



(12) 发明专利申请

(10) 申请公布号 CN 103261400 A

(43) 申请公布日 2013. 08. 21

(21) 申请号 201180058984. 3

(51) Int. Cl.

(22) 申请日 2011. 12. 13

C12N 1/20 (2006. 01)

(30) 优先权数据

C12P 7/06 (2006. 01)

61/424, 077 2010. 12. 17 US

C12P 7/10 (2006. 01)

(85) PCT申请进入国家阶段日

2013. 06. 07

(86) PCT申请的申请数据

PCT/US2011/064580 2011. 12. 13

(87) PCT申请的公布数据

W02012/082711 EN 2012. 06. 21

(83) 生物保藏信息

PTA-7858 2006. 09. 12

(71) 申请人 纳幕尔杜邦公司

地址 美国特拉华州威尔明顿

(72) 发明人 P. G. 凯米 W. D. 希茨

(74) 专利代理机构 中国专利代理(香港)有限公

司 72001

代理人 李慧惠 李炳爱

权利要求书1页 说明书24页

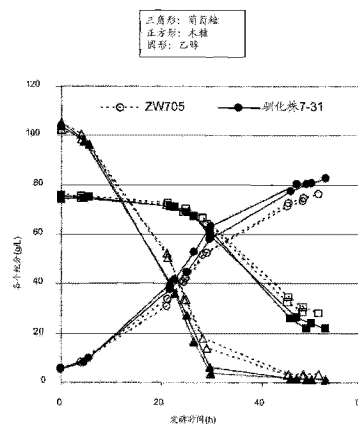
序列表35页 附图4页

(54) 发明名称

在生物质水解产物培养基中具有改善的乙醇生产的利用木糖的运动发酵单胞菌

(57) 摘要

使用驯化方法分离了在包含生物质水解产物的培养基中具有改善性能的利用木糖的、产生乙醇的运动发酵单胞菌(*Zymomonas mobilis*)菌株。据发现,独立分离的菌株在相同的编码区中具有独立的突变。在这一编码中的突变可经工程改造以产生改善的表型。



1. 重组的、利用木糖的、产生乙醇的发酵单胞菌(*Zymomonas*)属微生物,所述微生物在 zmo1432 开放阅读框中具有至少一种基因修饰。

2. 根据权利要求 1 所述的重组发酵单胞菌,其中所述具有至少一种基因修饰的 zmo1432 开放阅读框编码多肽,所述多肽具有的氨基酸序列与如 SEQ ID NO:2 中所示的氨基酸序列具有至少约 95% 的同一性,其中与缺乏所述基因修饰的发酵单胞菌在相同条件下生产的产量相比,所述突变蛋白的存在提高发酵期间所述发酵单胞菌在生物质水解产物培养基中生产的乙醇产量。

3. 根据权利要求 2 所述的重组发酵单胞菌,其中所述基因修饰是在所述 zmo1432 编码区中的突变,所述突变引起选自下列的氨基酸取代:1)在 SEQ ID NO:2 中,在位置 366 处的精氨酸对苏氨酸的取代;和 2)在 SEQ ID NO:2 中,在位置 117 处的苯丙氨酸对丝氨酸的取代。

4. 根据权利要求 1 所述的重组发酵单胞菌,与缺乏所述 zmo1432 开放阅读框中的至少一种基因修饰的发酵单胞菌相比,所述重组发酵单胞菌对生物质水解产物具有更高的耐受性。

5. 多核苷酸,所述多核苷酸编码具有选自 SEQ ID NO:3 和 SEQ ID NO:4 的氨基酸序列的蛋白。

6. 多肽,所述多肽具有选自 SEQ ID NO:3 和 SEQ ID NO:4 中的氨基酸序列。

7. 制备重组发酵单胞菌产乙醇生物的方法,包括:

a) 提供利用木糖的、产生乙醇的发酵单胞菌属微生物,所述微生物包含 zmo1432 开放阅读框,所述 zmo1432 开放阅读框编码具有如 SEQ ID NO:2 中所示的氨基酸序列的多肽;以及

b) 在 a) 的 zmo1432 开放阅读框中引入突变,使得所述突变的开放阅读框的表达表达了具有选自 SEQ ID NO:3 和 SEQ ID NO:4 的氨基酸序列的多肽。

8. 生产乙醇的方法,包括:

a) 提供权利要求 1 的重组发酵单胞菌;

b) 提供包含木糖的生物质水解产物培养基;以及

c) 使 a) 的发酵单胞菌在 b) 的生物质水解产物培养基中生长,其中产生乙醇。

在生物质水解产物培养基中具有改善的乙醇生产的利用木糖的运动发酵单胞菌

[0001] 政府权利声明

[0002] 本发明由美国政府支持,以资助号 DE-C36-07G017056 受到能源部资助。美国政府拥有本发明的必然权利。

技术领域

[0003] 本发明涉及微生物学和发酵领域。更具体地,本发明分离并表征了在生物质水解产物培养基中具有改善的生长和乙醇生产的突变的发酵单胞菌(*Zymomonas*)菌株。

背景技术

[0004] 通过微生物生产乙醇提供了一种替代化石燃料的可供选择的能源,因此成为当前研究的重要领域。希望产生乙醇以及其它有用产物的微生物能够在不影响人类食品供应的培养基中生长并产生乙醇,诸如避免使用通过玉米谷粒产生的糖类。由于纤维素生物质加工中的进展,可将高浓度的葡萄糖、木糖和其它糖类释放于发酵中使用的生物质水解产物中。这样,生物质向乙醇的转化为通过使用可再生非食品资源提供化石燃料替代品来改善环境影响提供了很大的可能性。

[0005] 已经对在天然情况下并不利用木糖的运动发酵单胞菌(*Zymomonas mobilis*)和其它细菌产乙醇生物进行针对木糖利用的基因工程改造,从而通过使用生物质水解产物中的更多糖类来改善生长和乙醇生产。但是,在包含生物质水解产物的培养基中的生长和乙醇生产通常并不理想,这是由于对微生物具有抑制性的乙酸酯和其它化合物的存在所导致的。在共同拥有并且共同未决的美国专利公布 US20110014670A1 中公开了制备对培养基中的乙酸酯和乙醇更具耐受性的改善的利用木糖的发酵单胞菌菌株的方法以及通过所述方法而分离的菌株。

[0006] 在经预处理的生物质的水解产物中有可能存在的单化合物的毒性作用描述于 Delegenes 等人((1996) *Enzymes and Microbial Technology* 19:220-224) 中。木糖发酵型运动发酵单胞菌对条件化稀酸黄杨半纤维素水解产物的驯化描述于 Lawford 等人((1999), *Applied Biochemistry and Biotechnology* 77:191-204) 中。

[0007] 仍然需要在生物质水解产物培养基中进行发酵期间具有改善的乙醇生产的分离的利用木糖的发酵单胞菌产乙醇生物菌株以及产生改善菌株的基因工程方法。

发明内容

[0008] 本发明提供了在生物质水解产物培养基中具有改善的生长和乙醇生产的重组的利用木糖的发酵单胞菌菌株。此外,本发明提供了制备在水解产物培养基中使用的改善的发酵单胞菌菌株的方法以及使用所述菌株制备乙醇的方法。

[0009] 因此,本发明提供了重组的、利用木糖的、产生乙醇的发酵单胞菌属微生物,所述微生物在 zmo1432 开放阅读框中具有至少一种基因修饰。

[0010] 在一个方面,本发明提供了多肽和编码所述多肽的多核苷酸,所述多肽具有选自 SEQ ID NO:3 和 SEQ ID NO:4 的氨基酸序列。

[0011] 在另一方面,本发明提供了制备重组发酵单胞菌产乙醇生物的方法,其包括:

[0012] a) 提供利用木糖的、产生乙醇的发酵单胞菌属微生物;

[0013] b) 提供多核苷酸,所述多核苷酸编码具有选自 SEQ ID NO:3 和 SEQ ID NO:4 的氨基酸序列的蛋白;以及

[0014] c) 将 b) 的多核苷酸引入 a) 的微生物中;其中内源性 zmo1432 编码区受到破坏。

[0015] 在替代性的方面,本发明提供了制备重组发酵单胞菌产乙醇生物的方法,其包括:

[0016] a) 提供利用木糖的、产生乙醇的发酵单胞菌属微生物,所述微生物包含 zmo1432 开放阅读框,所述 zmo1432 开放阅读框编码具有如 SEQ ID NO:2 中所示的氨基酸序列的多肽,以及

[0017] b) 在 a) 的 zmo1432 开放阅读框中引入突变,使得所述突变的开放阅读框的表达表达了具有选自 SEQ ID NO:3 和 SEQ ID NO:4 的氨基酸序列的多肽。

[0018] 在另一方面,本发明提供了生产乙醇的方法,其包括:

[0019] a) 提供本发明的重组发酵单胞菌;

[0020] b) 提供包含木糖的生物质水解产物培养基;以及

[0021] c) 使 a) 的发酵单胞菌在 b) 的生物质水解产物培养基中生长,其中产生乙醇。

[0022] 附图说明、生物保藏和序列描述

[0023] 申请人已经按照国际承认的用于专利程序目的的微生物保藏的布达佩斯条约的有关条款进行了以下生物保藏:

[0024] 保藏菌株信息

[0025]

保藏人鉴定参考	国际保藏所命名	保藏日期
发酵单胞菌(<i>Zymomonas</i>) ZW658	ATCC No PTA-7858	2006 年 9 月 12 日

[0026] 图 1 示出了玉米棒水解产物随时间发酵的示图,其在发酵单胞菌菌株和驯化株 7-31 之间比较了葡萄糖和木糖利用以及乙醇生产。

[0027] 图 2 示出了玉米棒水解产物随时间发酵的示图,其在发酵单胞菌菌株 ZW705、驯化株 7-31 和驯化株 5-6 之间比较了葡萄糖和木糖利用以及乙醇生产。

[0028] 图 3 示出了由 zmo1432 编码的蛋白和新洋葱伯克霍尔德菌(*Burkholderia cenocepacia*)的 Bcemn03_1426 蛋白之间的比对。

[0029] 图 4 示出了由 zmo1432 编码的蛋白和大肠杆菌(*E. coli*)的 AaeB 蛋白之间的比对。

[0030] 下列序列符合 37C.F.R. § 1.821-1.825(“对含有核酸序列和/或氨基酸序列公开的专利申请的要求—序列规则”),并且与世界知识产权组织(WIPO)标准 ST. 25 (2009)和 EPO 和 PCT 的序列列表要求(规则 5.2 和 49.5(a-bis),以及行政指导的 208 节和附录 C)相一致。用于核苷酸和氨基酸序列数据的符号和格式遵循如 37C.F.R. § 1.822 所示的规定。

[0031] SEQ ID NO:1 是在公开的运动发酵单胞菌基因组序列(NCBI 参考号: NC_006526.2)中命名为 zmo1432 的编码区的核苷酸序列。

[0032] SEQ ID NO:2 是在公开的运动发酵单胞菌基因组序列(NCBI 参考号: NC_006526.2) 中命名为 zmo1432 的编码区编码的氨基酸序列。

[0033] SEQ ID NO:3 是 SEQ ID NO:2 的氨基酸序列,但在位置 366 处的氨基酸变为精氨酸。

[0034] SEQ ID NO:4 是 SEQ ID NO:2 的氨基酸序列,但在位置 117 处的氨基酸变为苯丙氨酸。

[0035] SEQ ID NO:5 是由 fusC (登录号:P24128)编码的新洋葱伯克霍尔德菌蛋白的氨基酸序列。

[0036] SEQ ID NO:6 是产酸克雷伯菌(*Klebsiella oxytoca*)的镰刀菌酸解毒蛋白的氨基酸序列(登录号:Q48403)。

[0037] SEQ ID NO:7 是由 Bcenmc03_1426 编码的新洋葱伯克霍尔德菌蛋白的氨基酸序列。SEQ ID NO:8 是大肠杆菌 AaeB 蛋白的氨基酸序列。

[0038] SEQ ID NO:9 是未成熟的 Xyn3 的氨基酸序列,其并入了对应于位置 1 至 16 的预测信号序列。

[0039] SEQ ID NO:10 是未成熟的 Fv3A 的氨基酸序列,其并入了对应于位置 1 至 23 的预测信号序列。

[0040] SEQ ID NO:11 是未成熟的 Fv43D 的氨基酸序列,其并入了对应于位置 1 至 20 的预测信号序列。

[0041] SEQ ID NO:12 是未成熟的 Fv51A 的氨基酸序列,其并入了对应于位置 1 至 19 的预测信号序列。

具体实施方式

[0042] 本发明描述了对利用木糖的发酵单胞菌细胞在生物质水解产物培养基中的驯化,用以改善生长和乙醇生产。在驯化株中对突变的表征鉴定出了所述改善所特有的突变,其可在未驯化的利用木糖的发酵单胞菌菌株进行工程改造而得。这些菌株用于更为有效地制造乙醇,从而产生乙醇以作为化石燃料替代品。

[0043] 下述的缩写和定义将用于解释说明书和权利要求。

[0044] 如本文所用,术语“包含”、“包括”、“具有”或“含有”,或者其任何其他变型旨在包括非排他的包括。例如,包含一系列元素的组合物、混合物、工艺、方法、制品或设备不必仅限于那些元素,而可以包括其他未明确列出的元素,或此类组合物、混合物、工艺、方法、制品或设备固有的元素。此外,除非有相反的明确说明,“或”是指包含性的“或”,而不是指排他性的“或”。例如,以下任何一个均表示满足条件 A 或 B:A 为真实(或存在)的且 B 为虚假(或不存在)的,A 为虚假(或不存在)的且 B 为真实(或存在)的,以及 A 和 B 都是真实(或存在)的。

[0045] 同样,涉及元素或组分例证(即出现)次数的位于本发明元素或组分前的不定冠词“一个”或“一种”旨在是非限制性的。因此,应将“一个”或“一种”理解为包括一个或至少一个,并且元素或组分的词语单数形式也包括复数形式,除非有数字明显表示单数。

[0046] 如本文所用,“一个或多个利用木糖的发酵单胞菌细胞”是指经遗传工程改造以表达产生使用木糖作为用于发酵的碳水化合物来源的能力的酶的一个或多个菌株细胞。

[0047] 术语“驯化株”是指经选择在特定的条件下生长来改善其在这些条件下生长并产生产物的能力的微生物。

[0048] 如本文所用,“对应的未驯化菌株”是指原始的利用木糖的发酵单胞菌菌株,其是使用本文所公开的生物质水解产物驯化方法由其制备改善的菌株的菌株。

[0049] 如本文所用,“饲养生长培养基”是指添加到连续培养器皿中的培养基。

[0050] 术语“木质纤维素生物质”或“生物质”是指任何木质纤维素材料并且包括含有纤维素、半纤维素、木质素、淀粉、寡糖和 / 或单糖的材料。生物质还可包含附加组分,如蛋白质和 / 或脂类。生物质可来源于单一来源,或者生物质可包括来源于一种以上来源的混合物;例如,生物质可包括玉米棒和玉米秸秆的混合物,或草和叶片的混合物。木质纤维素生物质包括但不限于生物能源作物、农业残留物、市政固体垃圾、工业固体垃圾、来自造纸业的淤渣、庭院垃圾、木材和林业垃圾。生物质的例子包括但不限于玉米棒、作物残留物如玉米壳、玉米秸秆、小草、小麦秸秆、大麦秸秆、干草、稻秆、柳枝稷、废纸、甘蔗渣、高粱植物材料、大豆植物材料、从谷物的研磨中获得的组分、树、枝、根、叶、木片、锯末、灌木和灌丛、蔬菜、水果和花。

[0051] 如本文所用,“生物质水解产物”和“纤维素水解产物”是指通常通过预处理和糖化过程由生物质产生的产物,所述生物质是纤维素材料。可发酵的糖存在于水解产物中,也存在其它产物。

[0052] 如本文所用,“生物质水解产物培养基”是指包含至少约 50% 的通过纤维素和 / 或木质纤维素生物质制备的水解产物的培养基。水解产物通过生物质的糖化来制备,通常先经过预处理。除所述水解产物外,所述培养基还包括用于生物催化剂生长和生产的确定组分

[0053] “Xyn3”是来自里氏木霉(*Trichoderma reesei*)木聚糖酶 GH10 家族。当所述酶作用于预处理的生物质或分离的半纤维素时,根据其在木乙糖酶存在下催化增加的木糖单体产量的能力的观察结果所间接显示,Xyn3 (SEQ ID NO:9) 具有木聚糖内切酶活性。

[0054] “Fv3A”是来自轮枝镰孢菌(*Fusarium verticillioides*)GH3 家族酶。Fv3A (SEQ ID NO:10)在,例如,使用 p-硝基苯基- β -木吡喃糖苷、木乙糖、混合的线型木聚糖低聚物、来自半纤维素的具支链的阿糖基木聚糖低聚物或者经稀氨水预处理的玉米棒作为底物的酶分析法中据显示具有 β -木糖苷酶活性。

[0055] “Fv43D”是来自轮枝镰孢菌 GH43 家族酶。Fv43D (SEQ ID NO:11)在,例如,使用 p-硝基苯基- β -木吡喃糖苷、木乙糖和 / 或混合的线型木聚糖低聚物作为底物的酶分析法中据显示具有 β -木糖苷酶活性。

[0056] “Fv51A”是来自轮枝镰孢菌 GH51 家族酶。Fv51A (SEQ ID NO:12)在,例如,使用 4-硝基苯基- α -L-阿拉伯呋喃糖苷作为底物的酶分析法中据显示具有 L- α -阿拉伯呋喃糖酶活性。

[0057] “基因”指表达特定蛋白质或功能性 RNA 分子的核酸片段,其可任选地包括位于编码序列之前的调控序列(5' 非编码序列)和之后的调控序列(3' 非编码序列)。“天然基因”或“野生型基因”指具有其自身调控序列的天然存在的基因。“嵌合基因”指不是天然基因的任何基因,包含在天然情况下不是一起存在的调控序列和编码序列。因此,嵌合基因可包含源自不同来源的调控序列和编码序列,或者源自相同来源、但排列方式与天然存在的排

列方式不同的调控序列和编码序列。“内源性基因”指位于生物的基因组内它的天然位置的天然基因。“外来基因”指正常情况下不存在于宿主生物中的基因,它通过基因转移导入宿主生物。外来基因可以包含插入到非天然生物体内的天然基因或嵌合基因。

[0058] 术语“基因构建体”是指编码表达一种或多种特定蛋白质或功能性 RNA 分子的核酸片段。在基因构建体中基因可以是天然的、嵌合的、或外来的。通常基因构建体应包含“编码序列”。“编码序列”指编码特定氨基酸序列的 DNA 序列。

[0059] “启动子”或“启动控制区”指能够控制编码序列或功能 RNA 的表达的 DNA 序列。一般来讲,编码序列位于启动子序列的 3' 端。启动子可整个源于天然基因,或者由源于天然存在的不同启动子的不同元件构成,或者甚至包含合成的 DNA 片段。本领域内的技术人员应当理解,不同的启动子可在不同的组织或细胞类型中,或者在不同的发育阶段,或者响应不同的环境条件而引导基因的表达。通常将在大多数细胞类型中、在大多数情况下引起基因表达的启动子称为“组成型启动子”。

[0060] 如本文所用,术语“表达”指衍生自基因的编码(mRNA)或功能性 RNA 的转录和稳定积聚。表达还可以指 mRNA 向多肽的翻译。“过表达”是指基因产物在转基因生物体内的生产,其超过了在正常或非转化生物体内的生产。

