(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 104334716 A
(43)申请公布日 2015.02.04

(21)申请号 201280070558.6

(22)申请日 2012.12.21

(30)优先权数据
11306767.2 2011.12.23 EP
61/579,808 2011.12.23 US

(85)PCT国际申请进入国家阶段日
2014.08.22

(86)PCT国际申请的申请数据
PCT/EP2012/076571 2012.12.21

(87)PCT国际申请的公布数据
WO2013/092965 EN 2013.06.27

(83)生物保藏信息
CNCM I-4157 2009.05.07
CNCM I-4155 2009.05.07

(71)申请人 德意诺夫
地址 法国巴黎市

(72)发明人 让-保罗·莱奥内蒂

(74)专利代理机构 北京友联知识产权代理事务所(普通合伙) 11343
代理人 尚志峰 汪海屏

(51)Int.Cl.
*C12N 1/20*(2006.01)
*C12N 9/04*(2006.01)
*C12N 9/10*(2006.01)
*C12N 9/24*(2006.01)
*C12N 9/26*(2006.01)
*C12N 9/42*(2006.01)
*C12N 15/74*(2006.01)
*C12P 7/06*(2006.01)
*C12N 9/00*(2006.01)

权利要求书2页 说明书20页
序列表50页 附图4页

(54)发明名称

具有重构转录单元的细菌及其用途

(57)摘要

本发明涉及重组细菌及其用途,特别是用于生产乙醇。本发明还涉及用于生产这样的细菌的方法,以及适用于这样的生产的核酸构建体。本发明具体涉及具有重构生物质降解单元的细菌。

CN 104334716 A

1. 一种异常球菌属 (Deinococcus) 或相关细菌，其包含插入其基因组中的重构生物质降解转录单元，所述单元包括在单个启动子控制下的至少 2 个基因，所述至少 2 个基因编码在生物质降解或水解为可发酵糖中涉及的独特生物质降解酶。

2. 根据权利要求 1 所述的细菌，其中所述生物质降解酶选自淀粉酶、阿拉伯呋喃糖酶、纤维素酶、木聚糖酶、漆酶、α - 葡糖苷酸酶和酯酶，优选选自淀粉酶、阿拉伯呋喃糖酶和纤维素酶。

3. 根据权利要求 1 或 2 所述的细菌，其中所述单元包含在单个启动子控制下的三个基因，所述三个基因的每一个编码优选选自淀粉酶、阿拉伯呋喃糖酶和纤维素酶的不同生物质降解酶。

4. 根据权利要求 1、2 或 3 所述的细菌，其中所述单元包含，在所述单个启动子控制下，2 个不同淀粉酶基因，或 1 淀粉酶基因和 1 个阿拉伯呋喃糖酶基因，或 2 个淀粉酶基因和 1 个阿拉伯呋喃糖酶基因。

5. 根据权利要求 1 至 4 中任一项所述的细菌，其中所述淀粉酶 α 淀粉酶，优选选自包含 SEQ ID NO:2 或 4（氨基酸序列）的淀粉酶。

6. 根据权利要求 1 至 5 中任一项所述的细菌，其中所述纤维素酶是内切纤维素酶或外切纤维素酶，优选选自包含 SEQ ID NO:8 或 10（氨基酸序列）的纤维素酶。

7. 根据权利要求 1 至 6 中任一项所述的细菌，其中所述阿拉伯呋喃糖酶包含 SEQ ID NO:6（氨基酸序列）。

8. 根据权利要求 1 至 7 中任一项所述的细菌，其中所述单元进一步包含至少一个在第二启动子控制下的额外基因。

9. 根据权利要求 8 所述的细菌，其中所述单元包含 (i) 在单个启动子控制下，2 个淀粉酶基因，或 1 个淀粉酶基因和 1 个阿拉伯呋喃糖酶基因，或 2 个淀粉酶基因和 1 个阿拉伯呋喃糖酶基因 ；和 (ii) 在第二启动子控制下，至少一个纤维素酶基因，优选 1 个内切纤维素酶基因和 1 个外切纤维素酶基因。

10. 根据权利要求 1 至 9 中任一项所述的细菌，其中所述细菌是异常球菌属细菌并且所述生物质降解酶起源于异常球菌属细菌。

11. 根据前述权利要求中任一项所述的细菌，其中所述单个或第二启动子选自异常球菌 pTufA 启动子、pTufB 启动子或 pGroESL 启动子。

12. 根据前述权利要求中任一项所述的细菌，其中所述重构生物质降解转录单元不含有非异常球菌来源的核酸。

13. 根据前述权利要求中任一项所述的细菌，其中所述重构生物质降解转录单元插入在所述细菌的基因组中代替内源性基因，优选代替内源性淀粉酶基因。

14. 根据前述权利要求中任一项所述的细菌，其中所述细菌进一步包含重组醇生产转录单元，优选重组乙醇生产转录单元。

15. 根据权利要求 14 所述的细菌，其中所述重组醇生产转录单元包含至少一个编码醇脱氢酶和 / 或丙酮酸脱羧酶的核酸序列。

16. 根据权利要求 14 或 15 所述的细菌，其中所述醇生产转录单元包含至少一个编码 NADP- 依赖性 ADH 的核酸序列。

17. 根据权利要求 14 至 16 中任一项所述的细菌，其中所述醇生产转录单元插入在所述

细菌的基因组中，优选代替内源性磷酸乙酰转移酶基因。

18. 一种异常球菌或相关细菌，其中所述细菌含有失活的磷酸乙酰转移酶基因。

19. 一种异常球菌或相关细菌，其中所述细菌包含至少一个编码NADP-依赖性ADH的重组核酸序列。

20. 根据前述权利要求中任一项所述的细菌，其是中度嗜热菌 (D.geothermalis)，解纤维素异常球菌 (D.cellulolysiticus)，耐辐射异常球菌 (D.radiodurans)，解蛋白异常球菌 (D.proteolyticus)，抗放射异常球菌 (D.radiopugnans)，嗜放射异常球菌 (D.radiophilus)，D.grandis，印度异常球菌 (D.indicus)，产气异常球菌 (D.frigens)，D.saxicola，马里科帕部落异常球菌 (D.maricopensis)，大理石异常球菌 (D.marmoris)，沙漠异常球菌 (D.deserti)，默氏异常球菌 (D.murrayi)，空气异常球菌 (D.aerius)，网文异常球菌 (D.aerolatus)，嗜气异常球菌 (D.aerophilus)，雅瑟留异常球菌 (D.aetherius)，D.alpinitundrae，高地异常球菌 (D.altitudinis)，D.apachensis，水生异常球菌 (D.aquaticus)，D.aquatilis，D.aquiradiocola，水活异常球菌 (D.aquivivus)，淤泥异常球菌 (D.caeni)，D.claudionis，无花果异常球菌 (D.ficus)，戈壁异常球菌 (D.gobiensis)，D.hohokamensis，D.hopiensis，D.misasensis，纳瓦河部落异常球菌 (D.navajonensis)，D.papagonensis，D.peraridilitoris，D.pimensis，D.piscis，D.radiomollis，玫瑰色异常球菌 (D.roseus)，D.sonorensis，乌鲁木齐异常球菌 (D.wulumuqiensis)，D.xibeiensis，新疆异常球菌 (D.xinjiangensis)，D.yavapaiensis 或云微所异常球菌 (D.yunweiensis)。

21. 一种组合物，其包含根据前述权利要求中任一项所述的细菌和至少一种其他细菌。

22. 根据权利要求 1 至 20 中任一项所述的异常球菌或相关细菌的酶提取物。

23. 一种生物催化剂，其包含根据权利要求 1 至 19 中任一项所述的细菌，或根据权利要求 22 所述的提取物。

24. 一种用于转化生物质的方法，包括将生物质暴露于根据权利要求 1 至 20 中任一项所述的异常球菌或相关细菌，或权利要求 22 所述的提取物，或权利要求 21 所述的组合物。

25. 一种用于生产醇，优选乙醇的方法，包括将糖或生物质暴露于根据权利要求 1 至 20 中任一项所述的异常球菌或相关细菌，或权利要求 22 所述的提取物，或权利要求 21 所述的组合物。

26. 根据权利要求 1 至 20 中任一项所述的异常球菌或相关细菌，或权利要求 22 的提取物，或权利要求 21 所述的组合物用于生产醇，优选乙醇的用途。

# 具有重构转录单元的细菌及其用途

[0001] 本发明涉及重组细菌、其制备及其用途。更具体地，本发明涉及具有重构转录单元的细菌及其用于生物质转化和/或生物燃料（特别是乙醇）的生产的用途。本发明还涉及细菌的核酸构建体、混合培养物和组合物，或其分离的提取物，以及生产生物乙醇的方法。

[0002] 引言

[0003] 生物燃料可以通过大量工艺步骤从生物质材料生产，包括生物质降解和发酵，利用例如化学、物理和/或生物学处理和催化剂。通常，生物燃料生产需要预处理生物质以至少部分地水解半纤维素，除去木质素并且使纤维素非晶化，以使纤维素酶可以接近它们的底物。此外，为了有效地将糖类转化为乙醇，微生物应表现出特定性能如高乙醇收率和生产力，对酸、乙醇和抑制剂的高耐受性，并且在简单生长介质中和在野生工艺条件下具有活性。生物燃料如乙醇从木质纤维素生产将具有丰富、多样且低成本原料的进一步优势。然而，这也需要大量的加工以使可利用的糖类通过通常用来生产乙醇的微生物发酵。

[0004] 没有天然的微生物，包括细菌或酵母菌，满足全部这些要求。

[0005] 自从最近几十年，已经选择和操作微生物以改进其用于生产乙醇的性能。革兰氏阴性菌如大肠杆菌(Escherichia coli)、奥克西托克雷白杆菌(Klebsiella oxytoca)和运动发酵单胞菌(Zymomonas mobilis)，革兰氏阳性菌如解纤维素梭菌(Clostridium cellulolyticum)或干酪乳杆菌(Lactobacillus casei)，并且若干酵母菌菌株已被改造用于从纤维素底物进行乙醇生产。特别地，生物合成基因（如PDC或ADH基因）已被克隆入细菌菌株中。而且，竞争性路径已被改变。

[0006] 这些微生物仍然显示缺陷。尤其是，潜在乙醇生产微生物如运动发酵单胞菌和酿酒酵母(Saccharomyces cerevisiae)通常不能水解复合糖如木质纤维 素。运动发酵单胞菌(Z.mobilis)不是很好适合于生物质转化，因为它仅发酵葡萄糖、果糖和蔗糖。而且，它显示对乙酸的极低耐受性。对于酿酒酵母，最佳温度通常为约37℃，这在其中温度可以显著增加的大型工业培养设施中可能不是最佳的。

[0007] 遗传改变的革兰氏阳性或土芽孢杆菌(Geobacillus)菌株已被提及（参见WO95/27064和WO2006/131734）。然而。从工业前景来看，对于这些菌株还没有满意的代谢物生产。此外，土芽孢杆菌菌株产生孢子，这对于工业用途来说是显著缺点。

[0008] 获得高乙醇收率意味着找到以很少副产物生产乙醇并且代谢所有主要的糖类如葡萄糖、木糖、阿拉伯糖、半乳糖和甘露糖的菌株。而且，合适的工业过程将要求酶和培养条件相对于pH和温度是相容的。

[0009] 异常球菌是由Anderson和合作者在1956年分离的革兰氏阳性菌。这种极端微生物有机体耐受由UV和电离辐射或由交联剂（丝裂霉素C）导致的DNA损害并且耐受失水。WO01/023526显示异常球菌对辐射的不寻常耐受性并且进一步提出它们的改造和在生物修复中的用途。WO2009/063079显示异常球菌属细菌可以抵抗溶剂并且转化生物而生产乙醇。WO2010/130806进一步公开重组异常球菌菌株，其中已插入乙醇生物合成基因。这些重组菌株确实在乙醇生产中表现出改进的性能。

[0010] 本发明现在公开改进的细菌的进一步传代，具有更高且显著的生物质降解性质。

更具体地，本发明公开了具有改进的生物质降解和生物燃料生产性能的改造的异常球菌或相关细菌。这些细菌已由发明人改造而含有重构的功能转录单元和 / 或修饰的代谢路径，导致显著改善的生物学性能。有利地，这些细菌已用由发明人分离和表征的特定基因改造，或者由发明人修饰以改善它们的表达，导致生产非 GMO、改进的细菌。这些细菌与工业培养条件相容，利用生物质材料，并且已被改造而保持整合基因的适当表达用于最佳乙醇生产。这些细菌、其提取物或包含其的组合物，特别适用用于改性生物质并生产生物燃料。

## 发明内容

[0011] 本发明的目的涉及包含插入在其基因组中的重构生物质降解转录单元 (reconstructed biomass degradation transcriptional unit) 的异常球菌或相关细菌。更具体地，这些细菌包含重构生物质降解转录单元，其包含在单个启动子控制下的至少 2 个基因，所述至少 2 个基因编码独特 (distinct) 的生物质降解酶。

[0012] 在一个优选实施方式中，重构生物质降解转录单元不含有非异常球菌遗传物质，其改善它们的活性并且增加监管认可。

[0013] 在本发明的另一个优选实施方式中，所述细菌进一步包重组醇生产转录单元，优选重组乙醇生产转录单元。

[0014] 在另一个优选实施方式中，所述细菌含有改变的乙醇竞争性生物合成路径。更具体地，本发明的特定目的在于包含选自以下的失活内源性基因的异常球菌或相关细菌：磷酸乙酰转移酶基因，丙氨酸脱氢酶基因，葡萄糖脱氢酶基因，磷酸烯醇丙酮酸羧激酶基因，磷酸烯醇丙酮酸羧化酶基因和 / 或苹果酸脱氢酶基因。

[0015] 本发明的另一个目的涉及包含如上所述的细菌和至少一种其他细菌的组合物。

[0016] 本发明的另一个目的涉及包含如上所述的细菌和培养基的组合物。

[0017] 本发明的另一个目的涉及如上所述的细菌的酶提取物。

[0018] 本发明还涉及包含如上所述的细菌或其提取物的生物催化剂。

[0019] 本发明还在于一种用于转化生物质的方法，包括将生物质暴露于如上所述的细菌、或提取物、或组合物。

[0020] 本发明的另一个目的是一种用于生产醇，特别是乙醇的方法，包括将糖或生物质暴露于如上所述的细菌、或提取物、或组合物，并且优选地，收集所产生的醇。

[0021] 本发明还涉及如上所述的细菌用于生产醇，特别是乙醇的用途。

[0022] 本发明还涉及一种用于生产异常球菌或相关细菌，或其原种 ( 又称祖先，ancestor) 的方法，所述方法包括：

[0023] a) 提供亲本异常球菌或相关细菌；

[0024] b) 同时或相继地在单个启动子控制性向所述亲本细菌的基因组中插入至 少 2 个基因，所述至少 2 个基因编码不同的生物质降解酶，和

[0025] c) 选择表达所述至少 2 个基因的 b) 的细菌。

[0026] 本发明还涉及一种核酸，其包含选择 SEQ ID NO:11, 13, 16, 18, 23 和 24 的序列。

[0027] 本发明还涉及一种分离的蛋白质，其包含选自 SEQ ID NO:12 和 14 的氨基酸序列，或其功能性变体或片段。

## 附图说明

[0028] 图 1 :用来构建具有重构生物质降解单元的异常球菌的插入步骤的图示。

[0029] 图 2 :重构细胞的淀粉分解活性。

[0030] 图 3 :重构细胞的淀粉酶活性。在含有 2L 介质的 5L 摇瓶中两个平行地进行培养。在左图 :作为时间的函数的 OD600 测量结果。在右图 :从无细胞上清液 ( 蓝色和红色 ) 和细胞团粒 ( 紫色和绿色 ) 测得的作为时间的函数的 Cer α 活性。在含有 1mM $CaCl_2$ 的 100mM MOPS 缓冲液 (pH 7.0) 中在 45℃确定活性。

[0031] 图 4 :野生 M23-3A 菌株和重构细胞中的 α - 淀粉酶活性的本土化。通过 SDS-PAGE 和通过酶谱技术 (pH 7.0) 分析浓缩 (20x) 的培养上清 (2 = M23-3A,3 = DG_4)、细胞团粒 (4 = M23-3A,5 = DG_4) 和细胞团粒的 Triton X-100 提取物 (6 = M23-3A,7 = DG_4)。在没有沸腾样品下在非还原性条件下进行 SDS-PAGE。

[0032] 图 5 :在全麦 3% 上培养期间的生物质和葡萄糖定量。

[0033] 图 6: 在全麦 3% 上培养期间的乙醇和有机酸定量。

[0034] 图 7: 在淀粉乳 20% 上培养期间的生物质和糖类定量。

[0035] 图 8: 在淀粉乳 20% 上培养期间的乙醇和有机酸定量。

## 具体实施方式

[0036] 本发明涉及异常球菌或相关细菌及其用于转化生物质和 / 或生产生物燃料或其他代谢物的用途。

[0037] 本发明将通过参考一下定义被最好理解 :

[0038] 定义

[0039] 在本发明的内容中,关于酶的术语“来源于细菌”表明该酶已分离自这样的细菌,或者该酶包含分离自或表征自这样的细菌的酶的氨基酸序列的所有或生物学活性部分。术语“来源于异常球菌细菌或相关细菌”进一步包括合成自在异常球菌或相关细菌中鉴别的核酸或氨基酸序列的任何重组、合成和 / 或任选修饰的酶 ( 例如,化学、酶促、物理地修饰 )。

[0040] 异常球菌属细菌是指异常球菌属的任何细菌,如但不限于,中度嗜热菌 (D.geothermalis), 解纤维素异常球菌 (D.cellulolysiticus), 耐辐射异常球菌 (D.radiodurans), 解蛋白异常球菌 (D.proteolyticus), 抗放射异常球菌 (D.radiopugnans), 嗜放射异常球菌 (D.radiophilus),D.grandis, 印度异常球菌 (D.indicus), 产气异常球菌 (D.frigens),D.saxicola, 马里科帕部落异常球菌 (D.maricopensis), 大理石异常球菌 (D.marmoris), 沙漠异常球菌 (D.deserti), 默氏异常球菌 (D.murrayi), 空气异常球菌 (D.aerius), 网文异常球菌 (D.aerolatus), 嗜气异常球菌 (D.aerophilus), 雅瑟留异常球菌 (D.aetherius),D.alpinitundrae, 高地异常球菌 (D.altitudinis),D.apachensis, 水生异常球菌 (D.aquaticus),D.aquatilis,D.aquiradiocola, 水活异常球菌 (D.aquivivus), 淤泥异常球菌 (D.caeni),D.claudionis, 无花果异常球菌 (D.ficus), 戈壁异常球菌 (D.gobiensis),D.hohokamensis,D.hopiensis,D.misasensis, 纳瓦河部落异常球菌 (D.navajonensis),D.papagonensis,D.peraridilitoris,D.pimensis,D.piscis,D.radiomollis, 玫瑰色异常球菌 (D.roseus),D.sonorensis, 乌鲁木齐异常球菌 (D.wulumuqiensis),D.xibeiensis,

