构建物。

9. 根据权利要求 8 所述的宿主生物,其中所述宿主生物选自真菌、细菌和真核细胞。

10. 根据权利要求 9 所述的宿主生物,其中所述宿主生物选自枯草芽孢杆菌、埃希氏大肠杆菌、里氏木霉、啤酒酵母、黑曲霉和地衣芽孢杆菌。

11. 根据权利要求 9 所述的宿主生物,其中所述宿主生物选自芽孢杆菌种。

12. 组合物,其包括编码支链淀粉酶的表达构建物,所述表达构建物包括 SEQ ID NO:5 的核酸序列。

13. 组合物,其包括编码 SEQ ID NO:6 的肽的表达构建物。

14. 产生支链淀粉酶的方法,其包括:

a. 提供:i) 表达构建物,其可操作地包括编码权利要求 1 所定义的具有支链淀粉酶活性的肽分子的核苷酸序列;ii) 宿主生物;和 iii) 培养工具;

b. 将所述表达构建物转染至所述宿主生物,用以产生转染的宿主生物;

c. 在所述培养工具中培养所述转染的宿主生物,培养时间长度和条件足以产生支链淀粉酶。

15. 根据权利要求 14 的方法,其中所述核苷酸序列为 SEQ ID NO:5 的序列。

16. 根据权利要求 14 所述的方法,其中所述支链淀粉酶被从所述培养工具中分离。

17. 根据权利要求 14 所述的方法,其中所述宿主生物选自真菌、细菌和真核细胞。

18. 根据权利要求 17 所述的方法,其中所述宿主生物选自枯草芽孢杆菌、埃希氏大肠杆菌、里氏木霉、啤酒酵母、黑曲霉和地衣芽孢杆菌。

19. 根据权利要求 17 所述的方法,其中所述宿主生物选自芽孢杆菌种。

20. 产生支链淀粉酶的方法,其包括:

a. 提供:i) 用表达构建物转染的宿主生物,所述表达构建物可操作地编码核苷酸序列,所述核苷酸序列编码 SEQ ID NO:6 的肽;和 ii) 培养工具;

b. 在所述培养工具中培养所述转染的宿主生物,培养时间长度和条件足以产生支链淀粉酶。

21. 根据权利要求 20 所述的方法,其中所述支链淀粉酶被从所述培养工具中分离。
22. 根据权利要求 20 所述的方法,其中所述宿主生物选自真菌、细菌和真核细胞。
23. 根据权利要求 22 所述的方法,其中所述宿主生物选自枯草芽孢杆菌、埃希氏大肠杆菌、里氏木霉、啤酒酵母、黑曲霉和地衣芽孢杆菌。
24. 根据权利要求 22 所述的方法,其中所述宿主生物选自芽孢杆菌种。
25. 核苷酸序列,其编码权利要求 1 所定义的具有支链淀粉酶活性的肽分子。
26. 权利要求 25 的核苷酸序列,其为 SEQ ID NO:5 的序列。
27. 组合物,其包括地衣芽孢杆菌菌株 BML612PUL_E99Q_E103Q Ori1 和 BML612PUL_E99Q_E103Q Ori2 中的一个或多个。

具有增加产率的支链淀粉酶变体

相关申请的交叉引用

本申请要求 2007 年 2 月 23 日提交的美国临时申请 US 60/903, 247 和 2006 年 8 月 23 日提交的美国临时申请 US 60/839, 735 的优先权, 其中每个通过全部引用并入本文。

发明领域

[0001] 本发明涉及酶促肽支链淀粉酶的新型变体, 编码所述新型肽的基因序列, 包括那些基因序列的表达载体以及表达新型支链淀粉酶变体的生物。而且, 本发明涉及这些新型支链淀粉酶肽在纺织、发酵、食品和其它工业中的应用。

发明背景

[0002] 支链淀粉酶是在许多工业应用, 特别是在食品和饮料工业中被发现有用的酶。支链淀粉酶是淀粉脱支酶, 并在淀粉水解产物的脱支 (用于改进面团性质)、 β -极限葡聚糖 (β -limit dextrans) 的脱支 (用于酿造啤酒和强麦酒 (ales)) 以及从例如玉米、马铃薯、小麦、木薯和水稻生产糖浆中是有效的。支链淀粉酶是分类为 EC 3. 2. 1. 41 的酶, 并且这种酶的特征是它们水解例如支链淀粉和普鲁分支葡聚糖中 α -1, 6-糖苷键的能力。

[0003] 支链淀粉酶是细菌, 特别是芽孢杆菌属的产物。用于工业应用的支链淀粉酶的生产并非没有问题。支链淀粉酶可被同样由细菌产生的各种蛋白酶快速降解, 从而使大量支链淀粉酶的回收低效且昂贵。本领域技术人员已经设计了方法, 以通过限制培养物中支链淀粉酶的降解来增加产量。例如, 我们以前已经示出, 从支链淀粉酶生产菌株删除 AprL 和 Mpr 基因 (它们表达蛋白酶) 对于经济表达活性支链淀粉酶是必要的。而且, 发酵时间限制在 51-60 小时, 以限制支链淀粉酶产物的蛋白水解性降解和活性损失。Svendsen 设计了支链淀粉酶变体, 所述支链淀粉酶变体改变酶的三维构象, 以增加酶的热稳定性或改变酶如何降解其底物 (见美国专利 6, 350, 599 和 6, 838, 257 以及美国申请第 2004/0082028 号)。

[0004] 最近, 我们已经示出, 支链淀粉酶降解的时间被测定, 结果如下: 在 30 和 50 小时之间, 观察全长支链淀粉酶分子部分剪切成分别缺少 N-端 98 和 102 个氨基酸的截短的分子。剪切分别发生在谷氨酸残基 E99 和 E103 的 N-末端, 并且可以通过 HPLC 可见。意料不到地, 1-98 和 1-102 被截短的支链淀粉酶分子保留支链淀粉酶活性, 并且, 甚至更意料不到的是, 表明这种活性高于全长支链淀粉酶的活性。51 小时后, 支链淀粉酶分子进一步降解导致活性降低, 最后所有活性消失。

[0005] 尽管如此, 支链淀粉酶肽和编码这些肽的核苷酸序列的设计仍有改进的余地。因此, 所需要的是更高效地生产支链淀粉酶的化合物和方法, 例如, 通过限制蛋白水解性降解, 增加发酵滴度 (fermentation titer) 或增加支链淀粉酶活性。

发明概述

[0006] 本发明涉及支链淀粉酶——一种肽酶——的新的和非显而易见形式。本发明的支

链淀粉酶包括新的修饰,所述新的修饰导致生产滴度和/或耐受降解(例如,由蛋白酶引起的酶促降解)方面的优越性能和/或在靶向底物材料的分解中比亲本(“野生型”)肽更具有活性。

[0007] 在这方面,本发明涉及肽 SEQ ID NO:2、SEQ ID NO:4 和 SEQ ID NO:6 的支链淀粉酶,分别如图 7(b)、8(b) 和 9(b) 所示。本发明还涉及编码氨基酸序列 SEQ ID NOS:2、4 和 6 的核苷酸序列。各自的核苷酸序列还分别作为 SEQ ID NOS:1、3 和 5 在图 7(a)、8(a) 和 9(a) 中示出。如本领域所熟知,遗传密码冗余,多个核苷酸密码子编码相同氨基酸。因此,本发明又涉及编码肽 SEQ ID NOS:2、4 和 6 的任何可选的核苷酸序列。本领域普通技术人员能够基于本说明书和本领域知识的教导,确定编码本发明肽序列的核苷酸序列。

[0008] 本发明涉及编码本发明肽的表达构建物。本发明不限于任何特定的表达构建物,只要其能够表达本发明的肽。在这方面,合适的表达构建物的非限制性实例示意地显示在图 4-6 的 (a) 部分。

[0009] 如在本说明书的详述和实施例部分中的更详细的解释,在一种实施方式中,通过密码子优化技术来修饰编码亲本支链淀粉酶肽(来自 *Bacillus deramificans*) 的序列,以产生密码子优化的核苷酸序列 [SEQ ID NO:1],所述优化的核苷酸序列编码野生型(即,亲本)支链淀粉酶氨基酸序列 [SEQ ID NO:2] 的复制品。编码氨基酸序列 [SEQ ID NO:2] 的核苷酸序列在两个方向上被克隆至地衣芽孢杆菌 (*B. licheniformis*) 整合载体 pICatH 的 XhoI 位点,产生表达构建物的 Ori1(pICatH-PUL-Ori1) 和 Ori2(pICatH-PUL-Ori2) 形式。

[0010] 在另一种实施方式中,编码亲本支链淀粉酶肽的序列被修饰,以产生支链淀粉酶肽,其中 N-端的 104 个氨基酸已被删除。在一种实施方式中,编码这种新型支链淀粉酶的核苷酸序列也是密码子优化的 [SEQ ID NO:3]。由该构建物, PULm 104 表达的肽在 SEQ ID NO:4 中给出(见,图 8(b))。编码 [SEQ ID NO:4] 的核苷酸序列在两个方向上被克隆到地衣芽孢杆菌整合载体 pICatH 的 XhoI 位点,产生表达构建物的 Ori1(pICatH-PULm104-Ori1) 和 Ori2(pICatH-PULm104-Ori2) 形式。

[0011] 在另一种实施方式中,编码亲本支链淀粉酶肽的序列被改变,以将 99 和 103 位上的氨基酸残基从谷氨酸 (E) 置换为谷氨酰胺 (Q),从而使所产生的肽更能抵抗在这些位置的蛋白水解降解。在一种实施方式中,编码这种新型支链淀粉酶的核苷酸序列也是密码子优化的 [SEQ ID NO:5]。由该核苷酸序列 PUL_E99Q_E103Q 表达的肽在 SEQ ID NO:6 中给出。编码 [SEQ ID NO:6] 的核苷酸序列在两个方向上被克隆到地衣芽孢杆菌整合载体 pICatH 的 XhoI 位点,产生表达构建物的 Ori1(pICatH-PUL_E99Q_E103Q-Ori1) 和 Ori2(pICatH-PUL_E99Q_E103Q-Ori2) 形式。

[0012] 本发明还涉及将本发明的表达构建物转染到适合的宿主生物。本发明不限于任何具体宿主生物。宿主生物可以是,例如,微生物、真核细胞或组织培养物、植物细胞或组织培养物、或真菌细胞或组织培养物。在优选的实施方式中,宿主生物为微生物。优选的宿主生物包括但不限于:芽孢杆菌种 (*Bacillus* sp.) (特别是枯草芽孢杆菌 (*Bacillus subtilis*))、地衣芽孢杆菌 (*B. licheniformis*) 和 *B. deramificans*), 埃希氏大肠杆菌, 里氏木霉, 啤酒酵母或黑曲霉。在最优选的实施方式中,宿主生物是地衣芽孢杆菌。

[0013] 本发明涉及从培养本发明的宿主生物的培养基中分离和纯化本发明的肽。在这点上,本发明不限于任何特定的分离和纯化技术,只要其产生 10% 的最低纯度。在更优选的实

施方式中,分离和纯化的肽的最低纯度为 25%,在甚至更优选的实施方式中,分离和纯化的肽的最低纯度为 50%。还在更优选的实施方式中,分离和纯化的肽的最低纯度为 75%。在最优选的实施方式中,本发明分离和纯化的肽的最低纯度为 90%。最低纯度可通过总干重百分比或本领域已知的其它适合的方法测量。

[0014] 本发明不限于任何分离和纯化本发明的肽的具体纯化方法。本领域已知的任何肽纯化方法都是适合的。合适的纯化方法非限制性实例包括亲和层析、沉淀、大小排阻层析、薄层层析、电泳、大小过滤等。

[0015] 本领域普通技术人员将认识到本发明的支链淀粉酶生物学活性片段可用于替代本发明上下文中的全长序列或等价物。“生物学活性片段”意图包括本发明序列的任何类似物、模拟、平截、缺失和 / 或置换。本发明的支链淀粉酶的模拟肽和支链淀粉酶活性结构域可应用结构数据进行计算机设计,如本领域已知的。另外,在本发明一种实施方式中,考虑本发明支链淀粉酶核苷酸序列的类似物和突变可通过定向分子进化产生。所述定向分子进化的技术在本领域是已知的(见,例如,Stemmer 等人的美国专利第 5,605,793 号或 Short 的美国专利第 6,537,776 号,其通过引用在此并入)。与本发明的肽相比,由定向分子进化产生的蛋白质将具有更小、更大或相同的作为支链淀粉酶起作用的能力。

[0016] 在本发明另一种实施方式中,本发明的肽被用作融合蛋白,其具有例如,其它的结构性或功能性肽结构域。这种结构域可以,例如,赋予融合蛋白其它的酶促能力或将肽限定在表面。

[0017] 本发明的肽还包括那些由于存在多个基因、可选的转录事件、可选的 RNA 剪接事件和可选的翻译和翻译后事件而出现的肽。多肽可在体系例如培养的细胞中表达,当表达的支链淀粉酶在天然细胞中表达时所述体系导致出现基本上相同的翻译后修饰,或在这样的体系中表达,当在天然细胞中表达时所述体系导致出现的翻译后修饰删除。

[0018] 本发明还包括组合物,所述组合物包括一种或多种支链淀粉酶肽(或编码它的核酸)和一种或多种另外的成分,例如:载体、稀释剂或溶剂。所述另外的成分可以是使组合物可用于体外、体内、制药或兽医应用的成分。

[0019] 在另一方面,本发明提供基本上纯的核酸,所述核酸具有或包括编码肽的核苷酸序列,其氨基酸序列包括或者是本发明支链淀粉酶肽的序列。

[0020] 在优选的实施方式中,本主题支链淀粉酶核酸将包括转录调控序列,例如,转录启动子或转录增强子序列中的至少一个,其可操作地连接至支链淀粉酶基因序列,例如使支链淀粉酶基因序列适合用作表达载体。