[0061] 如本文所用,术语“转化”指将核酸片段转移至宿主生物体内,导致在基因上稳定遗传。转化的核酸可以是宿主细胞中保留的质粒形式,或者一些转化的核酸可以被整合进宿主细胞基因组中。含有转化核酸片段的宿主生物被称为“转基因”或“重组”或“转化”生物体。

[0062] 如本文所用,术语“质粒”和“载体”是指通常携带有不属于细胞中心代谢的部分的基因的染色体外元件,并且常常是环状双链 DNA 分子的形式。这类元件可以是源自任何来源的自主复制序列、基因组整合序列、噬菌体或单链或双链 DNA 或 RNA 的核苷酸序列(线性或环状),其中多个核苷酸序列已连接或重组进入一种独特构建体中,该独特构建体能够将所选基因产物的启动子片段和 DNA 序列与相应的 3' 末端非翻译序列一起引入细胞中。

[0063] 术语“可操作地连接”指单个核酸片段上的核酸序列的关联,使得其中一个核酸序列的功能受到另一个核酸序列的影响。例如,当启动子能够影响编码序列的表达(即,该编码序列受到该启动子的转录控制)时,则该启动子与该编码序列可操作地连接。编码序列可以按有义或反义的取向可操作地连接至调控序列。

[0064] 术语“选择性标记”指一种鉴定因子,通常是抗生素或化学药品抗性基因,该因子能基于标记基因的效应,即,对抗生素的抗性进行选择,其中所述效应用于追踪遗传的受关注核酸和 / 或用于鉴定遗传了受关注核酸的细胞或生物。

[0065] 如本领域已知的,术语“% 同一性”是两条或更多条多肽序列之间或两条或更多条多核苷酸序列之间的关系,该关系通过对序列进行比较确定。在本领域中,“同一性”还表示多肽或多核苷酸序列之间序列关联的程度,根据具体情况,它由这些序列的序列串之间的匹配程度确定。“同一性”和“相似性”可容易地通过已知方法计算出来,所述的方法包括但不限于以下文献中所描述的那些:1.) Computational Molecular Biology (Lesk, A. M. 编辑) Oxford University :NY (1988);2.) Biocomputing: Informatics and Genome Projects (Smith, D. W. 编辑) Academic :NY (1993);3.) Computer Analysis of Sequence Data, 部分 I (Griffin, A. M. 和 Griffin, H. G. 编辑) Humana :NJ (1994);4.) Sequence Analysis in

Molecular Biology (von Heinje, G. 编辑) Academic (1987); 和 5.) Sequence Analysis Primer (Gribskov, M. 和 Devereux, J. 编辑) Stockton :NY (1991)。

[0066] 设定确定同一性的优选方法来用于给出待测试序列之间的最佳匹配。确定同一性和相似性的方法在可公开获得的计算机程序中编成了代码。序列比对和 % 同一性计算可以用 LASERGENE 生物信息学计算软件包 (LASERGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI)) 中的 MegAlign™ 程序进行。序列的多重比对使用 “Clustal 比对方法” 进行, 该方法涵盖若干个不同的算法, 包括对应于称为 Clustal V 比对方法的 “Clustal V 比对方法”, 该方法描述于 Higgins 和 Sharp, (CABIOS. 5:151-153 (1989); Higgins, D. G. 等人, Comput. Appl. Biosci., 8:189-191 (1992)) 中, 并且可以在 LASERGENE 生物信息学计算软件包 (DNASTAR Inc.) 中的 MegAlign™ 程序中找到的比对方法。对于多重比对, 默认值对应于 GAP PENALTY=10、以及 GAP LENGTH PENALTY=10。用 Clustal 方法进行成对比对和蛋白质序列的百分比同一性计算的默认参数为 KTUPLE=1、GAP PENALTY=3、WINDOW=5、以及 DIAGONALS SAVED=5。对于核酸, 这些参数为 KTUPLE=2、GAP PENALTY=5、WINDOW=4、以及 DIAGONALS SAVED=4。在用 Clustal V 程序进行序列比对后, 有可能通过观察相同程序中的 “序列距离” 表来获得 “% 同一性”。此外, 也可以使用 “Clustal W 比对方法”, 该方法相当于称为 Clustal W (描述于 Higgins 和 Sharp, CABIOS ;5:151-153 (1989); Higgins, D. G. 等人, Comput. Appl. Biosci. 8:189-191 (1992) 中), 并且可以在 LASERGENE 生物信息学计算软件包 (DNASTAR Inc.) 中的 MegAlign™ v6.1 程序中找到的比对方法。用于多重比对的默认参数 (GAP PENALTY=10、GAP LENGTH PENALTY=0.2、Delay Divergen Seqs(%)=30、DNA Transition Weight=0.5、Protein Weight Matrix=Gonnet 系列、以及 DNA Weight Matrix=IUB)。在使用 Clustal W 程序对序列进行比对之后, 可通过查看同一程序中的 “序列距离” 表来获得 “% 同一性”。

[0067] 制备在水解产物培养基中具有改善的发酵的发酵单胞菌菌株

[0068] 本发明提供了重组的、利用木糖的、产生乙醇的发酵单胞菌微生物, 所述微生物在 zmo1432 开放阅读框中具有至少一种基因修饰。这一突变的效果在于表达了改善该菌株在水解产物培养基中的性能的多肽, 由此提高了该菌株对水解产物中的多种生长抑制物的耐受并且提高了乙醇产量。发酵活性的改善已被关联于发酵单胞菌基因组 (NCBI 参考号: NC_006526.2) 的 zmo1432 区中的突变, 其在本文定义为 SEQ ID NO:1, 该序列编码了 SEQ ID NO:2 的多肽。

[0069] 因此在此提出, 可使发酵单胞菌菌株具有在水解产物培养基中的改善的发酵, 所述菌株通过在开放阅读框 (ORF) 内引入至少一种基因修饰而能够利用木糖作为碳源并且产生乙醇, 所述开放阅读框编码了在修饰前与 SEQ ID NO:2 具有至少约 95% 的氨基酸同一性的蛋白质的基因。所述蛋白质可与 SEQ ID NO:2 具有 95%、96%、96%、98%、99% 或 100% 的同一性。可制成所述蛋白质中的任何基因修饰, 所述基因修饰在水解产物培养基存在的情况下提高了由携带所述突变蛋白的菌株进行的乙醇生产。乙醇生产中的提高通过在相同的发酵条件下与缺乏所述基因修饰的发酵单胞菌菌株的生产进行比较而测定。通过本文实例 3 中描述的方法, 本领域的技术人员可容易地对在所述 ORF 中具有基因修饰的菌株进行分析以估测在生物质水解产物存在情况下的提高的乙醇生产。所述的改善菌株对生物质水解产物以及生物质水解产物中存在的抑制物具有更高的耐受性, 其中的耐受性是指与具有较少

或不含水解产物的培养基中相比,所述菌株在具有标明水平的水解产物(耐受度水平)的培养基中进行类似地生长并且产生乙醇的能力。通过与缺乏所述基因修饰的发酵单胞菌菌株进行比较而测定更高的耐受性。

[0070] 在一个实施例中,所述基因修饰导致了在 SEQ ID NO:2 的氨基酸序列的位置 366 处的变异,其以精氨酸取代苏氨酸。在另一个实施例中,所述基因修饰导致了在 SEQ ID NO:2 的氨基酸序列的位置 117 处的变异,其以苯丙氨酸取代丝氨酸。可在所述核苷酸序列中制出使密码子 366 变成编码精氨酸或使密码子 117 变成编码苯丙氨酸的任何变化。编码精氨酸的密码子是 CGT、CGC、CGA、CGG、AGA 和 AGG。编码苯丙氨酸的密码子是 TTT 和 TTC。

[0071] 在一个实施例中,所述基因修饰处于与 ORF zmo1432 具有至少 95% 的序列同一性的编码区内,所述 ORF zmo1432 得名于公开的运动发酵单胞菌基因组序列(NCBI 参考号: NC_006526.2),其为具有 SEQ ID NO:1 的核苷酸 1446603 至 1448633 的互补序列。可对 SEQ ID NO:1 的 ORF 称呼代用名,但是在任何情况下其可通过与 SEQ ID NO:1 的序列进行比较而鉴定为 zmo1432。在发酵单胞菌菌种或菌株中鉴定为 zmo1432 的编码区序列中可能存在某些变异。因此,鉴定为 zmo1432 的编码区可与 SEQ ID NO:1 具有至少约 95%、96%、97%、98%、99% 或 100% 的序列同一性,并且具有基因修饰的编码区在基因修饰前可与 SEQ ID NO:1 具有至少约 95%、96%、97%、98%、99% 或 100% 的序列同一性。在修饰前,这一编码区是基因修饰的目标。

[0072] 可通过本领域的技术人员已知的任何方法在发酵单胞菌菌株中产生至少一种所描述的基因修饰,所述菌株能够利用木糖作为碳源并且产生乙醇。这类方法包括通过下文以及在本文实例 1 和 2 中所述的驯化,通过化学诱变和筛选,以及通过基因工程。

[0073] 用于在靶编码区中引入基因修饰的基因工程可通过包括使用双交换同源重组来将内源性靶编码区替换为携带上述突变的相同编码区在内的方法来进行。在同源重组中,侧连于靶整合位点的 DNA 序列结合奇放线菌素抗性基因或其它选择性标记而安置,并且对突变序列的替换导致了所述选择性标记的插入以及突变序列替换进入所述靶基因组位点。所述选择性标记处于编码区外从而使编码区在产物内表达。此外,选择性标记可结合位点特异性重组结合位点,以便在对应的位点特异性重组酶表达后将所述抗性基因从基因组中切除。使用 EPICENTRE® 的 EZ::Tn 体外转座系统进行的转座特别适用于整合替换的突变序列,其在美国专利公布 US-2009-0246846-A1 的实例 1 和 6 中使用。

[0074] 作为另外一种选择,可通过诸如插入、突变或删除这样的操作在细胞中破坏靶内源性编码区的表达,并且可如上所述将表达携带突变的编码区的基因引入该细胞。在发酵单胞菌中用于使编码区失活的双交换同源重组的一个例子描述于美国专利 7,741,119 中。所引入的基因可以是包括其天然启动子的内源性基因(具有突变编码区),或者所引入的基因可以是嵌合基因,其包含可操作地连接的启动子和与所破坏的编码区具有至少约 95%、96%、97%、98% 或 99% 的同一性的编码区。嵌合基因中通常包括 3' 终止区。可用于发酵单胞菌的启动子包括 ZmPgap 和发酵单胞菌烯醇酶基因的启动子。

[0075] 可将所引入的基因保持在质粒中,或者使用例如同源重组、定点整合或随机整合的方法整合进基因组中。用于引入的基因通常构建在载体中或转移入载体中以用于进一步的操作。载体是本领域熟知的。特别可用于在发酵单胞菌中表达的是能在大肠杆菌和发酵单胞菌中复制的载体,例如描述于美国专利 5,514,583 中的 pZB188。载体可包括在细胞

中自主复制的质粒和携带载体整合进细菌基因组的质粒。用于 DNA 整合的质粒可包括转座子、与靶细菌基因组同源的核酸序列区或其它支持整合的序列。其它类型的载体还可是转座体,其使用,例如,可通过 **EPICENTRE[®]** 市售获得的系统而产生。如何为所需靶宿主和所需功能选择适当的载体是众所周知的。

[0076] 宿主菌株

[0077] 能够利用木糖作为碳源并且产生乙醇的任何发酵单胞菌菌株均可可是本发明的起始菌株。将这类菌株用于在水解产物培养基中的驯化,用于化学诱变和筛选,或者进行基因工程以产生本文所述的改善菌株。可进行工程改造以表达木糖至乙醇的发酵通路的发酵单胞菌菌株(诸如运动发酵单胞菌(*Z. mobilis*))是特别有用的。内源性基因可提供部分代谢途径,或者可以通过任何已知的基因操纵技术进行改变以提供具有酶活性的蛋白用于木糖代谢。例如,内源性转酮醇酶可在建立木糖代谢途径时补充其它导入的酶活性。通常为了木糖代谢中所涉及的四种酶的表达,可如美国专利 5,514,583 所述(其以引用方式并入本文)中所述将四个基因引入发酵单胞菌菌株如运动发酵单胞菌中。这些基因包括编码木糖异构酶的基因,该酶催化木糖转化成木酮糖,以及木酮糖激酶,该酶磷酸化木酮糖以形成 5-磷酸木酮糖。此外,转酮醇酶和转醛醇酶,戊糖磷酸化途径中的两种酶,将 5-磷酸木酮糖转化成连接戊糖代谢与 Entner-Doudoroff 糖酵解途径的中间体,该途径使木糖代谢成乙醇。编码这些酶的 DNA 序列可从能够代谢木糖的多种微生物中的任何一种获得,例如肠细菌和一些酵母以及真菌。编码区的来源包括黄单胞菌、克雷伯氏菌、埃希氏菌、红细菌、黄杆菌、醋杆菌、葡糖杆菌、根瘤菌、农杆菌、沙门氏菌、假单胞菌和发酵单胞菌。特别有用的是大肠杆菌的编码区。

[0078] 将编码 DNA 序列可操作地连接至运动发酵单胞菌细胞中表达的启动子如运动发酵单胞菌甘油醛-3-磷酸脱氢酶启动子(GAP 启动子)和运动发酵单胞菌烯醇酶启动子(ENO 启动子)。编码区可从启动子开始单个表达,或者两个或更多个编码区可接合到一个操纵子中,从相同启动子开始一起表达。可将所得的嵌合基因导入发酵单胞菌并保持在质粒中,或者使用例如同源重组、定点整合、或随机整合的方法整合进基因组中。特别有用的利用木糖的菌株包括 ZM4 (pZB5) (描述于美国专利 U. S. 5,514,583、U. S. 6,566,107 和 U. S. 5,571,2133 中并且以引用的方式并入本文)、8b (美国专利申请 U. S. 2003/0162271; Mohagheghi 等人, (2004) *Biotechnol. Lett.* 25 :321-325) 以及 ZW658 (ATCC PTA-7858)、ZW800、ZW801-4、ZW801-5 和 ZW801-6 (描述于共同拥有且共同未决的美国专利申请公布号为 U. S. 2008-0286870A1 中,该专利以引用方式并入本文)。

[0079] 还可将受到另外工程改造以利用其它非天然底物的糖的发酵单胞菌菌株用于本发明方法中。一个例子是针对阿拉伯糖利用而工程化的运动发酵单胞菌的菌株,其描述于美国专利 5,843,760 中,该专利以引用方式并入本文。

[0080] 驯化

[0081] 为进行驯化,将利用木糖的发酵单胞菌菌株(上文所述的起始菌株)连续地生长于包含生物质水解产物的培养基中。生物质水解产物通过生物质的糖化而产生。所述生物质通常在糖化前进行了预处理。本领域技术人员可用任何已知的方法处理生物质,在水解产物中产生可发酵糖。通常所述生物质采用物理的和 / 或化学的处理,以及酶促糖化进行预处理。物理的和化学的处理可包括碾磨、铣削、切割、碱处理如用氨或氢氧化钠、以及酸处

理。低氨预处理是特别有用的,其中生物质与包含氨的水溶液接触,形成生物质-氨水混合物,其中氨的浓度足够维持生物质-氨水混合物的碱性 pH,但相对于生物质干重其小于约 12 重量%,并且生物质的干重至少约为 15 重量%固体(相对于生物质-氨水混合物的重量),如共同申请和共同拥有的美国专利申请公布 US-20070031918-A1 所公开的,其以引用方式并入本文。

[0082] 酶解糖化通常利用酶组合物或共混物来降解纤维素和 / 或半纤维素并且产生水解产物,其包含糖,诸如葡萄糖、木糖和阿拉伯糖。糖化酶见 Lynd, L. R. 等人的综述。(Microbiol. Mol. Biol. Rev., 66:506-577, 2002)。使用至少一种酶,并且通常使用糖化酶共混物,其包括一种或更多种糖苷酶。糖苷酶水解二糖、低聚糖和多糖的醚键,并且存在于广义“水解酶”(EC3.) 的酶分类 EC3. 2. 1. x (Enzyme Nomenclature 1992, Academic Press, San Diego, CA, 以及增补 1 (1993)、增补 2 (1994)、增补 3 (1995)、增补 4 (1997) 和增补 5 [分别在 Eur. J. Biochem., 223:1-5, 1994; Eur. J. Biochem., 232:1-6, 1995; Eur. J. Biochem., 237:1-5, 1996; Eur. J. Biochem., 250:1-6, 1997; 和 Eur. J. Biochem., 264:610-650 1999 中])。本发明的方法中可用的糖苷酶能根据它们水解的生物物质组分进行分类。可用于本发明方法中的糖苷酶包括纤维素水解糖苷酶(例如,纤维素酶、内切葡聚糖酶、外切葡聚糖酶、纤维二糖水解酶、 β -葡萄糖苷酶)、半纤维素水解糖苷酶(例如木聚糖酶、内切木聚糖酶、外切木聚糖酶、 β -木糖苷酶、阿拉伯木聚糖酶、甘露聚糖酶、半乳糖酶、果胶酶、葡萄糖醛酸糖苷酶)和淀粉水解糖苷酶(例如淀粉酶、 α 淀粉酶、 β 淀粉酶、葡萄糖淀粉酶、 α -葡萄糖苷酶、异淀粉酶)。此外,将其它活性加入糖化酶聚生体(如肽酶(EC3. 4. x. y)、脂肪酶(EC3. 1. 1. x 和 3. 1. 4. x)、木素酶(EC1. 11. 1. x)或阿魏酸酯酶(EC3. 1. 1. 73))中以促进从生物物质的其它组分中释放多糖是有益的。本领域熟知生产多糖水解酶的微生物常常表现出某种活性,如降解纤维素的能力,该活性由具有不同底物特异性的若干种酶或一组酶催化。因此,来自微生物的“纤维素酶”可包括一组酶、一种或更多种酶或所有酶,它们都可用于纤维素降解活性。取决于获取酶时利用的纯化方案,商业或非商业酶制剂,如纤维素酶,可包括多种酶。

[0083] 糖化酶可商购获得。这类糖化酶包括,例如, **Spezyme[®]** CP 纤维素酶、**Multifect[®]** 木聚糖酶、**Accelerase[®]** 1500、以及 **Accellerase[®]** DUET (Danisco U. S. Inc., Genencor International, Rochester, NY)。此外,糖化酶可以是未经纯化的并以细胞提取物或全细胞制剂的形式提供。可使用已经进行工程改造以表达一种或多种糖化酶的重组微生物制备所述酶。

[0084] 用于糖化的另外酶类包括,例如,糖基水解酶,诸如 GH3、GH39、GH43、GH55、GH10 和 GH11 家族的成员。GH 是一组酶,其水解两个或更多个碳水化合物间的糖苷键,或碳水化合物与非碳水化合物基团间的糖苷键。GH 家族基于序列类似性进行分类,并且所述分类可在碳水化合物-活性酶(CAZy) 数据库中得到(Cantarel 等人(2009) Nucleic Acids Res. 37 (数据库专刊):D233-238)。这些酶的某一些能够作用于多种底物并且作为糖化酶表现出效能。糖苷水解酶家族 3 (“GH3”) 的酶具有大量的已知活性,包括,例如, β -葡萄糖苷酶(EC:3. 2. 1. 21); β -木糖苷酶(EC:3. 2. 1. 37); N-乙酰基 β -氨基葡萄糖苷酶(EC:3. 2. 1. 52); 葡聚糖 β -1, 3-葡萄糖苷酶(EC:3. 2. 1. 58); 纤维糊精酶(EC:3. 2. 1. 74); 外切-1, 3-1, 4-葡