新疆异常球菌 (D.xinjiangensis),D.yavapaiensis 或云微所异常球菌 (D.yunweiensis) 细菌。优选的异常球菌属细菌是中度嗜热菌 (D.geothermalis), 解纤维素异常球菌 (D.cellulolysiticus), 沙漠异常球菌 (D.deserti), 默氏异常球菌 (D.murrayi) 和耐辐射异常球菌 (D.radiodurans)。

[0041] 与异常球菌“相关的”细菌是指这样的细菌,其 (i) 含有 16S rDNA,在使用引物 GTTACCCGGAATCACTGGGCGTA(SEQ ID NO:26) 和 GGTATCTACGCATTCCACCGCTA(SEQ ID NO:25) 扩增后,其产生约 158 个碱基对的片段和 / 或 (ii) 抵抗 4mJ/$cm^2$ 的 UV 处理。在特别的实施方式中,异常 球菌相关细菌是具有 16S rDNA 分子的细菌,其在序列上与异常球菌 16S rDNA 序列至少 70%,优选至少 80% 相同。

[0042] “基因”是指编码蛋白质的任何核酸。术语基因涵盖 DNA,如 cDNA 或 gDNA,以及 RNA。基因可以首先通过例如重组、酶促和 / 或化学技术制备,并且随后在宿主细胞或体外系统中复制。基因通常包含编码所需蛋白质的开放读框。基因可以含有另外的序列,如转录终止子、信号肽、IRES、内含子等。优选地,基因不含有内含子。

[0043] “转录单元”在本发明范围内是指一组在一个启动子控制下的至少一种基因。

[0044] 关于序列、核酸或细菌中的单元的术语“重构”或“重组”表明该序列、核酸或单元不是天然存在于细菌中并且已经组装和 / 或插入到所述细菌或其原种中。在重构或重组单元中,序列优选与它们在其中组装或插入的细菌具有相同起源。例如,异常球菌细菌中的重构单元优选基本上包含来源于异常球菌细菌的核酸。

[0045] 关于蛋白质或酶的术语“片段”是指它们的任何片段,其包含所述蛋白质的至少约 10,15,20,25,40,50 或甚至更多个,优选至少 60 好人连续氨基酸。最优选的片段是功能性的,通过它们本身或当融合至或与另一个多肽结合时。蛋白质的片段还指蛋白质的成熟形式 ( 即,在蛋白质的 N- 末端不含有信号肽 )。

[0046] 关于蛋白质或酶的术语“变体”是指任何这样的蛋白质,其与参照蛋白质表现出至少 50% 氨基酸序列同一性,甚至更优选至少 60%,70%,80% 或 90% 的氨基酸序列同一性,并且保留参照蛋白质的活性。序列同一性 ( 同源性 ) 的程度可以使用任何计算机程序和相关参数,包括 BLAST 2.2.2 或 FASTA 版本 3.0t78,在缺省参数下确定。优选的变体与参照序列具有至少 90% 的同一性水平,最优选为至少 92,95, 或 97% 的同一性水平。在一个优选实施方式中,相比于参照蛋白质,变体至多包含 1 至 50,1 至 40,1 至 30,1 至 25,1 至 20,1 至 15,1 至 10, 或 1 至 5 个修饰的 ( 例如,缺失的,取代的或插入的 ) 氨基酸残基。如果蛋白质表现出至少 20%,优选至少 30% 并且更优选至少 50% 的参照蛋白质的酶活性,则它们可以称为变体。

[0047] 根据本发明的术语“生物质”通常是指尤其包含纤维素和 / 或木聚糖的生物质。因此术语生物质包括生物起源的有机材料,包括植物或动物有机材料。生物质可以是未加工的或预处理的。生物质通常是含纤维素或木聚糖的有机植物或动物材料。生物质的实例包括但不限于林业产品,包括不适用于木材或纸张生产的成熟树木、有机废物、农业产品,如草、农作物和动物粪便,以及水产品,如藻类和海草。生物质的实例包括来源于各种类型的植物的木头和植物材料,包括芒 (miscanthus), 麻 (hemp), 甜菜 (sugarbeet), 小麦、玉米, 杨木 (poplar), 柳树 (willow), 高粱 (sorghum), 甘蔗 (sugarcane), 以及各种各样的树物种,范围为桉树至油棕。生物质的来源包括但不限于植物残体、硬木或软木枝干、玉米棒子

(cobs)、秸秆、草、树叶、种子、纸张等。(参见例如 Sun et al.,Bioresource Technology 83(2002)1-11)。术语生物质还涵盖转化的生物质或二次生物质,其基本上含有水解的预处理生物质产品。

[0048] 在本发明的范围内,“改性”生物质包括其任何改性,包括转化、降解、水解、转换和加工。术语“改性”生物质通常涵盖导致可发酵糖产生的生物质的任何改性。改性通常还涵盖生物质的生物聚合物的水解。

[0049] “生物质降解酶”是这样的酶,其参与生物质(或生物质的组分)降解成降解产物。生物质降解酶优选是有助于生物质降解或水解成可发酵糖的酶。这样的酶的实例包括淀粉酶,纤维素酶,阿拉伯呋喃糖酶(如例如,α-L-阿拉伯呋喃糖酶),木聚糖酶,漆酶,α-葡糖苷酸酶和酯酶,如阿魏酸酯酶或乙酰基木聚糖酯酶。

[0050] 术语“醇”或“生物醇”更具体地是指包含 1 至 5 个碳原子,优选选自 1 至 4 个碳原子的直链或支链醇、二醇或三醇。“醇”的具体和优选实例包括选自甲醇、乙醇、丙醇、异丙醇、丙二醇、丁醇、2,3-丁二醇、1,4-丁二醇、异丁醇或丙三醇的 $C_{1-4}$ 醇,更优选乙醇。

[0051] 根据本发明的术语“生物燃料”包括但不限于植物油、生物柴油、生物醇、生物气、合成气和固体生物燃料。

[0052] 生物质降解转录单元

[0053] 已证实当被应激破坏时,异常球菌属细菌具有重新装配其全部或部分基因组的能力。异常球菌属细菌从生物质生产生物能产品的能力披露在 WO2009/063079。本发明现在证实,这些细菌的性能可以通过改造代谢路径而改善。更具体地,本发明提供具有重构的生物质降解转录单元的新型细菌。本发明证实,这样的单元可以在没有改变细胞生长的情况下插入到异常球菌或相关细菌的基因组中。本发明证实,若干基因从单个启动子的表达是可行的并且允许更好地调控细胞中的酶水平。本发明进一步描述在单个插入位点包含具有 5 个基因的功能性重构生物质降解单元的细菌,其表现出改进的性能。本发明进一步描述这样的用异常球菌来源的核酸构建的重新改造的细菌,并且其不含有异源性遗传物质。

[0054] 这些细菌代表有价值的产品和生物催化剂并且特别适合于以改进的能力改性生物质。

[0055] 本发明的生物质降解转录单元是指优选地核酸分子,其包含在单个启动子控制下放置的至少两个不同基因,所述两个不同基因编码两种不同的生物质降解酶。本发明公开了在异常球菌细菌的基因组内的一个单个位置中在独特启动子控制下插入若干基因。这样的构造提供该基因的最佳表达并且不会影响细菌的生长。生物质降解酶可以选自淀粉酶,木聚糖酶,纤维素酶,漆酶,阿拉伯呋喃糖酶,α-葡糖苷酸酶和酯酶。酯酶的优选实例包括但不限于阿魏酸酯酶或乙酰基木聚糖酯酶。阿拉伯呋喃糖酶的具体实例包括,但不限于,α-L-阿拉伯呋喃糖酶。在一个优选实施方式中,生物质降解酶选自淀粉酶,纤维素酶和阿拉伯呋喃糖酶。

[0056] 在一个优选实施方式中,所述单元包含在单个启动子控制下的三个基因。本发明这样的单元的具体实例包含:

[0057] -2 个不同的淀粉酶基因,或

[0058] -1 个淀粉酶基因和 1 个阿拉伯呋喃糖酶基因,或

[0059] -2 个不同的淀粉酶基因和 1 个阿拉伯呋喃糖酶基因。

[0060] 淀粉酶参与多糖，特别是淀粉的水解。淀粉是由通过 1-4 和 1-6 糖苷键结合在一起的大量葡萄糖单元构成的碳水化合物。术语“淀粉酶”包括具有 α- 淀粉酶，β-淀粉酶，葡糖淀粉酶，α-葡糖苷酶或支链淀粉酶（糖基水解酶）活性的多肽。α-淀粉酶具有水解淀粉中的内部 α-1,4-糖苷键以产生更小分子量的麦芽糖糊精的能力。葡糖淀粉酶具有水解由 a-1,4-和 a-1,6-糖苷键连接的葡萄糖聚合物的能力。葡糖淀粉酶具有从葡聚糖释放 β-D-葡萄糖的能力。

[0061] 用于本发明中的淀粉酶基因优选是编码 α 淀粉酶的基因。更优选地，淀粉酶基因编码来源于异常球菌细菌的淀粉酶。在这方面，申请人已从异常球菌属细菌鉴别了具有改进的性能的新型淀粉酶基因，其序列表示在 SEQ ID NO:2 和 4（氨基酸序列）中。

[0062] 在一个优选实施方式中，本发明的转录单元包含编码包含 SEQ ID NO:2 或 4 的氨基酸序列或其片段或变体的 α 淀粉酶的基因。

[0063] 由发明人进行的功能性分析已经证实，SEQ ID NO 2 和 4 的 α 淀粉酶显示出不同的底物特异性，使得通过将这两种淀粉酶结合到本发明的细菌中，增加活性谱。

[0064] 纤维素酶是这样的酶，其催化纤维素或半纤维素（硬木和软木的主要组分）的水解。纤维素酶可以为不同类型，如内切葡聚糖酶，内切纤维素酶，纤维二糖水解酶 (CBH) 或纤维二糖苷酶，或 β-葡糖苷酶（纤维二糖酶；BGL）。用于本发明的纤维素酶基因优选是编码内切纤维素酶或外切纤维素酶的基因。更优选地，纤维素酶基因编码来源于异常球菌细菌的纤维素酶。在这方面，申请人已从异常球菌属细菌鉴别了具有改进的性能的新型纤维素酶基因，其序列表示在 SEQ ID NO:8 和 10（氨基酸序列）中。

[0065] 在一个优选实施方式中，本发明的转录单元包含编码包含 SEQ ID NO:8 或 10 的氨基酸序列、或其片段或变体的纤维素酶的基因。

[0066] 阿拉伯呋喃糖酶基因可以编码阿拉伯呋喃糖酶，更优选是来源于异常球菌细菌的阿拉伯呋喃糖酶的基因。在这方面，申请人已从异常球菌属细菌鉴别了具有改进的性能的新型阿拉伯呋喃糖酶基因，其序列表示在 SEQ ID NO:5（核酸）和 6（氨基酸）中。

[0067] 在一个优选实施方式中，本发明的转录单元包含编码包含 SEQ ID NO:6 的氨基酸序列、其片段或变体的阿拉伯呋喃糖酶的基因。

[0068] 本发明的生物质降解转录单元的实例包括在单个启动子控制下的选自以 下的至少两个基因：

[0069] -编码包含 SEQ ID NO:2 或 4 的氨基酸序列、或其片段或变体的 α 淀粉酶的基因；

[0070] -编码包含 SEQ ID NO:6 的氨基酸序列或其片段或变体的阿拉伯呋喃糖酶的基因；和

[0071] -编码包含 SEQ ID NO:8 或 10 的氨基酸序列或其片段或变体的纤维素酶的基因。

[0072] 本发明的生物质降解转录单元的更具体的实例描述于实施例中，并且包含在单个启动子控制下并且以 5’至 3’的顺序的以下基因：

[0073] -编码包含 SEQ ID NO:2 的氨基酸序列、或其片段或变体的 α 淀粉酶的基因，和编码包含 SEQ ID NO:4 的氨基酸序列、或其片段或变体的 α 淀粉酶的基因；或

[0074] -编码包含 SEQ ID NO:2 的氨基酸序列、或其片段或变体的 α 淀粉酶的基因，和编码包含 SEQ ID NO:6 的氨基酸序列、或其片段或变体的阿拉伯呋喃糖苷酶的基因；或

[0075] -编码包含 SEQ ID NO:2 的氨基酸序列、或其片段或变体的 α 淀粉酶的基因，编码

包含 SEQ ID NO:4 的氨基酸序列、或其片段或变体的 α 淀粉酶的基因，和编码包含 SEQ ID NO:6 的氨基酸序列、或其片段或变体的阿拉伯呋喃糖苷酶的基因。

[0076] 本发明证实，上述基因构造允许整合基因的适当表达用于最佳乙醇生产。

[0077] 所述单元可以包含可替换或另外的基因，其编码例如漆酶，木聚糖酶或酯酶。这样的其他的基因的具体实例，由申请人分离自异常球菌或相关细菌，披露于 PCT/EP2011/069669 和 PCT/EP2011/069670，通过引用结合。

[0078] 转录单元中的基因的表达由单个启动子调控。所述启动子与宿主的同源的（例如来自异常球菌基因的启动子）或异源的（例如，来自不同气源，如不同细菌、噬菌体、合成或杂交启动子等）。优选的启动子是同源的。在这方面，各种启动子已被研究和用于基因表达。合适的异常球菌启动子的实例包括来自翻译延伸因子 Tu 基因 tufA(DR0309) 和 tufB(DR2050) 的 PtufA 和 PtufB 启动子，位于 pI3 中的 resU 基因的启动子，和 groESL 操纵子的启动子区域 PgroESL(Lecointe et al,2004 ;Meima et al,2001)，或这样的启动子的衍生物。

[0079] 本发明证实，转录单元中的基因的合适水平的表达当在选自或来源于 PtufA,PtufB 或 PgroESL 的启动子控制下放置所述基因时获得。

[0080] 转录单元可以进一步包含额外的调控序列，如例如终止子和 / 或增强子。

[0081] 另外，转录单元可以包含在不同启动子控制下的另外的基因。在这方面，有可能在转录单元中插入编码生物质降解酶的额外基因，其由相同或不同启动子调控。

[0082] 在一个特别和优选的实施方式中，本发明的转录单元包含在第二启动子控制下的至少一个额外的生物质降解基因。在另一个优选实施方式中，转录单元包含在启动子控制下的 2 个额外的生物质降解基因。本发明证实，这样的重构生物质降解单元可以在异常球菌菌株的基因组中的一个位置中组装。本发明证实，操纵子中的构造为乙醇生产提供最佳基因表达水平。本发明证实，这样的细菌可存活且稳定。因此这些细菌代表用于生物燃料生产的非常有潜力的生物催化剂。

[0083] 在一个特别的实施方式中，本发明涉及包含重构生物质降解转录单元的异常球菌或相关细菌，其中所述单元包含 ：

[0084] (i) 在单个启动子控制下，2 个淀粉酶基因，或 1 个淀粉酶基因和 1 个阿拉伯呋喃糖酶基因，或 2 个淀粉酶基因和 1 阿拉伯呋喃糖酶基因 ；和

[0085] (ii) 在第二启动子控制下，至少一个纤维素酶基因，优选 1 个内切纤维素酶基因和 1 个外切纤维素酶基因。

[0086] 基因的这些特别组合物和安排允许在细菌中的改进的表达和活性。

[0087] 转录单元可以在体外制备然后插入到细胞中，或者通过基因的相继插入在细胞中进行构建。常规重组 DNA 技术可以用于 DNA 操纵，克隆，和细胞转化。特别地，核酸可以插入到细菌的基因组中，或者作为（自发地）复制分子插入，例如，在质粒、游离基因 (episome)、人工染色体等上。

[0088] 在一个优选实施方式中，转录单元被整合到细菌的基因组中。为此，将构建体克隆入适用于整合到异常球菌细菌的基因组中的一个或多个整合盒 中。这样的整合盒通常包含连接于（或侧悬）允许整合（优选位点特异性整合）的一个或多个序列的核酸。这样的序列可以例如与基因组的靶向区域同源的核酸序列，允许通过横跨该区域进行整合。

[0089] 插入可以靶向基因组的非必需（例如，非编码区域）中。然而，在一个优选实施方式中，插入靶向细菌基因组内的特定基因。在这方面，本发明的特别细菌包含整合到其基因组中的重构生物质降解转录单元，代替编码淀粉酶的全部或部分内源性基因。本发明证实内源性基因由本发明的转录单元替代进一步改进该细菌的性能。在这种情况下，术语“部分基因”是指基因的任何部分，其缺失足以引起细胞中的基因失活。

[0090] 各种技术可以用来将核酸插入到异常球菌或相关细菌中，如例如披露于WO2010/130806。特别地，它们可以通过自然转化（其可以在氯化钙存在下进一步被增强）或电穿孔插入。

[0091] 可替换的克隆位点包括，优选，代替地，选自磷酸乙酰转移酶基因、丙氨酸脱氢酶基因、葡萄糖脱氢酶基因、磷酸烯醇丙酮酸羧激酶基因、磷酸烯醇丙酮酸羧化酶基因或苹果酸脱氢酶基因的基因。

[0092] 在可替换实施方式中，尽管较不是优选的，但是转录单元可以被克隆入合适载体中，其可以在异常球菌中复制。典型的载体含有，除了克隆的插入物外，选择基因（例如，抗生素耐受性，染料等）和有效在异常球菌中复制的起源。这样的载体的实例包括pMD66,pI3,pRAD1和pUE30。pMD66是用于含有pI3的12kb片段的耐辐射异常球菌(D.radiodurans)和大肠杆菌的大型载体(27kb)(Daly et al,1994)。pI3由Masters和Minton(1992)描述。pRAD1是含有用于耐辐射异常球菌的最小复制子的耐辐射异常球菌-大肠杆菌穿梭质粒(Meima and Lidstrom,2000)。pUE30是来源于抗放射异常球菌(D.radiopugnans)的菌株的内源性质粒，其能够在异常球菌中复制（参见US2003/0175977）。

[0093] 本发明还涉及一种用于生产异常球菌或相关细菌，或其原种的方法，所述方法包括：

[0094] a)提供亲本异常球菌或相关细菌；

[0095] b)在所述亲本细菌的基因组中同时或相继地插入在单个启动子控制下的 至少2个基因，所述至少2个基因编码不同的生物质降解酶，和

[0096] c)选择表达所述至少2个基因的b)的细菌。

[0097] 具有插入的核酸的细菌可以根据本身已知的技术选择。基因的表达可以使用（例如，定量地）PCR验证并且这些酶的产生可以通过蛋白质印迹(Western blot)或通过本领域本身已知的酶测定验证并在实施例中举例说明。

[0098] 如在实验部分中公开的，已经生产了含有重构生物质降解转录单元的若干异常球菌属细菌。这些细菌可以培养，是存活的并且稳定地含有插入的单元。稳定性是优选的以使超过95％的细菌在2个生长周期后仍然含有所述单元。本发明的另外的优势在于，生物质降解转录单元可以完全有异常球菌来源的遗传物质形成。作为结果，细菌更稳定，更有效，不是GMO，并且更适于极端培养条件。而且，发明人鉴别和表征的特定酶表现出潜在和补充活性，这赋予细菌显著的活性。