[0021] 而在进一步优选的实施方式中,编码本发明支链淀粉酶肽的核酸,在严格条件下与核酸探针杂交,所述核酸探针对应于 SEQ ID NO :1、3 或 5 的至少 12 个连续核苷酸,更优选地对应于 SEQ ID NO :1、3 或 5 的至少 20 个连续核苷酸。

[0022] 本发明的另一种优选实施方式提供本文描述的支链淀粉酶在多种工业装置中的应用。例如,本发明涉及本发明的新型支链淀粉酶变体在食品原料(包括,但不限于各种生面团和糖浆)、饮料(包括且不限于各种酿造饮料,如啤酒和强麦酒)、生物乙醇和众多本领域普通技术人员已知的其它产品的生产中的应用。

附图简述

[0023] 图 1 显示在本方案中设计的支链淀粉酶方法和支链淀粉酶变体的示意图。

[0024] 图 2 示出存在于本发明制造的支链淀粉酶表达菌株中的构建体的分子结构,其与菌株 BMP 139 中的支链淀粉酶表达构建物相比较。

[0025] 图 3 示出 pICatH 载体的示意图。pICatH 载体包含温度敏感的复制起点 (ori pE194, 用于在芽孢杆菌 (*Bacillus*) 中复制)、ori pBR322 (用于在大肠杆菌 (*E. coli*) 中扩增)、用于筛选的新霉素抗性基因和具有重复的天然地衣芽孢杆菌 (*B. licheniformis*) 氯霉素抗性基因 (cat)——用于筛选、染色体整合和盒扩增。

[0026] 图 4 示出 pICatH-PUL Ori1 构建物的示意图。

[0027] 图 5 示出 pICatH-PULm104Ori1 构建物的示意图。

[0028] 图 6 示出 pICatH-PUL_E99Q_E103Q Ori 构建物的示意图。

[0029] 图 7A-B 示出密码子优化的“野生型”支链淀粉酶 (PUL) 的 (a) 核酸序列 [SEQ ID NO :1] 和 (b) 氨基酸序列 [SEQ ID NO :2]。

[0030] 图 8A-B 示出 PULm104 支链淀粉酶的 (a) 核酸序列 [SEQ ID NO :3] 和 (b) 氨基酸序列 [SEQ ID NO :4]。

[0031] 图 9A-B 示出 PUL_E99Q_E103Q 支链淀粉酶的 (a) 核酸序列 [SEQ ID NO :5] 和 (b) 氨基酸序列 [SEQ ID NO :6]。

定义部分

[0032] 应当注意,如本说明书和所附权利要求书中使用的,单数形式“一”(a)、“一”(an)和“该”(“the”)包括复数指代物,除非上下文另外明确说明。因此,例如,提及含有“化合物”(“a compound”)的组合物包括两种或多种化合物的混合物。还应注意,术语“或”通常使用的意义包括“和/或”,除非上下文另有明确说明。

[0033] 术语“支链淀粉酶”是指一种特定种类的葡聚糖酶,降解普鲁分支葡聚糖的淀粉分解内酶。它是由例如克雷伯菌属 (*Klebsiella*) 的革兰氏阴性细菌作为细胞外的细胞表面锚定脂蛋白产生的。但是,革兰氏阳性细菌产生支链淀粉酶作为分泌蛋白。I 型支链淀粉酶特定地攻击 α -1,6 键,而 II 型支链淀粉酶也能够水解 α -1,4 键。它还可由某些其它细菌和古细菌产生。支链淀粉酶在生物技术中用作去污剂。支链淀粉酶 (EC 3.2.1.41) 还被称为普鲁分支葡聚糖-6-葡聚糖水解酶(脱支化酶)。普鲁分支葡聚糖被认为是由 α -1,6-糖苷键连接的麦芽三糖单位的链。支链淀粉酶水解切割普鲁分支葡聚糖(α -葡聚糖多糖)。

[0034] 术语“密码子优化”是指提高表达水平的技术,其通过用编码相同氨基酸但被宿主生物更有效处理的密码子置换编码序列中的核苷酸密码子进行。不同物种之间的密码子偏好性可显著不同。为提高外源蛋白在特定表达系统(细菌、真菌、酵母、昆虫、植物或哺乳动物细胞)中的表达水平,很重要的是调整外源蛋白密码子的频率,以匹配宿主表达系统密码子的频率。一个典型的实例是 GFP(绿色荧光蛋白, green fluorescent protein),其被优化以在哺乳动物细胞中获得高水平表达。因此,可将密码子优化用于在多种宿主生物中表达本发明蛋白——在所述宿主生物中这种序列可能不能有效表达,如果进行表达的话。

[0035] 术语“表达载体”如本文使用是指重组 DNA 分子,其含有所期望的编码序列和适合的核酸序列,它们对于可操作连接的编码序列在具体宿主生物中表达是必需的。原核生物

表达所必需的核酸序列包括启动子,任选地,操作子序列,核糖体结合位点和可能的其它序列。

[0036] 如本文使用的“异源启动子”是指天然地不与基因或纯化的核酸连接的启动子。术语“启动子”、“启动子元件”或“启动子序列”如本文使用是指当连接至感兴趣的核苷酸序列时,能够控制感兴趣的核苷酸序列转录为 mRNA 的 DNA 序列。启动子典型地,虽然不是必要地,位于感兴趣的核苷酸序列的 5' (即上游),启动子控制到 mRNA 的转录并提供用于 RNA 聚合酶特异结合的位点和用于转录起始的其它转录因子。

[0037] 应用于调控元件的术语“细胞类型特异的”或“宿主生物特异的”或等同术语是指这样的调节元件,其能够指导感兴趣的核苷酸序列在特定类型的细胞或生物中选择性表达——所述相同的感兴趣核苷酸序列在不同类型细胞或生物中、相同组织内相对地没有表达。当被应用于调节元件时,术语“细胞类型特异的”或“宿主生物特异的”也指这样的调控元件,其能够促进感兴趣的核苷酸序列分别在单一组织或生物内的区域中选择性表达。

[0038] 肽的“分离”、“纯化制备”或“基本上纯的制备物”如本文使用是指这样的多肽,其已被鉴定并从至少一种通常与其天然状态结合或从其实际来源获得的污染物中分离。所述至少一种其它污染物可以是,例如,与其一起天然发生的其它蛋白、脂肪和核酸。优选地,多肽还从例如抗体或凝胶基质的物质分离,所述凝胶基质例如聚丙烯酰胺,用于纯化所述多肽。优选地,多肽占纯化制备物干重的至少 10、20、50、70、80 或 95%。优选地,制备物包含:足够的多肽用于蛋白测序;至少 1、10 或 100 毫克的多肽;至少 1、10 或 100 毫克的多肽。

[0039] “纯化的细胞制备物”如本文使用是指在植物或动物细胞的情况下,细胞的体外制备物而不是全部完整的植物或动物。在培养的细胞或微生物细胞的情况下,其由至少 10% 和更优选地 50% 的目标细胞制备物组成。

[0040] “基本上纯的核酸”,例如,基本上纯的 DNA,是指这样的核酸,其是如下的一种或两种:不与一个或两个序列例如编码序列直接邻近,其与所述编码序列在核酸所来源的自然的天然发生的基因组中直接邻近(即,一个在 5' 端而一个在 3' 端);或其基本上没有这样的核酸序列——其在所述核酸来源的生物中存在。术语包括,例如,重组 DNA,其整合到载体中,例如,整合到自主复制的质粒或病毒中,或到原核生物或真核生物的基因组 DNA 中,或作为独立于其它 DNA 序列的单独分子(例如,由 PCR 或限制内切酶处理产生的 cDNA 或基因组 DNA 片段)存在。另外,术语“分离的”,当用于涉及核酸时,如在“分离的核酸序列”中是指这样的核酸序列,其被鉴定并从至少一种通常与其以天然状态结合或从其实际来源获得的污染物核酸中分离。分离的核酸是以不同于在自然中发现的形式或设置存在的核酸。相反地,未分离的核酸是指核酸例如 DNA 和 RNA,它们以天然状态存在。例如,给定的 DNA 序列(例如,基因)在宿主细胞染色体上发现,与比邻基因(neighboring gene)相邻;RNA 序列,如编码特定蛋白的特定 mRNA 序列,在细胞中发现,作为与众多编码众多蛋白质的其它 mRNA 一起的混合物。然而,包含例如 SEQ ID NO:1 的分离的核酸序列包括,例如,在细胞中的核酸,所述细胞通常含有 SEQ ID NO:1,其中所述核酸序列在不同于天然细胞的染色体内或染色体外位置,或另外地,由与在自然中发现的核酸序列不同的核酸序列侧翼。分离的核酸序列可以单链或双链形式存在。当分离的核酸序列被用于表达蛋白质时,所述核酸序列将包含(最少)有义或编码链的至少一部分(即,核酸序列可以是单链的)。可选地,它可包含有义和反义链(即,核酸序列可以是双链的)。

[0041] “同源”如本文使用是指两个多肽分子之间或两个核酸分子之间的序列相似性。当两个对比序列的一个位置被相同碱基或氨基酸单体亚单位占据时,例如,如果两个 DNA 分子中的每个 DNA 分子的一个位置被腺嘌呤占据,那么,在那个位置上分子同源。两个序列之间的同源性百分比是两个序列共有的匹配或同源位置的数目除以对比位置的数目 $\times 100$ 的函数。例如,如果在两个序列中 10 分之 6 的位置匹配或同源,那么两个序列是 60% 同源。例如, DNA 序列 ATTGCC 和 TATGGC 具有 50% 同源性。通常,比较在比对两个序列时进行,以给出最大同源性。

[0042] 术语“肽(一个或多个)”、“蛋白(一种或多种)”和“多肽(一种或多种)”于此可互换使用。

[0043] 术语“蛋白酶”是指蛋白质或蛋白质的多肽结构域或多肽,它们源于微生物例如真菌、细菌,或源于植物或动物,并且具有催化蛋白骨架上一个或多个不同位置上的肽键的切割的能力。

[0044] 优选地,根据本发明的支链淀粉酶蛋白是分离的或纯化的。纯化或分离的意思是由于将支链淀粉酶从某些或所有自然发生的、与其自然结合的组分中分离,支链淀粉酶蛋白从其自然状态发生改变。这种分离或纯化的完成可以通过本领域公认的分选技术,例如离子交换层析、亲和层析、疏水分选、透析、蛋白酶处理、硫酸铵沉淀或其它蛋白质盐沉淀、离心、大小排阻层析、过滤、微过滤、在梯度上的凝胶电泳或分选进行,从而去除在最终组合物中不期望的完整细胞、细胞碎片、杂质、外来蛋白或酶。进一步可能的是随后将组分加入到含有支链淀粉酶的组合物,这可提供额外优势,例如,活化剂、抗抑制剂、希望的离子、控制 pH 的化合物或其它酶。优选地,根据本发明的支链淀粉酶蛋白通过重组方法产生。

[0045] 如本文使用的“微生物”是指细菌、真菌、病毒、原生动物和其它微生物或微小生物。在本发明中,微生物用作外源肽表达的宿主。

[0046] 如本文使用的“衍生物”、“变体”或“修饰的肽、多肽或蛋白质”表示来源于前体蛋白(例如天然蛋白)的蛋白质,其是通过在 C- 和 N- 端的任一端或两端加入一个或多个氨基酸,在氨基酸序列的一个或多个不同位点置换一个或多个氨基酸,在蛋白质任一端或两端或者在氨基酸序列中的一个或多个位点缺失一个或多个氨基酸,或在氨基酸序列中的一个或多个位点插入一个或多个氨基酸而得到。优选地如下获得支链淀粉酶衍生物制备物:修饰编码天然蛋白的 DNA 序列,将该 DNA 序列转化至适合的宿主,并且表达修饰的 DNA 序列以形成支链淀粉酶衍生物。本发明的“衍生物”包括肽,其包括与前体氨基酸序列(例如野生型或天然状态的支链淀粉酶)相比改变了的氨基酸序列,其中所述肽保留前体支链淀粉酶特有的支链淀粉酶性质,但是在某些具体方面具有改变的性质。例如,支链淀粉酶衍生物可具有增加的最适 pH,对酶促降解或其它降解增加的抗性,增加的酶促有效性,增加的温度或氧化稳定性,但保留其特有的酶促修饰活性。类似地,根据本发明的衍生物包括蛋白质、或其它底物、结合结构域,其已被添加或修饰以改变其底物结合能力。可考虑的是根据本发明的衍生物源于编码支链淀粉酶衍生物的 DNA 片段,其中表达的支链淀粉酶衍生物的功能性活性被保留。衍生物进一步包括改变支链淀粉酶特性的化学修饰。

[0047] 通常地,支链淀粉酶衍生物将具有至少约 50%、70% 或 85% 氨基酸序列同一性,优选地至少约 85% 氨基酸序列同一性,更优选地至少约 90% 氨基酸序列同一性,甚至更优选地至少约 95% 氨基酸序列同一性,还更优选地 98% 氨基酸序列同一性。优选地,任何氨

氨基酸置换是使用 L-氨基酸的“保守氨基酸置换”，其中一种氨基酸被另一种生物学相似的氨基酸置换。保守氨基酸置换是保持被取代氨基酸的通常的电荷、疏水性 / 亲水性、和 / 或空间排列体积 (steric bulk) 的那些置换。保守置换的实例是在下列组之间的那些 :GIy/Ala, Val/Ile/Leu, Lys/Arg, Asn/Gln, Glu/Asp, Ser/Cys/Thr 和 Phe/Trp/Tyr。衍生物可例如只有 1 至 10 个氨基酸残基不同, 如 6-10 个, 只有 5 个、只有 4、3、2 或甚至 1 个氨基酸残基不同。表 1 于此示例性说明在本领域公认的氨基酸置换。