聚糖酶(EC:3.2.1);和/或 β -半乳糖苷酶(EC3.2.1.23)的活性。糖苷水解酶家族39(“GH39”)酶还具有大量已知活性,包括,例如 α -L-艾杜糖醛酸酶(EC:3.2.1.76)和/或 β -木糖苷酶(EC:3.2.1.37)的活性。糖苷水解酶家族43(“GH43”)的酶具有大量已知活性,包括,例如,L- α -阿拉伯呋喃糖酶(EC3.2.1.55); β -木糖苷酶(EC3.2.1.37);聚阿拉伯糖内切酶(EC3.2.1.99);和/或半乳糖 1,3- β -半乳糖苷酶(EC3.2.1.145)的活性。糖苷水解酶家族51(“GH51”)的酶已知具有,例如,L- α -阿拉伯呋喃糖酶(EC3.2.1.55)和/或内切葡聚糖酶(EC3.2.1.4)的活性。糖苷水解酶家族10(“GH10”)详细的描述于Schmidt等人,1999,Biochemistry38:2403-2412和Lo Leggio等人,2001,FEBS Lett509:303-308中,并且糖苷糖水解酶家族11(“GH11”)描述于Hakouvainen.,1996,Biochemistry35:9617-24中。

[0085] 在本发明的驯化方法中,将利用木糖的发酵单胞菌在提高比例的生物质水解产物的存在下连续生长于生长培养基中,如本文实例1和2中所述。以周期性间隔收取样品并分析其在水解产物培养基中的性能,包括木糖和葡萄糖利用和乙醇生产。在包含提高比例的水解产物以及其它胁迫组分(诸如乙酸酯和乙醇)的培养基中进行的多轮驯化可用于产生具有改善性能的菌株,诸如上文所述的菌株。

[0086] 改善的利用木糖的菌株的发酵

[0087] 经工程改造的具有本文所述的至少一种基因修饰的利用木糖且产生乙醇的发酵单胞菌菌株可用于发酵以产生乙醇。使运动发酵单胞菌接触包含生物质水解产物的培养基,所述生物质水解产物包括糖,其包含葡萄糖和木糖。所述糖的至少一部分来源于经预处理和糖化的纤维素或木质纤维素生物质。培养基中可包括另外的糖和/或其它培养基组分。

[0088] 当糖的浓度高至足以抑制生长时,所述培养基包括山梨醇、甘露糖醇、或它们的混合物,这公开于共有的美国专利7,629,156中。半乳糖醇或核糖醇可代替或与山梨醇或甘露糖醇组合。运动发酵单胞菌培养基中生长,所述培养基中进行了发酵并且产生了乙醇。发酵在不补充空气、氧气、或其它气体(这可包括各种条件如厌氧、微氧、或微需氧发酵)的情况下运行至少约24小时,并且可运行30小时或更长时间。达到最大乙醇产量的时间是不确定的,取决于发酵条件。发酵可在介于约30°C和约37°C之间的温度和约4.5至约7.5的pH下进行。

[0089] 可以在实验室规模的发酵罐中以及扩规发酵中(其中生产了商业规模量的乙醇),在包含包括木糖在内的混合糖培养基中培养本发明的运动发酵单胞菌。当需要商业生产乙醇时,可使用多种培养方法。例如,由本发明的运动发酵单胞菌菌株进行的大规模生产可以通过分批培养方法和连续培养方法进行。经典的分批培养方法是封闭系统,其中培养基的组成在培养开始时设定并且在培养过程期间不进行人工改变。因此,在培养过程开始时,用所需的生物接种培养基,并且不向系统中添加任何物质可以发生生长或代谢活性。然而,通常来说,“分批”培养是指碳源的添加是成批的,但经常试图控制诸如pH和氧浓度之类的因素。在分批系统中,代谢产物和生物质组成持续改变直至培养结束时。在分批培养物内,细胞缓慢通过静态延滞期到达高速生长对数期,并最后达到稳定期,此时生长速率减缓或终止。如果不加以处理,稳定期中的细胞将最终死亡。处于对数期的细胞通常负责乙醇的批量生产。

[0090] 标准分批式系统的一种变型是补料分批系统。分批补料培养方法也适用于本发明运动发酵单胞菌菌株的培养,并且其包括典型的分批系统,不同的是底物随着培养进行以递增方式添加。补料-分批系统中实际底物浓度的测量是困难的,并因此根据可测量因素例如 pH 和废气如 CO₂ 的分压的改变来评估。分批和分批补料培养方法在本领域内是常用的且熟知的,并且例子可见于 Biotechnology :A Textbook of Industrial Microbiology, Crueger, Crueger 和 Brock, 第二版(1989) Sinauer Associates, Inc., Sunderland, MA, 或 Deshpande, Mukund V., Appl. Biochem. Biotechnol., 36, 227, (1992), 将其以引用方式并入本文。

[0091] 乙醇的商业生产也可通过连续培养进行。连续培养是一种开放式系统,其中将培养基连续加入生物反应器里,并同时移出等量条件培养基用于加工。连续培养一般使细胞维持在其中细胞主要处于对数生长期的恒定高液相密度。作为另外一种选择,连续培养可以用固定化细胞来进行,其中连续添加碳和营养素,且连续从细胞团块中取出有价值的产物、副产物或废弃物。细胞固定可使用范围广泛的固体载体进行,所述固体载体由本领域技术人员已知的天然材料和 / 或合成材料组成。

[0092] 尤其适用于乙醇生产的发酵方法如下所述。将所需本发明运动发酵单胞菌菌株在包含半复合培养基的摇瓶中生长,培养温度为 30℃ 至约 37℃,在回旋振荡器中以约 150rpm 的转速振荡培养,然后将其转移到包含类似培养基的 10L 种子发酵罐中。种子培养物在种子发酵罐中厌氧生长直至 OD₆₀₀ 介于 3 和 10 之间,然后将其转移到生产发酵罐中,其中的发酵参数经最优化用于乙醇生产。从种子罐转移到生产罐的典型种菌体积在约 2% 至约 20%v/v 的范围内。正如本领域的技术人员所知,用于本发明菌株的发酵培养基包含至少约 50% 的生物质水解产物并且可补充以其它营养物。所述培养基中存在终浓度约 5mM 的山梨醇和甘露醇。

[0093] 使用苛性碱溶液(如氢氧化铵、氢氧化钾、或氢氧化钠)和硫酸或磷酸将发酵控制在 pH5.0-6.0。将所述发酵罐的温度控制在 30℃ -35℃ 下。为使发泡最小化,根据需要向所述容器内加入消泡剂(任何类别-基于硅氧烷的、基于有机物的,等等)。可任选地将诸如卡那霉素这样的抗生素(所述菌株中存在针对其的抗生素抗性标记或者所述菌株对其有抗性)用于最小化污染。

[0094] 上述任何条件组和这些条件中在本领域为人所熟知的附加变化是通过利用木糖的发酵单胞菌菌株产生乙醇的合适条件。

[0095] 优选实施例说明

[0096] 据本文公开,在生物质水解产物存在的情况下具有改善的葡萄糖和木糖的利用和乙醇生产的利用木糖的发酵单胞菌菌株可通过在生物质水解产物培养基中进行的驯化和筛选过程而获得。葡萄糖和木糖利用以及乙醇生产中的提高是通过与利用木糖的对应的未驯化菌株的葡萄糖和木糖利用而测量的。所对应的未驯化菌株是用作为水解产物驯化过程的起始菌株的菌株。

[0097] 运动发酵单胞菌菌株是分离自水解产物驯化实验,如本文实例 1 和 2 中所述。两种分离的菌株命名为驯化株 5-6 (或 AR25-6) 和驯化株 7-31 (或 AR37-31),并且受到进一步表征。当如本文实例 3 中所述在玉米棒水解产物中生长时,这些菌株对葡萄糖和木糖的利用比对应的未驯化菌株 ZW705 更迅速。在发酵运行 21 小时后,已经利用了约 20% 更多的

葡萄糖。此外,在 21 小时生产了约 22% 更多的乙醇,而在 52 小时使用了约 21% 更多的木糖并且产生了 5% 更多的乙醇。一般来讲,具有驯化株 5-6 或驯化株 7-31 菌株的新型基因变化(其在下文所述)的菌株中的葡萄糖利用、木糖利用和乙醇生产中的精确百分比提高应取决于众多因素。这些因素包括发酵条件以及所述菌株除本文公开的新型基因变化以外的遗传特性。

[0098] 对所述基因组进行测序以鉴定驯化株 5-6 或驯化株 7-31 菌株的基因组中的任何变化。通过将这些基因组序列与对应的其实菌株 ZW705 (描述于美国专利公布 US20110014670A1 中)、野生型菌株 ZW1 (ATCC31821) 和公开的运动发酵单胞菌菌株 ATCC31821 的序列(Seo 等人, Nat. Biotech. 23:63-82005 ;NCBI 参考号 :NC_006526. 2)之间进行的比较,测定了两种菌株在相同的编码区中均具有单一的新型突变。这一编码区鉴定为处于公开的运动发酵单胞菌基因组序列(NCBI 参考号 :NC_006526. 2)中的 zmo1432 开放阅读框(ORF)并且是具有 SEQ ID NO:1 的核苷酸 1446603 至 1448633 的互补序列。

[0099] 驯化株 5-6 在 SEQ ID NO:1 的位置 1097 处具有从 C 到 G 的突变。这导致了 zmo1432 编码的蛋白质(SEQ ID NO:2)的密码子 366 从编码苏氨酸的 ACA 变成编码精氨酸的 AGA, 产生了 SEQ ID NO:3 的蛋白质,其中精氨酸取代了苏氨酸 No. 366。驯化株 7-31 在 SEQ ID NO:1 的位置 350 处具有从 C 到 G 的突变。这导致了 zmo1432 编码的蛋白质(SEQ ID NO:2)的密码子 117 从编码丝氨酸的 TCT 变成编码苯丙氨酸的 TTT,产生了 SEQ ID NO:4 的蛋白质,其中苯丙氨酸取代了丝氨酸 No. 117。

[0100] 假定的编码区 zmo1432 (或 ORF) 在运动发酵单胞菌全基因组序列(NCBI 参考号 :NC_006526. 2)中标注为编码“镰刀菌酸抗性蛋白”。其标注为具有镰刀菌酸抗性蛋白保守区,所述保守区是蛋白质基序 :PFAM:PF04632 (Wellcome Trust Sanger Institute, Genome Research Limited, Hinxton, England)。这一基序存在于与镰刀菌酸抗性关联的蛋白质中,其来自洋葱伯克霍尔德菌(*Burkholderia cepacia*) (Swiss-Prot::P24128 (SEQ ID NO:5);Utsumi 等人(1991) Agric. Biol. Chem. 55:1913-1918 ;所述生物体由洋葱假单胞菌(*Pseudomonas cepacia*) 重新命名)和产酸克雷伯菌(*Klebsiella oxytoca*) (Swiss-Prot::Q48403 (SEQ IDNO:6);Toyoda 等人(1991) J Phytopathol. 133:165-277), 所述蛋白质可能是膜转运蛋白。

[0101] 洋葱伯克霍尔德菌染色体 1 全序列中名为 Bcenm03_1426 区标注为编码假定的镰刀菌酸抗性转运蛋白(登录号 :YP_001764723 ;Copeland 等人于 2/27/2008 提交)。该蛋白质(SEQ ID NO:7)与 zmo1432 编码的蛋白质具有相似性,如本文实例 4 中所述。此外,大肠杆菌蛋白 AaeB (SEQ IDNO:8) (其为芳族羧酸外排泵(VanDyk 等人, J. Bact. 186:7196-7204 (2004))与 zmo1432 编码的蛋白质具有相似性,如本文实例 4 中所述。因此,在生长于水解产物培养基的利用木糖的发酵单胞菌菌株中,发酵单胞菌的转运蛋白可能是改善葡萄糖和木糖利用以及乙醇生产的突变的靶。

[0102] 实例

[0103] 本发明将在下面的实例中得到进一步阐述。应该理解,这些实例尽管说明了本发明的优选实施例,但仅是以例证的方式给出的。从上文的讨论和这些实例中,本领域的技术人员能够确定本发明的特性,并且在不脱离其实质和范围的情况下,能对本发明进行各种变化和修改以适应不同的用途和条件。

[0104] 缩写的含义如下：“kb”是指千碱基，“bp”是指碱基对，“nt”是指核苷酸，“hr”是指小时，“min”是指分钟，“sec”是指秒钟，“d”是指天数，“L”是指升，“ml”是指毫升，“ μ L”或“ μ l”是指微升，“ μ g”是指微克，“ng”是指纳克，“mM”是指毫摩，“ μ M”是指微摩，“nm”是指纳摩，“ μ mol”是指微摩尔，“pmol”是指皮摩尔，“OD600”是指 600nm 下的光密度。

[0105] 一般方法

[0106] 恒浊器

[0107] 驯化在恒浊器中进行(美国专利 6,686,194 ;Heurisko USA, Inc.Newark, DE),其为连续流动培养装置,其中培养物中的细胞浓度通过控制培养基到培养物的流动而保持恒定,使得培养物的浊度保持在指定的狭窄限制内。培养物在连续培养装置中可利用两种培养基生长,一种为静息培养基(培养基 A),一种为刺激培养基(培养基 B)。培养物在生长室中的静息培养基上生长至浊度设置点,然后将其以设置以保持细胞密度的稀释速率稀释。通过每 10 分钟加入一次限定体积的培养基来进行稀释。当恒浊器进入培养基刺激模式时,基于在前一次培养基添加后返回设置点的速率选择加入刺激培养基或静息培养基。生长室中培养基的稳态浓度为培养基 A 和培养基 B 的混合;根据从各个培养基提取的速率决定两种培养基的比例,其为设置的稀释速率以允许维持设置的细胞密度。代表生长室中种群的细胞样品以每周的间隔从恒浊器外流液中回收(在捕获室)。所述细胞样品在 MRM3G6 培养基上生长一次并在 -80°C 作为甘油原液保存。

[0108] 酶

[0109] **Spezyme[®]** CP-100、**Multifect[®]** CX12L 和 **Accellerase[®]** 1500 来自于 Danisco U. S. Inc., Genencor (Rochester, NY)。

[0110] Novozyme188 获自 Novozymes (2880Bagsvaerd, Denmark)。

[0111] 本文糖化加工中使用的另外酶类是糖基水解酶(GH) Xyn3、Fv3A、Fv51A 和 Fv43D。Xyn3 (SEQ ID NO:9)来自里氏木霉的 GH10 木聚糖酶家族;Fv3A (SEQ ID NO:10)来自轮枝镰孢菌 GH3 酶家族;Fv43D (SEQ ID NO:11)来自轮枝镰孢菌 GH43 酶家族;并且 Fv51A (SEQ ID NO:12)来自轮枝镰孢菌 GH51 酶家族。

[0112] 培养基

[0113] 首先通过对玉米棒粉进行稀氨水预处理来制备驯化中使用的玉米棒水解产物。将水平的 Littleford Day130L 反应容器加载研磨的玉米棒,所述反应容器包含围绕容器主体和一侧的传递蒸汽的夹套(Littleford Day, Inc., Florence, KY)。在靠近容器顶部导入 29 重量%的氢氧化铵溶液和水以提供相对于生物质干重的 6 重量%的 NH_3 之前,应用真空使容器内压力降至 0.1atm。在所述容器顶部附近引入蒸气以将内部容器温度升高至 145°C 。保持这一温度 20 分钟。在预处理结束时,将反应器压力降至大气压。随后施用真空(大约低于 1atm)以降低温度至低于 60°C 。

[0114] 随后使用酶聚生体来处理经预处理的玉米棒,从而实现使用酶混合物来对纤维素和半纤维素聚合物进行酶法水解,所述酶混合物包含:Spezyme CP-100, 34mg 蛋白/g 经预处理的玉米棒中的葡聚糖;Multifect CX12L, 12.5mg 蛋白/g 经预处理的玉米棒中的木聚糖;Novozymes188, 6.6mg 蛋白/g 经预处理的玉米棒中的葡聚糖。在 25% (重量/体积)的经预处理的玉米棒干物质、pH5.3 和 47°C 下进行所述水解反应。连续搅拌所述反应并运行

72 小时。

[0115] 使用初次连续离心从所述水解产物中除去固体,从而使该混合物部分地澄清。在 $18,000 \times G$ 下对所述部分澄清的混合物再次离心 20 分钟,随后首先通过 0.45 微米滤器,继之以 0.22 微米滤器,从而产生澄清的经过滤消毒的水解产物。

[0116] 通过 HPLC 分析测定葡萄糖、木糖和乙酸酯的各自浓度。所述澄清的水解产物具有 68g/L 的葡萄糖浓度、46g/L 的木糖浓度和 5g/L 的乙酸酯浓度。对所述澄清的水解产物补充以 6.2g/L 的乙酸铵来提高总乙酸铵浓度,从而使该浓度处于 11 至 12g/L 的范围内。在注明的情况下,加入 0.5% 的酵母提取物(Difco 酵母提取物, Becton, Dickinson and Co., Sparks, MD)来提供额外的营养物。将这一培养基标为 HYAc/YE。将 HYAc/YE 培养基的 pH 调节到 5.8,并且对该培养基进行过滤灭菌。

[0117] 使用上述方法制备了用于 1L 发酵测试的水解产物培养基,不同之处在于水解中使用的酶组合物变为包含 **Accellerase[®]** 1500、Xyn3、Fv3A、Fv51A 和 Fv43D 的酶共混物,其以 21.3mg 蛋白/g 葡聚糖 + 木聚糖加入所述水解反应。将所述水解产物不经澄清处理而用于 1L 规模的发酵。

[0118] 附加的培养基

[0119] 每升 MRM3 包含:酵母提取物(10g)、 KH_2PO_4 (2g) 和 $MgSO_4 \cdot 7H_2O$ (1g)

[0120] MRM3G6.5X4.5NH₄Ac12.3 是包含 65g/L 葡萄糖、45g/L 木糖、12.3g/L 乙酸铵的 MRM3

[0121] G5 或 MRM3G5 是包含 50g/L 葡萄糖的 MRM3

[0122] G10 或 MRM3G10 是包含 100g/L 葡萄糖的 MRM3

[0123] MRM3G2 是包含 20g/L 葡萄糖的 MRM3

[0124] MRM3X2 是包含 20g/L 木糖的 MRM3

[0125] 半 YEMaxSM:10g/L 的 Difco 酵母提取物、2g/L 的 KH_2PO_4 、5g/L 的 $MgSO_4 \cdot 7H_2O$ 、10mM 的山梨醇、150g/L 的葡萄糖

[0126] 对于图 2 显示的平板研究,对 MRM3G2 (20g/L 的葡萄糖)和 MRM3X2 (20g/L 的木糖)补充以 1.5% 琼脂,加热至糖溶解,冷却至 45°C 并且倒入培养皿。

[0127] 冷冻原液培养物的制备

[0128] 对菌株 ZW705、每周的恒浊器样品群体培养物和分离的突变体克隆制备了冷冻原液培养物。通过将原始冷冻原液生长于 G5 或 G10 培养基并且使用为制备新冷冻原液而产生的生物质来制备另外的冷冻原液。