[0099] 醇生产转录单元

[0100] 在一个优选实施方式中，除了生物质降解转录单元之外，本发明的细菌包含重组醇生产转录单元。因此这样的细菌表达优化的基因组合以从生物质生产生物燃料或代谢物。

[0101] 重组醇生产转录单元优选包含编码醇脱氢酶 (adh) 和 / 或丙酮酸脱羧酶 (pdc) 的至少一个基因。

[0102] 丙酮酸脱羧酶 (PDC, EC:4.1.1.1) 催化丙酮酸至乙醛和二氧化碳的单氧化脱羧。醇脱氢酶 (ADH, EC:1.1.1.1) 催化乙醛至乙醇的转化。ADH 和 / 或 PDC 基因插入到异常球菌中已在由申请人提交的在先申请中报道 (WO2010/130806)。为了产生或改进这个代谢路径，编码 PDC 和 / 或 ADH 的基因现在已被克隆并且成功地引入到本发明的具有重构生物质降解单元的异常球菌或相关细菌中。

[0103] 更具体地，已制备编码功能性 PDC 的基因。这样的核酸分子可以包含天然或合成或突变的 PDC 基因的序列的全部或部分，只要改核酸分子编码催化丙酮酸至乙醛和二氧化碳的单氧化脱羧的蛋白质。PDC 存在于植物、真菌和酵母菌中，但在细菌中很稀少。在耐辐射异常球菌基因组中没有发现完全测 序的明显的 PDC。PDC 基因已在各种菌株中鉴别，如在运动发酵单胞菌 (Brau and Sahm, 1986 ; Conway et al, 1987a ; Neale et al, 1987) 中，在巴氏醋酸杆菌 (Acetobacter pasteurianus) (Genbank:AF368435) (Chandra et al, 2001) 中，在胃八叠球菌 (Sarcina ventriculi) (Genbank:AF354297) (Lowe and Zeikus, 1992) 中和在棕榈发酵细菌 (Zymobacter palmae) (Genbank:AF474145) (Raj et al, 2002) 中。

[0104] 在一个优选实施方式中，PDC 核酸包含细菌 PDC 基因的全部或部分的序列。在特定实施方式中，所述核酸包含来自运动发酵单胞菌的 PDC 基因的序列 (ZmPDC, ZMO1360)。ZmPDC 基因序列包含 1707 个碱基对并且表示在 SEQ ID NO:15 中。修饰的基因序列，优化了在异常球菌中的表达，已由发明人合成，其表示在 SEQ ID NO:16。

[0105] ADH 基因已克隆子不同有机体，包括但不限于运动发酵单胞菌 (Ingram et al, 1987)，短乳杆菌 (Lactobacillus brevis) (Liu et al, 2007)，或嗜热脂肪芽孢杆菌 (Geobacillus stearothermophilus) (Genbank:Z25544) (Talarico et al, 2005)。来自异常球菌的 ADH 基因也已被分离并由发明人克隆。这些基因和编码的 ADH 酶的序列披露于 SEQ ID NO:11-14。

[0106] 为了生产编码 ADH 活性的细菌，制备编码功能性 ADH 的基因，其可以包含天然或合成或突变的 ADH 基因的序列的全部或部分，只要核酸分子编码催化乙醛至乙醇的转化的蛋白质。

[0107] 在一个优选实施方式中，ADH 核酸包含细菌 ADH 基因的全部或部分的序列。在一个特定实施方式中，所述核酸包含来自运动发酵单胞菌的 adh 基因的序列 (ZmADH, ZMO1596)。ZmADH II 包含 1152 个碱基对并且其序列描述于 WO2010/130806 ( 还参见 SEQ ID NO:17)。修饰的基因序列，优化在异常球菌中的表达，已由发明人合成，其存在于 SEQ ID NO:18。

[0108] 在另一个优选实施方式中，所述核酸编码包含 SEQ ID NO:12 或 14 的氨基酸序列的全部或部分的 ADH。

[0109] 根据另一个特定实施方式中，醇生产转录单元包含编码 NADP- 依赖性 ADH 或 PDC 的至少一个基因。

[0110] 代谢通量比分析证实，一些异常球菌表现出令人惊讶的主要通过戊糖磷 酸途径消耗葡萄糖的性能，表明 NADPH 的重要部分作为氧化还原电势产生。为了利用此 NADPH 池，可以使用 NADP- 依赖性 ADH 基因。在这方面，乙醇单元可以包含来源于穆尔氏菌属 (Moorella sp) 的 ADH 基因 HUC 22-1 ( 例如，包含 SEQ ID NO:19 的全部或功能性部分 ) 或

来自发酵单胞菌属 (Zymomonas) 的 Tzadh( 包含 SEQ ID NO:20 的全部或功能性部分 )。本发明证实,这样的基因在异常球菌属细菌中的表达是特别有用的,因为这些细菌具有高水平的 NADPH 通量。

[0111] 因此本发明的一个目的还在于异常球菌或相关细菌,其中所述细菌包含编码 NADP- 依赖性 ADH 或 PDC 的至少一个重组核酸序列。

[0112] 本发明的另一个目的是异常球菌或相关细菌,其中所述细菌包含编码 NADP- 依赖性 ADH 的至少一个重组核酸序列和编码 NAD- 依赖性 ADH 的至少一个重组核酸序列。

[0113] 核酸可以插入到细菌的基因组中,或作为 ( 自发地 ) 复制分子例如插入在质粒、游离基因、人工染色体等上,如上所述。在一个优选实施方式中,乙醇生产转录单元插入在所述细菌的基因组中,代替内源性基因,更优选代替磷酸乙酰转移酶基因。

[0114] 在这方面,本发明的目的在于异常球菌或相关细菌,其中所述细菌含有失活的磷酸乙酰转移酶基因 (pta)。乙酰辅酶 A(AcetylCoA),其在糖酵解期间产生,可以用于乙醇生产以及用于作为第二竞争性路径的乙酸形成。通过缺失 pta 基因,乙酸生产应被减少 ( 或消除 ) 而有利于乙醇生产。本发明证实,这种基因可以在没有改变细菌生长的情况下缺失,同时改善细菌生产生物燃料的能力。

[0115] 适当的 PDC 或 ADH 的表达可以利用定量 PCR 验证并且这些酶的产生可以通过蛋白质印迹或通过本领域本身已知的酶测定验证。PDC 活性可以通过分析 $NAD^+$ 的还原测量并且 ADH 活性可以通过分析由于这些酶的活性所致的 $NAD^+$ 的还原或 NADH 的氧化测量 (Conway 等 ,1987a 和 b)。

[0116] 竞争性路径的消除

[0117] 本发明的细菌的性能可以同消除或改变细胞中的竞争性反应后路径而进 一步改进。

[0118] 本发明人现在已生产新型细菌,其中参与竞争性路径的基因已被失活。通过在所需路径中使用更多的其能量,这样的细菌在从生物质生产生物燃料方面表现出进一步改善的效力。更具体地,本发明的特定目的在于这样的异常球菌或相关细菌,其包含选自以下的失活内源性基因 :磷酸乙酰转移酶基因 , 丙氨酸脱氢酶基因 , 葡萄糖脱氢酶基因 , 磷酸烯醇丙酮酸羧激酶基因 , 磷酸烯醇丙酮酸羧化酶基因 , 和 / 或苹果酸脱氢酶基因。

[0119] 在一个特别的实施方式中,靶基因全部或部分地被缺失,并且不编码功能性蛋白质。靶基因可以通过同源性重组、基因替代、靶向突变、或本领域本身已知的任何其他技术在所述细菌或其原种中失活。

[0120] 在一个优选实施方式中,基因通过缺失所述基因的至少一部分而失活,其可以通过异源性核酸 ( 例如,选择标记物 ) 替代。

[0121] 在一个优选实施方式中,本发明的细菌缺少所述基因的一部分,优选其至少 100 个连续的核苷酸,更优选至少 200,300,400 或 500 个。在实施例中,已经生产缺陷型异常球菌菌株,其缺少整个磷酸乙酰转移酶基因。这种菌株已使用包含乙醇生产转录单元 ( 其由于所述基因的蛋白质同源的两个区域侧悬 ) 的特定构建体通过双交换 (double crossing-over) 制备。典型的同源性区域应是足够长以允许杂交和交换,例如,多于 200 个核苷酸,优选多于 300 个核苷酸,典型地为 300 至 7000 个。这样的构建体代表本发明的特别目的。

[0122] 在这方面，本发明还涉及一种用于生产如上所述的异常球菌细菌或其原种的方法，所述方法包括：

[0123] -提供（亲本）异常球菌细菌；

[0124] -处理所述细菌以使选自以下的内源性基因失活：磷酸乙酰转移酶基因，丙氨酸脱氢酶基因，葡萄糖脱氢酶基因，磷酸烯醇丙酮酸羧激酶基因，磷酸烯醇丙酮酸羧化酶基因，和/或苹果酸脱氢酶基因，和

[0125] -选择具有失活的所述基因的细菌。

[0126] 培养物、组合物和用途

[0127] 本发明的细菌可以制备自异常球菌或相关细菌的任何物种。实例在本申请上文列出。它们是选自以下的物种的优选异常球菌属细菌：中度嗜热菌(D.geothermalis)，解纤维素异常球菌(D.cellulolysiticus)，沙漠异常球菌(D.deserti)，默氏异常球菌(D.murrayi)和耐辐射异常球菌(D.radiodurans)。

[0128] 适用于本发明的亲本菌株的示例性实例包括但不限于例如中度嗜热菌(D.geothermalis)DSM 11300(DRH05)，中度嗜热菌(D.geothermalis)DSM 11301(DRH06)，中度嗜热菌(D.geothermalis)DSM 11302(DRH07)，中度嗜热菌(D.geothermalis)HAMBI 2481(DRH37)，中度嗜热菌(D.geothermalis)HAMBI 2480(DRH38)，中度嗜热菌(D.geothermalis)HAMBI 2411(DRH39)，中度嗜热菌(D.geothermalis)HAMBI 2791(DRH41)，中度嗜热菌(D.geothermalis)M36-7D；耐辐射异常球菌(D.radiodurans)R1(ATCC 13939)；默氏异常球菌(D.murrayi)M11-9D(CNCM I-4155)(保藏日期：2009年5月7日，保藏单位：法国国家培养物和微生物保藏中心(CNCM)，保藏地址：巴斯德研究院，25，RUE DU DOCTEUR ROUX,75724巴黎CEDEX 15)；或默氏异常球菌(D.murrayi)M13-1A(CNCM I-4157)(保藏日期：2009年5月7日，保藏单位：法国国家培养物和微生物保藏中心(CNCM)，保藏地址：巴斯德研究院，25，RUE DU DOCTEUR ROUX,75724巴黎CEDEX 15)。

[0129] 异常球菌的其他菌株在公众保藏中心保藏，或者可以由本领域技术人员获得，其可以用来实现本发明。

[0130] 本发明的细菌可以在任何合适培养基和/或装置中培养和/或保持。这样的介质的实例包括复合葡萄糖介质或如在实施例中公开的确定介质，如例如，限定介质蔗糖、限定介质淀粉。

[0131] 本发明的另一个目的涉及一种组合物，其包含如上所述的细菌和至少一种其他细菌。

[0132] 本发明的另一个目的涉及一种组合物，其包含如上所述的细菌和培养基。

[0133] 本发明的另一个目的涉及如上所述的细菌的酶提取物。所述酶提取物至 少含有由所述重构单元编码的生物质降解酶。

[0134] 本发明还涉及生物催化剂，其包含如上所述的细菌或其提取物。

[0135] 本发明还在于一种用于转化生物质的方法，包括将生物质暴露于如上所述的细菌、或提取物或组合物。

[0136] 本发明的另一个目的是一种用于生产生物燃料的方法，包括将糖或生物质暴露于如上所述的细菌、或提取物或组合物。优选地，所述方法包括收集所生产的生物燃料的步

骤。

[0137] 本发明还涉及如上所述的细菌用于生产生物燃料或代谢物的用途。

[0138] 底物可以是任何培养基或各种类型的生物质或来源于其的产物。特别地，生物燃料可以从可再生资源，尤其是植物或动物生物质，或者从市政和工业废物生产。

[0139] 更优选地，本发明的方法用于生产乙醇。

[0140] 本发明的方法可以在转换的反应器中进行。“反应器”是指特别设计用来实施本发明的用于生物质转化的常用发酵罐或任何设备或系统，并因此特别是由以下组成：生物反应器，生物过滤器，旋转生物接触器，和其他气相和 / 或液相生物反应器，尤其是适用于处理生物质或生物质衍生物的那些。根据本发明可以使用的设备可以连续或间歇地使用。

[0141] 在反应器中，为了实现本发明的方法，使用本发明的至少一种细菌，或其细菌提取物，同时所述反应器被排列和供应以使物理化学条件可以设置并在其中保持以使所述细菌在可操作用于所考虑的应用，并且使得任选地，细菌生长是可能的且优选在其中被促进。

[0142] 所述方法在需氧、厌氧下或在微需氧下进行，这取决于底物和细菌。本发明的一个优点涉及本发明的细菌抵抗紧张条件的能力，包括在培养基中存在乙醇。本发明的方法因此优选在约 40℃以上的温度下，特别是在 40-70℃的温度下、在酸性 pH 条件和 / 或在乙醇存在下进行。

[0143] 本方面的其他方面和优点将在以下实施例中披露，这些实施例应被视为举例说明性的并且不限制本申请的范围。

[0144] 实施例

[0145] A. 材料和方法

[0146] 细菌菌株和生长条件：

[0147] 使用大肠杆菌 (E.coli) 菌株 SCS110, JM109 或 DH5α 来繁殖质粒。将它们在 Luria-Bertani(LB) 肉汤 ( 每升：胰蛋白胨 10g, 酵母膏 5g, 氯化钠 10g) 中在 37℃和 200RPM 下培养。固体介质通过加入琼脂 1.5％制备。

[0148] 异常球菌属细菌在 PGY 中在 45℃和 200RPM 下培养。PGY 介质的组成为如下，每升：蛋白胨 (10g), 酵母膏 (5g) 和葡萄糖 (1g)。固体介质为，每升：蛋白胨 (10g), 酵母膏 (5g), 葡萄糖 (1g) 和琼脂 (15g)。

[0149] 当需要时，LB 或 PGY 介质补充适当抗生素：

[0150] - 氯霉素，用于中度嗜热菌 (D.geothermalis) 的终浓度为 3μg/ml，并且用于大肠杆菌的终浓度为 30μg/ml

[0151] - 博来霉素，用于中度嗜热菌 (D.geothermalis) 的终浓度为 6μg/ml for 中度嗜热菌 (D.geothermalis)，并且用于大肠杆菌的终浓度为 10μg/ml

[0152] - 氨苄西林，用于大肠杆菌的终浓度为 100μg/ml

[0153] 转化：

[0154] 大肠杆菌转化使用来自 Stratagene 的商业竞争性细胞 SCS110 或来自 Promega 的 JM10 进行。

[0155] 对于异常球菌细胞，在稳定期的新鲜培养物在 50ml 的 PGY 中稀释 100 倍。生长细胞直至指数期后期 ($OD_{600nm}$ = 0.8)；团粒再悬浮在适当提及的冰冷 2xPGY/10％ v/v 甘油 /30mM $CaCl_2$ 中。对于转化，向 100μl 细胞中加入所需量的质粒 DNA。混合物在冰上温育

30 分钟，转到 42℃持续 90 秒并返回到冰上持续 5 分钟。加入 200μl 的新鲜 2XPGY 介质并将转化株在 2 小时内在 200RPM 和 37℃下振荡。将它们系列稀释并铺到适当选择性的 PGY 板上。

[0156] DNA 操纵：

[0157] 从大肠杆菌细胞的质粒小量制备使用 QIAGEN minipreps DNA 纯化系统进行并且使用来自 Macherey-Nagel 的质粒 DNA 纯化 NucleoBond® Xtra Midi Plus EF 试剂盒进行。这些制备从稳定期的 3-100ml 的大肠杆菌培养物 进行。

[0158] 从异常球菌的基因组 DNA 提取使用来自 Qiagen 的 DNeasy® 血液和组织商业试剂盒进行。这些制备从 5ml 稳定期培养物进行。

[0159] 寡核苷酸通过 Eurogentec 合成。用于 PCR 扩增的聚合酶是来自 Finzyme 的 PHUSION Hi-Fidelity 聚合酶，以及来自 Novagen 的 KOD Xtreme-hot start DNA 聚合酶用于重叠 PCR。PCR 片段使用来自 Promega 的 Wizard SV Gel 和 PCR Clean-Up System 试剂盒清洁。

[0160] 遗传物质通过琼脂糖凝聚电泳分离。DNA 用来自 Eppendorf 的 Biophotometer 定量。

[0161] DNA 插入物通过 Genecust Europe 合成并克隆入适当载体。

[0162] 遗传性插入到异常球菌染色体的方法：

[0163] DNA 片段插入到异常球菌的染色体中使用同源性重组机制进行。插入的盒设计如下：必须被插入的区域的 DNA 序列由与染色体靶标上游或下游的序列同源的 500bp 区域侧悬（参见图 1）。

[0164] 对于前 2 步（编码基因 12_1103 的内生性 α-淀粉酶的缺失和 pTufB-amy1 的插入），插入盒由 pMD66 热敏性穿梭载体携带，转化到异常球菌中并且使用高温暴露（52℃持续 4 天）以允许染色体插入和质粒失去。

[0165] 对于所有其他插入 / 缺失，将插入盒克隆入 puc57 自杀载体中并转化入异常球菌中。已将关注的区域结合到染色体张的转化株在含有适当抗生素的 PGY 介质上选择。

[0166] 正确的插入 / 缺失通过在基因组 DNA 上的 PCR、标记物替代（当可能时，氯霉素至博来霉素或博来霉素至氯霉素）、以及关注区域上的修饰染色体的测序进行检查。

[0167] 图 1 是用来插入生物质降解基因和检查正确插入的第一次插入步骤和方法的示意图。

[0168] 醇脱氢酶活性测试：

[0169] 将在无水乙醇中的 2.5mg/ml 的 4ml 副品红 (pararosaniline)(Sigma) 加入到含有 50mg 亚硫酸钠的 200ml LB 琼脂中 (Conway et al,1987b)。将在 TGY 琼脂板（如果需要，补充有适当抗生素）上生长的 2 天龄的耐辐射异常球菌 (D.radiodurans) 细胞在指示剂板上铺板并在 37℃温育 2 至 3 小时。

[0170] 代谢物生产：

[0171] 这种方法使得能够评价遗传修饰的微生物从生物质或生物质衍生物生产关注的代谢物的能力。

[0172] 该试验在 30℃进行。

[0173] 从在复合介质葡萄糖中制备的预培养物（在稳定期），接种 6ml 的富集介质（以

1% v/v 接种）。

[0174] 测试的富集培养基是复合介质葡萄糖 (Complex Medium Glucose)、确定介质蔗糖 (Defined Medium Sucrose)、确定介质淀粉 (Defined Medium Starch)。