[0048] 如在此使用的, 支链淀粉酶的“天然序列”或支链淀粉酶的“野生型”序列包括多肽, 其具有与源于天然的本菌株的支链淀粉酶相同的氨基酸序列, 或与来自从其制备修饰或衍生支链淀粉酶的支链淀粉酶相同的氨基酸序列, 例如, 由本发明的亲本菌株 *B. deramificans* (BMP139) 表达的支链淀粉酶。这种天然序列的支链淀粉酶可从自然界分离或可通过重组或合成方法产生。在一种实施方式中, 术语“野生型”或“天然序列”的支链淀粉酶是指这样的支链淀粉酶肽, 其中本发明变体从其衍生, 并且为图 7b 中的 SEQ ID NO :2。

[0049] 如此处使用的, 关于此处鉴定的氨基酸或核苷序列的“百分比 (%) 序列同一性”被定义为候选序列中与支链淀粉酶序列中的氨基酸残基或核苷相同的氨基酸残基或核苷的百分比, 如有必要, 在经过比对序列和引入空位 (gap) 后, 以获得最大百分比序列同一性, 并且不考虑将任何保守置换作为序列同一性的部分。进行序列比对和确定序列同一性的方法是普通技术人员已知的, 其可以进行而无不适当的实验, 并且同一性数值的计算可以明确获得。参见, 例如, Ausubel 等人, eds. (1995) *Current Protocols in Molecular Biology*, 第 19 章 (Greene Publishing and Wiley-Interscience, New York); 和 ALIGN 程序 (Dayhoff(1978) in *Atlas of Protein Sequence and Structure 5: Suppl. 3* (National Biomedical Research Foundation, Washington, D. C.)). 许多算法可用于对比序列和确定序列同一性, 并且包括, 例如同源性比对算法, Needleman 等人, (1970) *J. Mol. Biol.* 48 :443; 局部同源性算法, Smith 等人, (1981) *Adv. Appl. Math.* 2 :482; 搜索相似方法, Pearson 等人, (1988) *Proc. Natl. Acad. Sci.* 85 :2444; Smith-Waterman 算法 (*Meth. Mol. Biol.* 70 :173-187 (1997)); 以及 BLASTP、BLASTN 和 BLASTX 算法 (见 Altschul 等人, (1990) *J. Mol. Biol.* 215 :403-410)。使用这些算法的计算程序也是可用的, 包括但不限于 :ALIGN 或 Megalign (DNASTAR) 软件, 或 WU-BLAST-2 (Altschul 等人, *Meth. Enzym.* , 266 :460-480 (1996)); 或 GAP、BESTFIT、BLAST (Altschul 等), 见上文, FASTA 和 TFASTA, 可获自 Genetics Computing Group (GCG) package, Version 8, Madison, Wis., USA; 和 Intelligenetics, Mountain View, Calif 的 PC/Gene 程序中的 CLUSTAL。本领域普通技术人员可确定用于测定比对的适合参数, 包括为获得被比较的序列长度上的最大比对所需的算法。优选地, 使用由程序确定的默认参数测定序列同一性。具体地, 可通过 Smith-Waterman 同源性搜索算法测定序列同一性 (*Meth. Mol. Biol.* 70 :173-187 (1997)), 如使用仿射空位搜索在 MSPRCH 程序 (Oxford Molecular) 中执行的, 所述仿射空位搜索具有下列搜索参数: 空位开放罚分 (gap open penalty) 为 12 并且空位延伸罚分 (gap extension penalty) 为 1。优选地, 成对氨基酸比较可使用 Genetics Computer Group, Inc., Madison, Wis. 的 GCG 序列分析软件包的 GAP 程序进行, 使用 blosum62 氨基酸置换矩阵, 其空位量为 12 并且长度为 2。对于两个氨基酸序列的最适比对, 变体氨基酸序列的连续片段相对于参考氨基酸序列可具有额外的氨基酸残基或缺失的氨基酸残基。用于与参考氨基酸序列比较的连续片段

包括至少 20 个连续的氨基酸残基,并可以是 30、40、50 或更多个氨基酸残基。与衍生物的氨基酸序列中包括空位有关的增加的序列同一性的修正可通过指定空位罚分进行。

[0050] 如本文使用,“表达构建物”(或“表达载体”)是指 DNA 构建物,其包括被可操作连接到适合的控制序列的 DNA 序列,所述适合的控制序列能够影响 DNA 在合适宿主内的表达。这种控制序列可包括:影响转录的启动子、任选的控制转录的操作子序列、编码 mRNA 上合适的核糖体结合位点的序列以及控制转录和翻译终止的序列。本发明不限于使用任何具体表达构建物。不同细胞类型优选地使用不同的表达载体。例如,用于枯草芽孢杆菌的载体的优选启动子是 AprE 启动子;用于 *Bacillus deramificans* 的载体的优选启动子是 amyL 启动子;用于大肠杆菌 (*E. coli*) 的优选启动子是 Lac 启动子;用于啤酒酵母的优选启动子是 PGK1;用于黑曲霉的优选启动子是 glaA;以及用于里氏木霉的优选启动子是 cbhI。载体可以是质粒、噬菌体粒或仅仅是可能的基因组插入。

[0051] 一旦被转化(或转染)至适合的宿主,表达构建物可不依赖于宿主基因组进行复制和发挥功能,或者可以在适合条件下,整合到宿主本身的基因组中。在本说明书中,术语“质粒”、“载体”和“表达构建物(一种或多种)”有时互换使用。然而,本发明意图包括其它形式的表达载体,其发挥等同功能并且其是或成为本领域已知的。因此,大量的宿主/表达载体组合可用于表达本发明的 DNA 序列。本说明书的其它地方未提及的对本发明有用的表达载体,例如,可由染色体片段、非染色体和合成的 DNA 序列组成,例如各种已知的 SV40 衍生物和已知的细菌质粒,例如,来自大肠杆菌 (*E. coli*) 的质粒,其包括 col E1、pCR1、pBR322、pMb9、pUC 19 和其衍生物;更宽宿主范围的质粒,例如 RP4;噬菌体 DNA,例如噬菌体 λ 的众多衍生物,例如 NM989;其它 DNA 噬菌体,例如 M13 和丝状单链 DNA 噬菌体;酵母质粒,例如 2 μ 质粒或其衍生物,用于真核细胞的载体,例如用于动物细胞的载体和源于质粒和噬菌 DNA 组合的载体,如被修饰从而使用噬菌体 DNA 或其它表达控制序列的质粒。

[0052] 使用本发明表达载体的表达技术是本领域已知的并通常描述在,例如 Sambrook 等人,MOLECULAR CLONING:A LABORATORY MANUAL, 第二版, Cold Spring Harbor Press(1989)。通常,包括本发明 DNA 序列的这种表达载体被转化至单细胞宿主,其通过整合事件直接插入至特定物种的基因组中进行(参见,例如 Bennett & Lasure, MORE GENE MANIPULATIONS IN FUNGI, Academic Press, San Diego, pp. 70-76(1991) 和其引用的描述靶向基因组插入到真菌宿主的文章)。

[0053] 术语“可操作连接”、“可操作组合”和“可操作顺序”如本文使用是指连接核酸序列,以使它们发挥它们预期的功能。例如,可操作连接启动子序列至感兴趣的核苷酸序列是指连接启动子序列和感兴趣的核苷酸序列,以使启动子序列能够指导感兴趣的核苷酸序列的转录和/或感兴趣的核苷酸序列编码的多肽的合成的方式进行。类似地,将具有年龄-相关调节活性的核酸序列可操作连接到启动子序列和到感兴趣的核苷酸序列是指,连接具有年龄-相关调节活性的核酸序列、启动子序列和感兴趣的核苷酸序列,以使具有年龄-相关调节活性的核酸序列能够在一段时间改变感兴趣的核苷酸序列转录至 mRNA 的水平 and / 或由感兴趣的核苷酸序列编码的多肽的合成的方式进行。

[0054] 如本文使用“宿主生物”、“宿主菌株”或“宿主细胞”是指对于包括根据本发明 DNA 的表达载体适合的宿主。用于本发明的宿主细胞通常是原核或真核宿主,其包括任何可转化的微生物,其中可获得表达。在本发明的内容中,例如,宿主菌株可以是枯草芽孢

杆菌 (*Bacillus subtilis*)、*Bacillus deramiftcans* (或其它芽孢杆菌种)、埃希氏大肠杆菌 (*Escherichia coli*)、里氏木霉 (*Trichoderma reesei*)、啤酒酵母 (*Saccharomyces cerevisiae*) 或黑曲霉 (*Aspergillus niger*)。以应用重组 DNA 技术构建的载体来转化或转染宿主细胞。这种转化的宿主细胞能够复制编码支链淀粉酶和其衍生物或变体 (突变体) 的载体或表达所期望的肽产物或二者兼有。

[0055] 术语“培养”或“培养条件 (一种或多种)”当用于生长宿主生物群体时,是指培养容器、培养基和培养条件,其适合宿主生物的生长,并且在用本发明核苷酸序列及其变体转染的宿主生物的情况下,适合产生本发明的支链淀粉酶。本发明不限于任何具体培养或培养条件,只要前述是令人满意的。

[0056] 如本文使用,“功能性附着”或“可操作连接”是指调节区,例如启动子、终止子、分泌信号或增强子区,其附着到或连接至结构基因并控制该基因的表达。

[0057] 如本文使用,物质 (例如多核苷酸或蛋白质) “源自”微生物是指该物质是该微生物本身的。

[0058] “木霉 (*Trichoderma*)”或“木霉菌种 (*Trichoderma* sp.)”是指任何真菌菌株,其以往被分类为木霉 (*Trichoderma*) 或目前被分类为木霉 (*Trichoderma*)。优选地,该物种是长柄木霉 (*Trichoderma longibrachiatum*)、里氏木霉 (*Trichoderma reesei*) 或绿色木霉 (*Trichoderma viride*)。在本发明中,木霉菌种可作为本发明的宿主生物使用。

[0059] 如本文所述,本发明的一个方面以“基本上纯的” (或重组的) 核酸和 / 或这种核酸的等价物为特征,所述核酸包括编码支链淀粉酶多肽的核苷酸序列。术语核酸如本文使用可包括片段和等价物。术语“等价物”是指编码功能上等同的多肽或功能上等同的蛋白质的核苷酸序列。等同的核苷酸序列包括由于一个或多个核苷酸置换、添加或缺失而不同的序列,例如等位变体,并且包括与 SEQ ID NO:1 中示出的支链淀粉酶的核苷酸序列不同的序列,原因在于遗传密码的简并性。

[0060] 除非另有所述,本发明实践采用细胞生物学、细胞培养、分子生物学、转基因生物学、微生物学、重组 DNA 和免疫学的常规技术,其在本领域的技术范围之内。这种技术描述于文献中。于此引用的所有出版物、专利和专利申请通过引用为所有目的全部引入于此。见例如: *Molecular Cloning A Laboratory Manual*, 2nd Ed., Sambrook, Fritsch 和 Maniatis 编辑 (Cold Spring Harbor Laboratory Press:1989); *DNACloning*, Volumes I 或 II (D. N. Glover 编辑, 1985); *Oligonucleotide Synthesis* (M. J. Gait ed., 1984); Mullis 等, 美国专利号: 4,683,195; *Nucleic Acid Hybridization* (B. D. Hames 和 S. J. Higgins 编辑, 1984); *Transformation And Translation* (B. D. Hames 和 S. J. Higgins 编辑, 1984); *Culture Of Animal Cells* (R. I. Freshney, Alan R. Liss, Inc., 1987); *Immobilized Cells And Enzymes* (IRL Press, 1986); B. Perbal, *A Practical Guide To Molecular Cloning* (1984); the treatise, *Methods In Enzymology* (Academic Press, Inc., N. Y.); *Gene Transfer Vectors For Mammalian Cells* (J. H. Miller and M. P. Calos 编辑, 1987, Cold Spring Harbor Laboratory); *Methods In Enzymology*, Vols. 154 and 155 (Wu 等人, 编辑), *Immunochemical Methods In Cell And Molecular Biology* (Mayer 和 Walker 编辑, Academic Press, London, 1987); *Handbook Of Experimental Immunology*, Volumes I-IV (D. M. Weir and C. C. Blackwell, 编辑, 1986); *Manipulating the Mouse Embryo*, (Cold Spring

Harbor Laboratory Press, Cold Spring Harbor, N. Y., 1986)。同样,关于蛋白酶的制备、表达、分离和使用方法的信息可通过美国专利第 6,768,001 号获得,其于此通过引用完整引入。

[0061] 在下列具体描述和权利要求中,本发明的其它特征和优势将明显可见。

发明详述

[0062] 本发明现仅使用下述定义和实例,通过引用方式详述于此。所有专利和出版物,包括在这些专利和出版物内公开的所有序列,于此明确地通过引用而并入。

[0063] 除非于此另外定义,如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含义。Singleton 等人的 DICTIONARY OF MICROBIOLOGY AND MOLECULAR BIOLOGY, 2D ED., John Wiley 和 Sons, New York (1994) 和 Hale & Marham, THE HARPER COLLINS DICTIONARY OF BIOLOGY, Harper Perennial, NY (1991) 提供给技术人员在本发明中使用的许多术语的通用字典。虽然任何类似或等同于那些在此所述的方法和材料可在本发明的实践或测试中使用,但是优选的方法和材料被描述。数值范围包括定义该范围的数值。除非另有所述,分别地,核酸以 5' 至 3' 方向从左向右书写;氨基酸序列以氨基至羧基的方向从左向右书写。对本领域的定义和术语,专业人员可具体注意 Sambrook 等人, 1989 和 Ausubel 等人, 1993。应该理解本发明不限于所描述的具体方法、方案和试剂,因为这些可变化。

[0064] 数值范围包括定义该范围的数值。

[0065] 除非另有所述,分别地,核酸以 5' 至 3' 方向从左向右书写;氨基酸序列以氨基至羧基的方向从左向右书写。

[0066] 本文提供的标题不限制本发明的各个方面或实施方式,其可通过作为整体参考说明书而理解。因此,紧接在下面定义的术语通过作为整体参考说明书而被更充分地定义。