[0129] 种菌接种、批式驯化和恒浊器培养物

[0130] 使用冷冻原液来接种过夜的 G5 和 G10 种菌培养物。使用种菌培养物来接种批式驯化培养物,其通过对所述种菌培养物培养基离心并且将细胞团块稀释于新鲜的驯化培养基而进行。对于恒浊器接种,将整个种菌团块重悬浮于 10-15ml 的恒浊器测试培养基中并将其用于接种恒浊器反应器或将 5ml 的过夜种菌用作为 10% 的接种物。

[0131] 培养物 OD₆₀₀ 测量

[0132] 为测量培养物的 OD,将样品稀释于 100g/L 的 xylose (稀释物和空白)中。在进行 OD 测量之前将经稀释的培养物放置 15 分钟。全部的 OD 测量均在 600nm 下进行。

[0133] HPLC 分析

[0134] 使用 Waters Alliance HPLC 系统来实施 HPLC 分析。所用的色谱柱为具有 BioRad Micro-Guard Cartridge Cation-H (#125-0129, Bio-Rad, Hercules, CA) 的 Transgenomic ION-300 色谱柱 (#ICE-99-9850, Transgenomic, Inc)。使用 0.01N H₂SO₄ 作为溶剂, 在 75°C 和 0.4mL/min 的流速下运行所述色谱柱。采用外标校准曲线, 用折射率检测器测定原料糖和产物的浓度。

[0135] 菌株 ZW705 的描述

[0136] 运动发酵单胞菌菌株 ZW705 产自菌株 ZW804-1。ZW801-4 是重组利用木糖的运动发酵单胞菌菌株, 描述于共同拥有的美国专利 7, 741, 119 中, 该专利以引用方式并入本文。菌株 ZW801-4 来源于菌株 ZW800, 菌株 ZW800 来源于菌株 ZW658, 它们均描述于美国专利 7, 741, 119 中。ZW658 通过经由序贯转座事件将两个操纵子整合到 ZW1 (ATCC31821) 基因组中, 然后通过包含木糖的选择培养基筛选进行构建 (U. S. 7, 629, 156), 所述操纵子是 P_{gap}xy1AB 和 P_{gap}taltkt, 它们包含编码木糖异构酶、木酮糖激酶、转醛醇酶和转酮醇酶的四个木糖利用基因。ZW658 被保存为 ATCC PTA-7858。在 ZW658 中, 使用宿主媒介的基因双交换同源重组和作为选择性标记的奇放线菌素抗性使编码葡萄糖-果糖氧化还原酶的基因插入失活以产生 ZW800 (U. S. 7, 741, 119)。使用 Cre 重组酶, 通过位点特异性重组移除由 loxP 位点束缚的奇放线菌素抗性标记以产生 ZW801-4。

[0137] 运动发酵单胞菌菌株 ZW801-4 的培养物针对在包含乙酸钠的培养基的胁迫条件下的生长进行驯化, 从而产生了 ZW705, 正如共同拥有的美国专利公布 US20110014670A1 中的描述, 该专利以引用方式并入本文。ZW801-4 的连续培养在 250ml 搅拌的、pH 和温度受控的发酵罐 (Sixfors; Bottmingen, Switzerland) 中进行。发酵基础培养基是 5g/L 酵母提取物, 15mM 磷酸铵, 1g/L 硫酸镁, 10mM 山梨醇, 50g/L 木糖和 50g/L 葡萄糖。通过在 97 天中逐渐提高加到上述连续培养基中的乙酸钠浓度, 同时保持通过特定稀释速率测得的确定生长速率来影响在高浓度乙酸和氨的存在下的适应生长。将乙酸钠浓度提升至 160mM。通过加入磷酸铵至在连续培养 139 天的末期最终总铵离子浓度为 210mM 来实现铵离子浓度的进一步提高。通过接种单菌落并扩增一个选择的菌落来从适应种群中分离菌株 ZW705。

[0138] 实例 1

[0139] 使用自动细胞密度受控的连续培养设备针对玉米棒水解产物的驯化

[0140] 将基本方法中所述的恒浊器用于将运动发酵单胞菌培养物进行在玉米棒水解产物培养基中生长的驯化。将菌株 ZW705 的培养物生长于基本方法中所述的恒浊器中直至任意的浊度设定点, 其规定所述培养物使用了输入培养基中的全部葡萄糖和约一半的木糖以符合在设定的稀释速度下所述设定点的细胞密度。使用静息培养基, 其为 50%HYAc/YE 和 50%MRM3G6. 5X4. 5NH₄Ac12. 3。使用刺激培养基 HYAc/YE 根据基本方法中的所述运行恒浊器。在 6 周中每周从捕集腔 (trap chamber) 收取细胞样品。

[0141] 6 周连续培养后, 将每周保存的细胞原液的样品在 MRM3G6 中复苏并且在 33°C 下在 10ml 静态培养物中生长至约 10D600。将这些培养物用于接种 12ml HYAc/YE 培养基的培养物直至约 0.40D600nm, 并在 30°C 下生长该培养物。根据表 1 在不同时间收取样品, 对 OD600、糖消耗和乙醇生产进行分析。结果示于表 1 中。

[0142] 表 1: 对 12ml 的 HYAc/YE 发酵物中的每周恒浊器样品的分析

[0143]

培养物	取样时间	OD 600nm	剩余葡萄糖 (g/L)	剩余木糖 (g/L)	乙醇 (g/L)
ZW705	0	0.400	69.8	46.0	0.0
	24	0.760	52.8	45.4	8.4
	48	2.140	5.4	30.7	38.2
	72	2.300	2.3	16.0	45.1
第 1 周	0	0.417	69.8	46.0	0.0
	24	0.78	48.8	45.6	9.7
	48	1.430	4.8	34.2	35.6
	72	1.890	2.4	25.7	37.7
第 2 周	0	0.351	69.8	46.0	0.0
	24	0.810	53.2	46.1	7.7
	48	1.670	7.6	36.1	33.5
	72	1.710	2.7	26.6	39.3
第 3 周	0	0.395	69.8	46.0	0.0
	24	1.090	42.9	44.3	13.4
	48	1.920	2.9	27.3	40.2
	72	2.410	1.9	19.7	42.1
第 4 周	0	0.386	69.8	46.0	0.0
	24	1.070	42.2	44.2	13.7
	48	1.510	3.3	30.7	38.4
	72	2.010	2.1	25.2	39.4
第 5 周	0	0.328	69.8	46.0	0.0
	24	0.920	52.7	45.5	7.7
	48	1.350	8.5	36.2	34.1
	72	1.650	2.9	25.8	40.4
第 6 周	0	0.402	69.8	46.0	0.0
	24	0.600	55.5	45.9	6.9
	48	1.360	13.9	39.8	29.1
	72	1.530	4.1	31.6	37.5

[0144] 全部培养物保持了在木糖上生长的能力,但是木糖利用的速度在驯化第三周后显示出降低。从第 3 周结束时取样的细胞培养物分离了单菌落。通过将保存的甘油原液在 MRM3G5 中生长并随后涂布于 MRM3X2 平板来分离菌落。单菌落是印迹于 MRM3X2 和 MRM3G2 平板上的复制品。在两种碳源上选择大而密实的印片作为菌株并且作为冷冻甘油原液加以维持。根据上文所述的冷冻原液群体测试的描述,将选定的菌株生长于 12ml 的测试培养物上。六种菌株和 ZW705 的结果在表 2 中示出。

[0145] 表 2:对从生长于 12ml 的 HYAc/YE 发酵物中的 ZW705 的恒浊器驯化第 3 周分离的纯培养物的分析

[0146]

菌株	取样时间	OD 600nm	剩余葡萄糖 (g/L)	剩余木糖 (g/L)	乙醇 (g/L)
ZW705	0	0.383	74.9	44.4	0.0
	24	0.450	51.1	43.4	13.4
	48	1.350	5.7	35.1	31.4
	72	2.770	1.9	17.1	49.0
12-18X-1-10	0	0.333	74.9	44.4	0.0
	24	0.760	47.7	44.4	14.6
	48	2.530	3.1	28.3	36.7
	72	2.800	1.7	12.1	50.4
12-18X-2-36	0	0.400	74.9	44.4	0.0
	24	0.900	46.9	44.3	15.1
	48	2.220	2.8	24.2	38.9
	72	3.430	1.8	9.1	53.3
12-18X-5-34	0	3.820	74.9	44.4	0.0
	24	0.530	52.7	44.9	13.1
	48	2.430	4.3	30.3	35.7
	72	3.050	1.9	13.0	51.1
12-18X-6-9	0	0.401	74.9	44.4	0.0
	24	0.530	51.2	44.4	13.1
	48	2.590	3.2	28.7	36.6
	72	3.050	1.8	11.4	52.5
12-18X-7-43	0	0.351	74.9	44.4	0.0
	24	0.600	49.6	43.9	14.3
	48	2.200	3.1	26.9	38.3
	72	3.330	1.7	9.9	53.8
12-18X-8-19	0	0.409	74.9	44.4	0.0
	24	0.749	45.2	44.4	16.0
	48	2.230	3.4	27.3	34.8
	72	2.910	1.7	11.7	52.1

[0147] 全部选定菌株在 12ml 测试中产生了比进入驯化的菌株(ZW705)更多的乙醇。所述菌株利用了略多于亲本菌株的木糖。选择菌株 12-18X-2-36 用于另一回合的驯化。

[0148] 实例 2

[0149] 使用自动细胞密度受控的连续培养设备针对外加乙醇的玉米棒水解产物的驯化

[0150] 如上文针对菌株 ZW705 以及基本方法中所述,将菌株 12-18X-2-36 (描述于实例 1 中) 的培养物生长于恒浊器中,不同之处在于静息培养基是 HYAc/YE 并且刺激培养基是 HYAc/YE+9 重量%乙醇。将恒浊器运行 4 周,每周对腔室外向流取样。通过所述外向流样品制备冷冻细胞原液。在 MRM3G6 中复苏冷冻细胞原液以用于 12ml 水解产物发酵,所述发酵在与第一次恒浊器运行所描述的相同条件下进行,但具有约 0.5OD600 的起始密度。一组发酵在相同的 HYAc/YE 培养基中进行,另一组发酵在加入了 30g/L 培养基的乙醇的 HYAc/YE 培养基中进行。两组发酵的结果均在表 3 中示出。

[0151] 表 3 :对生长于 12ml 的 HYAc/YE 或 HYAc/YE+30g/L 乙醇发酵物中的每周恒浊器样品的分析

[0152]

		未添加乙醇的 HYAc/YE				具有 30g/L 乙醇的 HYAc/YE			
培养物	取样时间	OD 600nm	剩余葡萄糖 (g/L)	剩余木糖 (g/L)	乙醇 (g/L)	OD 600nm	剩余葡萄糖 (g/L)	剩余木糖 (g/L)	乙醇 (g/L)
ZW705	0	0.530	75.5	47.7	0.0	0.530	73.3	48.3	26.2
	24	0.890	50.1	48.0	13.7	0.470	63.6	46.8	27.3
	48	2.200	3.9	33.6	43.6	0.760	51.1	45.7	32.7
	72	2.500	1.8	19.7	50.3	1.060	39.0	45.7	39.6
第 1 周	0	0.513	75.5	47.7	0.0	0.515	73.3	48.3	26.2
	24	1.360	20.1	44.2	30.2	1.370	45.2	46.9	36.3
	48	2.200	2.1	27.3	47.7	1.810	14.7	42.5	50.1
	72	2.520	1.7	18.5	50.6	1.920	3.8	40.1	57.7
第 2 周	0	0.504	75.5	47.7	0.0	0.504	73.3	48.3	26.2
	24	1.140	32.4	47.1	23.2	0.950	51.6	46.8	34.1
	48	1.910	2.4	27.8	47.7	1.500	19.4	42.8	50.4
	72	2.290	1.7	18.5	51.4	1.780	5.7	38.0	60.5
第 3 周	0	0.527	75.5	47.7	0.0	0.527	73.3	48.3	26.2
	24	0.640	53.5	48.7	12.3	0.570	62.1	48.2	28.5
	48	1.420	7.3	37.7	39.5	0.600	51.9	46.1	33.2
	72	2.010	2.0	25.5	47.3	1.060	35.3	44.8	38.7
第 4 周	0	0.518	75.5	47.7	0.0	0.518	73.3	48.3	26.2
	24	0.730	25.5	45.2	27.8	0.480	64.3	47.7	25.5
	48	1.760	2.3	23.7	49.3	0.560	52.8	46.8	32.9
	72	1.890	1.8	15.8	52.6	1.080	38.4	45.2	40.4

[0153]

[0154] 在未向 HYAc/YE 培养基中添加乙醇的测试发酵中,全部培养物均类似于 ZW705 对照。在向 HYAc/YE 培养基中添加 30g/L 乙醇的发酵中,在第 1 周和第 2 周后收取的培养物在利用葡萄糖和产生乙醇方面远远优于 ZW705 以及第 3 周和第 4 周的样品。在以外加乙醇起始的测试中,在第 2 周储存的培养物具有最高的终末乙醇滴度并且其被选择用于分离单细胞来源菌株。将针对产生 12-18X-2-36 菌株的筛选而描述的菌株分离工序实施两次,并且两回合筛选的结果在表 4 和 5 中示出。筛选在未添加乙醇或添加了 40g/L 乙醇的 HYAc/YE 培养基中起始。

[0155] 表 4 :对从生长于 12ml HYAc/YE 或 HYAc/YE+ 乙醇的发酵物中的 12-18X-2-36 进行的恒浊器驯化第 2 周分离的纯培养物的分析(第 1 回合)

[0156]

培养物	取样 时间	未添加乙醇的 HYAc/YE				具有 40g/L 乙醇的 HYAc/YE			
		OD 600nm	剩余葡 萄糖 (g/L)	剩余 木糖 (g/L)	乙醇 (g/L)	OD 600nm	剩余葡 萄糖 (g/L)	剩余 木糖 (g/L)	乙醇 (g/L)
ZW705	0.0	0.5	77.7	50.9	0.0	0.50	75.1	49.2	28.1
	24.0	0.8	56.5	50.2	9.0	0.42	69.8	50.0	24.5
	48.0	1.8	12.6	41.6	28.2	0.75	62.6	49.6	27.5
	72.0	1.9	3.8	24.0	46.7	0.60	52.7	48.7	36.5
1-32	0.0	0.5	77.7	50.9	0.0	0.50	75.1	49.2	28.1
	24.0	0.8	55.3	49.6	12.0	0.41	65.3	48.0	35.2
	48.0	1.9	7.8	38.5	39.1	0.68	57.3	47.6	38.2
	72.0	2.4	2.6	21.2	51.6	0.72	50.8	47.5	41.0
1-44	0.0	0.5	77.7	50.9	0.0	0.50	75.1	49.2	28.1
	24.0	0.8	53.7	49.9	12.6	0.44	66.7	48.2	34.7
	48.0	1.9	6.5	37.5	40.1	0.75	60.8	48.5	35.5
	72.0	2.2	1.8	21.7	50.1	0.65	55.8	48.1	37.4
2-24	0.0	0.5	77.7	50.9	0.0	0.50	75.1	49.2	28.1
	24.0	0.7	59.4	49.7	9.8	0.41	67.6	48.6	32.0
	48.0	1.9	4.8	34.9	42.8	0.69	61.2	48.6	34.2
	72.0	2.3	2.5	18.6	51.3	0.64	55.9	48.6	36.9
2-47	0.0	0.5	77.7	50.9	0.0	0.50	75.1	49.2	28.1

[0157]

	24.0	0.9	52.1	49.5	13.5	0.49	66.3	48.7	30.8
	48.0	1.8	9.6	40.3	37.0	0.90	58.6	48.8	34.4
	72.0	1.8	2.7	27.6	46.3	0.61	51.8	48.4	38.2
3-8	0.0	0.5	77.7	50.9	0.0	0.50	75.1	49.2	28.1
	24.0	0.8	56.9	49.8	10.9	0.44	66.6	48.0	33.8
	48.0	2.0	50.0	34.3	43.0	0.76	58.3	47.9	37.2
	72.0	2.4	2.6	19.6	50.7	0.85	50.6	47.7	40.8
3-45	0.0	0.5	77.7	50.9	0.0	0.50	75.1	49.2	28.1
	24.0	0.6	60.4	950.0	9.2	0.44	67.6	48.4	34.3
	48.0	1.3	14.1	43.4	33.6	0.55	61.4	48.1	36.7
	72.0	1.8	2.6	19.6	45.1	0.52	57.2	47.9	38.0
4-17	0.0	0.5	77.7	50.9	0.0	0.50	75.1	49.2	28.1
	24.0	0.6	58.8	49.6	10.0	0.44	66.7	48.3	32.8
	48.0	1.4	5.7	38.0	40.4	0.60	59.7	48.4	34.8
	72.0	1.9	2.4	29.6	49.6	0.54	55.2	48.6	35.5
4-27	0.0	0.5	77.7	50.9	0.0	0.50	75.1	49.2	28.1
	24.0	0.8	49.8	49.4	14.6	0.52	63.6	48.4	33.7
	48.0	1.6	4.3	34.1	44.0	0.74	53.8	48.2	37.9
	72.0	2.1	2.6	23.7	48.1	0.68	47.5	48.1	40.3
5-6	0.0	0.5	77.7	50.9	0.0	0.50	75.1	49.2	28.1
	24.0	0.6	53.9	79.7	12.6	0.46	59.7	48.7	35.5
	48.0	1.8	5.8	36.9	41.9	0.90	47.7	48.8	40.3
	72.0	1.9	2.7	23.8	48.7	0.59	41.5	48.6	42.9
5-35	0.0	0.5	77.7	50.9	0.0	0.50	75.1	49.2	28.1
	24.0	0.8	52.7	49.7	13.1	0.42	65.8	48.2	33.3
	48.0	1.7	5.7	36.8	41.6	0.74	58.4	48.0	36.3
	72.0	1.9	2.4	22.7	49.6	0.80	53.6	48.1	38.1

[0158] 当以未添加乙醇起始时,大多数菌株利用了所述水解产物培养基中的全部葡萄糖,但并未利用全部的木糖,从而使总乙醇生产取决于木糖利用的程度。数个菌株利用了比未驯化亲本菌株 ZW705 更多的木糖并且产生了更多的乙醇。当将 40g/L 的乙醇加入起始培养基时,在全部菌株中生长和糖利用远为更低。在这一测试中,大多数情况下中木糖利用极低或者木糖完全未被消耗。三个菌株利用了比 ZW705 显著更多的葡萄糖并且产生了更多的乙醇。这些菌株中的最佳菌株是表 4 中示出的驯化株 5-6。

[0159] 表 5 :对从生长于 12ml HYAc/YE 或 HYAc/YE+ 乙醇的发酵物中的 12-18X-2-36 进行的恒浊器驯化第 2 周分离的纯培养物的分析(第 2 回合)

[0160]