[0175] 复合介质葡萄糖含有：蛋白胨 2g/L，酵母膏 5g/L 和葡萄糖 10g/L 在的渗透水中：溶液通过高压灭菌法灭菌（在 120℃下 15 分钟）。向该溶液中加入以下溶液：MOPS 缓冲液溶液 (10X)pH7[酸 MOPS 400mM，$NH_4Cl$ 200mM，NaOH 100mM，KOH 100mM，$CaCl_2$5μM，$Na_2SO_4$2.76mM，$MgCl_2$5.28mM]；微量营养物 (10000X)[$(NH_4)_6(Mo_7)$24 300mM，$H_3BO_3$4mM，$CoCl_2$0.3mM，$CuSO_4$0.1mM，$MnCl_2$2.5mM，$ZnSO_4$0.1mM]；$FeCl_3$(100X)2mM，在 $C_6H_5Na_3O_7$20mM 中；$K_2HPO_4$1g/L：溶液通过过滤灭菌 (0.2μm)。

[0176] 确定的介质含有：在渗透水中的碳源 10g/L：溶液通过高压灭菌法灭菌（在 120℃ 15 分钟）。向该溶液中加入以下溶液：MOPS 缓冲液溶液 (10X)pH7[酸 MOPS 400mM，$NH_4Cl$ 200mM，NaOH 100mM，KOH 100mM，$CaCl_2$5μM，$Na_2SO_4$2.76mM，$MgCl_2$5.28mM]；微量营养物 (10000X)[$(NH_4)_6(Mo_7)$24 300mM，$H_3BO_3$4mM，$CoCl_2$0.3mM，$CuSO_4$0.1mM，$MnCl_2$2.5mM，$ZnSO_4$0.1mM]；$FeCl_3$(100X)2mM，在 $C_6H_5Na_3O_7$20mM 中；$K_2HPO_4$1g/L：溶液通过过滤灭菌 (0.2μm)。

[0177] 向这些培养基中，除了野生型菌株，在接种前加入氯霉素：3μg/mL 培养基。

[0178] 在需氧和厌氧下进行培养 (Biomerieux，Genbag)。

[0179] 需氧条件下的培养物在搅拌下在 30℃在温育器中留置 7 天。然后将培 养物以 4000rpm 离心 10 分钟。过滤上清 (0.2μm)，倒入其他试管中，并放置在 -80℃。

[0180] 厌氧条件下的培养物在 30℃在温育器中留置 4 周。然后将培养物以 4000rpm 离心 10 分钟。过滤上清 (0.2μm)，倒入其他试管中，并放置在 -80℃。

[0181] 利用气相色谱 FID 分析 (Varian CP-WAX 57CB 25m*0.32mm 柱）来定量醇。有机酸通过液相色谱质谱 (MicrOTOF-QII Bruker) 或毛细管电泳 (5mM 2,6- 吡啶二甲酸 0.5mM 溴化十六烷基三甲铵；5.6pH 调节缓冲液 /61cm 长、50μm 直径毛细管 Agilent) 定量。残余葡萄糖通过偶联折射计的 HPLC(Phenomenex LUNA 3μm $NH_2$100A 150*4.6mm 柱，乙腈 /$H_2O$ 85:15 流动相）。

[0182] 在需氧条件下，在含有 3% - 或 6% - 的全麦中生长 7 天后，对不同菌株监测酸和乙醇生产。

[0183] B. 具有 2- 基因重构生物质降解转录单元的异常球菌菌株的改造。

[0184] 中度嗜热菌 (D.geothermalis)DSM 11300 用作亲本菌株。在单独的实验组中，对以下列出的其他异常球菌或相关细菌应用相同方案：

[0185] DRH05 中度嗜热菌 (D.geothermalis)DSM 11300

[0186] DRH06 中度嗜热菌 (D.geothermalis)DSM 11301

[0187] DRH07 中度嗜热菌 (D.geothermalis)DSM 11302

[0188] DRH37 中度嗜热菌 (D.geothermalis)HAMBI 2481

[0189] DRH38 中度嗜热菌 (D.geothermalis)HAMBI 2480

[0190] DRH39 中度嗜热菌 (D.geothermalis)HAMBI 2411

[0191] DRH41 中度嗜热菌 (D.geothermalis)HAMBI 2791

[0192] M36-7D 中度嗜热菌 (D.geothermalis)

[0193] DR1 耐辐射异常球菌 (D.radiodurans)ATCC 13939

[0194] M11-9D 默氏异常球菌 (D.murrayi)CNCM I-4155(保藏日期:2009 年 5 月 7 日,保藏单位:法国国家培养物和微生物保藏中心 (CNCM),保藏地址:巴斯德研究院,25,RUE DU DOCTEUR ROUX,75724 巴黎 CEDEX 15)

[0195] M13-1A 默氏异常球菌 (D.murrayi)CNCM I-4157(保藏日期:2009 年 5 月 7 日,保藏单位:法国国家培养物和微生物保藏中心 (CNCM),保藏地址:巴斯德研究院,25,RUE DU DOCTEUR ROUX,75724 巴黎 CEDEX 15)

[0196] B1.wt α-淀粉酶 12_1103(amy) 编码基因的缺失

[0197] 将在 pTufA 启动子控制下放置的含有博来霉素抗性基因 (bleo) 的 DNA 片段插入到亲本菌株的染色体中,通过同源性重组代替内生性 α-淀粉酶编码基因 12_1103,得到菌株 DG_06。使用与该内生性 α-淀粉酶编码基因的直接上游和下游的区域同源的 500bp 片段来进行插入。从 ATG 至终止密码子的整个 α-淀粉酶开放读框的完全代替通过 PCR 和测序确认。

[0198] B2. 在 amy 座位处,来自菌株 M23-3A 的 α-淀粉酶编码基因 M23r-3A.18_109(amy1) 的插入

[0199] 将具有氯霉素抗性基因且在 pTufB 启动子下,在操纵子中的带有 α-淀粉酶编码基因 M23r-3A.18_109(amy1-SEQ ID NO:1) 的 DNA 片段插入到菌株 DG_06 的染色体 (amy::ptufA-bleo) 中,代替 pTufA-bleo 并得到菌株 DG_02(amy::pTufB-amy1-cat)。使用 pTufA-bleo 盒直接上游和下游的区域同源的 500bp 片段进行插入,因此代替整个 pTufA-bleo 盒。通过标记物替代 (bleo 至 cat) 和 PCR 来检查插入。amy1 编码序列融合至 N-末端信号肽序列以允许蛋白质的分泌。

[0200] B3.amy1 基因的下游,来自菌株 M23-3A 的 α-淀粉酶编码基因 M23r-3A.305_673amy2) 的插入

[0201] 将含有 α-淀粉酶编码基因 M23r-3A.305_673amy2,SEQ ID NO:3) 和 pTufA-bleo 抗性盒的 DNA 片段插入到菌株 DG_02 的染色体 (amy::pTufB-amy1-cat) 中,在 amy1 基因的直接下游并代替 cat 抗性基因。使用与该 cat 基因直接上游和下游的区域同源的 500bp 片段进行插入,因此缺失整个 cat 基因。所得菌株 DG_04(amy::pTufB-amy1-amy2-pTufA-bleo) 带有在 pTufB 启动子 控制下放置的淀粉酶的操纵子。通过标记物替代 (cat 至 bleo) 和通过 PCR 确认的操纵子结构来检查插入。

[0202] B4. 重构细菌中的淀粉酶表达和淀粉分解活性

[0203] a-板上的淀粉分解测定

[0204] 细胞在富 PGY 介质中生长至指数期后期。然后将 2 至 8μl 的培养物(量根据 OD600nm 标准化)在含有 0.5%淀粉的最小培养基板上成斑并在 37℃温育。

[0205] 在 3 天后,使用革兰氏碘染色可视化淀粉分解晕。

[0206] 结果示于图 2 并显示在重构菌株中存在强淀粉酶活性。

[0207] b-淀粉分解菌株 DG_04 的表征

[0208] 含有 α-淀粉酶 M23r-2A.18_109 和 M23r-2A.305_673 的菌株 DG_04(amy::pTufB-M23r2A.18_109-M23r-2A.305_673-pTufA-bleo) 在含有 0.5%可溶淀粉的确定最小培养基 (Sigma S976) 上生长。培养基富集有 5mM $CaCl_2$ 和 6μg/ml 博来霉素 (Calbiochem 203408)。在含有 2L 介质的 5L 摇瓶中进行培养。每天取样并测量 OD600 和 α-淀粉酶活

性。

[0209] 如图 3 所示，观察到 α 淀粉酶活性，并且大多数活性从细胞表面检测到。

[0210] 测试细胞团粒的 Triton X-100 提取以释放位于细胞表面的 α－淀粉酶活性。将细胞悬浮在含有 2mM $CaCl_2$ 的 100mM MOPS 缓冲液中并加入 Triton X-100 至终浓度为 0.1％。样品良好混合并在室温温育 15 分钟。在温育后，通过离心除去细胞并分析上清的 α－淀粉酶活性。大约 45％的 α－淀粉酶活性可以通过 Triton 提取从细胞表面释放。

[0211] 浓缩的 DG_04 培养上清、细胞和 Triton 提取物通过 SDS-PAGE 和酶谱技术进一步分析（图 4）。淀粉分解异常球菌菌株 M23-3A 用作参照菌株。DG_4 在同工酶凝胶上显示 5 个活性带。在 100kDa 附近检测到两个主要带，而两个其他活性带位于 65kDa 附近。靶标 α－淀粉酶（$2^{nd}$ α－淀粉酶）的理论分子量 为 109kDa 并且一个主要带最可能代表靶标酶。在 50kDa 附近观察到一个弱带，其与 $1^{st}$ α－淀粉酶 (MW ＝ 53kDa) 的分子量对应。亲本菌株含有若干淀粉分解基因。然而，仅从两个插入的 α－淀粉酶基因鉴别了清楚的信号序列。因此，未鉴别的带可能起源于胞外酶的蛋白质分解降解或释放胞内活性的部分细胞溶解。凝胶伪像 (Gel artifacts) 和蛋白质聚集体也可能是由于分析中使用的非还原条件导致的同工酶结果的外延。亲本菌株的最高活性在细胞团粒样品（泳道 5）中检测到并且随着活性可以利用 Triton 提取物部分地从细胞表面释放，这表明 $2^{nd}$ α－淀粉酶的纯化将从 Triton 提取物进行。利用淀粉分解野生型菌株 M23-3A，在培养上清中检测到具有 50kDa 分子量的极弱活性（泳道 2），而在细胞团粒样品中观察到 100kDa 附近的两个活性带（泳道 4）。结果表明，在野生菌株 M23-3A 中，较小的 M23r-2A.18_109 α－淀粉酶（$1^{st}$ α－淀粉酶）位于上清中。相反，从细胞表明发现更大的 M23-3A.305_673 α－淀粉酶（$2^{nd}$ α－淀粉酶）。

[0212] C. 具有 3- 基因重构生物质降解单元的异常球菌菌株的改造

[0213] 含有来自菌株 DRH46 的 α-L- 阿拉伯呋喃糖酶 61_237 编码序列和 pgroESL-cat 氯霉素抗性盒的 DNA 片段插入到菌株 DG_04 的染色体 (amy::pTufB-amy1-amy2-pTufA-bleo) 中，amy2 基因的直接下游，去除 pTufA-bleo 并将 arabinofur 基因放置在与该淀粉酶相同的操纵子中。使用与 pTufA-bleo 盒的直接上游和下游的区域同源的 500bp 片段进行插入，因此代替整个 pTufA-bleo 盒。所得菌株 DG_16(amy::pTufB-amy1-amy2-arabinofur-pgroESL-cat) 检查标记物替代 (bleo 至 cat) 并且其操纵子结构通过 PCR 确认。

[0214] D. 具有 4- 基因重构生物质降解单元的异常球菌菌株的改造

[0215] －在 arabinofur 基因下游，来自菌株 DRH46 的内切纤维素酶编码基因 DRH46.66_2727(endocell) 的插入

[0216] 带有以下构建体 (pTufA-endocell-bleo) 的 DNA 片段，其中来自菌株 DRH046 的内切纤维素酶编码基因 66_2727 在具有 bleo 抗性基因的操纵子中并且在启动子 pTufA 控制下放置，在 arabinofur 基因的下游插入到菌株 DG_16 的染色体，得到菌株 DG-16B(amy::pTufB-amy1-amy2-arabinofur-pTufA-endocell-bleo)。

[0217] E. 具有 5- 基因重构生物质降解单元的异常球菌菌株的改造

[0218] 在编码转录单元 pgroESL-$sacB_{Bs}$-cat 的序列后，含有来自菌株 M1-3H 的外切纤维素酶编码基因 284_1-4(exocell) 的插入盒 [sacB, 来自枯草杆菌 (Bacillus subtilis) 的果聚糖蔗糖酶基因，在具有 cat 抗性基因的操纵子中并且在 pgroESL 启动子的控制下 ] 用于在具有 endocell 基因的操纵子中插入到 DG_16B 菌株的染色体中，产生菌株 DG_16C(amy

::pTufB-amy1-amy2-arabinofur-pTufA-endocell-exocell-pgroESL-sacB-cat)。

[0219] F. 具有醇生产转录单元和生物质降解单元的异常球菌属细菌的改造

[0220] 除了生物质降解单元外，将包含 PDC 和 ADH 基因的乙醇生产单元插入到细胞中。

[0221] F1. 在第一组实验中，含有具有来自运动发酵单胞菌的丙酮酸脱羧酶编码基因 ($pdc_{Zm}$) 和来自运动发酵单胞菌的醇脱氢酶编码基因 ($adh_{Zm}$) 的操纵子的盒（在 pgroESL 启动子控制下放置）和 pTufA-bleo 博来霉素抗性盒插入到菌株 DG_16 的染色体中，代替编码磷酸乙酰转移酶的内生性 pta 基因 (26_1230) 或代替含有磷酸乙酰转移酶基因 (pta, 26_1230) 和乙酸激酶基因 (ack, 26_1224) 的整个乙酸盐生产操纵子。所得菌株分别称为 DG_17(amy::pTufB-amy1-amy2-arabinofur ;pta::pgroESL-$pdc_{Zm}$-$adh_{Zm}$-pTufA-bleo) 和 DG_19。

[0222] 对于 pta 基因缺失，将与 pta 基因直接上游和下游的区域同源的 500bp 片段用作同源性区域。为了避免对邻接 ack 基因（在具有 pta 基因的操纵子中）的极性影响，保留来自 pta 的第一 ATG 密码子。

[0223] 对于代替整个乙酸生产操纵子的插入，将与 ack-pta 操纵子直接上游和下游的区域同源的 500bp 片段用作同源性区域。

[0224] 这些细菌水解淀粉的能力和性能如实验部分中所公开的那样确定。结果概括在下表 1 中：

[0225]

| 菌株 | 淀粉水解（%） |
|---|---|
| DG | 10 |
| DG_04 | 80 |
| DG_16 | 80 |
| DG_17 | 80 |

[0226] 这些结果证实，生物质降解单元的表达确实向细菌赋予改进的水解淀粉的能力。而且，所述细菌仍然存活且重构单元不影响生长性能。

[0227] 如下表 2 所示，虽然 DG 或 DG_16 在测试条件下都不产生乙醇，但是 DG_17 产生大量乙醇，不仅在试管中而且在 1L 反应器中。

[0228] 表 2

[0229]

| | 在小麦 3%上的 EtOH 生产 (%) | 在小麦 6%上的 EtOH 生产 (%) | 在小麦 6%上的 EtOH 生产(%) |
|---|---|---|---|
| **DG** | 0 | 0 | - |
| **DG_16** | 0 | 0 | - |
| **DG_17** | 0.1 | 0.08 | 0.04 |
| | 在试管中进行的实验 | | 在 1L-生物反应器中进行的实验 |

[0230] F2. 实验的另外步骤中，构建另一个乙醇生产单元并插入到细胞中。在这些实验中，在如 F1 中的类似策略后，使用 pdc(SEQ ID NO:16) 和 adh 基因 (SEQ ID NO:18) 的异常球菌 DRH05 密码子优化序列，从 DG-16 改造两个新菌株，得到菌株 DG_18(amy::pTufB-amy1-amy2-arabinofur ;pta::pgroESL-$pdc_{Zm}$*-$adh_{Zm}$*-pTufA-bleo) 和 DG_20(amy::pTufB-amy1-amy2-arabinofur ;pta/ack::pgroESL-$pdc_{Zm}$*-$adh_{Zm}$*-pTufA-bleo)。

[0231] F3.NADP- 依赖性 ADH 编码基因的插入

[0232] 代谢通量比分析证实，在亲本 DG 菌株中，多达 56％的葡萄糖通过戊糖 磷酸途径消耗，表明重要部分的 NADPH 作为氧化还原电势产生。为了将这种 NADPH 池用于 EtOH 生产，将来自穆尔氏菌属 (Moorella sp)HUC 22-1 的 NADP- 依赖性 ADH 编码基因在来自运动发酵单胞菌的 NAD- 依赖性 ADH 基因下游引入到菌株 DG_17 和 DG_18 中。

[0233] 这些新的菌株，称为 DG_21(amy::TufB-amy1-amy2-arabinofur ;pta::pgroESL-$pdc_{Zm}$-$adh_{Zm}$-$adh_{Mo}$-pTufA-kan) 和 DG_22(amy::pTufB-amy1-amy2-arabinofur ;pta/ack::pgroESL-$pdc_{Zm}$*-$adh_{Zm}$*-$adh_{Mo}$-pTufA-kan)，已通过在含卡那霉素板上选择菌株而获得。

[0234] 这些菌株表达两种不同类型的 ADH，一种依赖于 NAD 而一种依赖于 NADP。如下表 3 所示，共表达确实最大化乙醇生产。事实上，相比于 DG-17，作为两种 ADH 基因表达的结果，菌株 DG-21 在 1L 生物反应器中产生更多 5 倍的乙醇。

[0235] 表 3

[0236]

| | 在小麦 3%上的 EtOH 生产 (%) | 在小麦 6%上的 EtOH 生产 (%) | 在小麦 6%上的 EtOH 生产(%) |
|---|---|---|---|
| DG | 0 | 0 | - |
| DG_16 | 0 | 0 | - |
| DG_17 | 0.1 | 0.08 | 0.04 |
| DG_21 | - | - | 0.2 |
| | 在试管中进行的实验 | | 在 1L-生物反应器中进行的实验 |

[0237] F4. 通过在 arabinofur 基因下游，插入来自菌株 DRH46 的内切纤维素酶编码基因 DRH46.66_2727(endocell) 的生物质降解的完成