分子生物学

[0067] 在一种实施方式中,本发明提供在 amyL 启动子控制下异源基因的表达。因此,本发明依赖于重组遗传领域的常规技术。公开用于本发明的普通方法的基本文本包括 Sambrook 等人, Molecular Cloning, A Laboratory Manual (2nd ed. 1989); Kriegler, Gene Transfer and Expression: A Laboratory Manual (1990) 和 Ausubel 等人编辑, Current Protocols in Molecular Biology (1994))。

[0068] 包括丝状真菌的纤维素酶基因启动子序列的异源基因,在被转化到里氏木霉细胞用于复制和/或表达之前,通常被克隆到中间载体中。这些中间载体典型地是原核载体,例如质粒或穿梭载体。

[0069] 为了获得克隆基因的高水平表达,异源基因优选地位于与自然发生纤维素酶基因距启动子的距离大约相同的距离上。但是,如本领域已知的,该距离的一些变化可允许而不损失启动子功能。

[0070] 本领域普通技术人员可意识到自然启动子可通过置换、取代、添加或去除一种或多种核苷而不改变其功能来修饰。本发明实践包括并不限于这种对启动子的改变。

[0071] 表达载体/构建物典型地包含转录单位或表达盒,所述转录单位或表达盒包含所

有异源序列表达所需的附加元件。因此,典型的表达盒包含可操作地连接至异源核酸序列的启动子和转录物有效多聚腺嘌呤化作用所需的信号,核糖体结合位点和翻译终止。所述盒的附加元件可包括增强子和,如果基因组 DNA 用作结构基因,具有功能性剪接供体和受体位点的内含子。

[0072] 本发明的实践不受遗传构建物中启动子的选择的限制。但是,示例性启动子是里氏木霉 *cbh1*、*cbh2*、*egl*、*eg2*、*eg3*、*eg5*、*xln1* 和 *xln2* 启动子。

[0073] 除启动子序列之外,表达盒还应包含结构基因的转录终止区下游,以提供有效终止。终止区可从与启动子序列相同的基因获得,或可从不同基因获得。

[0074] 虽然任何真菌终止子在本发明中可能具有功能,但优选的终止子包括:来自构巢曲霉 (*Aspergillus nidulans*) *trpC* 基因的终止子 (Yelton, M. 等人, (1984) PNAS USA 81:1470-1474; Mullaney, E. J. 等人, (1985) MGG 199:37-45), 泡盛曲霉 (*Aspergillus awamori*) 或黑曲霉 (*Aspergillus niger*) 葡糖淀粉酶基因 (Nunberg, J. H. 等 (1984) Mol. Cell Biol. 4:2306, Boel, E. 等 (1984) EMBO J. 3:1581-1585) 和米黑毛霉 (*Mucor miehei*) 羧基蛋白酶基因 (EPO 公开号 0 215 594)。

[0075] 用于将遗传信息传递到细胞中的具体表达载体不是特别重要。任何用于在真核或原核细胞中表达的常规载体均可使用。标准细菌表达载体包括细菌噬菌体 λ 和 M13, 以及质粒如基于 pBR322 的质粒、pSKF、pET23D, 和融合表达系统, 例如 MBP、GST 和 LacZ。表位标记也可以被加到重组蛋白中, 以提供方便的分离方法, 例如 c-myc。

[0076] 通常包括在表达载体中的元件还包括复制子、编码抗生素抗性的基因, 以允许筛选含有重组质粒的细菌、和质粒的非重要区中的独特的限制位点, 以允许异源序列的插入。所选择的特定抗生素抗性基因不重要, 任何本领域已知的许多抗性基因都是适合的。原核序列被优选地选择, 因为它们不干扰里氏木霉 DNA 的复制或整合。

[0077] 本发明的转化方法可引起转化载体的全部或部分稳定整合到丝状真菌基因组中。但是, 也考虑导致维持自我复制的染色体外转化载体的转化。

[0078] 许多标准转染方法可用于产生表达大量异源蛋白的里氏木霉细胞系。某些已公开的、用于将 DNA 构建物引入到生产纤维素酶的木霉 (*Trichoderma*) 菌株的方法包括: Lorito, Hayes, DiPietro 和 Harman, 1993, Curr. Genet. 24:349-356; Goldman, VanMontagu 和 Herrera-Estrella, 1990, Curr. Genet. 17:169-174; Penttila, Nevalainen, Ratto, Salminen 和 Knowles, 1987, Gene 6:155-164, 用于曲霉 (*Aspergillus*), Yelton, Hamer 和 Timberlake, 1984, Proc. Natl. Acad. Sci. USA 81:1470-1474, 用于镰刀菌属 (*Fusarium*), Bajar; Podila 和 Kolattukudy, 1991, Proc. Natl. Acad. Sci. USA 88:8202-8212, 用于链霉菌 (*Streptomyces*), Hopwood 等, 1985, The John Innes Foundation, Norwich, UK 和用于芽孢杆菌 (*Bacillus*), Brigidi, DeRossi, Bertarini, Riccardi 和 Matteuzzi, 1990, FEMS Microbiol. Lett. 55:135-138)。

[0079] 但是, 可使用任何用于将外源核苷酸序列引入到宿主细胞的熟知方法。这些方法包括使用磷酸钙转染、凝聚胺 (polybrene)、原生质体融合、电转化、生物射弹、脂质体、显微注射、原生质载体 (plasmavector)、病毒载体和任何其它熟知的方法, 用于将克隆的基因组 DNA、cDNA、合成的 DNA 或其它外源遗传物质引入至宿主细胞 (见, 例如, Sambrook 等人, 上文)。还使用的是农杆菌 (*Agrobacterium*) - 介导的转染方法, 其在美国专利第 6, 255, 115

号中描述。仅仅必需的是使用的特定遗传工程方法能够成功地将至少一个基因引入到能够表达异源基因的宿主细胞中。

[0080] 本发明还涉及由微生物异源产生的支链淀粉酶。合适的细菌的实例是革兰氏阳性细菌,例如枯草芽孢杆菌 (*Bacillus subtilis*)、地衣芽孢杆菌 (*Bacillus licheniformis*)、迟缓芽孢杆菌 (*Bacillus lentus*)、短芽孢杆菌 (*Bacillus brevis*)、嗜热脂肪芽孢杆菌 (*Bacillus stearothermophilus*)、嗜碱芽孢杆菌 (*Bacillus alkalophilus*)、解淀粉芽孢杆菌 (*Bacillus amyloliquefaciens*)、凝结芽孢杆菌 (*Bacillus coagulans*)、环状芽孢杆菌 (*Bacillus circulans*)、灿烂芽孢杆菌 (*Bacillus lautus*)、巨大芽孢杆菌 (*Bacillus megaterium*)、苏云金芽孢杆菌 (*Bacillus thuringiensis*) 或铅青链霉菌 (*Streptomyces lividans*) 或鼠灰链霉菌 (*Streptomyces murinus*)、或革兰氏阴性细菌,例如大肠杆菌 (*E. coli*)。细菌的转化可通过例如原生质体转化或以已知方式本身使用感受态细胞来实现。

[0081] 通常,本发明包括通过表达掺入表达系统的 DNA——其已被转化至宿主细胞——来生产支链淀粉酶的方法。多种宿主 / 表达载体组合可以被用于表达本发明的 DNA 序列。许多原核和真核表达载体是商业上可得的。适合的表达载体的选择在本领域普通技术人员知识范围之内。载体可以是质粒、噬菌粒、或仅仅是可能的基因组插入。一旦转化至适合的宿主,载体可不依赖于宿主基因组进行复制和作用,或可以,在某些情况下,整合到宿主本身的基因组中。在本说明书中,质粒和载体有时可互换使用,因为质粒是目前载体最经常使用的形式。但是,本发明意图包括这种表达载体的其它形式,其发挥相等的功能并且是或成为本领域已知的。有用的表达载体例如包括染色体、非染色体和合成的 DNA 序列的片段,例如对于此目的有用的各种已知的质粒和噬菌体。另外,任何的各种表达控制序列通常被用于这些载体。

[0082] 在本发明中有用的宿主细胞通常是原核的或真核的宿主,包括在其中可获得根据本发明的支链淀粉酶表达的任何可转化的微生物。以使用重组 DNA 技术构建的载体转化或转染宿主细胞。这种转化的宿主细胞能够复制编码支链淀粉酶和其变体(突变体)的载体或表达所期望的支链淀粉酶。这些宿主可包括熟知的真核和原核宿主,例如埃希氏大肠杆菌 (*E. coli*)、假单胞菌 (*Pseudomonas*)、芽孢杆菌 (*Bacillus*)、链霉菌 (*Streptomyces*)、各种真菌、酵母的菌株和动物细胞。优选地,所述宿主在细胞外表达本发明的支链淀粉酶便于纯化和下游过程。

[0083] 在一些实施方式中,宿主细胞是芽孢杆菌属的成员;而在某些实施方式中,感兴趣的芽孢杆菌 (*Bacillus*) 菌株是工业芽孢杆菌 (*Bacillus*) 菌株。工业芽孢杆菌 (*Bacillus*) 菌株的实例包括,但不限于地衣芽孢杆菌 (*B. licheniformis*)、枯草芽孢杆菌 (*B. subtilis*)、迟缓芽孢杆菌 (*B. lentus*)、解淀粉芽孢杆菌 (*B. amyloliquefaciens*)。在另外的实施方式中,芽孢杆菌 (*Bacillus*) 宿主菌株选自迟缓芽孢杆菌 (*B. lentus*)、短芽孢杆菌 (*B. brevis*)、嗜热脂肪芽孢杆菌 (*B. stearothermophilus*)、嗜碱芽孢杆菌 (*B. alkalophilus*)、凝结芽孢杆菌 (*B. coagulans*)、环状芽孢杆菌 (*B. circulans*)、短小芽孢杆菌 (*B. pumilus*)、苏云金芽孢杆菌 (*B. thuringiensis*)、克劳氏芽孢杆菌 (*B. clausii*) 和巨大芽孢杆菌 (*B. megaterium*),以及芽孢杆菌 (*Bacillus*) 属的其它生物,如上所述。在某些实施方式中,枯草芽孢杆菌 (*B. subtilis*) 被使用。在其它实施方式中,使用地衣芽孢

杆菌 (*B. licheniformis*)。例如,美国专利第 5,264,366 和 4,760,025 (RE34,606) 号,和 US2002/0182734 (国际公开号 WO 02/14490) 描述了各种芽孢杆菌 (*Bacillus*) 宿主菌株,它们在本发明中有用,尽管其它适合的菌株被考虑用于本发明。优选地,使用蛋白酶阴性的芽孢杆菌 (*Bacillus*) 菌株 (基因缺失的,例如 Δapr 或 Δnpr 等)。

[0084] 转化芽孢杆菌 (*Bacillus*) 种的各种方法是已知的。实际上,用于改变包含质粒构建物的芽孢杆菌 (*Bacillus*) 的染色体和将质粒转化到大肠杆菌的方法是熟知的。在大多数方法中,质粒随后从大肠杆菌中分离并被转化到芽孢杆菌 (*Bacillus*) 中。然而,使用这种间插微生物例如大肠杆菌不是必须的,并且,在某些实施方式中,DNA 构建物通过原生质体或感受态细胞转化,被直接转化到感受态芽孢杆菌宿主中。本发明的突变体支链淀粉酶的表达和纯化可以通过本领域公认的用于进行这种过程的方法来实现。

[0085] 在表达载体被引入细胞后,培养转染的细胞,其培养条件是有助于在蛋白酶基因启动子序列控制下的基因表达。大量转化的细胞可如下述实施例所述进行培养。最后,使用标准技术从培养物中回收产物。

[0086] 因此,本发明于此提供所期望多肽的表达和增强分泌,其表达是在基因启动子序列控制下进行,包括自然发生淀粉酶基因、融合 DNA 序和各种异源构建物的表达。本发明还提供用于表达和分泌高水平的这种所期望的肽的方法。

蛋白质表达

[0087] 本发明蛋白质的生产通过培养用表达载体转化的细胞进行,其中表达载体含有在淀粉酶基因启动子序列控制下表达的基因。本发明特别适用于加强蛋白质的细胞内和/或细胞外产生。蛋白质可以是同源的或异源的。可由本发明生产的蛋白质包括但不限于:激素、酶、生长因子、细胞因子、抗体以及类似物。

[0088] 酶包括但不限于:水解酶,例如蛋白酶、酯酶、脂肪酶、酚氧化酶、透酶、淀粉酶、支链淀粉酶、纤维素酶、葡萄糖异构酶、漆酶和蛋白二硫键异构酶。

[0089] 适合表达所述基因的条件包括提供给培养物诱导饲养组合物,见例如 US-2004-0121446。蛋白质产生的最适条件根据宿主细胞的选择和待表达的蛋白质的选择而变化。本领域普通技术人员将通过常规实验或优化容易确定这些条件。

[0090] 感兴趣的蛋白,例如此处所述的支链淀粉酶,通常在表达后纯化或分离。感兴趣的蛋白可以用本领域普通技术人员已知的各种方法分离或纯化,这取决于在样本中存在的其它组分。标准纯化方法包括电泳技术、分子技术、免疫技术和层析技术,包括离子交换、疏水、亲和以及反相 HPLC 层析和聚焦层析。例如,感兴趣的蛋白可以使用标准抗感兴趣蛋白的抗体柱纯化。超滤和渗滤技术,与蛋白浓度联合也是有效的。对适合纯化技术的一般指导,见 Scopes, *Protein Purification* (1982)。必需的纯化程度将依赖于感兴趣蛋白的用途而变化。在某些情况下,纯化不是必需的。

本发明支链淀粉酶的类似物

[0091] 类似物可以在氨基酸序列或以不涉及序列的方式,或以此两种方式不同于“野生型”亲本支链淀粉酶或天然发生的支链淀粉酶。非序列修饰包括支链淀粉酶的体内或体外化学衍生。非序列修饰包括在乙酰化、甲基化、磷酸化、羧化或糖基化中的改变。