培养物	取样时间	未添加乙醇的 HYAc/YE				具有 40g/L 乙醇的 HYAc/YE			
		OD 600nm	剩余葡萄糖 (g/L)	剩余木糖 (g/L)	乙醇 (g/L)	OD 600nm	剩余葡萄糖 (g/L)	剩余木糖 (g/L)	乙醇 (g/L)
ZW705	0.0	0.5	77.8	51.5	0.0	0.50	74.9	49.1	30.6
	24.0	0.9	40.7	48.6	19.0	0.60	56.2	49.0	34.0
	48.0	2.5	2.9	26.5	48.3	1.14	28.4	46.9	48.8
	72.0	2.9	2.2	17.9	52.7	1.49	17.5	44.7	55.3
6-4	0.0	0.5	77.8	51.5	0.0	0.50	74.9	49.1	30.6
	24.0	0.9	19.0	46.0	30.6	0.79	54.4	49.0	34.9
	48.0	2.4	2.9	28.4	47.2	1.75	18.9	44.2	55.2
	72.0	2.6	2.1	18.4	51.8	1.98	9.2	39.8	61.7
6-43	0.0	0.5	77.8	51.5	0.0	0.50	74.9	49.1	30.6
	24.0	1.3	28.0	47.7	25.0	0.52	58.6	48.3	36.5
	48.0	2.7	2.7	20.9	50.5	0.65	44.9	47.6	44.0
	72.0	2.8	2.2	14.6	52.9	1.12	37.9	47.3	46.8
7-13	0.0	0.5	77.8	51.5	0.0	0.50	74.9	49.1	30.6
	24.0	1.6	17.5	46.6	29.7	0.51	57.0	48.2	38.0
	48.0	32.7	2.6	18.2	50.6	0.99	35.7	46.3	49.4
	72.0	3.3	2.2	11.8	53.3	1.99	25.9	44.9	53.8
7-31	0.0	0.5	77.8	51.5	0.0	0.50	74.9	49.1	30.6
	24.0	1.4	23.1	46.9	27.7	0.77	39.5	48.0	39.5
	48.0	2.5	2.7	19.2	52.3	1.35	22.3	45.2	56.7
	72.0	3.0	2.1	12.8	53.7	2.09	12.7	43.6	56.5
8-2	0.0	0.5	77.8	51.5	0.0	0.50	74.9	49.1	30.6
	24.0	1.5	22.1	47.5	20.7	0.59	54.8	48.3	37.8
	48.0	3.3	2.8	20.3	50.3	1.24	30.4	46.3	51.7
	72.0	3.4	2.3	14.8	51.8	1.14	20.4	44.8	55.0
8-22	0.0	0.5	77.8	51.5	0.0	0.50	74.9	49.1	30.6
	24.0	1.2	28.7	47.6	24.6	0.56	58.1	48.6	35.5
	48.0	2.3	2.8	19.2	51.5	0.99	33.6	46.4	49.2
	72.0	2.8	2.2	11.4	55.2	1.14	21.7	44.1	54.9
9-5	0.0	0.5	77.8	51.5	0.0	0.50	74.9	49.1	30.6
	24.0	1.3	32.0	49.2	23.0	0.67	57.9	48.2	37.0
	48.0	2.5	2.9	23.0	48.8	1.07	32.0	46.2	51.8
	72.0	2.7	2.3	14.7	52.4	1.68	22.2	44.3	55.6
9-21	0.0	0.5	77.8	51.5	0.0	0.50	74.9	49.1	30.6
	24.0	1.4	28.4	48.0	24.9	0.51	58.7	48.3	35.9
	48.0	2.6	2.7	21.1	50.4	0.88	36.4	47.0	48.4
	72.0	2.7	2.3	13.4	53.5	1.06	26.1	45.3	52.2
10-17	0.0	0.5	77.8	51.5	0.0	0.50	74.9	49.1	30.6

[0161]

	24.0	0.9	42.1	49.4	18.4	0.32	62.8	48.4	33.9
	48.0	2.4	3.0	24.7	49.0	0.43	51.7	47.8	41.2
	72.0	2.8	2.3	14.0	52.7	0.55	45.6	47.7	41.0
10-32	0.0	0.5	77.8	51.5	0.0	0.50	74.9	49.1	30.6
	24.0	1.1	37.6	48.9	20.1	0.50	59.3	48.4	34.1
	48.0	2.3	2.8	23.4	47.3	1.05	35.2	46.9	46.6
	72.0	2.4	2.3	15.7	51.0	1.34	22.6	45.0	52.3

[0162] 第二菌株分离实验中的结果类似。在未添加乙醇的测试中数个菌株利用了比 ZW705 更多的木糖,并且在添加乙醇的测试中利用了更多的葡萄糖。多个菌株在生长的最初 24 小时中利用了更多的葡萄糖,并且根据在最初 24 小时中 OD600nm 的测量还获得了更高的细胞量。这些菌株中,驯化株 7-31 被选用于进一步测试。

[0163] 实例 3

[0164] 水解产物驯化株的性能测试

[0165] 种菌发酵

[0166] 种菌发酵在 1L 的发酵罐内(Sartorius Stedim BIOSTAT)进行。将无菌的空发酵罐填充经过滤消毒的半 YEMaxSM。种菌发酵在 33°C 和 pH5.5 下实施,使用了未过滤的 4N NH₄OH 作为碱来控制于 pH5.5。使用足以在半 YEMaxSM 生长 7.5 小时达到约 2.5 的 OD600 的冷冻原液细胞接种种菌发酵物。将细胞稀释进入所述 1L 发酵罐中以给出约 0.025 的初始 OD600nm。一般情况下,在已经消耗了~120g/L 的葡萄糖的时候和/或 OD600 达到约 10 的时候收集种菌发酵物。对所述种菌发酵进行定期取样以监测生长,并且在约 18.5 小时收集。

[0167] 水解产物发酵

[0168] 水解产物发酵在 1L 的发酵罐内(Sartorius Stedim BIOSTAT)进行。将无菌的空发酵罐填充 450ml 的玉米棒水解产物,其如基本方法中所述而制备。水解产物发酵在 pH5.8 下实施,使用了未过滤的 4N NaOH 作为碱来调节 pH。水解产物发酵在 33°C 下起始。使用来自种菌发酵(参见上文)的 10 体积%(50ml)的种菌接种水解产物发酵物以产生约 1.0 的初始 OD。对所述水解产物发酵进行定期取样以监测反应进程。如基本方法中所述对样品分析葡萄糖、木糖和乙醇。

[0169] 两组水解产物发酵相隔一周来运行,各组发酵均以对照菌株 ZW705 和从驯化中选择的菌株驯化株 7-31(实例 2)而进行。所述发酵的时间过程在图 1 中示出。在所述发酵过程中,驯化株 7-31 利用葡萄糖比 ZW705 更快,并且利用了更多的总木糖以达到更高的乙醇滴度。

[0170] 如上文所述运行水解产物发酵并对其分析以比较对照菌株 ZW705、驯化株 7-31 以及驯化株 5-6(参见实例 2)。结果在图 2 中示出。驯化株 5-6 的性能等同于驯化株 7-31。通过在发酵起始时更快地利用葡萄糖并通过在发酵结束时利用了更多的木糖,两种驯化株均获得了更高的最终乙醇滴度。

[0171] 实例 4

[0172] 对驯化株的比较性 DNA 测序

[0173] 野生型运动发酵单胞菌(ZM4)的全基因组序列已被描述(Jeong-Sun Seo 等人, Nature Biotechnology. 23,2005)。对野生型起始菌株(ZW1 ;ATCC31821)、中间菌

株(ZW658 ;ATCC PTA-7858) 和 ZW705 通过高通量 454 技术(Shendure 与 Ji, Nature Biotech. 26:1135 (2008)) 进行了测序并与公开的野生型序列进行了比较。使用 Illumina 技术(Shendure 与 Ji, Nature Biotech 26:1135 (2008)) 对驯化株 5-6 和驯化株 7-31 进行了测序。通过野生型序列、ZW658 的序列以及对来自所述两种木糖利用操纵子的插入和 GFPR 基因的敲除的蓄意序列变化的经独立测定的插入位点(描述于基本方法中) 而得到了共有序列, 来自驯化株的序列可与所述共有序列进行比较。测序和基因组组装的描述如下。

[0174] 测序

[0175] 使用 Illumina/Solexa 和 Roche-454 测序技术对菌株进行测序。这些大规模平行测序方法提供了短读值的极高通量。对于 Illumina, 其给出了超过 2 亿的 100bp 长度的读值。对于 Roche-454, 输出的 1 百万读值长为 500bp。两种方法均实现了从基因组 DNA 片段的两端进行的测序以产生配对末端读值。

[0176] 来自驯化株 5-6 和 7-31 的数以亿计短读值与参照基因组序列进行比对, 所述参照基因组序列通过野生型、ZW658 和 ZW705 序列而制备。分析所得的比对来收集覆盖率和变异的信息。由于大量的读值与特定区比对, 所述比对是读值针对参照的堆积, 并且给定位点的覆盖深度是覆盖这一位点的读值数量。在某一位点如果共有碱基不同于参照碱基, 则该位点被称为具有单核苷酸多态性(SNP) 变异。

[0177] 在驯化株 5-6 和驯化株 7-31 序列与驯化起始菌株 ZW705 的序列的比较中, 对各驯化株均鉴定出一个单碱基变化, 或 SNP。驯化株 5-6 的 SNP 位于位置 1453116 处, 其处于名为 zmo1330 的基因的开放阅读框内。驯化株 7-31 的 SNP 位于相同的开放阅读框内, 但是该变化处于不同的位点(1453863)。驯化株 5-6 的变化处于编码区(SEQ ID NO:1) 的位置 1097 处, 并且为在该位置从 C 向 G 的变化。这一突变导致了氨基酸 366 的密码子从 ACA 向 AGA 的变化, 引起了在编码的蛋白质中氨基酸 366 从苏氨酸向精氨酸的变化。驯化株 7-31 的变化处于编码区(SEQ ID NO:1) 的位置 350 处, 并且为在该位置从 C 向 T 的变化。这一突变导致了氨基酸 117 的密码子从 TCT 向 TTT 的变化, 引起了在编码的蛋白质中氨基酸 117 从丝氨酸向苯丙氨酸的变化。

[0178] 将 zmo1330 编码的蛋白质用于对转运蛋白分类数据库(transporter classification data base) (Saier Lab Bioinformatics Group; Saier 等人(2009), Nucl. Acids Res., 37:D274-8) 的 BLAST 分析, 其将所述蛋白质鉴定为属于 aaeB 家族。这一家族的蛋白质特征在于具有预测为 5 至 6 次跨膜的疏水性 N 末端、相当长且疏水性较低的中段以及随后在 C 末端类似的 5 至 6 次跨膜。这一结果提示 zmo1330 编码的蛋白质是膜转运蛋白。

[0179] 在公开的运动发酵单胞菌基因组序列(Seo 等人, *ibid*; NCBI 参考号: NC_006526.2) 中, zmo1330 对应于标注为编码“镰刀菌酸抗性蛋白”的 zmo1432。在洋葱假单胞菌中对镰刀菌酸的抗性所需的蛋白组合由 Utsumi 等人鉴定(Agric. Biol. Chem. 55:1919-1918 (1991))。表现为产生镰刀菌酸抗性的操纵子的开放阅读框(ORF) 组中有由 Utsumi 等人(*ibid*) 命名的 fusB。fusB 编码的蛋白质的序列与来自新洋葱伯克霍尔德菌的 Bcenm03_1426 (SEQ ID NO:7) 相同, 其与 zmo1432 编码的蛋白质进行了比对。如图 3 中示出, 所述两种蛋白质序列比对具有三个小间隙并且具有 22% 同一性, 但具有 63% 的相似性(Clustal W 比对)。Bcenm03_1426 的序列与 AaeB (SEQ ID NO:8) 具有大致相同的

同一性和相似性, AaeB 是在大肠杆菌操纵子中编码的两个蛋白质中的较大一个, 所述两个蛋白质据显示为对氨基苯甲酸(pABA) 耐受性所需 (VanDyk 等人, J. Bact. 186:7196-7204 (2004))。Zmo1432 与大肠杆菌 aaeB 具有 17% 的同一性和 55% 的相似性。zmo1432 编码的蛋白质的比对以及与 aaeB 的比对在图 4 中示出。

[0001]

序列表

<110> E. I. du Pont de Nemours and Company

<120> 在生物质水解产物培养基中具有改善的乙醇生产的利用木糖的运动发酵单胞菌

<130> CL5332PCT

<160> 12

<170> PatentIn 版本 3.5

<210> 1

<211> 2031

<212> DNA

<213> 运动发酵单胞菌 (Zymomonas mobilis)

<400> 1

```
atgctgttta atttacggca ggcagctttt gcgcttaatt gttatatagc tgcaatgctc    60
ggcctctatg tctctatgcg gattgggctt gaacgtcctt tctgggcgat gacgactgtc    120
tatatcgta gtcacccct tacggggct atacgttcta aatccttta tcgcgtcatt    180
ggftccctcc ttggcgctac tttgttctg gccgttgtcc caaaatttga taatgcgccc    240
ctttttctct gtatgatttt agggctgtgg gcgtcttttt gtattttat tgttgtctc    300
gaccgttcac cgcgctctta tattttcttt ctfggtatcg ttacggcctc tgttatcggc    360
tttcttagcg tagaaaacce tatcaatgtt ttccatatag cctctcttcg acttcaggaa    420
atctgctttg gggttgtatc ggcaggtttt gttcattccg tgccttttcc ccattecgtc    480
agtaatcttt tatcgcgta attggatcag atttccatg attgcgaacg ctgggctaac    540
catgccttgg cggagatat gacggatatt gatgcgaaag accgccaaaa tctgaccgtt    600
gacctiacca atgtccattt tcttgggaca catatccct atgacacagc gggattgcgg    660
ccaacgcgta tggttttagc cgcgggtcag gatcagatta ttcttttgat gcctgttate    720
gccgctatgg aagatcggac aagagaaatt gacgatgccg gcggtatgtc ggaagagatt    780
accgcctatg tcgaatctgt tcggcaatgg gtcgcagatc cccccgttga tgatcgggca    840
gaagccagtc gtttgattgc gcgcggcaat gccttgggtg aaaaactaaa ggttgaaaat    900
```

[0002]

fggcgcgaac tgctcgagct aaatatgac gggcgactgc gtcattgat cgaagccttg 960
 caatcgacc gttgtgtgt cgaagccgtc agccatccag aagatcatcc acctgccatg 1020
 attgccgtt taagtagcgc ccacgcggt cgctcaatgc atcgcgacta tggtaggcg 1080
 gctttaaccg cttgacatt gttatggc attatggcct catccattt ctggatcatg 1140
 accagtggc caaatgggtc gaccgggtgc ctttggcgg cgatgtcgtt cggctctgcg 1200
 gccagggcag gtgacctgt taaacaaca gggcattatc tctcggggc tgcacggg 1260
 gtgattgtc ctggtttta tgtgttggc atcatgacac aaattcatga attgaattg 1320
 gtcattgtga cgatgttccc tgtgtgttc attatcggct atcttactgc cgatcagaac 1380
 tacctgccga ttgtcgacc ctttatggt gtctcaacc taacaatggc aatccatccc 1440
 gcttattctg ccgatttga gcttattt aacaatggt tggccattat taccggctgc 1500
 ggcatctccc ttgtcgttt caaagtaat gcggttatc gtgccgatgt gatggtcaga 1560
 aggttctgc aatcgggtg gcgtgatct tcggcaacct taaaacgacc gggcgcgccc 1620
 gatatcgtt attggtcaag ccgatgtc gatcgtatc gtcttatggc accgcgctt 1680
 tcagcaacag ggacagatca gaatgtatc cgcgatgat gtattcgcga ttgcgtatc 1740
 ggtatctgta tctgcgctt gcggcagct gctgctcgt ttgatgaaa tgcctcat 1800
 caaatatcca cttggcgca agccattgt ggctattat acgaattgc acggtcaccg 1860
 aatgctgaat ctccgatat cttttgaca gatatcatc gtgttattga tagctttgtg 1920
 gatctacata attcaatga tcgcccgaa gggctgacc cttggtcag cttacgtgc 1980
 aatatgttc ctgatgcacc aggggtcata aaacaacgaa gtccggcatg a 2031

<210> 2

<211> 676

<212> PRT

<213> 运动发酵单胞菌 (*Zymomonas mobilis*)

<400> 2

Met Leu Phe Asn Leu Arg Gln Ala Ala Phe Ala Leu Asn Cys Tyr Ile
 1 5 10 15

[0003]

Ala Ala Met Leu Gly Leu Tyr Val Ser Met Arg Ile Gly Leu Glu Arg
20 25 30

Pro Phe Trp Ala Met Thr Thr Val Tyr Ile Val Ser His Pro Leu Thr
35 40 45

Gly Ala Ile Arg Ser Lys Ser Phe Tyr Arg Val Ile Gly Ser Phe Leu
50 55 60

Gly Ala Thr Phe Val Leu Ala Val Val Pro Lys Phe Asp Asn Ala Pro
65 70 75 80

Leu Phe Leu Cys Met Ile Leu Gly Leu Trp Ala Ser Phe Cys Ile Phe
85 90 95

Ile Val Val Leu Asp Arg Ser Pro Arg Ser Tyr Ile Phe Phe Leu Gly
100 105 110

Ile Val Thr Ala Ser Val Ile Gly Phe Leu Ser Val Glu Asn Pro Ile
115 120 125

Asn Val Phe His Ile Ala Ser Leu Arg Leu Gln Glu Ile Cys Phe Gly
130 135 140

Val Val Ser Ala Gly Phe Val His Ser Val Leu Phe Pro His Ser Val
145 150 155 160

Ser Asn Leu Leu Ser Arg Gln Leu Asp Gln Ile Leu His Asp Cys Glu
165 170 175

Arg Trp Ala Asn His Ala Leu Ala Gly Asp Met Thr Asp Ile Asp Ala
180 185 190

Lys Asp Arg Gln Asn Leu Thr Val Asp Leu Thr Asn Val His Phe Leu
195 200 205

Gly Thr His Ile Pro Tyr Asp Thr Ala Gly Leu Arg Pro Thr Arg Met
210 215 220

[0004]

Ala Leu Ala Ala Val Gln Asp Gln Ile Ile Leu Leu Met Pro Val Ile
225 230 235 240

Ala Ala Met Glu Asp Arg Thr Arg Glu Ile Asp Asp Ala Gly Gly Met
245 250 255

Ser Glu Glu Ile Thr Ala Tyr Val Glu Ser Val Arg Gln Trp Val Ala
260 265 270

Asp Pro Pro Val Asp Asp Ala Ala Glu Ala Ser Arg Leu Ile Ala Arg
275 280 285

Gly Asn Ala Leu Gly Glu Lys Leu Lys Val Glu Asn Trp Arg Glu Leu
290 295 300

Leu Glu Leu Asn Met Ile Gly Arg Leu Arg His Leu Ile Glu Ala Leu
305 310 315 320

Gln Ser Thr Arg Leu Leu Val Glu Ala Val Ser His Pro Glu Asp His
325 330 335

Pro Pro Ala Met Ile Ala Ala Leu Ser Ser Ala His Arg Val Arg Ser
340 345 350

Met His Arg Asp Tyr Gly Met Ala Ala Leu Thr Ala Leu Thr Leu Phe
355 360 365

Met Val Ile Met Ala Ser Ser Ile Phe Trp Ile Met Thr Ser Trp Pro
370 375 380

Asn Gly Ser Thr Gly Cys Leu Leu Ala Ala Met Ser Phe Gly Leu Ser
385 390 395 400

Ala Gln Ala Gly Asp Pro Val Lys Gln Gln Gly His Tyr Leu Leu Gly
405 410 415

[0005]