[0238] 含有以下构建体 (pTufA-endocell-bleo) 的 DNA 片段，其中内切纤维素酶编码基因 DRH46.66_2727 在具有 bleo 抗性基因的操纵子中并且在 pTufA 启动子控制下放置，在 arabinofur 基因下游插入到菌株 DG_21 的染色体 (amy::pTufB-amy1-amy2-arabinofur ;pta::pgroESL-$pdc_{Zm}$-$adh_{Zm}$-$adh_{Mo}$-pTufB-kan) 中，得到菌株 DG_23(amy::pTufB-amy1-amy2-arabinofur-pTufA-endocell-bleo ;pta::pgroESL-$pdc_{Zm}$-$adh_{Zm}$-$adh_{Mo}$-pTufA-kan)。使用与 cat 基因直接上游和下游同源的 500bp 片段进行插入，由此替代整个 cat 基因。检查菌株的

标记物替代 (cat 至 bleo) 并通过 PCR 确认。

[0239] 还将内切纤维素酶编码基因插入到菌株 DG_22 中，得到菌株 DG_24(amy::pTufB-amy1-amy2-arabinofur-pTufA-endocell-bleo ;pta::pgorESL-$pdc_{Zm}$*-$adh_{Zm}$*-$adh_{Mo}$-pTufA-kan)。

[0240] F5. 在具有 endocell 基因的操纵子中，插入来自菌株 M1-3H 的外切纤维素酶编码基因 M1-3H.284_1-4(exocell)

[0241] 在编码转录单元 pgroESL-$sacB_{Bs}$-cat 的序列后，含有的外切纤维素酶编码基因 284_1-4(exocell) 的插入盒 A[sacB, 来自枯草杆菌的果聚糖蔗糖酶基因，在具有 cat 抗性基因的操纵子中并在pgroESL 启动子的控制下] 用于插入到 DG_23 和 DG_24 的染色体中，产生菌株 DG_25(amy::pTufB-amy1-amy2-arabinofur-pTufA-endocell-exocell-pgroESL-sacB-cat,pta::pgroESL-$pdc_{Zm}$-$adh_{Zm}$-$adh_{Mo}$-pTufB-kan) 和 DG_26(amy::pTufB-amy1-amy2-arabinofur-pTufA-endocell-exocell-pgroESL-sacB-cat,pta::pgroESL-$pdc_{Zm}$*-$adh_{Zm}$*-$adh_{Mo}$-pTufA-kan)。使用与 bleo 基因直接上游和下游的区域同源的 500bp 片段进行插入，因此代替整个 bleo 基因。所得菌株通过 PCR 检查器操纵子结构，以及标志物替代 (bleo 至 cat)。

[0242] G. DG_16 中的非 GMO 乙醇生产路径的插入

[0243] 具有来自异常球菌 misasensis(SEQ ID NO:11) 的乙醛脱氢酶编码基因 (DSM22328.88_978) 和来自解纤维素异常球菌 (Deinoccocus cellulosilyticus) DRH46(SEQ ID NO:13) 的醇脱氢酶编码基因 (DRH46.26_648) 的乙醇生产操纵子，在 pgroESL 启动子的控制下放置，并将 pTufA-bleo 博来霉素抗性盒插入到菌株 DG_16 的染色体中代替编码磷酸乙酰转移酶的内生性 pta 基因 (26_1230)。所得菌株称为 DG_29(amy::pTufB-amy1-amy2-arabinofur ;pta::pgroESL-$acdh_{DSM22328}$-$adh_{drh46}$-pTufA-bleo)。

[0244] H. 从抗生素抗性基因清洁最终修饰的菌株

[0245] 为了缺失 pgroESL-$sacB_{Bs}$-cat 转录单元，将带有侧悬 pgroESL-$sacB_{Bs}$-cat 的区域的序列的 DNA 片段插入到菌株 DG_25 和 DG_26 的染色体中。使用相同策略来从“乙醇”操纵子去除卡那霉素抗性基因。来自枯草杆菌的 sacB 基因用作计数 - 选择标记物，因为它编码对蔗糖的毒性。最终操纵子结构通过 PCR 和测序确认。

[0246] I. 通过缺失竞争性代谢路径优化乙醇生产

[0247] 使用以下检测 / 清洁盒系列地进行编码竞争性活性的 4 个靶基因的缺失。盒骨架含有在 pgroESL 启动子控制下的 $sacB_{Bs}$-cat 转录单元 ( 来自枯草杆菌的果聚糖蔗糖酶基因，在具有氯霉素 cat 抗性基因的操纵子中 )。使用博来霉素选择，该盒插入到 DG 菌株的染色体中代替靶基因 ( 一次一个 )。对于每次插入，使用第二 DNA 片段，其中侧悬靶标座位的 2 个染色体区域被组装，以获得清洁缺失，从染色体去除 pgroESL-sacB-cat 单元，并在含蔗糖介质中计数选择。

[0248] 该方法用于以下遗传型缺失 :

[0249] - 编码基因 31_22 的丙氨酸脱氢酶

[0250] - 编码基因 52_885 的葡萄糖脱氢酶 1(GDH1)

[0251] - 编码基因 22_1025 的葡萄糖脱氢酶 2(GDH2)

[0252] - 编码基因 52_1088 的磷酸烯醇丙酮酸羧激酶 (PEPCK)

[0253] - 编码基因 12_1095 的磷酸烯醇丙酮酸羧化酶 (PEPC)

[0254] -编码基因 34_424 的苹果酸脱氢酶

[0255] J. 发酵罐中的优化生产

[0256] 菌株 DG_21 在 Infors HT multifors 2-倍 1L 发酵罐系统中培养。用于两个容器的培养条件为 45℃温度、400rpm 搅拌速度、气流 65mL/min，在 7 的 pH 通过 1M NaOH 和 1,5M $H_3PO_4$ 维持。

[0257] 接种物在具有 10g/L 葡萄糖电脑复合介质上在摇瓶中培养。添加博来霉 素抗生素，终浓度为 6μg/mL。用于接种物培养的条件为 45℃和 150rpm。第一批细胞从甘油储液接种至各自含有 20mL 所述介质的 250mL Erlenmeyer 瓶。在 16h 后，将含有 200mL 新鲜介质的 1L Erlenmeyer 瓶以 $OD_{600nm}$ 0.05 接种并温育 22h。对来自 250mL Erlenmeyer 瓶的足量细胞悬浮液离心并再悬浮在含有 10mL 复合介质（具有 10g/L 葡萄糖）的接种瓶中。将细胞从接种瓶转移到具有注射器的容器中。

[0258] 培养介质通过将各组分称重到小瓶中并用冷自来水天长该瓶至 1L 标线。然后将介质倒入容器中并将 1mL 的 1:10 稀二氧化硅消泡剂加入懂啊混合物中以防止灭菌期间起泡。然后将容器在 121℃高压灭菌 25 分钟。每个容器中使用的培养介质描述如下：

[0259]

| 容器 | 底物 | $NH_4Cl$ (g/L) | $K_2HPO_4$ (g/L) | $CaCl_2, 2H_2O$ (g/L) |
|---|---|---|---|---|
| 1 | 小麦 3% | 1 | 1 | 0.5 |
| 2 | 淀粉乳 20% | 1 | 1 | 0.5 |

[0260] 通过使用无菌玻璃瓶，通过取样管，每天手动地进行取样。从这些样品，进行多种分析：

[0261] . 生物质定量通过实施 qPCR 方法进行。

[0262] . 游离葡萄糖使用 YSI 分析仪定量。

[0263] . 气相色谱 FID 分析 (Varian CP-WAX 57CB 25m*0.32mm 柱）用来定量醇。

[0264] . 有机酸通过液相色谱质谱 (MicrOTOF-QII Bruker) 定量。

[0265] 结果在以下提供。它们证实，在发酵罐中，DG_21 可以产生接近 2g/L 的乙醇滴度。而且，它们证实，生产的动力学非常主动，因为低葡萄糖消耗和高细胞密度可以达到，进一步记录了本发明的细菌的显著的性能。

[0266] 更具体地，图 5 示出了约 7g/L 的生物质细胞密度可以在小麦上培养后在稳定期达到，以及葡萄糖消耗甚至在稳定期还继续。

[0267] 图 6 示出了乙醇生产是线性的并且在 96 小时培养中达到大约 2g/L。当葡萄糖从介质中耗尽时，乙醇生产不再进一步增加并且观察到平台。相关地，琥珀酸盐和乙酸盐在上清中检测到，分别达到约 3.0g/L 和 1.2g/L。

[0268] 图 7 示出了，在淀粉乳上培养后，约 15g/L 的生物质细胞密度可以在稳 定期达到，以及葡萄糖消耗即使在稳定期中也继续。在 144h 后，剩余的总葡萄糖为约 20g/L，这表明葡萄糖的低消耗速率。

[0269] 图 8 示出了乙醇生产逐渐增加，并且在淀粉乳中在 144 小时培养中达到大约 2g/L。相关地，乙酸盐生产显著降低。

[0270] 因此，这些结果显示了细菌从生物质生产高水平乙醇的性能。

[0001]

序列表

<110> DEINOVE

<120> Bacteria with reconstructed transcriptional units and the uses thereof (具有重构转录单元的细菌及其用途)

<130> B1285PC

<160> 26

<170> PatentIn version 3.3

<210> 1
<211> 1458
<212> DNA
<213> Deinococcus sp.

<220>
<221> CDS
<222> (1)..(1458)

<400> 1

```
atg cgc cgc ctt ccc ctc ctc gcc gcg ctg ctc gcc tcg ctg gca ggc      48
Met Arg Arg Leu Pro Leu Leu Ala Ala Leu Leu Ala Ser Leu Ala Gly
1               5                  10                  15

gcg cag gcc tcc ccc acc ctc ccg tcc ttc gag ggg cag gtg atc tat      96
Ala Gln Ala Ser Pro Thr Leu Pro Ser Phe Glu Gly Gln Val Ile Tyr
            20                  25                  30

cag gtg atg cct gac cgg ttt ttt gac ggg aac aag gcg aac gac gca     144
Gln Val Met Pro Asp Arg Phe Phe Asp Gly Asn Lys Ala Asn Asp Ala
        35                  40                  45

ggg gtc gac cgc tcg gac cca cgt gcc tgg cac ggc ggc gat ctg gcg     192
Gly Val Asp Arg Ser Asp Pro Arg Ala Trp His Gly Gly Asp Leu Ala
    50                  55                  60

ggc ctc acg gcc aaa ctg ccc tac ctc cgg cag ctg ggg gcg acg gcg     240
Gly Leu Thr Ala Lys Leu Pro Tyr Leu Arg Gln Leu Gly Ala Thr Ala
65                  70                  75                  80

gta tgg ctg acg cca atc tac cgg cag cag acg gcc aac gcc ttc ggt     288
Val Trp Leu Thr Pro Ile Tyr Arg Gln Gln Thr Ala Asn Ala Phe Gly
                85                  90                  95

acc gcc ccc tac cac ggc tac tgg ccc gcc gac ttc cgc gac gtg gac     336
Thr Ala Pro Tyr His Gly Tyr Trp Pro Ala Asp Phe Arg Asp Val Asp
            100                 105                 110

cca cat ttc ggg aca ctg gcc gat ttt gtc ggt ttc gtc aag gcg gca     384
Pro His Phe Gly Thr Leu Ala Asp Phe Val Gly Phe Val Lys Ala Ala
        115                 120                 125
```

[0002]

```
cac ggg gcg ggg ctg cgc gtg gtc ctc gat cag gtg atc aac cat tac      432
His Gly Ala Gly Leu Arg Val Val Leu Asp Gln Val Ile Asn His Tyr
    130                 135                 140

ggc tac gag gcg gcg gcg gtc aag gaa cac ccg gcc tgg ttc aat ggg      480
Gly Tyr Glu Ala Ala Ala Val Lys Glu His Pro Ala Trp Phe Asn Gly
145                 150                 155                 160

aaa gca gcg tgc gac gct tcc ggc aac aag gac gtg aac tgc cca ctg      528
Lys Ala Ala Cys Asp Ala Ser Gly Asn Lys Asp Val Asn Cys Pro Leu
                165                 170                 175

gcg ggc ctg ccg gac ctc aag cag tcg aac ccc gag gtc agg gcg ctg      576
Ala Gly Leu Pro Asp Leu Lys Gln Ser Asn Pro Glu Val Arg Ala Leu
            180                 185                 190

ctg ctg ggc aac gcc gac ttc tgg cgg ggg cag ggg gtg gat ggt ttc      624
Leu Leu Gly Asn Ala Asp Phe Trp Arg Gly Gln Gly Val Asp Gly Phe
        195                 200                 205

cgc tac gac gcg atc aag aac gtg gag acg ccc ttt ctg aaa gag ctg      672
Arg Tyr Asp Ala Ile Lys Asn Val Glu Thr Pro Phe Leu Lys Glu Leu
    210                 215                 220

ttg gca cgc gac cgt gcc gcc ggg acg tgg acg ctg ggc gag tgg tac      720
Leu Ala Arg Asp Arg Ala Ala Gly Thr Trp Thr Leu Gly Glu Trp Tyr
225                 230                 235                 240

ggg gca gac acc ggg acg gtg gcc gac tgg cag cag gcc ggg ttc gac      768
Gly Ala Asp Thr Gly Thr Val Ala Asp Trp Gln Gln Ala Gly Phe Asp
                245                 250                 255

agt ctc ttc ctc ttc agc ctg caa cag gcg atg ggg cag agc ctg atg      816
Ser Leu Phe Leu Phe Ser Leu Gln Gln Ala Met Gly Gln Ser Leu Met
            260                 265                 270

ggc ggg cag ggc ctc agc cgg gtg gcg agc gtg ctg agc cgc caa ggc      864
Gly Gly Gln Gly Leu Ser Arg Val Ala Ser Val Leu Ser Arg Gln Gly
        275                 280                 285

gag ctg cca cgc ccc ggc gag gtg gcc ctc ttt ctg gac aac cac gat      912
Glu Leu Pro Arg Pro Gly Glu Val Ala Leu Phe Leu Asp Asn His Asp
    290                 295                 300

gtc ccg cgc ttt gcc cag ggc agc ctg ttc gag gac cag gcc cag gcc      960
Val Pro Arg Phe Ala Gln Gly Ser Leu Phe Glu Asp Gln Ala Gln Ala
305                 310                 315                 320

cgc acc cgc tac ggc ctg cgt gcg ctg atg acc ctg aag ggc gtg ccg     1008
Arg Thr Arg Tyr Gly Leu Arg Ala Leu Met Thr Leu Lys Gly Val Pro
                325                 330                 335

gtg ctc tgg cag ggc act gag att gcc atg cgc ggc ggt ccc gac ccc     1056
Val Leu Trp Gln Gly Thr Glu Ile Ala Met Arg Gly Gly Pro Asp Pro
            340                 345                 350
```

[0003]

```
gat aac cgc cgc gat atg cgc ttc gag aac gag tgg acc cct gcc gag      1104
Asp Asn Arg Arg Asp Met Arg Phe Glu Asn Glu Trp Thr Pro Ala Glu
        355                 360                 365

cgc cag gtc ttc gag acg gcc cgg gac gcc atc gcc gtc cgc cag gcc      1152
Arg Gln Val Phe Glu Thr Ala Arg Asp Ala Ile Ala Val Arg Gln Ala
    370                 375                 380

agc cgg gcc ctc agc atc ggg acc cag aag ctg ctc ccc aca ccc gcc      1200
Ser Arg Ala Leu Ser Ile Gly Thr Gln Lys Leu Leu Pro Thr Pro Ala
385                 390                 395                 400

tcc ctg gaa gac gac ctg ctc ctc ttc aca cgc gag gcg cag ggg gag      1248
Ser Leu Glu Asp Asp Leu Leu Leu Phe Thr Arg Glu Ala Gln Gly Glu
                405                 410                 415

cgc gtg ctg gtc gcc tgg cac aac ggc agg aac cgc aag acg tac agc      1296
Arg Val Leu Val Ala Trp His Asn Gly Arg Asn Arg Lys Thr Tyr Ser
            420                 425                 430

ctc cgc ctg agc gcc ctg gga ctg aaa gcg gag ccg cag gcc gtc acc      1344
Leu Arg Leu Ser Ala Leu Gly Leu Lys Ala Glu Pro Gln Ala Val Thr
        435                 440                 445

cgc agc ctc ttc gcc ggg cag gac gcc aag ctc agc gtg agc ggg ggc      1392
Arg Ser Leu Phe Ala Gly Gln Asp Ala Lys Leu Ser Val Ser Gly Gly
    450                 455                 460

tgg ctg cac ctg agt ctg cct gcg gag gac gcg gca gga ttt ggg ctg      1440
Trp Leu His Leu Ser Leu Pro Ala Glu Asp Ala Ala Gly Phe Gly Leu
465                 470                 475                 480

ggg gga agc acc cgg tga                                             1458
Gly Gly Ser Thr Arg
                485
```

<210> 2
<211> 485
<212> PRT
<213> Deinococcus sp.

<400> 2

```
Met Arg Arg Leu Pro Leu Leu Ala Ala Leu Leu Ala Ser Leu Ala Gly
1               5                   10                  15

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

Gln Val Met Pro Asp Arg Phe Phe Asp Gly Asn Lys Ala Asn Asp Ala
        35                  40                  45
```

[0004]

Gly Val Asp Arg Ser Asp Pro Arg Ala Trp His Gly Gly Asp Leu Ala
50 55 60

Gly Leu Thr Ala Lys Leu Pro Tyr Leu Arg Gln Leu Gly Ala Thr Ala
65 70 75 80

Val Trp Leu Thr Pro Ile Tyr Arg Gln Gln Thr Ala Asn Ala Phe Gly
85 90 95

Thr Ala Pro Tyr His Gly Tyr Trp Pro Ala Asp Phe Arg Asp Val Asp
100 105 110

Pro His Phe Gly Thr Leu Ala Asp Phe Val Gly Phe Val Lys Ala Ala
115 120 125

His Gly Ala Gly Leu Arg Val Val Leu Asp Gln Val Ile Asn His Tyr
130 135 140

Gly Tyr Glu Ala Ala Ala Val Lys Glu His Pro Ala Trp Phe Asn Gly
145 150 155 160

Lys Ala Ala Cys Asp Ala Ser Gly Asn Lys Asp Val Asn Cys Pro Leu
165 170 175

Ala Gly Leu Pro Asp Leu Lys Gln Ser Asn Pro Glu Val Arg Ala Leu
180 185 190

Leu Leu Gly Asn Ala Asp Phe Trp Arg Gly Gln Gly Val Asp Gly Phe
195 200 205

Arg Tyr Asp Ala Ile Lys Asn Val Glu Thr Pro Phe Leu Lys Glu Leu
210 215 220

Leu Ala Arg Asp Arg Ala Ala Gly Thr Trp Thr Leu Gly Glu Trp Tyr
225 230 235 240

Gly Ala Asp Thr Gly Thr Val Ala Asp Trp Gln Gln Ala Gly Phe Asp
245 250 255

Ser Leu Phe Leu Phe Ser Leu Gln Gln Ala Met Gly Gln Ser Leu Met
260 265 270

[0005]

Gly Gly Gln Gly Leu Ser Arg Val Ala Ser Val Leu Ser Arg Gln Gly
275 280 285

Glu Leu Pro Arg Pro Gly Glu Val Ala Leu Phe Leu Asp Asn His Asp
290 295 300

Val Pro Arg Phe Ala Gln Gly Ser Leu Phe Glu Asp Gln Ala Gln Ala
305 310 315 320

Arg Thr Arg Tyr Gly Leu Arg Ala Leu Met Thr Leu Lys Gly Val Pro
325 330 335

Val Leu Trp Gln Gly Thr Glu Ile Ala Met Arg Gly Gly Pro Asp Pro
340 345 350

Asp Asn Arg Arg Asp Met Arg Phe Glu Asn Glu Trp Thr Pro Ala Glu
355 360 365

Arg Gln Val Phe Glu Thr Ala Arg Asp Ala Ile Ala Val Arg Gln Ala
370 375 380

Ser Arg Ala Leu Ser Ile Gly Thr Gln Lys Leu Leu Pro Thr Pro Ala
385 390 395 400

Ser Leu Glu Asp Asp Leu Leu Leu Phe Thr Arg Glu Ala Gln Gly Glu
405 410 415

Arg Val Leu Val Ala Trp His Asn Gly Arg Asn Arg Lys Thr Tyr Ser
420 425 430

Leu Arg Leu Ser Ala Leu Gly Leu Lys Ala Glu Pro Gln Ala Val Thr
435 440 445

Arg Ser Leu Phe Ala Gly Gln Asp Ala Lys Leu Ser Val Ser Gly Gly
450 455 460

Trp Leu His Leu Ser Leu Pro Ala Glu Asp Ala Ala Gly Phe Gly Leu
465 470 475 480

Gly Gly Ser Thr Arg
485

[0006]

<210> 3
<211> 3066
<212> DNA
<213> Deinococcus sp.