[0092] 优选的类似物包括支链淀粉酶（或其生物学活性片段），其序列不同于“野生型”序列或自然序列，有一个或多个保守氨基酸置换或有一个或多个非保守氨基酸置换、缺失或插入，这些不同不破坏支链淀粉酶的生物学活性。保守置换典型地包括用一种氨基酸置换另一种具有相似性质的氨基酸，例如在下列组中的置换：缬氨酸，甘氨酸；甘氨酸，丙氨酸；缬氨酸，异亮氨酸，亮氨酸；天冬氨酸，谷氨酸；天冬酰胺，谷氨酰胺；丝氨酸，苏氨酸；赖氨酸，精氨酸；以及苯丙氨酸，酪氨酸。

[0093] 可根据例如下表进行保守置换，所述表描述了普遍接受的维恩图 (Venn diagram) 氨基酸分组。表 1

组		亚组	
疏水	F W Y H K M I L V A G C	芳香族	F W Y H
		脂肪族	I L V
极性	W Y H K R E D C S T N Q	带电的	H K R E D
		带正电	H K R
		带负电	E D
小的	V C A G S P T N D	极小	A G S

其它保守置换可从下表获得。表 2 保守氨基酸替换

氨基酸	代码	由以下任一替换
丙氨酸	A	D-Ala, Gly, β -Ala, L-Cys, D-Cys
精氨酸	R	D-Arg, Lys, D-Lys, 高-Arg, D-高-Arg, Met, Ile, D-Met, D-Ile, Orn, D-Orn
天冬酰胺	N	D-Asn, Asp, D-Asp, Glu, D-Glu, Gln, D-Gln
天冬氨酸	D	D-Asp, D-Asn, Asn, Glu, D-Glu, Gln, D-Gln
半胱氨酸	C	D-Cys, S-Me-Cys, Met, D-Met, Thr, D-Thr
谷氨酰胺	Q	D-Gln, Asn, D-Asn, Glu, D-Glu, Asp, D-Asp
谷氨酸	E	D-Gln, D-Asp, Asp, Asn, D-Asn, Gln, D-Gln
甘氨酸	G	Ala, D-Ala, Pro, D-Pro, β -Ala, Acp
异亮氨酸	I	D-Ile, Val, D-Val, Leu, D-Leu, Met, D-Met
亮氨酸	L	D-Leu, Val, D-Val, Ile, D-Ile, Met, D-Met
赖氨酸	K	D-Lys, Arg, D-Arg, 高-Arg, D-高-Arg, Met,

		D-Met, Ile, D-Ile, Orn, D-Orn
甲硫氨酸	M	D-Met, S-Me-Cys, Ile, D-Ile, Leu, D-Leu, Val, D-Val
苯丙氨酸	F	D-Phe, Tyr, D-Tyr, L-Dopa, His, D-His, Trp, D-Trp, 反-3,4 或 5- 苯基脯氨酸, 顺-3,4 或 5- 苯基脯氨酸
脯氨酸	P	D-Pro, L-1- 噻唑烷 (thioazolidine)-4- 羧酸, D- 或 L-1- 噻唑烷 -4- 羧酸
丝氨酸	S	D-Ser, Thr, D-Thr, 别-Thr, Met, D-Met, Met(O), D-Met(O), L-Cys, D-Cys
苏氨酸	T	D-Tyr, Ser, D-Ser, 别-Thr, Met, D-Met, Met(O), D-Met(O), Val, D-Val
酪氨酸	Y	D-Tyr, Phe, D-Phe, L-Dopa, His, D-His
缬氨酸	V	D-Val, Leu, D-Leu, Ile, D-Ile, Met, D-Met

[0094] 本发明中的其它类似物是带有增加肽稳定性的修饰的那些;这种类似物可含有,例如在肽序列中一个或多个非-肽键(其替换肽键)。还包括:这样的类似物,其包括除了天然存在的 L- 氨基酸之外的残基,例如 D- 氨基酸或非天然存在或合成的氨基酸,例如, α 或 β 氨基酸类似物和环状类似物。

其它实施方式

[0095] 本发明包括的是:等位变异;自然突变体;诱导突变体;由 DNA 编码的蛋白质,所述 DNA 在高或低严谨性条件下与编码 SEQ IDNO:1 的多肽的核酸杂交(对高和低严谨性的定义见 Current Protocols in Molecular Biology, John Wiley & Sons, New York, 1989, 6.3.1-6.3.6, 通过引用并入于此);和多肽,其通过抗血清特异性地结合支链淀粉酶,特别是通过抗血清结合于支链淀粉酶活性位点或结合结构域。

[0096] 本发明的核酸和多肽包括不同于在此公开的序列,原因在于公开序列中的测序错误。

[0097] 本发明还包括片段,优选地支链淀粉酶的生物学活性片段或类似物。生物学活性片段或类似物是具有任何体内或体外活性的物质,其特征在于 SEQ ID NOS:2、4 和 6 中示出的支链淀粉酶或其它天然存在的支链淀粉酶,例如,在此所述的生物学活性的一种或多种。特别优选的是这样的片段,其存在于体内,例如从转录后加工产生的或从可选地剪切 RNA 的翻译产生的片段。片段包括表达在天然或内源细胞的那些,例如,翻译后加工的结果,例如,去除氨基-端信号序列的结果以及在表达系统例如 CHO 细胞中制造的那些。特别优选的片段是这样的片段,例如活性片段,其由蛋白水解切割或可选的剪接事件产生。由于肽例

如支链淀粉酶通常表现出一定范围的生理性质,并且由于这种性质可归因于分子的不同部分,所以有用的支链淀粉酶片段或支链淀粉酶类似物是在对支链淀粉酶活性的任何生物学分析中显示出生物学活性的物质。最优选地,片段或类似物在任何体内或体外的支链淀粉酶分析中,具有 10%、40%、60%、70%、80%或至少 90%的支链淀粉酶活性 (SEQ ID NOS: 2、4 和 6)。一种制造这种支链淀粉酶类似物的方法包括通过定向分子进化合成支链淀粉酶类似物,如下所述。

[0098] 支链淀粉酶片段可用本领域普通技术人员已知的方法产生。显示支链淀粉酶生物学活性的候选片段的能力可以通过如本文所述的本领域普通技术人员已知的方法进行评估。还包括的支链淀粉酶肽含有这样的残基,其对于肽的生物学活性是不需要的或由可选的 mRNA 剪接或可选的蛋白加工事件产生。

[0099] 为了获得支链淀粉酶肽,编码支链淀粉酶的 DNA 可被引入到表达载体中,载体被引入适合所期望的蛋白表达的细胞,并通过现有技术方法回收和纯化肽。抗肽和蛋白的抗体可通过免疫动物,例如兔或小鼠,以及通过现有技术方法回收抗-支链淀粉酶抗体来制备。

发明的工业应用

[0100] 本发明在工业中具有许多实际应用,正如本文可考虑的,这种描述的意图是示例性的而非包含性的。

[0101] 在几种实施方式中,本发明考虑用于乙醇生产、烘焙、果汁生产、酿造、蒸馏、造酒、皮革、油类和脂肪、造纸和纸浆和动物饲料生产。

[0102] 在其它实施方式中,本发明考虑用作去垢剂和清洁产品的活性“生物学”组分。这里,蛋白酶、淀粉酶和脂肪酶被用于分解蛋白、淀粉和脂肪污点。本发明的实施方式包括测试酶与去污剂成分的相容性,其通过进行稳定性研究和在多种配方中测试它们而进行。

[0103] 在另一种实施方式中,本发明考虑用于纺织工业中,主要用于织物和衣物的整理。主要应用包括:从编织后织物的丝线中脱浆、去除浆(即,去除纤维的硬成分);生物抛光过程(Bio-polishing-a process),用于降低起球倾向,并给予织物更加平滑和光泽的外观;生物磨石过程(Bio-stoning-a process),其中小剂量的酶可代替用于石洗粗斜纹棉布以获得用旧外观的传统浮石。

[0104] 而在另一种实施方式中,本发明考虑酶促用途,用于将淀粉液化和糖化为葡萄糖并异构化为果糖。本发明可用于将大量的玉米和其它谷物转换为甜料,如高果糖玉米糖浆和麦芽糖糖浆。

实施例

[0105] 在下列实验性公开中,应用下列缩写词:eq(当量);M(摩尔浓度); μ M(微摩尔浓度);N(当量浓度);mol(摩尔数);mmol(毫摩尔数); μ mol(微摩尔数);nmol(纳摩尔数);g(克);mg(毫克);kg(千克); μ g(微克);L(升);ml(毫升); μ l(微升);cm(厘米);mm(毫米); μ m(微米);nm(纳米); $^{\circ}$ C(摄氏度);h(小时);min(分钟);sec(秒);msec(微秒);TLC(薄层层析);nt(核苷酸);Q(谷氨酰胺);E(谷氨酸);CAP(choloroamphenicol 氯霉素)。

[0106] 本发明在下列实施例中进一步描述,其不以任何方式试图限制本发明要求保护的
范围。附图表示被认为是本发明的说明书和描述的组成部分。于此引用的所有文献具体地
通过引用如同于此描述。下列实施例被提供用于说明,但不限制所要求保护的发明。

实施例 1

支链淀粉酶变体的设计

[0107] 设计下列支链淀粉酶变体(见图 1)。

[0108] “PUL”这是“野生型”*B. deramificans* 支链淀粉酶,与 BMP139 表达的分子相同。基
因已被密码子优化,由 amyL(LAT) 启动子驱动,并具有 amyL 信号序列。这种构建与存在于
BMP139 中的构建的差别如下(见图 2)。首先,与 BMP139 中的天然编码序列相比,该新构建
具有密码子优化的编码区。第二,相对于 BMP139 中大约 800nt,该新构建具有较短的 100nt
的 amyL 启动子区。第三,该新构建具有 amyL 终止子,而 BMP 139 具有 *B. deramificans* 支链
淀粉酶终止子。新和旧的构建都具有 amyL 信号序列并都表达相同的支链淀粉酶分子。虽
然由这种新构建表达的分子与目前产物相同,但其可以具有产生滴度方面的优势,这是密
码子优化的结果。

[0109] “PULm104”这是 *B. deramificans* 支链淀粉酶,其 N-端 104 个氨基酸已被去除。该
构建包括 amyL 启动子、amyL 信号序列、密码子优化的支链淀粉酶编码区,其缺乏编码成熟
支链淀粉酶 N-端 104 个氨基酸的序列,和 amyL 终止子。截短的支链淀粉酶 PULm104 类似
PULm98 和 PULm102 分子,所述 PULm98 和 PULm102 分子由在 E99 和 E103 进行 N-端剪切全长
支链淀粉酶产生。支链淀粉酶被删除可达氨基酸 104,以获得 amyL 信号序列和支链淀粉酶
序列 ASA-A 之间理想的信号肽酶靶共有序列。支持这种截短的支链淀粉酶变体的基本原因
是遵循之前意料不到的观察结果——与全长分子相比,在剪切的支链淀粉酶变体中可见较
高的特异性活性。

[0110] “PUL_E99Q_E103Q”这是 *B. deramificans* 支链淀粉酶,其中在 E99 和 E103 的蛋白
酶靶基序已被修饰为 Q99 和 Q103,目的是使支链淀粉酶分子可抵抗在 E99 和 E103 的剪切。
而且,在 51h 之后的支链淀粉酶降解依赖于在 E99 和 E103 的起始剪切的情况下,这种修饰
可被期望防止 51h 后的降解和活性降低。

实施例 2

质粒构建和转化

[0111] 合成两种密码子优化的支链淀粉酶构建物,一种编码“野生型”支链淀粉酶蛋白,
另一种编码 E99Q_E103Q 变体。

[0112] 两种都包括具有 amyL 启动子(以前已证明允许在 *E. coli* 中克隆;较长的启动子
延伸是致死的)、amyL 信号序列、密码子优化的支链淀粉酶(变体)序列和 amyL 终止子的
57 个核苷酸。这些构建物用作上述三种支链淀粉酶构建的 PCR-构建的模板:

[0113] “PUL”下列引物被用于从合成的“野生型”支链淀粉酶构建物扩增“野生型”支链淀
粉酶构建(XhoI 位点为粗体):Plat5-XhoI_FW:ccccgcctcgaggcttttcttttgaagaaaatatag

ggaaaatggtacttgttataaaattcggaatattttatacaatatcatatgt ttacattgaaagggg[SEQ ID NO. : 7]。Tlat-XhoI_RV :tggaatctcgaggttttatcctttaccttgtctcc[SEQ ID NO. :8]。

[0114] “PULm104”用于截短的支链淀粉酶的表达盒通过融合 PCR 产生。下列两个片段从合成的“野生型”支链淀粉酶构建物中扩增,并随后融合:

[0115] 覆盖 amyL 启动子和 amyL 信号序列的片段。

[0116] 覆盖截短的支链淀粉酶编码序列和 amyL 终止子的片段。

[0117] Ad A) 下列引物用于扩增 amyL 启动子和信号序列:Plat5-XhoI_FW[SEQ ID NO : 7](见上文)。ssLAT-PULm104_RV :gcgttgctgactgccggttttagcagctgctgaagctgcagaatg aggcagc[SEQ ID NO :9](融合引物;反向支链淀粉酶序列,其起始于密码子 105,标为粗体)。

[0118] Ad B) 下列引物用于扩增支链淀粉酶编码序列和 amyL 终止子:ssLAT-PULm104_FW :gctgcctcattctgcagcttcagcagctgctaaaccggcagtcagcaacgc[SEQ ID NO :10](融合引物:支链淀粉酶序列,其起始于密码子 105,标为粗体)。Tlat-XhoI_RV :[SEQ ID NO :8](见上文)。

[0119] 融合引物各自包含两个序列段,其在模板序列中分开 312nt(代表 N-端 104 个氨基酸)。A) 和 B) 所述的两个 PCR 片段在使用引物 Plat5-XhoI_FW 和 Tlat-XhoI_RV 的 PCR 反应中融合。

[0120] “PUL_E99Q_E103Q”E99Q_E103Q 支链淀粉酶变体的构建与“野生型”支链淀粉酶的构建物相同(见上文 1),采用 E99Q_E103Q 合成构建物作为模板。