Ala Val Ile Gly Val Ile Val Ala Gly Phe Tyr Val Phe Ala Ile Met
420 425 430

Thr Gln Ile His Glu Phe Glu Leu Val Met Leu Thr Met Phe Pro Val
435 440 445

Leu Phe Ile Ile Gly Tyr Leu Thr Ala Asp Gln Asn Tyr Leu Pro Ile
450 455 460

Val Arg Pro Phe Met Val Val Phe Asn Leu Thr Met Ala Ile His Pro
465 470 475 480

Ala Tyr Ser Ala Asp Phe Glu Leu Tyr Phe Asn Asn Gly Leu Ala Ile
485 490 495

Ile Thr Gly Cys Gly Ile Ser Leu Val Gly Phe Lys Val Met Arg Val
500 505 510

Ile Gly Ala Asp Val Met Val Arg Arg Leu Leu Gln Ser Gly Trp Arg
515 520 525

Asp Leu Ser Ala Thr Leu Lys Arg Pro Gly Ala Pro Asp Ile Val Asp
530 535 540

Trp Ser Ser Arg Met Leu Asp Arg Ile Gly Leu Met Ala Pro Arg Val
545 550 555 560

Ser Ala Thr Gly Thr Asp Gln Asn Val Ile Arg Asp Asp Gly Ile Arg
565 570 575

Asp Leu Arg Ile Gly Ile Cys Met Leu Arg Leu Arg Gln Leu Ala Ala
580 585 590

Arg Val Asp Glu Asn Val Arg His Gln Ile Ser Thr Leu Ala Gln Ala
595 600 605

Ile Ala Gly Tyr Tyr Asp Glu Leu Ser Arg Ser Pro Asn Ala Glu Ser
610 615 620

[0006]

Ser Asp Ile Ile Leu Thr Asp Ile Asp Arg Val Ile Asp Ser Phe Val
625 630 635 640

Asp Leu His Asn Ser Ile Asp Arg Arg Glu Gly Leu Thr Ala Leu Val
645 650 655

Ser Leu Arg Arg Asn Met Phe Pro Asp Ala Pro Gly Phe Ile Lys Gln
660 665 670

Arg Ser Pro Ala
675

<210> 3
<211> 676
<212> PRT
<213> 人工序列

<220>
<223> 突变的zmo1432编码的蛋白-在位置366处

<400> 3

Met Leu Phe Asn Leu Arg Gln Ala Ala Phe Ala Leu Asn Cys Tyr Ile
1 5 10 15

Ala Ala Met Leu Gly Leu Tyr Val Ser Met Arg Ile Gly Leu Glu Arg
20 25 30

Pro Phe Trp Ala Met Thr Thr Val Tyr Ile Val Ser His Pro Leu Thr
35 40 45

Gly Ala Ile Arg Ser Lys Ser Phe Tyr Arg Val Ile Gly Ser Phe Leu
50 55 60

Gly Ala Thr Phe Val Leu Ala Val Val Pro Lys Phe Asp Asn Ala Pro
65 70 75 80

Leu Phe Leu Cys Met Ile Leu Gly Leu Trp Ala Ser Phe Cys Ile Phe
85 90 95

[0007]

Ile Val Val Leu Asp Arg Ser Pro Arg Ser Tyr Ile Phe Phe Leu Gly
100 105 110

Ile Val Thr Ala Ser Val Ile Gly Phe Leu Ser Val Glu Asn Pro Ile
115 120 125

Asn Val Phe His Ile Ala Ser Leu Arg Leu Gln Glu Ile Cys Phe Gly
130 135 140

Val Val Ser Ala Gly Phe Val His Ser Val Leu Phe Pro His Ser Val
145 150 155 160

Ser Asn Leu Leu Ser Arg Gln Leu Asp Gln Ile Leu His Asp Cys Glu
165 170 175

Arg Trp Ala Asn His Ala Leu Ala Gly Asp Met Thr Asp Ile Asp Ala
180 185 190

Lys Asp Arg Gln Asn Leu Thr Val Asp Leu Thr Asn Val His Phe Leu
195 200 205

Gly Thr His Ile Pro Tyr Asp Thr Ala Gly Leu Arg Pro Thr Arg Met
210 215 220

Ala Leu Ala Ala Val Gln Asp Gln Ile Ile Leu Leu Met Pro Val Ile
225 230 235 240

Ala Ala Met Glu Asp Arg Thr Arg Glu Ile Asp Asp Ala Gly Gly Met
245 250 255

Ser Glu Glu Ile Thr Ala Tyr Val Glu Ser Val Arg Gln Trp Val Ala
260 265 270

Asp Pro Pro Val Asp Asp Ala Ala Glu Ala Ser Arg Leu Ile Ala Arg
275 280 285

[0008]

Gly Asn Ala Leu Gly Glu Lys Leu Lys Val Glu Asn Trp Arg Glu Leu
290 295 300

Leu Glu Leu Asn Met Ile Gly Arg Leu Arg His Leu Ile Glu Ala Leu
305 310 315 320

Gln Ser Thr Arg Leu Leu Val Glu Ala Val Ser His Pro Glu Asp His
325 330 335

Pro Pro Ala Met Ile Ala Ala Leu Ser Ser Ala His Arg Val Arg Ser
340 345 350

Met His Arg Asp Tyr Gly Met Ala Ala Leu Thr Ala Leu Arg Leu Phe
355 360 365

Met Val Ile Met Ala Ser Ser Ile Phe Trp Ile Met Thr Ser Trp Pro
370 375 380

Asn Gly Ser Thr Gly Cys Leu Leu Ala Ala Met Ser Phe Gly Leu Ser
385 390 395 400

Ala Gln Ala Gly Asp Pro Val Lys Gln Gln Gly His Tyr Leu Leu Gly
405 410 415

Ala Val Ile Gly Val Ile Val Ala Gly Phe Tyr Val Phe Ala Ile Met
420 425 430

Thr Gln Ile His Glu Phe Glu Leu Val Met Leu Thr Met Phe Pro Val
435 440 445

Leu Phe Ile Ile Gly Tyr Leu Thr Ala Asp Gln Asn Tyr Leu Pro Ile
450 455 460

Val Arg Pro Phe Met Val Val Phe Asn Leu Thr Met Ala Ile His Pro
465 470 475 480

Ala Tyr Ser Ala Asp Phe Glu Leu Tyr Phe Asn Asn Gly Leu Ala Ile
485 490 495

[0009]

Ile Thr Gly Cys Gly Ile Ser Leu Val Gly Phe Lys Val Met Arg Val
500 505 510

Ile Gly Ala Asp Val Met Val Arg Arg Leu Leu Gln Ser Gly Trp Arg
515 520 525

Asp Leu Ser Ala Thr Leu Lys Arg Pro Gly Ala Pro Asp Ile Val Asp
530 535 540

Trp Ser Ser Arg Met Leu Asp Arg Ile Gly Leu Met Ala Pro Arg Val
545 550 555 560

Ser Ala Thr Gly Thr Asp Gln Asn Val Ile Arg Asp Asp Gly Ile Arg
565 570 575

Asp Leu Arg Ile Gly Ile Cys Met Leu Arg Leu Arg Gln Leu Ala Ala
580 585 590

Arg Val Asp Glu Asn Val Arg His Gln Ile Ser Thr Leu Ala Gln Ala
595 600 605

Ile Ala Gly Tyr Tyr Asp Glu Leu Ser Arg Ser Pro Asn Ala Glu Ser
610 615 620

Ser Asp Ile Ile Leu Thr Asp Ile Asp Arg Val Ile Asp Ser Phe Val
625 630 635 640

Asp Leu His Asn Ser Ile Asp Arg Arg Glu Gly Leu Thr Ala Leu Val
645 650 655

Ser Leu Arg Arg Asn Met Phe Pro Asp Ala Pro Gly Phe Ile Lys Gln
660 665 670

Arg Ser Pro Ala
675

[0010]

<210> 4
<211> 676
<212> PRT
<213> 人工序列

<220>
<223> 具有117突变的zmo1432编码的蛋白

<400> 4

Met Leu Phe Asn Leu Arg Gln Ala Ala Phe Ala Leu Asn Cys Tyr Ile
1 5 10 15

Ala Ala Met Leu Gly Leu Tyr Val Ser Met Arg Ile Gly Leu Glu Arg
20 25 30

Pro Phe Trp Ala Met Thr Thr Val Tyr Ile Val Ser His Pro Leu Thr
35 40 45

Gly Ala Ile Arg Ser Lys Ser Phe Tyr Arg Val Ile Gly Ser Phe Leu
50 55 60

Gly Ala Thr Phe Val Leu Ala Val Val Pro Lys Phe Asp Asn Ala Pro
65 70 75 80

Leu Phe Leu Cys Met Ile Leu Gly Leu Trp Ala Ser Phe Cys Ile Phe
85 90 95

Ile Val Val Leu Asp Arg Ser Pro Arg Ser Tyr Ile Phe Phe Leu Gly
100 105 110

Ile Val Thr Ala Phe Val Ile Gly Phe Leu Ser Val Glu Asn Pro Ile
115 120 125

Asn Val Phe His Ile Ala Ser Leu Arg Leu Gln Glu Ile Cys Phe Gly
130 135 140

Val Val Ser Ala Gly Phe Val His Ser Val Leu Phe Pro His Ser Val
145 150 155 160

[0011]

Ser Asn Leu Leu Ser Arg Gln Leu Asp Gln Ile Leu His Asp Cys Glu
165 170 175

Arg Trp Ala Asn His Ala Leu Ala Gly Asp Met Thr Asp Ile Asp Ala
180 185 190

Lys Asp Arg Gln Asn Leu Thr Val Asp Leu Thr Asn Val His Phe Leu
195 200 205

Gly Thr His Ile Pro Tyr Asp Thr Ala Gly Leu Arg Pro Thr Arg Met
210 215 220

Ala Leu Ala Ala Val Gln Asp Gln Ile Ile Leu Leu Met Pro Val Ile
225 230 235 240

Ala Ala Met Glu Asp Arg Thr Arg Glu Ile Asp Asp Ala Gly Gly Met
245 250 255

Ser Glu Glu Ile Thr Ala Tyr Val Glu Ser Val Arg Gln Trp Val Ala
260 265 270

Asp Pro Pro Val Asp Asp Ala Ala Glu Ala Ser Arg Leu Ile Ala Arg
275 280 285

Gly Asn Ala Leu Gly Glu Lys Leu Lys Val Glu Asn Trp Arg Glu Leu
290 295 300

Leu Glu Leu Asn Met Ile Gly Arg Leu Arg His Leu Ile Glu Ala Leu
305 310 315 320

Gln Ser Thr Arg Leu Leu Val Glu Ala Val Ser His Pro Glu Asp His
325 330 335

Pro Pro Ala Met Ile Ala Ala Leu Ser Ser Ala His Arg Val Arg Ser
340 345 350

Met His Arg Asp Tyr Gly Met Ala Ala Leu Thr Ala Leu Thr Leu Phe
355 360 365

[0012]

Met Val Ile Met Ala Ser Ser Ile Phe Trp Ile Met Thr Ser Trp Pro
370 375 380

Asn Gly Ser Thr Gly Cys Leu Leu Ala Ala Met Ser Phe Gly Leu Ser
385 390 395 400

Ala Gln Ala Gly Asp Pro Val Lys Gln Gln Gly His Tyr Leu Leu Gly
405 410 415

Ala Val Ile Gly Val Ile Val Ala Gly Phe Tyr Val Phe Ala Ile Met
420 425 430

Thr Gln Ile His Glu Phe Glu Leu Val Met Leu Thr Met Phe Pro Val
435 440 445

Leu Phe Ile Ile Gly Tyr Leu Thr Ala Asp Gln Asn Tyr Leu Pro Ile
450 455 460

Val Arg Pro Phe Met Val Val Phe Asn Leu Thr Met Ala Ile His Pro
465 470 475 480

Ala Tyr Ser Ala Asp Phe Glu Leu Tyr Phe Asn Asn Gly Leu Ala Ile
485 490 495

Ile Thr Gly Cys Gly Ile Ser Leu Val Gly Phe Lys Val Met Arg Val
500 505 510

Ile Gly Ala Asp Val Met Val Arg Arg Leu Leu Gln Ser Gly Trp Arg
515 520 525

Asp Leu Ser Ala Thr Leu Lys Arg Pro Gly Ala Pro Asp Ile Val Asp
530 535 540

Trp Ser Ser Arg Met Leu Asp Arg Ile Gly Leu Met Ala Pro Arg Val
545 550 555 560

[0013]

Ser Ala Thr Gly Thr Asp Gln Asn Val Ile Arg Asp Asp Gly Ile Arg
565 570 575

Asp Leu Arg Ile Gly Ile Cys Met Leu Arg Leu Arg Gln Leu Ala Ala
580 585 590

Arg Val Asp Glu Asn Val Arg His Gln Ile Ser Thr Leu Ala Gln Ala
595 600 605

Ile Ala Gly Tyr Tyr Asp Glu Leu Ser Arg Ser Pro Asn Ala Glu Ser
610 615 620

Ser Asp Ile Ile Leu Thr Asp Ile Asp Arg Val Ile Asp Ser Phe Val
625 630 635 640

Asp Leu His Asn Ser Ile Asp Arg Arg Glu Gly Leu Thr Ala Leu Val
645 650 655

Ser Leu Arg Arg Asn Met Phe Pro Asp Ala Pro Gly Phe Ile Lys Gln
660 665 670

Arg Ser Pro Ala
675

<210> 5

<211> 346

<212> PRT

<213> 洋葱伯克霍尔德菌 (Burkholderia cepacia)

<400> 5

Met Ser Ala Met Thr Arg Val Ser Glu Val Ile Ile Gly Ile Val Ser
1 5 10 15

Ala Gly Val Val Ser Ala Leu Val Phe Pro Arg Tyr Thr Gly Glu Gln
20 25 30

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

[0014]

Ala Ser Ala Leu Ser Gly Gln Leu Asp Arg Ala His Ile Glu Thr Ile
50 55 60

His Thr Arg Phe Ala Tyr Val Val Gly Phe Glu Ala Ala Arg Ser Met
65 70 75 80

Ala Val Phe Glu Asp Pro Asp Thr Arg Met Arg Ser Gly Arg Leu Ala
85 90 95

Arg Leu Asn Ser Glu Phe Met Ser Ala Ser Ser Arg Phe His Ala Leu
100 105 110

His Gln Leu Met Asn Arg Leu His Ala Ala Gly Ala Gln Ala Ala Ile
115 120 125

Asp Ala Ile Glu Pro Tyr Phe Arg Glu Ile Ala Pro Leu Leu Thr Arg
130 135 140

Asn Gly Glu Pro Val Arg Thr Ser Ile Asp Ala Ala His Ser Ala Glu
145 150 155 160

Gln Leu Leu Ala Trp Arg Asp Ala Leu Pro Arg Arg Ile Arg Ala Thr
165 170 175

Arg Ala Glu Leu Glu Thr Gln Pro Asp Phe Pro Leu Leu Asp Phe Asp
180 185 190

Thr Ala Ala Glu Leu Leu Tyr Arg Phe Ile Thr Asp Leu Gln Glu Tyr
195 200 205

Ala Ala Thr Tyr Ala Ser Leu Ala Thr Ala Thr His Glu Arg Glu Arg
210 215 220

Trp Ile Glu Arg Tyr Glu Pro Arg Thr Asn Lys Thr Ala Ala Thr Ile
225 230 235 240

Ala Gly Ile Arg Thr Ala Thr Val Ile Leu Ala Leu Gly Trp Phe Trp

[0015]

245 250 255

Ile Glu Thr Ala Trp Pro Ser Gly Val Met Leu Val Leu Asn Ala Ala
260 265 270

Ala Thr Cys Ala Leu Ala Ser Ser Ala Pro Arg Pro Thr Ala Met Ala
275 280 285

Ala Gln Met Gly Met Gly Thr Ala Leu Ala Val Cys Thr Gly Phe Leu
290 295 300

Leu Thr Phe Gly Ile Tyr Pro Arg Ile Asp Gly Phe Val Leu Leu Cys
305 310 315 320

Ala Ala Leu Ala Pro Leu Leu Ala Ile Gly Ile Tyr Met Ser Leu Lys
325 330 335

Pro Lys Leu Ala Gly Tyr Gly Gly Ala Ile
340 345

<210> 6
<211> 234
<212> PRT
<213> 产酸克雷伯菌 (Klebsiella oxytoca)

<400> 6

Met Arg Arg Gln Asp Thr Phe Ser Ile Ala Ile Pro Ser Ser Ala Ala
1 5 10 15

Arg Thr Ala Leu Gly Ala Phe Val Ile Ile Leu Ser Gly Cys Leu Leu
20 25 30

Trp Ile Tyr Ser Ala Trp Pro Asp Gly Gly Thr Ala Val Ser Ile Leu
35 40 45

Gly Val Cys Cys Thr Leu Phe Gly Ser Phe Asp Thr Pro Ala Pro His
50 55 60

[0016]

Ile Val Lys Tyr Ile Ile Gly Ser Val Trp Gly Val Val Ile Ser Leu
65 70 75 80

Ile Tyr Ser Phe Ala Leu Leu Pro Pro Leu Ser Asp Phe Pro Val Leu
 85 90 95

Val Ala Val Leu Ala Pro Val Tyr Leu Leu Ala Gly Ser Leu Gln Ala
 100 105 110

Arg Pro Pro Thr Thr Phe Met Ala Met Gly Ile Thr Leu Thr Leu Pro
 115 120 125

Val Leu Cys Glu Leu Gly Ala Arg Tyr Ser Gly Asp Phe Ala Asp Ala
 130 135 140

Ala Asn Thr Ala Ile Ala Leu Phe Phe Ala Thr Gly Phe Ala Val Ile
145 150 155 160

Gly Met Ser Leu Leu Gln Thr Val Gln Ala Asp Ala Ala Ile Lys Arg
 165 170 175

Leu Leu Lys Leu Cys Gln Arg Asp Ile Arg Arg Ser Val Ser Gly Val
 180 185 190

Phe Lys Gly Asp Glu Thr His Trp Thr Asn Leu Met Ile Asp Arg Gly
 195 200 205

Ala Leu Leu Leu Pro Arg Leu Arg Ala Ala Gly Ser Pro Pro Pro Gly
 210 215 220

Arg Ser Ile Ala Trp Cys Thr Phe Cys Ala
225 230

<210> 7

<211> 734

<212> PRT

<213> 洋葱伯克霍尔德菌 (Burkholderia cepacia)

<400> 7

[0017]

Met Ser Ala Ser Ser Pro Ala Ser Thr Pro Ala Gly Gly Pro Phe Ala
1 5 10 15

Ala Trp Tyr Ala Ala Phe Gly Asp Trp Ala Arg Thr Asp Gly Ala Ala
20 25 30

Trp Leu Tyr Leu Phe Lys Ala Leu Leu Ala Ala Phe Ile Ala Thr Gly
35 40 45

Val Ser Met Arg Leu Asp Leu Pro Ala Pro Lys Thr Ala Met Thr Thr
50 55 60

Val Phe Ile Val Met Gln Pro Gln Ser Gly Ala Val Leu Ala Lys Ser
65 70 75 80

Phe Tyr Arg Val Ala Gly Thr Ile Phe Gly Leu Ile Ala Thr Leu Thr
85 90 95

Phe Val Gly Leu Phe Pro Gln Gln Pro Gln Leu Phe Leu Leu Ala Val
100 105 110

Ala Leu Trp Val Ala Leu Cys Thr Ala Gly Ala Ala Arg Asn Arg Asn
115 120 125

Phe Arg Ser Tyr Gly Phe Leu Leu Ala Gly Tyr Thr Thr Ala Leu Ile
130 135 140

Gly Leu Pro Ala Ser Gln His Pro Asp Gly Ala Phe Met Ser Ala Met
145 150 155 160

Thr Arg Val Ala Glu Ile Met Val Gly Ile Val Ser Ala Gly Val Val
165 170 175

Ser Ala Leu Val Phe Pro Arg Thr Thr Gly Glu Gln Met Arg Thr Thr
180 185 190

Val Arg Lys Arg Phe Gly Ser Phe Val Asp Tyr Val Ala Ala Ala Leu

[0018]