<220>
<221> CDS
<222> (1)..(3066)

<400> 3

```
atg aaa cgt ttc cag aag gtg ggt cgc agt ggc gcc ctg gcc gtc ctt      48
Met Lys Arg Phe Gln Lys Val Gly Arg Ser Gly Ala Leu Ala Val Leu
1               5                  10                  15

acg ttg gct ctg tcc gcc tgt ggc gtc ttg aag gcg ccc gag acg gga      96
Thr Leu Ala Leu Ser Ala Cys Gly Val Leu Lys Ala Pro Glu Thr Gly
            20                  25                  30

ggc aac act cgt gcc tgg cag gac gag gtg atc tac ttc gcc atg acc     144
Gly Asn Thr Arg Ala Trp Gln Asp Glu Val Ile Tyr Phe Ala Met Thr
        35                  40                  45

gac cgc ttc gcc aac ggg aac ccg gcc aac gac aac ggc ccg aac cgc     192
Asp Arg Phe Ala Asn Gly Asn Pro Ala Asn Asp Asn Gly Pro Asn Arg
    50                  55                  60

aat gag ggc gac cgg gcc gac cgg acc aac ccg ctc ggc tgg cac ggc     240
Asn Glu Gly Asp Arg Ala Asp Arg Thr Asn Pro Leu Gly Trp His Gly
65                  70                  75                  80

ggc gac ttc gcg ggg ctg aag gcg aag atc gag gag ggc tat ttc aag     288
Gly Asp Phe Ala Gly Leu Lys Ala Lys Ile Glu Glu Gly Tyr Phe Lys
                85                  90                  95

cgc atg ggc ttt acg gcc ctc tgg atc agc ccg gtg gtc ctg cag gtt     336
Arg Met Gly Phe Thr Ala Leu Trp Ile Ser Pro Val Val Leu Gln Val
            100                 105                 110

ccg gcc atc gag ggc ccg aag acc ggg ccg aac gcc ggg aag ctc ttc     384
Pro Ala Ile Glu Gly Pro Lys Thr Gly Pro Asn Ala Gly Lys Leu Phe
        115                 120                 125

gcg gga tac cac ggc tac tgg gcc gag gac ttt ttc aag gta gac cca     432
Ala Gly Tyr His Gly Tyr Trp Ala Glu Asp Phe Phe Lys Val Asp Pro
    130                 135                 140

cac ttc ggc acg ctg gac gag tac aag tcc ctc atc cag act gcg cac     480
His Phe Gly Thr Leu Asp Glu Tyr Lys Ser Leu Ile Gln Thr Ala His
145                 150                 155                 160

agg aac ggc atc aag gtg att cag gac att gtg gtc aac cac gcg ggc     528
Arg Asn Gly Ile Lys Val Ile Gln Asp Ile Val Val Asn His Ala Gly
                165                 170                 175

tac ggc gcc aca ctc acc aag acc aat cct gac tgg ttt cac acc cag     576
```

[0007]

```
Tyr Gly Ala Thr Leu Thr Lys Thr Asn Pro Asp Trp Phe His Thr Gln
            180                 185                 190

gct gaa tgc gac gcc agc acc aac aaa cgg gtg gac tgt ccg ctg gcg      624
Ala Glu Cys Asp Ala Ser Thr Asn Lys Arg Val Asp Cys Pro Leu Ala
        195                 200                 205

ggc ctg cct gac ttc aag cag gag cgg ccc gag gtc aca acg tac ctg      672
Gly Leu Pro Asp Phe Lys Gln Glu Arg Pro Glu Val Thr Thr Tyr Leu
    210                 215                 220

aac gac ttc gtg aac tcc tgg cgc aag gaa acc ggc atc gac ggg ctg      720
Asn Asp Phe Val Asn Ser Trp Arg Lys Glu Thr Gly Ile Asp Gly Leu
225                 230                 235                 240

cgg atc gac acc atg cag cac gtc tct gac agc tac tgg cag cag ttc      768
Arg Ile Asp Thr Met Gln His Val Ser Asp Ser Tyr Trp Gln Gln Phe
                245                 250                 255

ttt gcc gcg ggt ggg ccg ggg gac cct tcc aag atc tgg tcg gtc ggc      816
Phe Ala Ala Gly Gly Pro Gly Asp Pro Ser Lys Ile Trp Ser Val Gly
            260                 265                 270

gag gtg ttc aac ggt gat ccg gcc ttc ctg gcc cac tat atg gat gac      864
Glu Val Phe Asn Gly Asp Pro Ala Phe Leu Ala His Tyr Met Asp Asp
        275                 280                 285

ctc gga tcg ccc agc gtg ttc gat ttc gcg ctg tac ttc gcc atc aag      912
Leu Gly Ser Pro Ser Val Phe Asp Phe Ala Leu Tyr Phe Ala Ile Lys
    290                 295                 300

gat ggc ttg tcg agt gcg cgc ggc gac cta gga cgc ttg gcc gac gtg      960
Asp Gly Leu Ser Ser Ala Arg Gly Asp Leu Gly Arg Leu Ala Asp Val
305                 310                 315                 320

ttc gcg cgg gat ggt gcg tac cgg gac ccc aca cgg ctg acc acc ttc     1008
Phe Ala Arg Asp Gly Ala Tyr Arg Asp Pro Thr Arg Leu Thr Thr Phe
                325                 330                 335

gtg gac aac cac gac gtg ccc cgc ttc gtg agc gag gtg cag gag cgc     1056
Val Asp Asn His Asp Val Pro Arg Phe Val Ser Glu Val Gln Glu Arg
            340                 345                 350

ggc ggg aca gcg gcg cag gcg aac gag cgc ctt gac ctg gcc ctc agt     1104
Gly Gly Thr Ala Ala Gln Ala Asn Glu Arg Leu Asp Leu Ala Leu Ser
        355                 360                 365

ctg atc tat acc tcg cgc ggc aca ccg agc gtg tac cag ggc acg gag     1152
Leu Ile Tyr Thr Ser Arg Gly Thr Pro Ser Val Tyr Gln Gly Thr Glu
    370                 375                 380

atc gcg cag cct ggc ttg ggc gac ccc tac aac tac gcc acc ggc caa     1200
Ile Ala Gln Pro Gly Leu Gly Asp Pro Tyr Asn Tyr Ala Thr Gly Gln
385                 390                 395                 400

ggc aac cgc gag gac atg aac ttc ggg gcc ctc tcg cag agc agt atc     1248
```

[0008]

```
Gly Asn Arg Glu Asp Met Asn Phe Gly Ala Leu Ser Gln Ser Ser Ile
                405                 410                 415

gac gag cgg ctg gca gct ctc gcc gcg gca cgc gcg aag tac cgg gca  1296
Asp Glu Arg Leu Ala Ala Leu Ala Ala Ala Arg Ala Lys Tyr Arg Ala
            420                 425                 430

ctc aca cat ggc gtg cag cag gag ctg tgg cgg cca aac ggc ggg gcg  1344
Leu Thr His Gly Val Gln Gln Glu Leu Trp Arg Pro Asn Gly Gly Ala
        435                 440                 445

ccc atc ttc gcc cac cgc cgg att gtc acg gat ggt caa ggc gga cag  1392
Pro Ile Phe Ala His Arg Arg Ile Val Thr Asp Gly Gln Gly Gly Gln
    450                 455                 460

ccc gtc gtc gtc gtg atc aac aac ggc gac acg ccc gtg gac ctc tcc  1440
Pro Val Val Val Val Ile Asn Asn Gly Asp Thr Pro Val Asp Leu Ser
465                 470                 475                 480

act ctg agc ggg ggc ggt att ccg ctg ctg ggg acc ttc agc ggg acg  1488
Thr Leu Ser Gly Gly Gly Ile Pro Leu Leu Gly Thr Phe Ser Gly Thr
                485                 490                 495

gcg ctg aca gaa att acc ggg cga acc agc gac ctg agc gtg agc ggc  1536
Ala Leu Thr Glu Ile Thr Gly Arg Thr Ser Asp Leu Ser Val Ser Gly
            500                 505                 510

ggc caa ctc gta ggc acg gtt cct gcc cgc tcc gcg ctt gct gtc acg  1584
Gly Gln Leu Val Gly Thr Val Pro Ala Arg Ser Ala Leu Ala Val Thr
        515                 520                 525

gcc ccg gcg ggc agc ggc agc aca ggc acg gtg aac ccc agg ctg ccg  1632
Ala Pro Ala Gly Ser Gly Ser Thr Gly Thr Val Asn Pro Arg Leu Pro
    530                 535                 540

gag gtg acg gat ctc agt gcg aag gcc gga gac agc gcc gtg cag ctc  1680
Glu Val Thr Asp Leu Ser Ala Lys Ala Gly Asp Ser Ala Val Gln Leu
545                 550                 555                 560

acg tgg acg gcc tcc acg gac ctg aac gtc acc ggc tgc cgc gtc tac  1728
Thr Trp Thr Ala Ser Thr Asp Leu Asn Val Thr Gly Cys Arg Val Tyr
                565                 570                 575

gcc cgc acc ggg agc ggg cag gaa cgg ctc ctc aac ttc gcg ccg ctg  1776
Ala Arg Thr Gly Ser Gly Gln Glu Arg Leu Leu Asn Phe Ala Pro Leu
            580                 585                 590

ccc aag gac cag acc acg tac ctc gcc gca ggc att ccg aac gac cag  1824
Pro Lys Asp Gln Thr Thr Tyr Leu Ala Ala Gly Ile Pro Asn Asp Gln
        595                 600                 605

gaa acg acc ttc cgg gtg gtc acg gta gac gcg cag ggc gcc gag agt  1872
Glu Thr Thr Phe Arg Val Val Thr Val Asp Ala Gln Gly Ala Glu Ser
    610                 615                 620

cgg ggc gtc agc gtc aag gcc acg ccc agc agc aag aac acg gtc agg  1920
```

[0009]

```
Arg Gly Val Ser Val Lys Ala Thr Pro Ser Ser Lys Asn Thr Val Arg
625                 630                 635                 640

gtg act ttc acg gtg gac gcc cgc agc cag ggc aac ggc ccg atc gag  1968
Val Thr Phe Thr Val Asp Ala Arg Ser Gln Gly Asn Gly Pro Ile Glu
                645                 650                 655

ctg cgc cgc ttc gac acg ggc tcg cag ctt gag tac ccc atg acg cag  2016
Leu Arg Arg Phe Asp Thr Gly Ser Gln Leu Glu Tyr Pro Met Thr Gln
            660                 665                 670

gtg agc cgc ggc atc tgg aag acg gcg att gac ctc ccc ctc ttc cgc  2064
Val Ser Arg Gly Ile Trp Lys Thr Ala Ile Asp Leu Pro Leu Phe Arg
        675                 680                 685

gag atc aag ttt aag ttc ggc aac gac gga ccc gcc gcc aag aac agc  2112
Glu Ile Lys Phe Lys Phe Gly Asn Asp Gly Pro Ala Ala Lys Asn Ser
    690                 695                 700

ggc tac gag gca ccc ggc caa ccc gac cgc agc tat gtg gtg gga aca  2160
Gly Tyr Glu Ala Pro Gly Gln Pro Asp Arg Ser Tyr Val Val Gly Thr
705                 710                 715                 720

aat cct aac gtc tac acc ggc acc tat gac ttt att acc cag ccg gtg  2208
Asn Pro Asn Val Tyr Thr Gly Thr Tyr Asp Phe Ile Thr Gln Pro Val
                725                 730                 735

ccg cag acc acc atc gag ggc caa gtc aga gga gcg ggc aat ccc ctc  2256
Pro Gln Thr Thr Ile Glu Gly Gln Val Arg Gly Ala Gly Asn Pro Leu
            740                 745                 750

gcg aat gcg ttg gtc gaa gcg gtg acc gcc aac ccc gac ctg cac tat  2304
Ala Asn Ala Leu Val Glu Ala Val Thr Ala Asn Pro Asp Leu His Tyr
        755                 760                 765

gcg atg acc ttt ccg gac ggc aca tac aca ctg ttt gtt ccg gca ggg  2352
Ala Met Thr Phe Pro Asp Gly Thr Tyr Thr Leu Phe Val Pro Ala Gly
    770                 775                 780

acc cac aca ctg cag gcc aag gca ggc ggc tac gta gca gcc agc cgg  2400
Thr His Thr Leu Gln Ala Lys Ala Gly Gly Tyr Val Ala Ala Ser Arg
785                 790                 795                 800

cag gcg atc tcg ccg ggg acg ggc gca gac ttc aac ctg gcc cag gac  2448
Gln Ala Ile Ser Pro Gly Thr Gly Ala Asp Phe Asn Leu Ala Gln Asp
                805                 810                 815

ctg agc acc aag tac acc atc gac ggc aac ctg gcc gac tgg acg gcc  2496
Leu Ser Thr Lys Tyr Thr Ile Asp Gly Asn Leu Ala Asp Trp Thr Ala
            820                 825                 830

ccc aag gtg acg ctg caa agc ccg acc gag gga ggc ttc ggg ccc gac  2544
Pro Lys Val Thr Leu Gln Ser Pro Thr Glu Gly Gly Phe Gly Pro Asp
        835                 840                 845

aac aat tgg ttg aca ctc cag gcc gac agt gat gac cac tat ctg tac  2592
```

[0010]

```
Asn Asn Trp Leu Thr Leu Gln Ala Asp Ser Asp Asp His Tyr Leu Tyr
    850                 855                 860

ctc gcg tac acg tac cgg gtg aag gga aac agc gcg atc ctg tac ctg      2640
Leu Ala Tyr Thr Tyr Arg Val Lys Gly Asn Ser Ala Ile Leu Tyr Leu
865                 870                 875                 880

gac acc aag atg ggc ggt gcg gcc caa gcc gac aat ttc gag gct tgg      2688
Asp Thr Lys Met Gly Gly Ala Ala Gln Ala Asp Asn Phe Glu Ala Trp
                885                 890                 895

aag cgg gcg gcg acc ttc agt ggg agc atg ggg ggc gcc gac gcc ttt      2736
Lys Arg Ala Ala Thr Phe Ser Gly Ser Met Gly Gly Ala Asp Ala Phe
            900                 905                 910

gtt gcg cgg tac gaa aac cag atg gct caa ctg agg ctg ttt cag agc      2784
Val Ala Arg Tyr Glu Asn Gln Met Ala Gln Leu Arg Leu Phe Gln Ser
        915                 920                 925

gat act gcc acg ccc gag gtc aac acg ggc gac tac aag ttt gca gcg      2832
Asp Thr Ala Thr Pro Glu Val Asn Thr Gly Asp Tyr Lys Phe Ala Ala
    930                 935                 940

agc ggt acc ctg ccc gag cag acg gtg gaa ctg gcg atc ccg tgg aca      2880
Ser Gly Thr Leu Pro Glu Gln Thr Val Glu Leu Ala Ile Pro Trp Thr
945                 950                 955                 960

gca ctc ggc ctc agc gaa aaa cct gcg aac ggt gtg aac gtg gtg ggt      2928
Ala Leu Gly Leu Ser Glu Lys Pro Ala Asn Gly Val Asn Val Val Gly
                965                 970                 975

gga att ttc ggt ggc gac ggc tac ggc gcg ggc gac atc gtg ccc aat      2976
Gly Ile Phe Gly Gly Asp Gly Tyr Gly Ala Gly Asp Ile Val Pro Asn
            980                 985                 990

acc acc agt aca ccc ccc ggt gcc  aac acc att gga acg  gat gcc gaa    3024
Thr Thr Ser Thr Pro Pro Gly Ala  Asn Thr Ile Gly Thr  Asp Ala Glu
        995                 1000                1005

cag cgc  cgg gca acc ttc act  cag ccc ctc aac gtg  agg taa           3066
Gln Arg  Arg Ala Thr Phe Thr  Gln Pro Leu Asn Val  Arg
    1010                1015                1020
```

<210> 4
<211> 1021
<212> PRT
<213> Deinococcus sp.