[0121] 产生的片段在两个方向被克隆到地衣芽孢杆菌整合载体 pICatH 的 XhoI 位点(图 3)。这样产生六个构建物:pICatH-PUL-Ori1(图 4)、pICatH-PUL-Ori2、pICatH-PULm104-Ori1(图 5)、pICatH-PULm104-Ori2、pICatH-PUL_E99Q_E103Q-Ori1(图 6)和 pICatH-PUL_E99Q_E103Q-Ori2。

[0122] Ori1 和 Ori2 构建物相对于氯霉素-抗性(catH)基因具有相反方向的支链淀粉酶基因。图 4(a)、5(a) 和 6(a) 示出三个 Ori1 构建物的质粒图谱。

[0123] 所有六种构建物被转化到枯草芽孢杆菌中,并对在 AZCL-普鲁分支葡聚糖(Megazyme)覆盖物(0.1%在 100mM NaAc pH 5 中,1%琼脂)晕形成进行筛选。pICatH-PULm104 转化子产生比任何一种全长支链淀粉酶的转化子都大的晕。构建物被测序验证并使用原生质体转化将其转化到地衣芽孢杆菌宿主菌株 BML612 和 BML780 中。

实施例 3

整合到地衣芽孢杆菌基因组

[0124] 转化后,在含有 5 μ g/ml 氯霉素和 10 μ g/ml 新霉素的最低再生平板(minimal regeneration plate)筛选转化子。转化子被影印铺板至含有相同抗生素的两个心浸液琼脂(Heart Infusion-agar)平板(本领域普通技术人员已知的),其中一个用 AZCL-普鲁分支葡聚糖覆盖,以筛选支链淀粉酶阳性的转化子。与枯草芽孢杆菌的情况类似,PULm104 转化子显示出最大的晕。质粒被整合进地衣芽孢杆菌染色体上的 catH 基因座。因此,如表 2 所示,继续进行下列整合,进一步用于去除/扩增:表 2

BML612PUL Ori1
 BML612PUL Ori2
 BML612PULm104Ori1
 BML612PULm104Ori2
 BML612PUL_E99Q_E103Q Ori1
 BML612PUL_E99Q_E103Q Ori2
 BML780PUL Ori1
 BML780PUL Ori2
 BML780PULm104 Ori1
 BML780PULm104 Ori2
 BML780PUL_E99Q_E103Q Ori1
 BML780PUL_E99Q_E103Q Ori2

[0125] 质粒去除和盒扩增如下进行。无外源 DNA 的菌株 (“免除菌株”) 的获得是通过去除载体序列 (“环出”), 只留下整合进染色体的 catH- 支链淀粉酶表达盒。随后, 通过将所述菌株置于逐步增加的氯霉素浓度 (5、25、50、75 μ g/ml) 来扩增表达盒。在每个扩增步骤后, 通过用 AZCL- 普鲁分支葡聚糖覆盖影印平板来监控支链淀粉酶的产生。对每种菌株, 获得四个扩增水平: CAP5、CAP25、CAP50 和 CAP75。

实施例 4

评价地衣芽孢杆菌支链淀粉酶菌株

[0126] 在所有扩增水平的菌株都被挑取至单个大的含有 5 μ g/ml 氯霉素的心浸液琼脂 (Heart Infusion-agar) 平板, 每个菌株挑取 2 个, 并在 37°C 生长过夜。支链淀粉酶产生菌株 BMP 139 被包括作为基准。用 AZCL- 普鲁分支葡聚糖琼脂覆盖平板来显现支链淀粉酶活性。覆盖物在 37°C 温育 8h, 随后在室温温育 16h。结果概括于下表 3。

表 3

菌株	CAP5	CAP25	CAP50	CAP75	
BML612 PUL Ori1	+	+	+	++	
BML612 PUL Ori2	+	+	+	++	
BML612 PULm104 Ori1	+	++	++	+++	
BML612 PULm104 Ori2	+	+	+++	+++	
BML612PUL_E99Q_E103Q Ori1	+	+	++	++	
BML612PUL_E99Q_E103Q Ori2	+	+	++	++	
BML780PUL Ori1	+	+	+	++	

BML780PUL Ori2	+	+	++	++	
BML780PULm104 Ori1	++	++	+++	+++++	
BML780PULm104 Ori2	++	++	+++	+++++	
BML780PUL_E99Q_E103Q Ori1	++	++	++	++	
BML780PUL_E99Q_E103Q Ori2	+	++	++	+++	
BMP139					++

图例 :+ = 晕直径 7-9mm++ = 晕直径 10-12mm+++ = 晕直径 13-15mm++++ = 晕直径 16-18mm+++++ = 晕直径 19-21mm

[0127] 从评价可见,明显的是扩增导致滴度和/或性能的提高。更具体的结论是:1)N-端截短的PUL菌株(PULm 104)具有非常显著的优于全长支链淀粉酶菌株的性能。BML780 PULm 104 CAP75菌株产生的晕比BMP 139菌株和BML780PUL CAP75菌株的晕,具有增加2.5倍的表面(直径增加多于1.5倍)。这表明较短的分子在较高的滴度产生,或其活性相对于全长支链淀粉酶分子增加。2)PUL_E99Q_E103Q变体可略优于“野生型”PUL。BML780PUL_E99Q_E103Q菌株产生的晕稍微大于BML780PUL菌株产生的晕。3)BMP139产生菌株显示出与CAP75扩增的“野生型”PUL菌株相同的性质。因此,基于平板评价,密码子优化未产生增加的性能。4)BML780菌株通常具有比BML612菌株更好的性质。这种观察与以往在BML612背景下支链淀粉酶的降解数据相符。事实是此处构建的BML612支链淀粉酶菌株还显示合理的AZCL-清除,表明在平板上,支链淀粉酶甚至在BML612背景下也是相对稳定的。5)在Ori1和Ori2支链淀粉酶菌株之间未观察到明显的性质差异。

[0128] 本说明书提到的所有专利和公开都通过引用被并入于此。对本发明所记载的方法和系统的各种修改和改变对本领域普通技术人员将是显而易见的,并不背离本发明的范围和精神。虽然本发明已联系具体优选的实施方式进行描述,但是应当理解,所保护的本发明不应不适当地限于这些具体的实施方式。实际上,对进行本发明的所描述模式的各种修改,其对本领域普通技术人员来说是显而易见的,意图在所附权利要求的范围内。