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

[0019]

Ile Arg Thr Ala Thr Val Ile Leu Val Leu Gly Trp Phe Trp Ile Glu
405 410 415

Thr Ala Trp Pro Ser Gly Val Thr Met Thr Leu Thr Ala Ala Ala Thr
420 425 430

Cys Ala Leu Ala Ser Ser Thr Pro Arg Pro Thr Ala Met Ser Ala Gln
435 440 445

Met Gly Met Gly Thr Ala Leu Ala Val Cys Thr Gly Phe Leu Leu Thr
450 455 460

Phe Gly Ile Tyr Pro His Ile Asp Gly Phe Pro Leu Leu Cys Val Ala
465 470 475 480

Leu Ala Pro Leu Leu Ala Ile Gly Ile Phe Met Thr Leu Lys Pro Lys
485 490 495

Leu Ala Gly Tyr Gly Met Gly Tyr Leu Ile Phe Phe Ser Phe Leu Ala
500 505 510

Gly Pro Asp Asn Ile Thr His Tyr Asp Pro Thr Ser Phe Met Asn Asp
515 520 525

Ser Leu Ala Leu Val Leu Ala Met Leu Ala Ser Ala Ile Ala Phe Ala
530 535 540

Val Leu Phe Pro Pro Thr Ala Pro Trp His Lys Lys Arg Leu Phe Ala
545 550 555 560

Asp Leu Arg His Gln Ala Val Ala Ala Gly His Ala Arg Leu Ala Gly
565 570 575

Leu Arg Thr Arg Phe Glu Ser Gly Ala Arg Asp Leu Met Tyr Gln Ala
580 585 590

His Thr Leu Ser Ala Asp Gln Pro Asp Val Gln Arg Asp Ala Leu Arg

[0020]

595 600 605

Trp Met Phe Ala Val Leu Glu Thr Gly Asn Ala Thr Ile Asp Leu Arg
610 615 620

His Glu Leu Ala Thr Leu Pro Ala Asp Pro Arg Tyr Ala Pro Ala Met
625 630 635 640

Pro Trp Arg Arg Ala Ile Asp Thr Met Arg Thr Ala Leu Ser Ala Leu
645 650 655

Phe Thr Arg Pro Ser Ala Ala Arg Phe Asp Ala Thr Leu Ala Ala Thr
660 665 670

Asn Ala Ala Ile Asp Ala Thr Arg Gln Thr Leu Asp Ala Val Glu Pro
675 680 685

Ser Arg Asp Glu Arg His Arg Leu Gln Arg Ile Leu Ser His Leu His
690 695 700

Phe Val Arg Thr Ala Leu Leu Asp Pro Glu Ser Pro Leu Glu Pro Leu
705 710 715 720

Asn Arg Asn Arg Pro Val His Pro Gln Pro Gly Ala Ser Ser
725 730

<210> 8
<211> 655
<212> PRT
<213> 大肠杆菌 (Escherichia coli)

<400> 8

Met Gly Ile Phe Ser Ile Ala Asn Gln His Ile Arg Phe Ala Val Lys
1 5 10 15

Leu Ala Thr Ala Ile Val Leu Ala Leu Phe Val Gly Phe His Phe Gln
20 25 30

[0021]

Leu Glu Thr Pro Arg Trp Ala Val Leu Thr Ala Ala Ile Val Ala Ala
35 40 45

Gly Thr Ala Phe Ala Ala Gly Gly Glu Pro Tyr Ser Gly Ala Ile Arg
50 55 60

Tyr Arg Gly Phe Leu Arg Ile Ile Gly Thr Phe Ile Gly Cys Ile Ala
65 70 75 80

Gly Leu Val Ile Ile Ile Ala Met Ile Arg Ala Pro Leu Leu Met Ile
85 90 95

Leu Val Cys Cys Ile Trp Ala Gly Phe Cys Thr Trp Ile Ser Ser Leu
100 105 110

Val Arg Ile Glu Asn Ser Tyr Ala Trp Gly Leu Ala Gly Tyr Thr Ala
115 120 125

Leu Ile Ile Val Ile Thr Ile Gln Pro Glu Pro Leu Leu Thr Pro Gln
130 135 140

Phe Ala Val Glu Arg Cys Ser Glu Ile Val Ile Gly Ile Val Cys Ala
145 150 155 160

Ile Met Ala Asp Leu Leu Phe Ser Pro Arg Ser Ile Lys Gln Glu Val
165 170 175

Asp Arg Glu Leu Glu Ser Leu Leu Val Ala Gln Tyr Gln Leu Met Gln
180 185 190

Leu Cys Ile Lys His Gly Asp Gly Glu Val Val Asp Lys Ala Trp Gly
195 200 205

Asp Leu Val Arg Arg Thr Thr Ala Leu Gln Gly Met Arg Ser Asn Leu
210 215 220

Asn Met Glu Ser Ser Arg Trp Ala Arg Ala Asn Arg Arg Leu Lys Ala
225 230 235 240

[0022]

Ile Asn Thr Leu Ser Leu Thr Leu Ile Thr Gln Ser Cys Glu Thr Tyr
245 250 255

Leu Ile Gln Asn Thr Arg Pro Glu Leu Ile Thr Asp Thr Phe Arg Glu
260 265 270

Phe Phe Asp Thr Pro Val Glu Thr Ala Gln Asp Val His Lys Gln Leu
275 280 285

Lys Arg Leu Arg Arg Val Ile Ala Trp Thr Gly Glu Arg Glu Thr Pro
290 295 300

Val Thr Ile Tyr Ser Trp Val Ala Ala Ala Thr Arg Tyr Gln Leu Leu
305 310 315 320

Lys Arg Gly Val Ile Ser Asn Thr Lys Ile Asn Ala Thr Glu Glu Glu
325 330 335

Ile Leu Gln Gly Glu Pro Glu Val Lys Val Glu Ser Ala Glu Arg His
340 345 350

His Ala Met Val Asn Phe Trp Arg Thr Thr Leu Ser Cys Ile Leu Gly
355 360 365

Thr Leu Phe Trp Leu Trp Thr Gly Trp Thr Ser Gly Ser Gly Ala Met
370 375 380

Val Met Ile Ala Val Val Thr Ser Leu Ala Met Arg Leu Pro Asn Pro
385 390 395 400

Arg Met Val Ala Ile Asp Phe Ile Tyr Gly Thr Leu Ala Ala Leu Pro
405 410 415

Leu Gly Leu Leu Tyr Phe Leu Val Ile Ile Pro Asn Thr Gln Gln Ser
420 425 430

[0023]

Met Leu Leu Leu Cys Ile Ser Leu Ala Val Leu Gly Phe Phe Leu Gly
435 440 445

Ile Glu Val Gln Lys Arg Arg Leu Gly Ser Met Gly Ala Leu Ala Ser
450 455 460

Thr Ile Asn Ile Ile Val Leu Asp Asn Pro Met Thr Phe His Phe Ser
465 470 475 480

Gln Phe Leu Asp Ser Ala Leu Gly Gln Ile Val Gly Cys Val Leu Ala
485 490 495

Phe Thr Val Ile Leu Leu Val Arg Asp Lys Ser Arg Asp Arg Thr Gly
500 505 510

Arg Val Leu Leu Asn Gln Phe Val Ser Ala Ala Val Ser Ala Met Thr
515 520 525

Thr Asn Val Ala Arg Arg Lys Glu Asn His Leu Pro Ala Leu Tyr Gln
530 535 540

Gln Leu Phe Leu Leu Met Asn Lys Phe Pro Gly Asp Leu Pro Lys Phe
545 550 555 560

Arg Leu Ala Leu Thr Met Ile Ile Ala His Gln Arg Leu Arg Asp Ala
565 570 575

Pro Ile Pro Val Asn Glu Asp Leu Ser Ala Phe His Arg Gln Met Arg
580 585 590

Arg Thr Ala Asp His Val Ile Ser Ala Arg Ser Asp Asp Lys Arg Arg
595 600 605

Arg Tyr Phe Gly Gln Leu Leu Glu Glu Leu Glu Ile Tyr Gln Glu Lys
610 615 620

Leu Arg Ile Trp Gln Ala Pro Pro Gln Val Thr Glu Pro Val Asn Arg
625 630 635 640

[0024]

Leu Ala Gly Met Leu His Lys Tyr Gln His Ala Leu Thr Asp Ser
645 650 655

<210> 9

<211> 347

<212> PRT

<213> 里氏木霉 (Trichoderma reesei)

<400> 9

Met Lys Ala Asn Val Ile Leu Cys Leu Leu Ala Pro Leu Val Ala Ala
1 5 10 15

Leu Pro Thr Glu Thr Ile His Leu Asp Pro Glu Leu Ala Ala Leu Arg
20 25 30

Ala Asn Leu Thr Glu Arg Thr Ala Asp Leu Trp Asp Arg Gln Ala Ser
35 40 45

Gln Ser Ile Asp Gln Leu Ile Lys Arg Lys Gly Lys Leu Tyr Phe Gly
50 55 60

Thr Ala Thr Asp Arg Gly Leu Leu Gln Arg Glu Lys Asn Ala Ala Ile
65 70 75 80

Ile Gln Ala Asp Leu Gly Gln Val Thr Pro Glu Asn Ser Met Lys Trp
85 90 95

Gln Ser Leu Glu Asn Asn Gln Gly Gln Leu Asn Trp Gly Asp Ala Asp
100 105 110

Tyr Leu Val Asn Phe Ala Gln Gln Asn Gly Lys Ser Ile Arg Gly His
115 120 125

Thr Leu Ile Trp His Ser Gln Leu Pro Ala Trp Val Asn Asn Ile Asn
130 135 140

Asn Ala Asp Thr Leu Arg Gln Val Ile Arg Thr His Val Ser Thr Val

[0025]

145	150	155	160
Val Gly Arg Tyr Lys Gly Lys Ile Arg Ala Trp Asp Val Val Asn Glu			
165	170	175	
Ile Phe Asn Glu Asp Gly Thr Leu Arg Ser Ser Val Phe Ser Arg Leu			
180	185	190	
Leu Gly Glu Glu Phe Val Ser Ile Ala Phe Arg Ala Ala Arg Asp Ala			
195	200	205	
Asp Pro Ser Ala Arg Leu Tyr Ile Asn Asp Tyr Asn Leu Asp Arg Ala			
210	215	220	
Asn Tyr Gly Lys Val Asn Gly Leu Lys Thr Tyr Val Ser Lys Trp Ile			
225	230	235	240
Ser Gln Gly Val Pro Ile Asp Gly Ile Gly Ser Gln Ser His Leu Ser			
245	250	255	
Gly Gly Gly Gly Ser Gly Thr Leu Gly Ala Leu Gln Gln Leu Ala Thr			
260	265	270	
Val Pro Val Thr Glu Leu Ala Ile Thr Glu Leu Asp Ile Gln Gly Ala			
275	280	285	
Pro Thr Thr Asp Tyr Thr Gln Val Val Gln Ala Cys Leu Ser Val Ser			
290	295	300	
Lys Cys Val Gly Ile Thr Val Trp Gly Ile Ser Asp Lys Asp Ser Trp			
305	310	315	320
Arg Ala Ser Thr Asn Pro Leu Leu Phe Asp Ala Asn Phe Asn Pro Lys			
325	330	335	
Pro Ala Tyr Asn Ser Ile Val Gly Ile Leu Gln			
340	345		

[0026]

<210> 10

<211> 766

<212> PRT

<213> 轮枝镰孢菌 (*Fusarium verticilloides*)

<400> 10

Met Leu Leu Asn Leu Gln Val Ala Ala Ser Ala Leu Ser Leu Ser Leu
1 5 10 15

Leu Gly Gly Leu Ala Glu Ala Ala Thr Pro Tyr Thr Leu Pro Asp Cys
20 25 30

Thr Lys Gly Pro Leu Ser Lys Asn Gly Ile Cys Asp Thr Ser Leu Ser
35 40 45

Pro Ala Lys Arg Ala Ala Ala Leu Val Ala Ala Leu Thr Pro Glu Glu
50 55 60

Lys Val Gly Asn Leu Val Ser Asn Ala Thr Gly Ala Pro Arg Ile Gly
65 70 75 80

Leu Pro Arg Tyr Asn Trp Trp Asn Glu Ala Leu His Gly Leu Ala Gly
85 90 95

Ser Pro Gly Gly Arg Phe Ala Asp Thr Pro Pro Tyr Asp Ala Ala Thr
100 105 110

Ser Phe Pro Met Pro Leu Leu Met Ala Ala Ala Phe Asp Asp Asp Leu
115 120 125

Ile His Asp Ile Gly Asn Val Val Gly Thr Glu Ala Arg Ala Phe Thr
130 135 140

Asn Gly Gly Trp Arg Gly Val Asp Phe Trp Thr Pro Asn Val Asn Pro
145 150 155 160

Phe Lys Asp Pro Arg Trp Gly Arg Gly Ser Glu Thr Pro Gly Glu Asp
165 170 175

[0027]

Ala Leu His Val Ser Arg Tyr Ala Arg Tyr Ile Val Arg Gly Leu Glu
180 185 190

Gly Asp Lys Glu Gln Arg Arg Ile Val Ala Thr Cys Lys His Tyr Ala
195 200 205

Gly Asn Asp Phe Glu Asp Trp Gly Gly Phe Thr Arg His Asp Phe Asp
210 215 220

Ala Lys Ile Thr Pro Gln Asp Leu Ala Glu Tyr Tyr Val Arg Pro Phe
225 230 235 240

Gln Glu Cys Thr Arg Asp Ala Lys Val Gly Ser Ile Met Cys Ala Tyr
245 250 255

Asn Ala Val Asn Gly Ile Pro Ala Cys Ala Asn Ser Tyr Leu Gln Glu
260 265 270

Thr Ile Leu Arg Gly His Trp Asn Trp Thr Arg Asp Asn Asn Trp Ile
275 280 285

Thr Ser Asp Cys Gly Ala Met Gln Asp Ile Trp Gln Asn His Lys Tyr
290 295 300

Val Lys Thr Asn Ala Glu Gly Ala Gln Val Ala Phe Glu Asn Gly Met
305 310 315 320

Asp Ser Ser Cys Glu Tyr Thr Thr Thr Ser Asp Val Ser Asp Ser Tyr
325 330 335

Lys Gln Gly Leu Leu Thr Glu Lys Leu Met Asp Arg Ser Leu Lys Arg
340 345 350

Leu Phe Glu Gly Leu Val His Thr Gly Phe Phe Asp Gly Ala Lys Ala
355 360 365

[0028]

Gln Trp Asn Ser Leu Ser Phe Ala Asp Val Asn Thr Lys Glu Ala Gln
370 375 380

Asp Leu Ala Leu Arg Ser Ala Val Glu Gly Ala Val Leu Leu Lys Asn
385 390 395 400

Asp Gly Thr Leu Pro Leu Lys Leu Lys Lys Lys Asp Ser Val Ala Met
405 410 415

Ile Gly Phe Trp Ala Asn Asp Thr Ser Lys Leu Gln Gly Gly Tyr Ser
420 425 430

Gly Arg Ala Pro Phe Leu His Ser Pro Leu Tyr Ala Ala Glu Lys Leu
435 440 445

Gly Leu Asp Thr Asn Val Ala Trp Gly Pro Thr Leu Gln Asn Ser Ser
450 455 460

Ser His Asp Asn Trp Thr Thr Asn Ala Val Ala Ala Lys Lys Ser
465 470 475 480

Asp Tyr Ile Leu Tyr Phe Gly Gly Leu Asp Ala Ser Ala Ala Gly Glu
485 490 495

Asp Arg Asp Arg Glu Asn Leu Asp Trp Pro Glu Ser Gln Leu Thr Leu
500 505 510

Leu Gln Lys Leu Ser Ser Leu Gly Lys Pro Leu Val Val Ile Gln Leu
515 520 525

Gly Asp Gln Val Asp Asp Thr Ala Leu Leu Lys Asn Lys Lys Ile Asn
530 535 540

Ser Ile Leu Trp Val Asn Tyr Pro Gly Gln Asp Gly Gly Thr Ala Val
545 550 555 560

Met Asp Leu Leu Thr Gly Arg Lys Ser Pro Ala Gly Arg Leu Pro Val
565 570 575

[0029]

Thr Gln Tyr Pro Ser Lys Tyr Thr Glu Gln Ile Gly Met Thr Asp Met
580 585 590

Asp Leu Arg Pro Thr Lys Ser Leu Pro Gly Arg Thr Tyr Arg Trp Tyr
595 600 605

Ser Thr Pro Val Leu Pro Tyr Gly Phe Gly Leu His Tyr Thr Lys Phe
610 615 620

Gln Ala Lys Phe Lys Ser Asn Lys Leu Thr Phe Asp Ile Gln Lys Leu
625 630 635 640

Leu Lys Gly Cys Ser Ala Gln Tyr Ser Asp Thr Cys Ala Leu Pro Pro
645 650 655

Ile Gln Val Ser Val Lys Asn Thr Gly Arg Ile Thr Ser Asp Phe Val
660 665 670

Ser Leu Val Phe Ile Lys Ser Glu Val Gly Pro Lys Pro Tyr Pro Leu
675 680 685

Lys Thr Leu Ala Ala Tyr Gly Arg Leu His Asp Val Ala Pro Ser Ser
690 695 700

Thr Lys Asp Ile Ser Leu Glu Trp Thr Leu Asp Asn Ile Ala Arg Arg
705 710 715 720

Gly Glu Asn Gly Asp Leu Val Val Tyr Pro Gly Thr Tyr Thr Leu Leu
725 730 735

Leu Asp Glu Pro Thr Gln Ala Lys Ile Gln Val Thr Leu Thr Gly Lys
740 745 750

Lys Ala Ile Leu Asp Lys Trp Pro Gln Asp Pro Lys Ser Ala
755 760 765

[0030]

<210> 11

<211> 350

<212> PRT

<213> 轮枝镰孢菌 (*Fusarium verticillioides*)

<400> 11

Met Gln Leu Lys Phe Leu Ser Ser Ala Leu Leu Leu Ser Leu Thr Gly
1 5 10 15

Asn Cys Ala Ala Gln Asp Thr Asn Asp Ile Pro Pro Leu Ile Thr Asp
20 25 30

Leu Trp Ser Ala Asp Pro Ser Ala His Val Phe Glu Gly Lys Leu Trp
35 40 45

Val Tyr Pro Ser His Asp Ile Glu Ala Asn Val Val Asn Gly Thr Gly
50 55 60

Gly Ala Gln Tyr Ala Met Arg Asp Tyr His Thr Tyr Ser Met Lys Thr
65 70 75 80

Ile Tyr Gly Lys Asp Pro Val Ile Asp His Gly Val Ala Leu Ser Val
85 90 95

Asp Asp Val Pro Trp Ala Lys Gln Gln Met Trp Ala Pro Asp Ala Ala
100 105 110

Tyr Lys Asn Gly Lys Tyr Tyr Leu Tyr Phe Pro Ala Lys Asp Lys Asp
115 120 125

Glu Ile Phe Arg Ile Gly Val Ala Val Ser Asn Lys Pro Ser Gly Pro
130 135 140

Phe Lys Ala Asp Lys Ser Trp Ile Pro Gly Thr Tyr Ser Ile Asp Pro
145 150 155 160

Ala Ser Tyr Val Asp Thr Asn Gly Glu Ala Tyr Leu Ile Trp Gly Gly
165 170 175

[0031]