<400> 4

```
Met Lys Arg Phe Gln Lys Val Gly Arg Ser Gly Ala Leu Ala Val Leu
1               5                   10                  15

Thr Leu Ala Leu Ser Ala Cys Gly Val Leu Lys Ala Pro Glu Thr Gly
```

[0011]

20 25 30

Gly Asn Thr Arg Ala Trp Gln Asp Glu Val Ile Tyr Phe Ala Met Thr
35 40 45

Asp Arg Phe Ala Asn Gly Asn Pro Ala Asn Asp Asn Gly Pro Asn Arg
50 55 60

Asn Glu Gly Asp Arg Ala Asp Arg Thr Asn Pro Leu Gly Trp His Gly
65 70 75 80

Gly Asp Phe Ala Gly Leu Lys Ala Lys Ile Glu Glu Gly Tyr Phe Lys
85 90 95

Arg Met Gly Phe Thr Ala Leu Trp Ile Ser Pro Val Val Leu Gln Val
100 105 110

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

Ala Gly Tyr His Gly Tyr Trp Ala Glu Asp Phe Phe Lys Val Asp Pro
130 135 140

His Phe Gly Thr Leu Asp Glu Tyr Lys Ser Leu Ile Gln Thr Ala His
145 150 155 160

Arg Asn Gly Ile Lys Val Ile Gln Asp Ile Val Val Asn His Ala Gly
165 170 175

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

Ala Glu Cys Asp Ala Ser Thr Asn Lys Arg Val Asp Cys Pro Leu Ala
195 200 205

Gly Leu Pro Asp Phe Lys Gln Glu Arg Pro Glu Val Thr Thr Tyr Leu
210 215 220

Asn Asp Phe Val Asn Ser Trp Arg Lys Glu Thr Gly Ile Asp Gly Leu
225 230 235 240

Arg Ile Asp Thr Met Gln His Val Ser Asp Ser Tyr Trp Gln Gln Phe

[0012]

245 250 255

Phe Ala Ala Gly Gly Pro Gly Asp Pro Ser Lys Ile Trp Ser Val Gly
260 265 270

Glu Val Phe Asn Gly Asp Pro Ala Phe Leu Ala His Tyr Met Asp Asp
275 280 285

Leu Gly Ser Pro Ser Val Phe Asp Phe Ala Leu Tyr Phe Ala Ile Lys
290 295 300

Asp Gly Leu Ser Ser Ala Arg Gly Asp Leu Gly Arg Leu Ala Asp Val
305 310 315 320

Phe Ala Arg Asp Gly Ala Tyr Arg Asp Pro Thr Arg Leu Thr Thr Phe
325 330 335

Val Asp Asn His Asp Val Pro Arg Phe Val Ser Glu Val Gln Glu Arg
340 345 350

Gly Gly Thr Ala Ala Gln Ala Asn Glu Arg Leu Asp Leu Ala Leu Ser
355 360 365

Leu Ile Tyr Thr Ser Arg Gly Thr Pro Ser Val Tyr Gln Gly Thr Glu
370 375 380

Ile Ala Gln Pro Gly Leu Gly Asp Pro Tyr Asn Tyr Ala Thr Gly Gln
385 390 395 400

Gly Asn Arg Glu Asp Met Asn Phe Gly Ala Leu Ser Gln Ser Ser Ile
405 410 415

Asp Glu Arg Leu Ala Ala Leu Ala Ala Ala Arg Ala Lys Tyr Arg Ala
420 425 430

Leu Thr His Gly Val Gln Gln Glu Leu Trp Arg Pro Asn Gly Gly Ala
435 440 445

Pro Ile Phe Ala His Arg Arg Ile Val Thr Asp Gly Gln Gly Gly Gln
450 455 460

Pro Val Val Val Val Ile Asn Asn Gly Asp Thr Pro Val Asp Leu Ser

[0013]

465 470 475 480

Thr Leu Ser Gly Gly Gly Ile Pro Leu Leu Gly Thr Phe Ser Gly Thr
485 490 495

Ala Leu Thr Glu Ile Thr Gly Arg Thr Ser Asp Leu Ser Val Ser Gly
500 505 510

Gly Gln Leu Val Gly Thr Val Pro Ala Arg Ser Ala Leu Ala Val Thr
515 520 525

Ala Pro Ala Gly Ser Gly Ser Thr Gly Thr Val Asn Pro Arg Leu Pro
530 535 540

Glu Val Thr Asp Leu Ser Ala Lys Ala Gly Asp Ser Ala Val Gln Leu
545 550 555 560

Thr Trp Thr Ala Ser Thr Asp Leu Asn Val Thr Gly Cys Arg Val Tyr
565 570 575

Ala Arg Thr Gly Ser Gly Gln Glu Arg Leu Leu Asn Phe Ala Pro Leu
580 585 590

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

Glu Thr Thr Phe Arg Val Val Thr Val Asp Ala Gln Gly Ala Glu Ser
610 615 620

Arg Gly Val Ser Val Lys Ala Thr Pro Ser Ser Lys Asn Thr Val Arg
625 630 635 640

Val Thr Phe Thr Val Asp Ala Arg Ser Gln Gly Asn Gly Pro Ile Glu
645 650 655

Leu Arg Arg Phe Asp Thr Gly Ser Gln Leu Glu Tyr Pro Met Thr Gln
660 665 670

Val Ser Arg Gly Ile Trp Lys Thr Ala Ile Asp Leu Pro Leu Phe Arg
675 680 685

Glu Ile Lys Phe Lys Phe Gly Asn Asp Gly Pro Ala Ala Lys Asn Ser

[0014]

690 695 700

Gly Tyr Glu Ala Pro Gly Gln Pro Asp Arg Ser Tyr Val Val Gly Thr
705 710 715 720

Asn Pro Asn Val Tyr Thr Gly Thr Tyr Asp Phe Ile Thr Gln Pro Val
725 730 735

Pro Gln Thr Thr Ile Glu Gly Gln Val Arg Gly Ala Gly Asn Pro Leu
740 745 750

Ala Asn Ala Leu Val Glu Ala Val Thr Ala Asn Pro Asp Leu His Tyr
755 760 765

Ala Met Thr Phe Pro Asp Gly Thr Tyr Thr Leu Phe Val Pro Ala Gly
770 775 780

Thr His Thr Leu Gln Ala Lys Ala Gly Gly Tyr Val Ala Ala Ser Arg
785 790 795 800

Gln Ala Ile Ser Pro Gly Thr Gly Ala Asp Phe Asn Leu Ala Gln Asp
805 810 815

Leu Ser Thr Lys Tyr Thr Ile Asp Gly Asn Leu Ala Asp Trp Thr Ala
820 825 830

Pro Lys Val Thr Leu Gln Ser Pro Thr Glu Gly Gly Phe Gly Pro Asp
835 840 845

Asn Asn Trp Leu Thr Leu Gln Ala Asp Ser Asp Asp His Tyr Leu Tyr
850 855 860

Leu Ala Tyr Thr Tyr Arg Val Lys Gly Asn Ser Ala Ile Leu Tyr Leu
865 870 875 880

Asp Thr Lys Met Gly Gly Ala Ala Gln Ala Asp Asn Phe Glu Ala Trp
885 890 895

Lys Arg Ala Ala Thr Phe Ser Gly Ser Met Gly Gly Ala Asp Ala Phe
900 905 910

Val Ala Arg Tyr Glu Asn Gln Met Ala Gln Leu Arg Leu Phe Gln Ser

[0015]

915 920 925

Asp Thr Ala Thr Pro Glu Val Asn Thr Gly Asp Tyr Lys Phe Ala Ala
930 935 940

Ser Gly Thr Leu Pro Glu Gln Thr Val Glu Leu Ala Ile Pro Trp Thr
945 950 955 960

Ala Leu Gly Leu Ser Glu Lys Pro Ala Asn Gly Val Asn Val Val Gly
965 970 975

Gly Ile Phe Gly Gly Asp Gly Tyr Gly Ala Gly Asp Ile Val Pro Asn
980 985 990

Thr Thr Ser Thr Pro Pro Gly Ala Asn Thr Ile Gly Thr Asp Ala Glu
995 1000 1005

Gln Arg Arg Ala Thr Phe Thr Gln Pro Leu Asn Val Arg
1010 1015 1020

<210> 5
<211> 1545
<212> DNA
<213> Deinococcus sp.

<220>
<221> CDS
<222> (1)..(1545)

<400> 5
atg cgc cgc ctg ccc ctc ctc gcc gcc ctg ctc gcc tcg ctg gcc ggc 48
Met Arg Arg Leu Pro Leu Leu Ala Ala Leu Leu Ala Ser Leu Ala Gly
1 5 10 15

gcg cag gcc aaa aaa gcc cag att ctt ctt gac acc cac aga acc atc 96
Ala Gln Ala Lys Lys Ala Gln Ile Leu Leu Asp Thr His Arg Thr Ile
20 25 30

agc gaa atc agc cac tac atc ttt ggt gga ttc gcc gag cac atg ggc 144
Ser Glu Ile Ser His Tyr Ile Phe Gly Gly Phe Ala Glu His Met Gly
35 40 45

cgc tgc atc tac gag ggc atc tac gac ccc caa agc cct ctg agc gac 192
Arg Cys Ile Tyr Glu Gly Ile Tyr Asp Pro Gln Ser Pro Leu Ser Asp
50 55 60

gag aac ggc atc cgc agg gat gtg atg gac gcc ctg aag gaa ctc aat 240
Glu Asn Gly Ile Arg Arg Asp Val Met Asp Ala Leu Lys Glu Leu Asn

[0016]

```
65                  70                  75                  80

ttc cgt tcc atc cgt tac ccc ggg ggc aac ttc gtg tca ggg tac aac      288
Phe Arg Ser Ile Arg Tyr Pro Gly Gly Asn Phe Val Ser Gly Tyr Asn
                85                  90                  95

tgg gaa gac gga att ggc ccc agg gaa aac cgc ccg gtc aag cgc gat      336
Trp Glu Asp Gly Ile Gly Pro Arg Glu Asn Arg Pro Val Lys Arg Asp
            100                 105                 110

ctg gcc tgg agg agc atc gaa acc aac cag ttt ggc acg gat gaa ttc      384
Leu Ala Trp Arg Ser Ile Glu Thr Asn Gln Phe Gly Thr Asp Glu Phe
        115                 120                 125

atg aag gtc tgc gct gaa ctg aag acc gaa ccc atg atg gcc gtg aac      432
Met Lys Val Cys Ala Glu Leu Lys Thr Glu Pro Met Met Ala Val Asn
    130                 135                 140

ctg ggc acc gga agt att cag gac gcg gcc aac atc gtc gaa tac tgc      480
Leu Gly Thr Gly Ser Ile Gln Asp Ala Ala Asn Ile Val Glu Tyr Cys
145                 150                 155                 160

aac ctc gaa ggc ggc acc cat tac agc gac ctg cgc atc aaa aac ggt      528
Asn Leu Glu Gly Gly Thr His Tyr Ser Asp Leu Arg Ile Lys Asn Gly
                165                 170                 175

gct gaa aaa cct tat ggt gtg aag ttc tgg tgt ctg ggg aac gag atg      576
Ala Glu Lys Pro Tyr Gly Val Lys Phe Trp Cys Leu Gly Asn Glu Met
            180                 185                 190

gat ggt ccc tgg cag gtg gga cag ctt tct gca gag gat tac agt aag      624
Asp Gly Pro Trp Gln Val Gly Gln Leu Ser Ala Glu Asp Tyr Ser Lys
        195                 200                 205

aaa gcc gtg cag gct gca aag gcc atg aag ctg atc gat cct tcc att      672
Lys Ala Val Gln Ala Ala Lys Ala Met Lys Leu Ile Asp Pro Ser Ile
    210                 215                 220

caa ctg att gcc tgc ggt tcc tcc tcc agc ctc atg aac tcc tac ccc      720
Gln Leu Ile Ala Cys Gly Ser Ser Ser Ser Leu Met Asn Ser Tyr Pro
225                 230                 235                 240

gag tgg gac cgc atc gtg ctg gaa gag acc tgg gac cag atc gat tac      768
Glu Trp Asp Arg Ile Val Leu Glu Glu Thr Trp Asp Gln Ile Asp Tyr
                245                 250                 255

ctc tcg atg cac tac tat gcc agc aac cgg gag gag gac act gcc agt      816
Leu Ser Met His Tyr Tyr Ala Ser Asn Arg Glu Glu Asp Thr Ala Ser
            260                 265                 270

tac ctc gcc tat acc cgt gaa ttc gaa gac cac ctg caa acc ctg gcc      864
Tyr Leu Ala Tyr Thr Arg Glu Phe Glu Asp His Leu Gln Thr Leu Ala
        275                 280                 285

gcc acc atc cgt tac gtg aaa gcc aag aaa cgc agc cag aaa gac gtg      912
Ala Thr Ile Arg Tyr Val Lys Ala Lys Lys Arg Ser Gln Lys Asp Val
```

[0017]

```
     290                 295                 300

ttc ctc tcc tgg gat gaa tgg aac gtc tgg tac cgc gaa atg aac ggc      960
Phe Leu Ser Trp Asp Glu Trp Asn Val Trp Tyr Arg Glu Met Asn Gly
305                 310                 315                 320

aac ggc gag tgg cag cag gcc ccc cac atc ctg gaa gag gtc tac aac     1008
Asn Gly Glu Trp Gln Gln Ala Pro His Ile Leu Glu Glu Val Tyr Asn
                325                 330                 335

ctt gaa gat gcg ctg gtg gtg gcc cag tgg atg aat gtc ctc ctg aag     1056
Leu Glu Asp Ala Leu Val Val Ala Gln Trp Met Asn Val Leu Leu Lys
            340                 345                 350

cac agc aat gtg ctg aag atg gcc tcc atc gca cag gtt gtc aat gtg     1104
His Ser Asn Val Leu Lys Met Ala Ser Ile Ala Gln Val Val Asn Val
        355                 360                 365

atc gct ccc atc atg acc aga cgg gat ggc atg ttc aaa cag acc atc     1152
Ile Ala Pro Ile Met Thr Arg Arg Asp Gly Met Phe Lys Gln Thr Ile
    370                 375                 380

tat tat cct ttc ctg gtg ttc agc aaa cac gct tct ggt cag gcg ctc     1200
Tyr Tyr Pro Phe Leu Val Phe Ser Lys His Ala Ser Gly Gln Ala Leu
385                 390                 395                 400

agc ctg cat gtg gcc tcc gac cag tac gag acg aaa aaa cac ggc ctc     1248
Ser Leu His Val Ala Ser Asp Gln Tyr Glu Thr Lys Lys His Gly Leu
                405                 410                 415

gtg aac ctg ctc gat gcc agt gcc agt ttt gat gcc agc cag aac gaa     1296
Val Asn Leu Leu Asp Ala Ser Ala Ser Phe Asp Ala Ser Gln Asn Glu
            420                 425                 430

ggg gct gtt ttt ctg gtc aac cgc agc cag gat gaa gaa ctc gaa acc     1344
Gly Ala Val Phe Leu Val Asn Arg Ser Gln Asp Glu Glu Leu Glu Thr
        435                 440                 445

gaa atc gtc ttt cag ggc cgt gtt ccc act tcc gtg cgc gtg gcc cac     1392
Glu Ile Val Phe Gln Gly Arg Val Pro Thr Ser Val Arg Val Ala His
    450                 455                 460

cag ctt gct ggc agc gac ccc aaa gcc cac aac tcc ttc gag gag cct     1440
Gln Leu Ala Gly Ser Asp Pro Lys Ala His Asn Ser Phe Glu Glu Pro
465                 470                 475                 480

gaa aag ctc acc ctg cag acg att gaa gca ggg gag atc aaa gac ggc     1488
Glu Lys Leu Thr Leu Gln Thr Ile Glu Ala Gly Glu Ile Lys Asp Gly
                485                 490                 495

aaa ctc gtg ctg aag ctt cct gcc ctg tcc ttt tcg gca gtg gtg ctg     1536
Lys Leu Val Leu Lys Leu Pro Ala Leu Ser Phe Ser Ala Val Val Leu
            500                 505                 510

gac tac taa                                                        1545
Asp Tyr
```

[0018]

<210> 6
<211> 514
<212> PRT
<213> Deinococcus sp.

<400> 6

Met Arg Arg Leu Pro Leu Leu Ala Ala Leu Leu Ala Ser Leu Ala Gly
1 5 10 15

Ala Gln Ala Lys Lys Ala Gln Ile Leu Leu Asp Thr His Arg Thr Ile
20 25 30

Ser Glu Ile Ser His Tyr Ile Phe Gly Gly Phe Ala Glu His Met Gly
35 40 45

Arg Cys Ile Tyr Glu Gly Ile Tyr Asp Pro Gln Ser Pro Leu Ser Asp
50 55 60

Glu Asn Gly Ile Arg Arg Asp Val Met Asp Ala Leu Lys Glu Leu Asn
65 70 75 80

Phe Arg Ser Ile Arg Tyr Pro Gly Gly Asn Phe Val Ser Gly Tyr Asn
85 90 95

Trp Glu Asp Gly Ile Gly Pro Arg Glu Asn Arg Pro Val Lys Arg Asp
100 105 110

Leu Ala Trp Arg Ser Ile Glu Thr Asn Gln Phe Gly Thr Asp Glu Phe
115 120 125

Met Lys Val Cys Ala Glu Leu Lys Thr Glu Pro Met Met Ala Val Asn
130 135 140

Leu Gly Thr Gly Ser Ile Gln Asp Ala Ala Asn Ile Val Glu Tyr Cys
145 150 155 160

Asn Leu Glu Gly Gly Thr His Tyr Ser Asp Leu Arg Ile Lys Asn Gly
165 170 175

Ala Glu Lys Pro Tyr Gly Val Lys Phe Trp Cys Leu Gly Asn Glu Met
180 185 190

[0019]

Asp Gly Pro Trp Gln Val Gly Gln Leu Ser Ala Glu Asp Tyr Ser Lys
195 200 205

Lys Ala Val Gln Ala Ala Lys Ala Met Lys Leu Ile Asp Pro Ser Ile
210 215 220

Gln Leu Ile Ala Cys Gly Ser Ser Ser Ser Leu Met Asn Ser Tyr Pro
225 230 235 240

Glu Trp Asp Arg Ile Val Leu Glu Glu Thr Trp Asp Gln Ile Asp Tyr
245 250 255

Leu Ser Met His Tyr Tyr Ala Ser Asn Arg Glu Glu Asp Thr Ala Ser
260 265 270

Tyr Leu Ala Tyr Thr Arg Glu Phe Glu Asp His Leu Gln Thr Leu Ala
275 280 285

Ala Thr Ile Arg Tyr Val Lys Ala Lys Lys Arg Ser Gln Lys Asp Val
290 295 300

Phe Leu Ser Trp Asp Glu Trp Asn Val Trp Tyr Arg Glu Met Asn Gly
305 310 315 320

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

Leu Glu Asp Ala Leu Val Val Ala Gln Trp Met Asn Val Leu Leu Lys
340 345 350

His Ser Asn Val Leu Lys Met Ala Ser Ile Ala Gln Val Val Asn Val
355 360 365

Ile Ala Pro Ile Met Thr Arg Arg Asp Gly Met Phe Lys Gln Thr Ile
370 375 380

Tyr Tyr Pro Phe Leu Val Phe Ser Lys His Ala Ser Gly Gln Ala Leu
385 390 395 400

Ser Leu His Val Ala Ser Asp Gln Tyr Glu Thr Lys Lys His Gly Leu
405 410 415

[0020]

Val Asn Leu Leu Asp Ala Ser Ala Ser Phe Asp Ala Ser Gln Asn Glu
420 425 430

Gly Ala Val Phe Leu Val Asn Arg Ser Gln Asp Glu Glu Leu Glu Thr
435 440 445

Glu Ile Val Phe Gln Gly Arg Val Pro Thr Ser Val Arg Val Ala His
450 455 460

Gln Leu Ala Gly Ser Asp Pro Lys Ala His Asn Ser Phe Glu Glu Pro
465 470 475 480

Glu Lys Leu Thr Leu Gln Thr Ile Glu Ala Gly Glu Ile Lys Asp Gly
485 490 495

Lys Leu Val Leu Lys Leu Pro Ala Leu Ser Phe Ser Ala Val Val Leu
500 505 510

Asp Tyr

<210> 7
<211> 2418
<212> DNA
<213> Deinococcus sp.