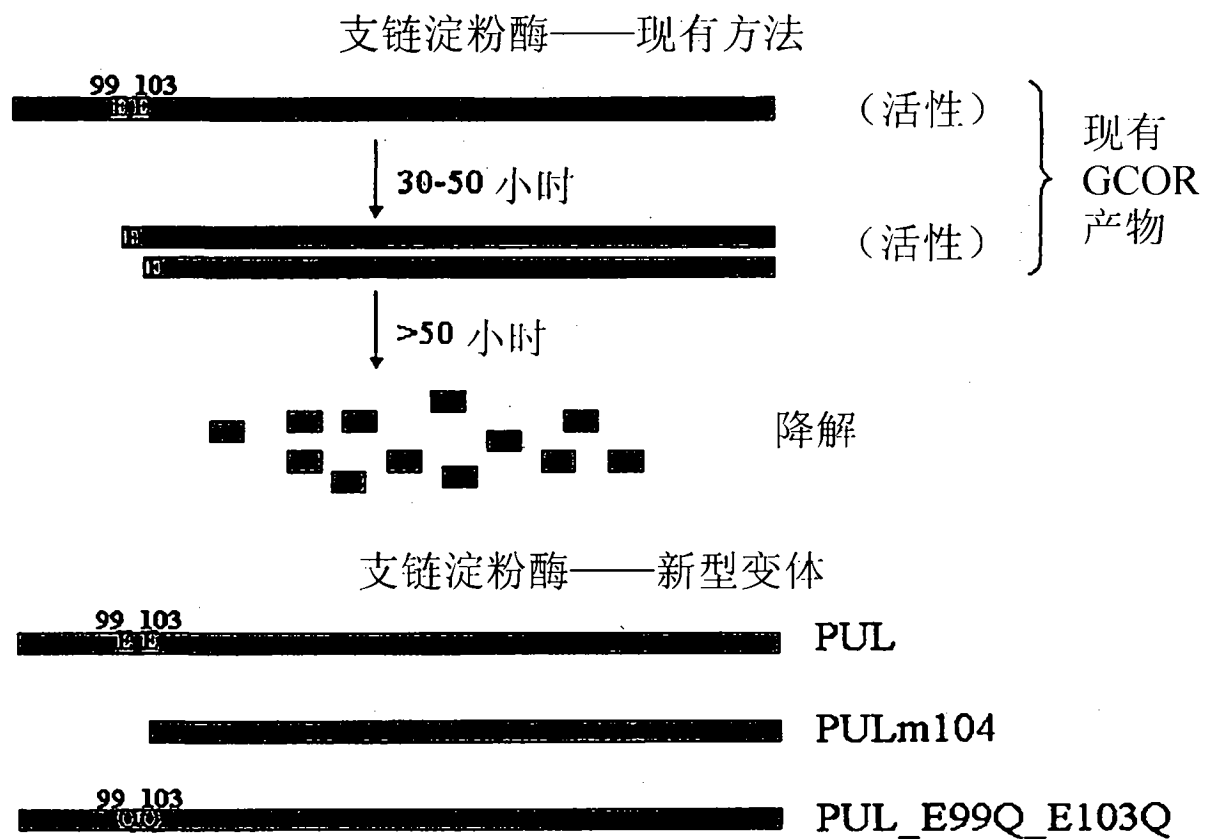


图 1

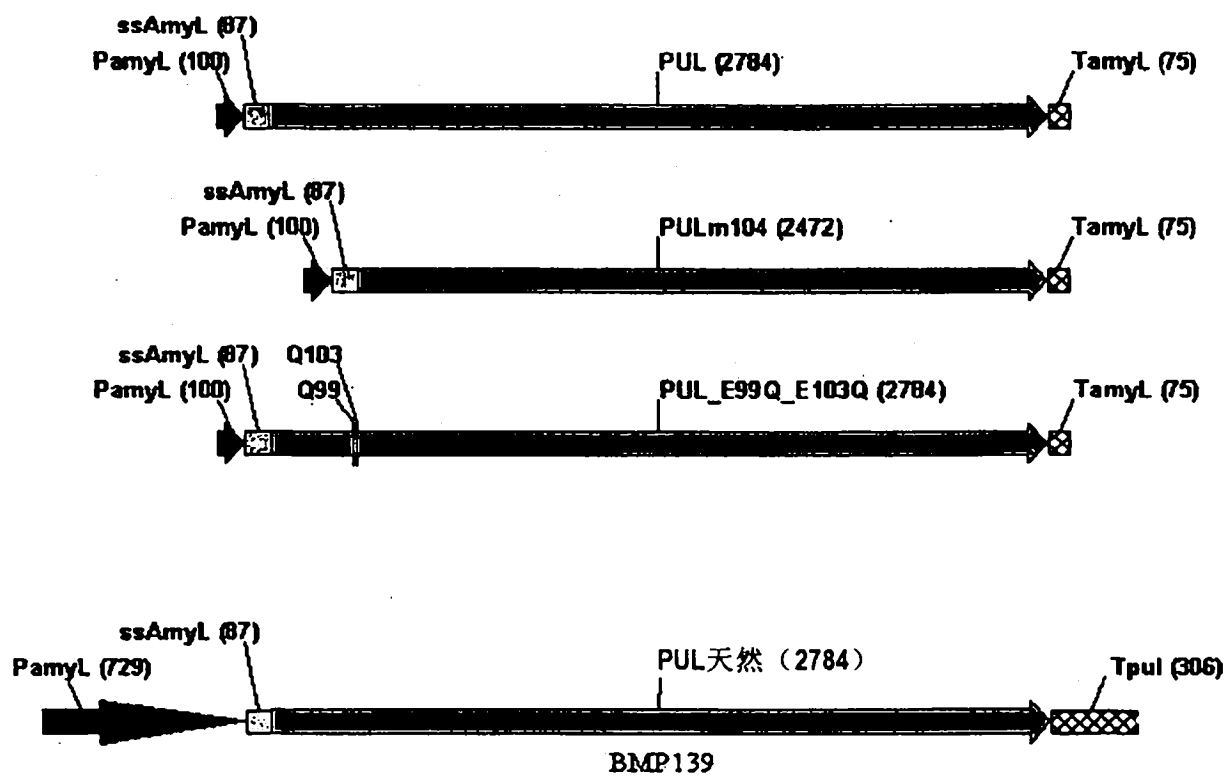


图 2

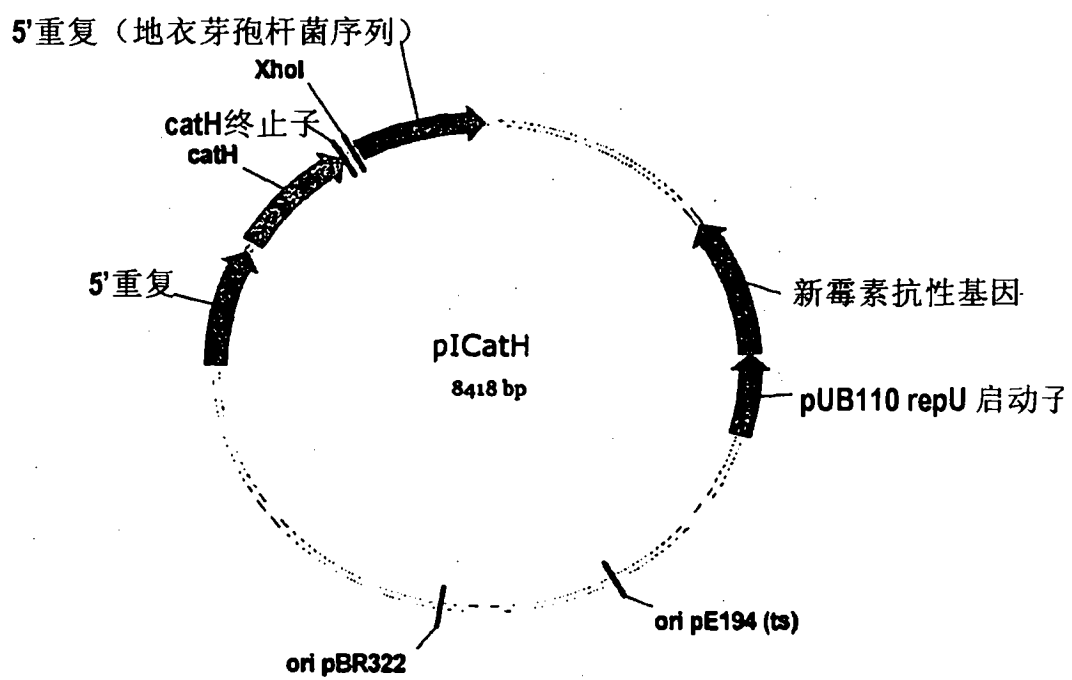


图 3

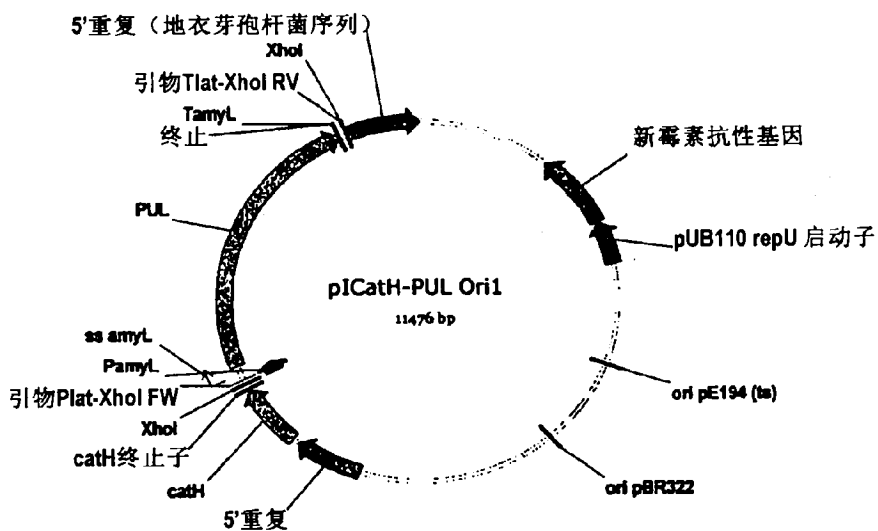


图 4

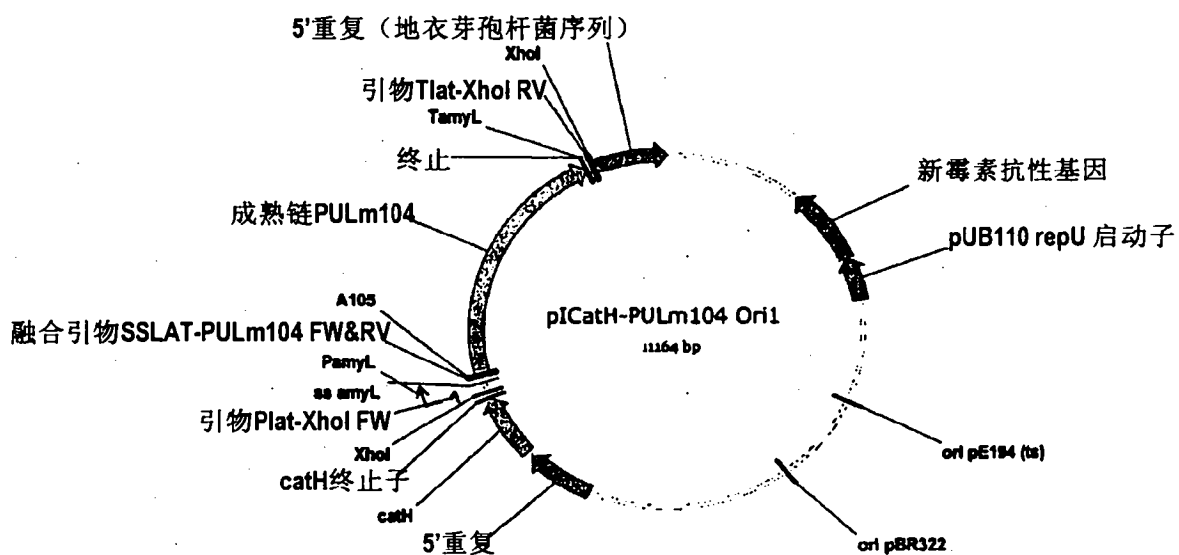


图 5

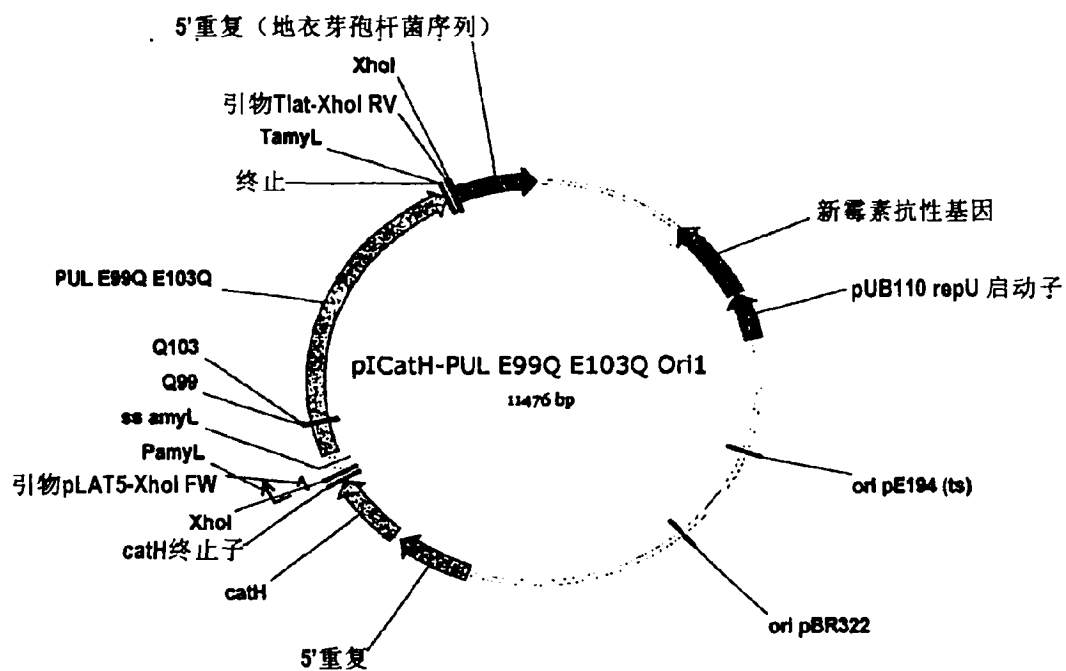


图 6

PUL野生型DNA核苷酸序列

粗体=amyL信号序列

黑体=PUL成熟编码序列

atgaaacaacaaaaacggctttacgcccgaattgctgacgctgttattttgcgctcatcttcttctgtgcc
tcattctgcagcttcagcagatggcaatacgcagcagatcatcgctccattattttagaccggcgaggag
attatcaaccgtggagcctttggatgtggccgaaagatggaggaggagcggaatatgattttaaccag
ccggcagattcacttggagcagtcgcctcagcagatattccgggaaatccgagccaagtccgcatcat
cgtcagaacacaggattggacgaaagatgtcagcgccgatcgctatatcgatctgagcaaaggcaatg
aagtctggctggctcgaaggcaacagccagatcttttatagcgaaaaagacgcccgaagatgctgctaaa
ccggcagtcagcaacgcttatctggatgccagcaaccaagtccctgggtcaaactgagccaaccgctgac
acttggagaaggagcgagcggatttacgggtccatgatgacacggcgaaacaaagatatcccggtcacga
gcgttaaagatgctagcctgggccaagatgtcacagcagttctggcgggcacggttcaacatatcttt
ggcggatcagattgggacccggataatcacagcagcgtgctgaaaaaagtccacgaacaacctgtatca
gttttagcggagatctgccggaaggcaactatcaatataaagtgcacctgaacgatagctggaacaatc
cgagctatccgagcgataacatcaatctgacagtcgccggcaggcggagcacatgtcacggttagctat
atcccgagcacacatgccgtctatgacacgatcaacaacccgaacgcccgatcttcaagtgcgaagcgg
cgtaaaaacggatctgggtcacagtcacattgggagaagatccggatgtcagccatacactgagcatcc
aaacggatggctatcaagcgaaacaagtcatcccgagaaacgctcctgaacagcagccagttatttat
agcggcgatgatctgggcaacacgtatacacaaaaagcgacgagctttaaagtttgggcgcccgaacag
cacacaagtcacagtcctgctgtatgattcagcaacaggcagcgctcacaaaaatcgctcccgatgacag
catcaggacatggagtctgggaagcgacgggtcaacaaaaacctggaaaactggattatatgtatgaa
gtcacgggccaaggatcaacaagaacagcgggtcgatccgtatgctacagcaatcgccccgaatggaac
aagaggcatgatcgctcgatctggcaaaaacagaccggcaggctggaatagcgataaacatatcacgc
cgaaaaacatcgaagatgaagtcatctatgaaatggacgtccgggatttttagcatcgatccgaacagc
ggcatgaaaaacaaaggcaaatatctggcgctgacggaaaaagggaacaaaaggcccgataacgtcaa
aacaggcatcgatagcctgaaacaactgggcatcacacatgtccaactgatgccgggtctttgctagca
atagcgtcgatgaaacggaccgcacacaagataactggggctatgacccgagaaattatgatgtcccg
gaaggccaatatgccacgaacgccaatggaaacgcccggatcaaagaatttaaagaaatggctcctgag
ccttcatagagaacatatcggcgtcaacatggacgtcgtctataaccatacgtttgccacacagatca
gcgactttgataaaatcggtgccggaatattattatcggacggatgacgcccggcaattatacgaatggc
agcggcacaggaaatgaaatcgccgcccgaagaccgatgggtccagaaatttatcatcgacagccttaa
atattgggtcaacgaatatcatatcgacggctttcgctttgatctgatggcgctgctgggcaagata
caatgagcaaaagcggcagcgaacttcatgctatcaatccgggcatcgctctttatggagaaccgtgg
acaggaggaacatcagcactgccggatgatcaactgctgacaaaaggcgcccaaaaagggaatgggagt
cgccgtctttaacgacaacctgagaaatgcctggatggcaacggtttttgatagcagcgcaccaaggat
ttgctacaggagcgacaggactgacagatgccatcaaaaatggcgtcgaaggcagcatcaacgatttt
acaagcagcccgggagaaacgatcaattatgtcacgagccatgacaactatacgtgtgtgggacaaaat
cgctctgagcaaccggaatgatagcgaagcggaccggatcaaatggatgaactggcacaagcagtcg
tcatgacatcacaaggcgtcccgtttatgcaaggcggagaagaaatgctgagaacgaaaggcggcaac
gacaacagctataatgccggcgatgccgtcaatgaatttgactggagccggaaagcacaatatccgga
cgtctttaactattattcaggacttatccatctgagactggaccatccggcggttagaatgacgacgg
cgaacgaaatcaacagccatcttcagtttctgaacagcccggaaaaatacggtcgcctatgaactgacg
gaccatgtgaacaaagacaaatggggcaacatcatcgctcgtttataacccgaacaaaacgggtcgccac
aatcaatcttccgagcggcaaatgggcaatcaatgccacaagcggcgaagtgggagaagcacactgg
gacaagcagaaggatcagtcgaagtcccggaatcagcatgatgatccttcataagaagtcagcccg
gaccacggcaaaaaa

图 7A

PUL野生型蛋白质序列

粗体=amyL信号序列

黑体=PUL成熟序列

MRQQKRLYARLLTLLFALIFLLPHSAASADGNTTTIIVHYFRPAGDYQFWSLWMWPKDGGGAEYDFNQ
PADSLGAVASADIPGNPSQVGIIVRTQDWT KDVSADRYIDLSKGNEVWLVEGNSQIFYSEKDAEDAAK
PAVSNAYLDASNQVLVKLSQPLTLGEGASGFTVHDDTANKDIPVTSVKDASLGQDVTAVLAGTFQHIF
GGSDWAPDNHSTLLKKVTNNLYQFSGDLPEGNYQYKVALNDSWNNPSYPSDNINLTVPAGGAHVTFSY
IPSTHAVYDTINPNADLQVESGVKTDLVTVTLGEDPDVSHTLSIQTDGYQAKQVIPRNVLNSSQYYY
SGDDLGNITYTQKATTFKVWAPTSTQVNVLLYDSATGSVTKIVPMTASGHGVWEATVNQNLENWYYMYE
VTGQGSTRTAVDPYATAIAPNGTRGMIVDLAKTDPAGWNSDKHITPKNIEDEVIYEMDVRDFSIDPNS
GMKNKGKYLALTEKGTGKPDNVKTGIDSLKQLGITHVQLMPVFASNSVDETDPDQDNWGYDPRNYDVP
EGQYATNANGNARIKEFKEMVLSLHREHIGVNMDVVYNHTFATQISDFDKIVPEYYYRTDDAGNYTNG
SGTGNEIAAERPVMQKFIIDSLKYWVNEYHIDGFRFDLMALLGKDTMSKAASELHAINPGIALYGEPW
TGGTSALPDDQLLTGKAQKGMGVAVFNDNLRNALDGNVFDSSAQGFATGATGLTDAIKNGVEGSINDF
TSSPGETINYVTSHDNYTLWDKIALSNPNDSEADRIKMDLAQAVVMTSQGVPFMQGGEMLRTKGGN
DNSYNAGDAVNEFDWSRKAQYPDVFNYYSGLIHLRLDHPAFRMTTANEINSHLQFLNSPENTVAYELT
DHVNKDKWGNIIVVYNPNKTVATINLPSGKWAINATSGKVGESTLGQAEGSVQVPGISMMILHQEVSP
DHGKK