Ile Trp Gly Gly Gln Leu Gln Ala Trp Gln Asp His Lys Thr Phe Asn
180 185 190

Glu Ser Trp Leu Gly Asp Lys Ala Ala Pro Asn Gly Thr Asn Ala Leu
195 200 205

Ser Pro Gln Ile Ala Lys Leu Ser Lys Asp Met His Lys Ile Thr Glu
210 215 220

Thr Pro Arg Asp Leu Val Ile Leu Ala Pro Glu Thr Gly Lys Pro Leu
225 230 235 240

Gln Ala Glu Asp Asn Lys Arg Arg Phe Phe Glu Gly Pro Trp Val His
245 250 255

Lys Arg Gly Lys Leu Tyr Tyr Leu Met Tyr Ser Thr Gly Asp Thr His
260 265 270

Phe Leu Val Tyr Ala Thr Ser Lys Asn Ile Tyr Gly Pro Tyr Thr Tyr
275 280 285

Gln Gly Lys Ile Leu Asp Pro Val Asp Gly Trp Thr Thr His Gly Ser
290 295 300

Ile Val Glu Tyr Lys Gly Gln Trp Trp Leu Phe Phe Ala Asp Ala His
305 310 315 320

Thr Ser Gly Lys Asp Tyr Leu Arg Gln Val Lys Ala Arg Lys Ile Trp
325 330 335

Tyr Asp Lys Asp Gly Lys Ile Leu Leu Thr Arg Pro Lys Ile
340 345 350

<210> 12

<211> 660

<212> PRT

<213> 轮枝镰孢菌 (Fusarium verticillioides)

[0032]

<400> 12

Met Val Arg Phe Ser Ser Ile Leu Ala Ala Ala Cys Phe Val Ala
1 5 10 15

Val Glu Ser Val Asn Ile Lys Val Asp Ser Lys Gly Gly Asn Ala Thr
20 25 30

Ser Gly His Gln Tyr Gly Phe Leu His Glu Asp Ile Asn Asn Ser Gly
35 40 45

Asp Gly Gly Ile Tyr Ala Glu Leu Ile Arg Asn Arg Ala Phe Gln Tyr
50 55 60

Ser Lys Lys Tyr Pro Val Ser Leu Ser Gly Trp Arg Pro Ile Asn Asp
65 70 75 80

Ala Lys Leu Ser Leu Asn Arg Leu Asp Thr Pro Leu Ser Asp Ala Leu
85 90 95

Pro Val Ser Met Asn Val Lys Pro Gly Lys Gly Lys Ala Lys Glu Ile
100 105 110

Gly Phe Leu Asn Glu Gly Tyr Trp Gly Met Asp Val Lys Lys Gln Lys
115 120 125

Tyr Thr Gly Ser Phe Trp Val Lys Gly Ala Tyr Lys Gly His Phe Thr
130 135 140

Ala Ser Leu Arg Ser Asn Leu Thr Asp Asp Val Phe Gly Ser Val Lys
145 150 155 160

Val Lys Ser Lys Ala Asn Lys Lys Gln Trp Val Glu His Glu Phe Val
165 170 175

Leu Thr Pro Asn Lys Asn Ala Pro Asn Ser Asn Asn Thr Phe Ala Ile
180 185 190

[0033]

Thr Tyr Asp Pro Lys Gly Ala Asp Gly Ala Leu Asp Phe Asn Leu Ile
195 200 205

Ser Leu Phe Pro Pro Thr Tyr Lys Gly Arg Lys Asn Gly Leu Arg Val
210 215 220

Asp Leu Ala Glu Ala Leu Glu Gly Leu His Pro Ser Leu Leu Arg Phe
225 230 235 240

Pro Gly Gly Asn Met Leu Glu Gly Asn Thr Asn Lys Thr Trp Trp Asp
245 250 255

Trp Lys Asp Thr Leu Gly Pro Leu Arg Asn Arg Pro Gly Phe Glu Gly
260 265 270

Val Trp Asn Tyr Gln Gln Thr His Gly Leu Gly Ile Leu Glu Tyr Leu
275 280 285

Gln Trp Ala Glu Asp Met Asn Leu Glu Ile Ile Val Gly Val Tyr Ala
290 295 300

Gly Leu Ser Leu Asp Gly Ser Val Thr Pro Lys Asp Gln Leu Gln Pro
305 310 315 320

Leu Ile Asp Asp Ala Leu Asp Glu Ile Glu Phe Ile Arg Gly Pro Val
325 330 335

Thr Ser Lys Trp Gly Lys Lys Arg Ala Glu Leu Gly His Pro Lys Pro
340 345 350

Phe Arg Leu Ser Tyr Val Glu Val Gly Asn Glu Asp Trp Leu Ala Gly
355 360 365

Tyr Pro Thr Gly Trp Asn Ser Tyr Lys Glu Tyr Arg Phe Pro Met Phe
370 375 380

Leu Glu Ala Ile Lys Lys Ala His Pro Asp Leu Thr Val Ile Ser Ser
385 390 395 400

[0034]

Gly Ala Ser Ile Asp Pro Val Gly Lys Lys Asp Ala Gly Phe Asp Ile
405 410 415

Pro Ala Pro Gly Ile Gly Asp Tyr His Pro Tyr Arg Glu Pro Asp Val
420 425 430

Leu Val Glu Glu Phe Asn Leu Phe Asp Asn Asn Lys Tyr Gly His Ile
435 440 445

Ile Gly Glu Val Ala Ser Thr His Pro Asn Gly Gly Thr Gly Trp Ser
450 455 460

Gly Asn Leu Met Pro Tyr Pro Trp Trp Ile Ser Gly Val Gly Glu Ala
465 470 475 480

Val Ala Leu Cys Gly Tyr Glu Arg Asn Ala Asp Arg Ile Pro Gly Thr
485 490 495

Phe Tyr Ala Pro Ile Leu Lys Asn Glu Asn Arg Trp Gln Trp Ala Ile
500 505 510

Thr Met Ile Gln Phe Ala Ala Asp Ser Ala Met Thr Thr Arg Ser Thr
515 520 525

Ser Trp Tyr Val Trp Ser Leu Phe Ala Gly His Pro Met Thr His Thr
530 535 540

Leu Pro Thr Thr Ala Asp Phe Asp Pro Leu Tyr Tyr Val Ala Gly Lys
545 550 555 560

Asn Glu Asp Lys Gly Thr Leu Ile Trp Lys Gly Ala Ala Tyr Asn Thr
565 570 575

Thr Lys Gly Ala Asp Val Pro Val Ser Leu Ser Phe Lys Gly Val Lys
580 585 590

[0035]

Pro Gly Ala Gln Ala Glu Leu Thr Leu Leu Thr Asn Lys Glu Lys Asp
595 600 605

Pro Phe Ala Phe Asn Asp Pro His Lys Gly Asn Asn Val Val Asp Thr
610 615 620

Lys Lys Thr Val Leu Lys Ala Asp Gly Lys Gly Ala Phe Asn Phe Lys
625 630 635 640

Leu Pro Asn Leu Ser Val Ala Val Leu Glu Thr Leu Lys Lys Gly Lys
645 650 655

Pro Tyr Ser Ser
660

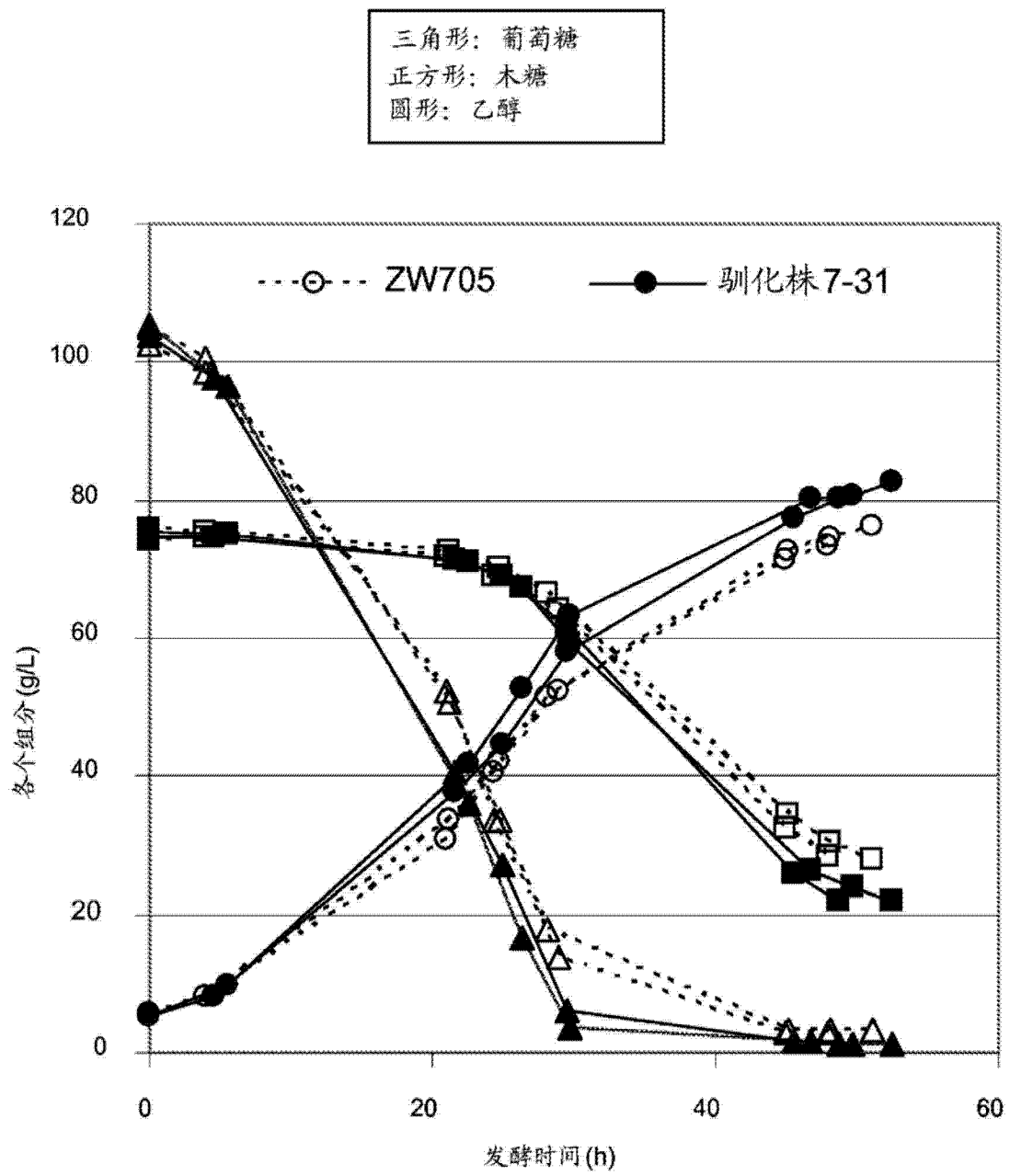


图 1

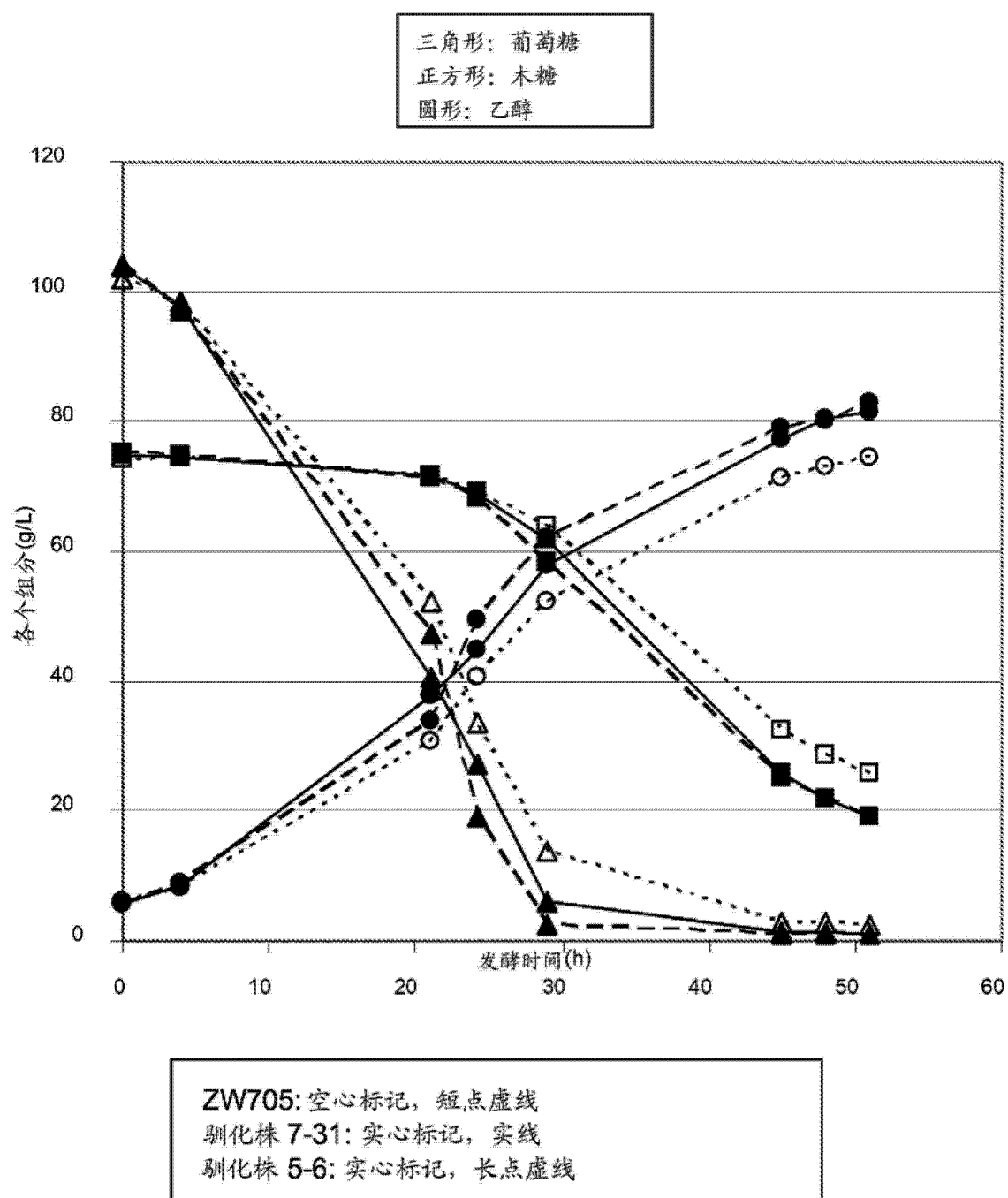


图 2

图 3

```

SEQ ID NO:2   zmo1432   ---MLFNLRQAAFALNCYIAAMLGLYVSMRIGLERPFWAMTTVYIVSH-----PLT
SEQ ID NO:8   EcaaeB    MGIFSIANQHIRFAVKLATAIVLALFVGHFQLETPRWAVLTAATVAAGPAPAAGGEPYS
                  : :  ::  **:  *  *.*:.*::  ** * **:  *  **:  * :

zmo1432       GAIRSKSFYRVIGSFLGATFVLAVPKFDNAPLFLCMLGLWASFCIFIVLDRSPRSYI
EcaaeB        GAIRYRGFLRIIGTFIGCIAGLVIIAMIRAPLLMILVCCIWAGFCTWISSLVRIENSYA
***  . *  *.*:.*:  *.:  . :.*:  :  :.*. *  *  *  *  .**

zmo1432       FFLGIVTASVIGFSLVENPINVFHIASLRLQEICFGVVSAGFVHSVLFPHSVSNLLSRQL
EcaaeB        WLAGY TALIIVITIQPEPLTPQFAVERCSEIVIGIVCAIMADLLFSPRSIKQEV DREL
:  *  ** : * :  :*:  . :*:  *  .** :*:.*  .:  :  *:.*:  :.*:

zmo1432       DQILHDCERWANHALAGDMTDIDAKDRQNLTVDLTNVHFLGTHIPYDTAGLRPTRMALAA
EcaaeB        ESLLVAQYQLMQLCIKHGDGEVVDKAWGDLVRRTTALQGMRSNLMMESSRWARANRLKA
.:.*  :  : .:  . :  *  :*.  *  :  :  :  :  :  :  :  *  *

zmo1432       VQDQIILLMPVIAAMEDRTREIDDAGGMSEITAYVESVRQWVADPPVDDAAEASRLIAR
EcaaeB        IN--TSLTLITQSCETYLIQNTREPLITDTFREFFD-----TPVETAQDVHKQLKR
:  :  : *  :  : *  :  .  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

zmo1432       GNALGEKLVENWRELLELNMIGRLRLHLEALQSTRLLVEAVSHPEDHPPAMIAALSSAH
EcaaeB        -----LRRVIAWTGERETPVT--IYSWVAAATRYQLLKRGVISNTKINATEEEILQGEF
: *  *  *  :  :  :  *  :.*  . *  .  :  :  *  .

zmo1432       RVRSMHRDYGMAALTALTFLMVI MASSIFWIMTSWPNGSTGCLLAAMSFGLSAQAGDPVK
EcaaeB        EVKVESAERHHAMVNEFWRTTLCILGTLFWLWTGWTSGSGAMVIAVVTSLAMRLNPRM
.*:  :  *  .:  :  :  :.*:  *.*. **  . :  :  *  .*:  :  : *

zmo1432       QQGHYLLGAVIGVIVAGFYVFAIMTQIHEFELVMLTMFPVLFIIIGYLTADQNYLPPIVRPF
EcaaeB        VAIDFIYGTLAALPLGLLYFLV IIPNTQQSMLLLCISLAVLGFFLGIEVQKRRLGSMGAL
.:  *:  :  :  :  :  :.*:.*:  :  :  :  :  :  :  :  :  :  :  :  :  :  :

zmo1432       MVVFNLTMAIH PAYSADFELYFNGLAIITGCGISLVGFKVMRVIGADVMVRLLQSGWR
EcaaeB        ASTINII VLDNP-MTFHFSQFLDSALGQIVGCVLAF TVILLVRDKSRDRTGRVLLNQFVS
.:*:  :  *  :  . *  :  :.*. *.* **  :  :  :  :  :  :  :  :  :  :  :

zmo1432       DLSATLKRP GAPDIVDWSSRMLDRIGLMA PRVSATGTDONVIRDDGIRDLRIGICMLRLR
EcaaeB        AAVSAMTTNVARRKENHLPALYQQLFLLMKNKFG-----DLPKFRLALTMI IAH
:  :  .  *  :  .  :  :  :  *  :  :  :  :  :  :  :  :  :  :  :  :  :

zmo1432       QLAARVDENVRHQISTLAQAIAGYYDELSRSPNAESSDIILTDIDRVIDSFVDLHNSIDR
EcaaeB        QRLRDAPIPVNEDLSAFHRQMRTADHVISARSDDKRRRYFGQLLEELEYQEKLRIWQA
*  .  *.:.*:  :  :  :  *.:  :  .  :  :  :  :  :  :  :  :  :  :

zmo1432       REGLTALVSLRRNMFPDAPGFIKQRSPA
EcaaeB        PPQVTEPVHRLAGMLHKYQHALTDS---
: *  *  .*:  .  :.:

```

图 4