图 7B

PULm104 核苷酸序列

粗体=amyL信号序列

黑体=PULm104成熟编码序列

atgaaacaacaaaaacggctttacgcccgattgctgacgctgttattttgcgctcatcttcttgcgtgcc
tcattctgcagcttcagcagctgctaaaccggcagtcagcaacgcttatctggatgccagcaaccaag
tcttgggtcaaactgagccaaccgctgacacttggagaaggagcgcgagcggatttacgggtccatgatgac
acggcgaacaaagatatcccggtcacgagcgttaaagatgctagcctgggccaagatgtcacagcagcgt
tctggcgggcacgcttcaacatatctttggcggtacagattgggcaccggataatcacagcacgctgc
tgaaaaaagtcacgaacaacctgtatcagtttagcggagatctgccggaaggcaactatcaatataaa
gtcgccctgaacgatagctggaacaatccgagctatccgagcgataacatcaatctgacagctcccggc
aggcgggagcacatgtcacgttttagctatatcccgagcacacatgccgtctatgacacgatcaacaacc
cgaacgccgatcttcaagtcgaaagcggcgtcaaaacggatctgggtcacagtcacattggggagaagat
ccggatgtcagccatacactgagcatccaaacggatggctatcaagcgaacaagtcaccccgagaaa
cgtcctgaacagcagccagctattattatagcggcgatgatctgggcaacacgtatacacaaaaagcga
cgacgtttaaagtttgggcgcgcgacaagcacacaagtcacagtcctgctgtatgattcagcaacaggc
agcgtcacaaaaatcgtcccgatgacagcatcaggacatggagtctgggaagcgacgggtcaacccaaaa
cctggaaaactgggtatttatgtatgaagtcacgggccaaggatcaacaagaacagcgggtcgatccgt
atgctacagcaatcgccccgaatggaacaagaggcatgatcgtcgatctggcaaaaacagaccgggca
ggctggaatagcgataaacatatcacgccgaaaaacatcgaagatgaagtcacatctatgaaatggacgt
ccgggatttttagcatcgatccgaacagcggcatgaaaaacaaaggcaaatatctggcgctgacggaaa
aaggaaacaaaaggcccgataaacgtcaaaacaggcatcgatagcctgaaacaactgggcatcacacat
gtccaactgatgccggtctttgctagcaatagcgtcgatgaaacggaccgcacacaagataaactgggg
ctatgacccgagaaattatgatgtcccggaaggccaatatgccacgaacgccaatggaaacgcccggga
tcaaagaatttaaagaaatgggtcctgagccttcatagagaacatatcggcgtcaacatggacgtcgtc
tataaccatacgtttgccacacagatcagcgactttgataaaatcgtgccggaatatattatcggac
ggatgacgcgggcaattatacgaatggcagcggcacaggaaatgaaatcgccgccgaaagaccgatgg
tccagaaatttatcatcgacagccttaaatattgggtcaacgaatatcatatcgacggctttcgcttt
gatctgatggcgctgctgggcaaagatacaatgagcaaagcggcgagcgaacttcagtctatcaatcc
gggcatcgctctttatggagaaccgtggacaggaggaacatcagcactgccggatgatcaactgctga
caaaaggcgcccaaaaaggaatgggagtcgccgtctttaacgacaacctgagaaatgcctggatggc
aacgtttttgatagcagcgcccaaggatttgctacaggagcgacaggactgacagatgccatcaaaaa
tggcgtcgaaggcagcatcaacgattttacaagcagccgggagagacgatcaattatgtcacgagcc
atgacaactatacgtgtgggacaaaatcgctctgagcaacccgaatgatagcgaagcggaccggatc
aaaatggatgaactggcacaagcagtcgtcatgacatcacaaggcgtcccgtttatgcaaggcggaga
agaaatgctgagaacgaaaggcggcaacgacaacagctataatgccggcgatgccgtcaatgaatttg
actggagccggaaagcacaatatccggacgtctttaactattattcaggacttatccatctgagactg
gaccatccggcggttagaatgacgacggcgaacgaaatcaacagccatcttcagtttctgaacagccc
ggaaaatacggctcgctatgaactgacggaccatgtgaacaaagacaaatggggcaacatcatcgtcg
tttataaaccgaaacaaaacggctcgccacaatcaatcttcgagcggcaaatgggcaatcaatgccaca
agcggcaaggttgagaaagcacactgggacaagcagaaggatcagtcgaagtcccggggaatcagcat
gatgatccttcacaaagtcagccccgaccacggcaaaaaa

图 8A

PULm104蛋白质序列

粗体=amyL信号序列

黑体=PULm104成熟序列

MKQQKRLYARLLTLLFALIFLLPHSAASAAAKPAVSNAYLDASNQVLVKLSQPLTLGEGASGFTVHDD
TANKDIPVTSVKDASLGQDVTAVLAGTFQHIFGGSDWAPDNHSTLLKKVTNNLYQFSGDLPEGNYQYK
VALNDSWNNPSYPSDNINLTVPAGGAHVTFSYIPSTHAVYDTINNPADLQVESGVKTDLVTVTLGED
PDVSHTLSIQTDGYQAKQVI PRNVLNSSQYYSGDDLGNTRYTQKATTFK VWAPTSTQVNVLLYDSATG
SVTKIVPMTASGHGVWEATVNQNLENWYMYEVTGQGSTRTAVDPYATAIAPNGTRGMIVDLAKTDPA
GWNSDKHITPKNIEDEVIYEMDVRDFSIDPNSGMKNKGKYLALTEKGTGKPDNVKTGIDSLKQLGITH
VQLMPVFASNSVDETDPTQDNWGYDPRNYDVPEGQYATNANGNARIKEFKEMVLSLHREHIGVNMDVV
YNHTFATQISDFDKIVPEYYYRTDDAGNYTNGSGTGNEIAAERPMVQKFIIDSLKYWVNEYHIDGFRF
DLMALLGKDTMSKAASELHAINPGIALYGEPWTGGTSALPDDQLLTKGAQKGMGVAVFNDNLRNALDG
NVFDSSAQGFATGATGLTDAIKNGVEGSINDFTSSPGETINYVTSHDNYTLWDKIALSNPNDSEADRI
KMDELAQAVVMTSQGVPFMQGGEMLRTKGGNDNSYNAGDAVNEFDWSRKAQYPDVFNYYSGLIHLRL
DHPAFRMTTANEINSHLQFLNSPENTVAYELTDHVNKDKWGNIIVVYNPNKTVATINLPSGKWAINAT
SGKVGESTLGQAEGSVQVPGISMMILHQEVSPDHGKK

图 8B

PUL E99Q_E103Q 核苷酸序列

粗体=amyL信号序列

黑体=PUL E99Q E103Q成熟编码序列

atgaaacaacaaaaacggctttacgcccgaattgctgacgctggttatttgcgctcatcttcttgcgtgcc
tcattctgcagcttcagcagatggcaatacgcgcgacgatcatcgctccattattttagaccggcgaggag
attatcaaccgtggagcctttggatgtggccgaaagatggaggaggagcggaatatgattttaaccag
ccggcagattcacttggagcagtcgcctcagcagatattccgggaaatccgagccaagtcggcatcat
cgtcagaacacaggattggacgaaagatgtcagcgccgatcgctatatcgatctgagcaaaggcaatg
aagctctggctggtcgaaggcaacagccagatcttttatagcgaaaaagacgccgaagatgctgctaaa
ccggcagtcagcaacgcttatctggatgccagcaaccaagtcctgggtcaaactgagccaaccgctgac
acttggagaaggagcgagcggatttacggtccatgatgacacggcggaacaaagatatcccggtcacga
gcgttaaagatgctagcctgggccaagatgtcacagcagttctggcgggcacgtttcaacatatcttt
ggcggatcagattgggacccggataatcacagcacgctgctgaaaaaagtcacgaacaacctgtatca
gttttagcggagatctgccggaaggcaactatcaatataaagtcgccctgaacgatagctggaacaatc
cgagctatccgagcgataacatcaatctgacagtcgccggcaggcgagcacatgtcacgtttagctat
atcccgagcacacatgccgtctatgacacgatcaacaaccggaacgccgatcttcaagtcgaaagcgg
cgtcaaaacggatctgggtcacagtcacattgggagaagatccggatgtcagccatacactgagcatcc
aaacggatggctatcaagcgaaacaagtcatcccgagaaacgctcctgaacagcagccagttattat
agcggcgatgatctgggcaacacgtatacacaacaaagcgacgacgctttaaagttggggcgccgacaag
cacacaagtcacgtcctgctgtatgattcagcaacaggcagcgtcacaaaaatcgtcccgatgacag
catcaggacatggagtctgggaagcgacgggtcaacccaaaacctggaaaactggtattatatgtatgaa
gtcacggggccaaggatcaacaagaacagcggtcgatccgtatgctacagcaatcgccccgaatggaac
aagaggcatgatcgctcgatctggcaaaaacagaccggcgaggtggaatagcgataaacatatcacgc
cgaaaaacatcgaagatgaagtcattatgaaatggacgtccgggatttttagcatcgatccgaacagc
ggcatgaaaaacaaaggcaaatatctggcgctgacggaaaaagggaacaaaaggcccgataacgtcaa
aacaggcatcgatagcctgaaacaactgggcatcacacatgtccaactgatgccggtctttgctagca
atagcgtcgatgaaacggacccgacacaagataactggggctatgacccgagaaattatgatgtcccg
gaaggccaatatgccacgaacgccaatggaaacgcccgatcaaagaatttaaagaaatggtcctgag
ccttcatagagaacatatcggcgtcaacatggacgtcgctctataaccatacgttttgccacacagatca
gcgactttgataaaatcgtgccggaatattattatcgagcggatgacgcgggcaattatacgaatggc
agcggcacaggaaatgaaatcgccgcccgaagaccgatgggtccagaaatttatcatcgacagccttaa
atattgggtcaacgaatatcatatcgacggccttctcgctttgatctgatggcgctgctgggcaaagata
caatgagcaaaagcggcgagcgaacttcattgctatcaatccgggcatcgctctttatggagaaccgtgg
acaggaggaacatcagcactgccggatgatcaactgctgacaaaaggcgcccaaaaagggaatgggagt
cgccgtctttaacgacaacctgagaaatgccctggatggcaacgtttttgatagcagcgcccaaggat
ttgctacaggagcgacaggactgacagatgccatcaaaaatggcgtcgaaggcagcatcaacgat
ttacaagcagcccggggagaaacgatcaattatgtcacagagccatgacaactatacgtgtgggacaaaat
cgctctgagcaaccggaatgatagcgaagcggaccggatcaaaatggatgaactggcacaagcagtcg
tcatgacatcacaaggcgtcccgctttatgcaaggcggagaagaaatgctgagaacgaaaggcggcaac
gacaacagctataatgccggcgatgccgtcaatgaatttgactggagccggaaagcacaatatccgga
cgtctttaactattattcaggacttatccatctgagactggaccatccggcgtttagaatgacgacgg
cgaacgaaatcaacagccatcttcagtttctgaacagcccggaataacggtcgcctatgaactgacg
gaccatgtgaacaaagacaaatggggcaacatcatcgctcgctttataaccggaacaaaacggtcgccac
aatcaatcttccgagcggcaaatgggcaatcaatgccacaagcggcaaaagttggagaaagcacactgg
gacaagcagaaggatcagccaagtcccggaatcagcatgatgatccttcataagaagtcagccccg
gaccacggcaaaaaa

图 9A

PUL E99Q_E103Q蛋白质序列

粗体=amyL信号序列

黑体=PUL E99Q E103Q成熟序列

MRQQKRLYARLLTLLFALIFLLPHSAASADGNTTTTIIVHYFRPAGDYQPWSLWMWPKDGGGA
EYDFNQPADSLGAVASADIPGNPSQVGIIVRTQDWT KDVSADRYIDLSKGNEVWLVEGNSQI
FYSQKDAQDAAKPAVSNA YLDASNQVLVKLSQPLTLGEGASGFTVHDDTANKDIPVTSVKDA
SLGQDVTAVLAGTFQHIFGGSDWAPDNHSTLLKKVTNNLYQFSGDLPEGNYQYKVALNDSWN
NPSYP SDNINLTVPAGGAHVTF SYIPSTHAVYDTINNPNADLQVESGVKTDLVTVTLGEDPD
VSHTLSIQTDGYQAKQVIPRNVLNSSQYYYSGDDLGN TYTQKATTFK VWAPTSTQVNVLLYD
SATGSVTKIVPMTASGHGVWEATVNQNLENWYMYEVTGQGSTRTAVDPYATAIAPNGTRGM
IVDLAKTDPAGWNSDKHITPKNIEDEVIYEMDVRDFSIDPNSGMKNKGKYLALTEKGTKGPD
NVKTGIDSLKQLGITHVQLMPVFASNSVDETDPTQDNWGYDPRNYDVPEGQYATNANGNARI
KEFKEMVLSLHREHIGVNMDVVYNHTFATQISDFDKIVPEYYYRTDDAGNYTNGSGTGNEIA
AERPMVQKFIIIDSLKYWVNEYHIDGFRFDLMALLGKDTMSKAASELHAINPGIALYGE PWTG
GTSALPDDQLLTKGAQKGMGVAVFNDNLRNALDGNVFDSSAQGFATGATGLTDAIKNGVEGS
INDFTSSPGETINYVTSHDNYTLWDKIALSNPNDSEADRIKMDELAQAVVMTS QGVPFMQGG
EEMLRTKGGNDNSYNAGDAVNEFDWSRKAQYPDVFNYYSGLIHLRLDHPAFRMTTANEINSH
LQFLNSPENTVAYELTDHVNKDKWGNIIVVYNPNKTVATINLPSGKWAINATSGKVGESTLG
QAEGSVQVPGISMMILHQEVS PDHGKK

图 9B