<220>
<221> CDS
<222> (1)..(2148)

<400> 7
atg tgc ctg gag tgc ttt ttc agg gtc ctc cag gcg ttg aag gcc aca 48
Met Cys Leu Glu Cys Phe Phe Arg Val Leu Gln Ala Leu Lys Ala Thr
1 5 10 15

cct gtc ctc ctg agg ggc cct tgt tgc acc ttc agg agg ttc aat ccc 96
Pro Val Leu Leu Arg Gly Pro Cys Cys Thr Phe Arg Arg Phe Asn Pro
20 25 30

atg atg cca acc cct tca aaa gtc ctg cct gcc agt gtg ctg ctg atg 144
Met Met Pro Thr Pro Ser Lys Val Leu Pro Ala Ser Val Leu Leu Met
35 40 45

gtc tcc ctg ctg acc agc tcc tgt aat ctg ttc cag cca cct gca ccc 192
Val Ser Leu Leu Thr Ser Ser Cys Asn Leu Phe Gln Pro Pro Ala Pro
50 55 60

[0021]

```
aac tgc acc ccc aag acc acc gga gcc acc gtt cca gca ggg gac tac      240
Asn Cys Thr Pro Lys Thr Thr Gly Ala Thr Val Pro Ala Gly Asp Tyr
65                  70                  75                  80

gat cct gca gca agt gaa aaa gcc ttc ccg gac ctg ctg acc aca gct      288
Asp Pro Ala Ala Ser Glu Lys Ala Phe Pro Asp Leu Leu Thr Thr Ala
                85                  90                  95

gcc cgc aaa ccc tca cag gcc gga gcc ctg caa ctc gtg cag caa gac      336
Ala Arg Lys Pro Ser Gln Ala Gly Ala Leu Gln Leu Val Gln Gln Asp
            100                 105                 110

tgc atg gtt acg ctg gca gac agc agc ggc aaa ccc atc cag ctt cgt      384
Cys Met Val Thr Leu Ala Asp Ser Ser Gly Lys Pro Ile Gln Leu Arg
        115                 120                 125

ggc atg agc acc cac ggc ctg caa tgg tac ccg gaa atc gtc aat gac      432
Gly Met Ser Thr His Gly Leu Gln Trp Tyr Pro Glu Ile Val Asn Asp
    130                 135                 140

aat gcc ttc aag gcc ctg gcc aac gac tgg ggg tcc aat gtc ttc cgt      480
Asn Ala Phe Lys Ala Leu Ala Asn Asp Trp Gly Ser Asn Val Phe Arg
145                 150                 155                 160

ctg gcc ctg tac gtc gga gaa ggg gga tac gcc acc aaa ccc gaa ctg      528
Leu Ala Leu Tyr Val Gly Glu Gly Gly Tyr Ala Thr Lys Pro Glu Leu
                165                 170                 175

aaa caa aaa gtc att gaa ggc att gat ttt gcc att gcc aac gac atg      576
Lys Gln Lys Val Ile Glu Gly Ile Asp Phe Ala Ile Ala Asn Asp Met
            180                 185                 190

tac gtg att gtg gac tgg cac gtg cat gcc cct ggc gac ccc aac gca      624
Tyr Val Ile Val Asp Trp His Val His Ala Pro Gly Asp Pro Asn Ala
        195                 200                 205

gac gtg tac acc aat gcc aaa ccg ctg gag ttc ttc aag tcc atc gcc      672
Asp Val Tyr Thr Asn Ala Lys Pro Leu Glu Phe Phe Lys Ser Ile Ala
    210                 215                 220

cag aag tac ccc aac aac aag cac atc att tac gag gtc gcc aac gaa      720
Gln Lys Tyr Pro Asn Asn Lys His Ile Ile Tyr Glu Val Ala Asn Glu
225                 230                 235                 240

ccc aac ccc ggt cag gct cca ggg gtc agc aat gac gct gaa ggc tgg      768
Pro Asn Pro Gly Gln Ala Pro Gly Val Ser Asn Asp Ala Glu Gly Trp
                245                 250                 255

aag aag atc aag tcc tac gca gag ccc atc atc aag atg ctg cgg gac      816
Lys Lys Ile Lys Ser Tyr Ala Glu Pro Ile Ile Lys Met Leu Arg Asp
            260                 265                 270

ctg ggc aac aag aac atc gtg att gtc ggg acc ccc aac tgg agc cag      864
Leu Gly Asn Lys Asn Ile Val Ile Val Gly Thr Pro Asn Trp Ser Gln
        275                 280                 285
```

[0022]

```
cgc ccg gat ctg gct gcc gac aac ccc atc aaa gac agt gcc acc ctt      912
Arg Pro Asp Leu Ala Ala Asp Asn Pro Ile Lys Asp Ser Ala Thr Leu
    290                 295                 300

tac act gtg cac ttc tac acc ggc acc cac atg ccc tcc acc aac ctg      960
Tyr Thr Val His Phe Tyr Thr Gly Thr His Met Pro Ser Thr Asn Leu
305                 310                 315                 320

gca gac cgg ggc aac gtg atg agc aat gcc cgt tac gcc ctg gag cac     1008
Ala Asp Arg Gly Asn Val Met Ser Asn Ala Arg Tyr Ala Leu Glu His
                325                 330                 335

ggt gtg ggc gtg ttc tcc acc gag tgg ggg gtc agc gag gcc agc gga     1056
Gly Val Gly Val Phe Ser Thr Glu Trp Gly Val Ser Glu Ala Ser Gly
            340                 345                 350

aac aac gga cct ttc ctc aaa gaa gcc gac gtc tgg ctg cag ttc ctc     1104
Asn Asn Gly Pro Phe Leu Lys Glu Ala Asp Val Trp Leu Gln Phe Leu
        355                 360                 365

aac aaa cac aac atc agc tgg gtc aac tgg tct ctc acc aac aag gcc     1152
Asn Lys His Asn Ile Ser Trp Val Asn Trp Ser Leu Thr Asn Lys Ala
    370                 375                 380

gag act tct gca gcc ttc ctg ccg ttc ccc aac cag acc agc ctt gat     1200
Glu Thr Ser Ala Ala Phe Leu Pro Phe Pro Asn Gln Thr Ser Leu Asp
385                 390                 395                 400

ccc ggg gca gac agg ctg tgg acc ccc agc gag ctc acc ctg tcc ggg     1248
Pro Gly Ala Asp Arg Leu Trp Thr Pro Ser Glu Leu Thr Leu Ser Gly
                405                 410                 415

gaa tac gtt cgg gcg cgc atc aaa ggg acc aaa tac cag ccc att gac     1296
Glu Tyr Val Arg Ala Arg Ile Lys Gly Thr Lys Tyr Gln Pro Ile Asp
            420                 425                 430

cgc act gcc ttc act gag gtg gcc ttc aat ttt gac aac gac acc acc     1344
Arg Thr Ala Phe Thr Glu Val Ala Phe Asn Phe Asp Asn Asp Thr Thr
        435                 440                 445

cag ggt ttt gcc ctg aac ccg gac agt gga gtc aaa ggg atc acg gtc     1392
Gln Gly Phe Ala Leu Asn Pro Asp Ser Gly Val Lys Gly Ile Thr Val
    450                 455                 460

agc gca gag aac aag atg ctg aaa ctc agc ccc ctg agc gga agc aat     1440
Ser Ala Glu Asn Lys Met Leu Lys Leu Ser Pro Leu Ser Gly Ser Asn
465                 470                 475                 480

gac gtt tca gca ggc aac ttc tgg gcc aac gcg cgc ttc tct gca gat     1488
Asp Val Ser Ala Gly Asn Phe Trp Ala Asn Ala Arg Phe Ser Ala Asp
                485                 490                 495

gga acc agc cag cat ccc aac ctg cgg ggt gcg aag agc atg agc atg     1536
Gly Thr Ser Gln His Pro Asn Leu Arg Gly Ala Lys Ser Met Ser Met
            500                 505                 510
```

[0023]

```
gat gtg tat gtg cct gcc ccc acc aaa gtc tcc gtg gcc gct gtg ccc      1584
Asp Val Tyr Val Pro Ala Pro Thr Lys Val Ser Val Ala Ala Val Pro
        515                 520                 525

cag agc agc aag gat ggc tgg acc aac cct gcg cgt gct gtg atc gtg      1632
Gln Ser Ser Lys Asp Gly Trp Thr Asn Pro Ala Arg Ala Val Ile Val
    530                 535                 540

aac gca gac cag ttt gtg aaa cag gca gat ggc aag tac aaa gcc acc      1680
Asn Ala Asp Gln Phe Val Lys Gln Ala Asp Gly Lys Tyr Lys Ala Thr
545                 550                 555                 560

gtc acc ctg tct gac gaa gat gct ccc aac ctg aaa ctc att gcc gaa      1728
Val Thr Leu Ser Asp Glu Asp Ala Pro Asn Leu Lys Leu Ile Ala Glu
                565                 570                 575

gat gag acc gac aat gtg ctc tcg aac ctg atc ctc ttc att ggc acc      1776
Asp Glu Thr Asp Asn Val Leu Ser Asn Leu Ile Leu Phe Ile Gly Thr
            580                 585                 590

gag agc cag gaa gcc aat gac acc gtg tgg atc gac aac atc acc ttc      1824
Glu Ser Gln Glu Ala Asn Asp Thr Val Trp Ile Asp Asn Ile Thr Phe
        595                 600                 605

tct ggg gac cgt gtg gtg gtc cct gtg gaa cat gat ccc atc ggc acc      1872
Ser Gly Asp Arg Val Val Val Pro Val Glu His Asp Pro Ile Gly Thr
    610                 615                 620

gcc acc ctg ccc tcc aca ttt gaa gac agc acc cgc cag ggc tgg gac      1920
Ala Thr Leu Pro Ser Thr Phe Glu Asp Ser Thr Arg Gln Gly Trp Asp
625                 630                 635                 640

tgg gcc gga gaa tct ggc gtt aaa act gca ctg aag atc cag acc gcc      1968
Trp Ala Gly Glu Ser Gly Val Lys Thr Ala Leu Lys Ile Gln Thr Ala
                645                 650                 655

aat gcg tca aaa gcc ctg tcc tgg gat gtc atc tac cct gat gtg aaa      2016
Asn Ala Ser Lys Ala Leu Ser Trp Asp Val Ile Tyr Pro Asp Val Lys
            660                 665                 670

cct gcc gat ggc tgg gcc tct gct ccc cgt ctg gtg ctg gag aaa tcc      2064
Pro Ala Asp Gly Trp Ala Ser Ala Pro Arg Leu Val Leu Glu Lys Ser
        675                 680                 685

aac ctt acc cgc ggt gcc aac aag tac ctc gct ttc gac ctg tac ctg      2112
Asn Leu Thr Arg Gly Ala Asn Lys Tyr Leu Ala Phe Asp Leu Tyr Leu
    690                 695                 700

aag cca gat cgg gcc agc aaa ggg acc ctc tcc gtc aacctggctt           2158
Lys Pro Asp Arg Ala Ser Lys Gly Thr Leu Ser Val
705                 710                 715

ttggtcctcc gagcctgggg tactgggcac aggccagcga gaacgtcgac atcgacttga   2218

ccacgctggg ggccatgacc aaaactgccg atggactgta ccgcattgca gggaaattcg   2278
```

[0024]

atctggacaa gatcaatgac aacaaggtga ttgctgcaga cactgttctg ggcaaaatca 2338

ccctggtggt cgccgatgtg aacagcgatt acgccgggaa gatgttcctg gacaacgtgc 2398

gtttcaccaa cgaaccctaa 2418

<210> 8
<211> 716
<212> PRT
<213> Deinococcus sp.

<400> 8

Met Cys Leu Glu Cys Phe Phe Arg Val Leu Gln Ala Leu Lys Ala Thr
1 5 10 15

Pro Val Leu Leu Arg Gly Pro Cys Cys Thr Phe Arg Arg Phe Asn Pro
20 25 30

Met Met Pro Thr Pro Ser Lys Val Leu Pro Ala Ser Val Leu Leu Met
35 40 45

Val Ser Leu Leu Thr Ser Ser Cys Asn Leu Phe Gln Pro Pro Ala Pro
50 55 60

Asn Cys Thr Pro Lys Thr Thr Gly Ala Thr Val Pro Ala Gly Asp Tyr
65 70 75 80

Asp Pro Ala Ala Ser Glu Lys Ala Phe Pro Asp Leu Leu Thr Thr Ala
85 90 95

Ala Arg Lys Pro Ser Gln Ala Gly Ala Leu Gln Leu Val Gln Gln Asp
100 105 110

Cys Met Val Thr Leu Ala Asp Ser Ser Gly Lys Pro Ile Gln Leu Arg
115 120 125

Gly Met Ser Thr His Gly Leu Gln Trp Tyr Pro Glu Ile Val Asn Asp
130 135 140

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

Leu Ala Leu Tyr Val Gly Glu Gly Gly Tyr Ala Thr Lys Pro Glu Leu

[0025]

165 170 175

Lys Gln Lys Val Ile Glu Gly Ile Asp Phe Ala Ile Ala Asn Asp Met
180 185 190

Tyr Val Ile Val Asp Trp His Val His Ala Pro Gly Asp Pro Asn Ala
195 200 205

Asp Val Tyr Thr Asn Ala Lys Pro Leu Glu Phe Phe Lys Ser Ile Ala
210 215 220

Gln Lys Tyr Pro Asn Asn Lys His Ile Ile Tyr Glu Val Ala Asn Glu
225 230 235 240

Pro Asn Pro Gly Gln Ala Pro Gly Val Ser Asn Asp Ala Glu Gly Trp
245 250 255

Lys Lys Ile Lys Ser Tyr Ala Glu Pro Ile Ile Lys Met Leu Arg Asp
260 265 270

Leu Gly Asn Lys Asn Ile Val Ile Val Gly Thr Pro Asn Trp Ser Gln
275 280 285

Arg Pro Asp Leu Ala Ala Asp Asn Pro Ile Lys Asp Ser Ala Thr Leu
290 295 300

Tyr Thr Val His Phe Tyr Thr Gly Thr His Met Pro Ser Thr Asn Leu
305 310 315 320

Ala Asp Arg Gly Asn Val Met Ser Asn Ala Arg Tyr Ala Leu Glu His
325 330 335

Gly Val Gly Val Phe Ser Thr Glu Trp Gly Val Ser Glu Ala Ser Gly
340 345 350

Asn Asn Gly Pro Phe Leu Lys Glu Ala Asp Val Trp Leu Gln Phe Leu
355 360 365

Asn Lys His Asn Ile Ser Trp Val Asn Trp Ser Leu Thr Asn Lys Ala
370 375 380

Glu Thr Ser Ala Ala Phe Leu Pro Phe Pro Asn Gln Thr Ser Leu Asp

[0026]

385 390 395 400

Pro Gly Ala Asp Arg Leu Trp Thr Pro Ser Glu Leu Thr Leu Ser Gly
405 410 415

Glu Tyr Val Arg Ala Arg Ile Lys Gly Thr Lys Tyr Gln Pro Ile Asp
420 425 430

Arg Thr Ala Phe Thr Glu Val Ala Phe Asn Phe Asp Asn Asp Thr Thr
435 440 445

Gln Gly Phe Ala Leu Asn Pro Asp Ser Gly Val Lys Gly Ile Thr Val
450 455 460

Ser Ala Glu Asn Lys Met Leu Lys Leu Ser Pro Leu Ser Gly Ser Asn
465 470 475 480

Asp Val Ser Ala Gly Asn Phe Trp Ala Asn Ala Arg Phe Ser Ala Asp
485 490 495

Gly Thr Ser Gln His Pro Asn Leu Arg Gly Ala Lys Ser Met Ser Met
500 505 510

Asp Val Tyr Val Pro Ala Pro Thr Lys Val Ser Val Ala Ala Val Pro
515 520 525

Gln Ser Ser Lys Asp Gly Trp Thr Asn Pro Ala Arg Ala Val Ile Val
530 535 540

Asn Ala Asp Gln Phe Val Lys Gln Ala Asp Gly Lys Tyr Lys Ala Thr
545 550 555 560

Val Thr Leu Ser Asp Glu Asp Ala Pro Asn Leu Lys Leu Ile Ala Glu
565 570 575

Asp Glu Thr Asp Asn Val Leu Ser Asn Leu Ile Leu Phe Ile Gly Thr
580 585 590

Glu Ser Gln Glu Ala Asn Asp Thr Val Trp Ile Asp Asn Ile Thr Phe
595 600 605

Ser Gly Asp Arg Val Val Val Pro Val Glu His Asp Pro Ile Gly Thr

[0027]

610 615 620

Ala Thr Leu Pro Ser Thr Phe Glu Asp Ser Thr Arg Gln Gly Trp Asp
625 630 635 640

Trp Ala Gly Glu Ser Gly Val Lys Thr Ala Leu Lys Ile Gln Thr Ala
645 650 655

Asn Ala Ser Lys Ala Leu Ser Trp Asp Val Ile Tyr Pro Asp Val Lys
660 665 670

Pro Ala Asp Gly Trp Ala Ser Ala Pro Arg Leu Val Leu Glu Lys Ser
675 680 685

Asn Leu Thr Arg Gly Ala Asn Lys Tyr Leu Ala Phe Asp Leu Tyr Leu
690 695 700

Lys Pro Asp Arg Ala Ser Lys Gly Thr Leu Ser Val
705 710 715

<210> 9
<211> 2388
<212> DNA
<213> Deinococcus sp.

<220>
<221> CDS
<222> (1)..(2388)

<400> 9
gtg aaa ctc ggc gag gac acg acc gcc ccg tac gag ttc acg gtg aac 48
Val Lys Leu Gly Glu Asp Thr Thr Ala Pro Tyr Glu Phe Thr Val Asn
1 5 10 15

gcc gac cct ggc ctg aac ggc acg cac gtc tac tcc gcg cag gcg gtg 96
Ala Asp Pro Gly Leu Asn Gly Thr His Val Tyr Ser Ala Gln Ala Val
20 25 30

gcg ggc gac gcg gcc ggg atc tcc gcg ccg gtc agc gtg cag atc cgc 144
Ala Gly Asp Ala Ala Gly Ile Ser Ala Pro Val Ser Val Gln Ile Arg
35 40 45

atc gcg gac acc cgc acc acc gaa ctg ctc agc aac ggg gac ttc agc 192
Ile Ala Asp Thr Arg Thr Thr Glu Leu Leu Ser Asn Gly Asp Phe Ser
50 55 60

cag ggc ctg aac ccc tgg tgg act gcc gga acg gcc gcc agc acg acc 240
Gln Gly Leu Asn Pro Trp Trp Thr Ala Gly Thr Ala Ala Ser Thr Thr

[0028]

65 70 75 80

ggc ggt gag acc tgc ctg aac atc acg cag ccg ggc agc aac ccc tgg 288
Gly Gly Glu Thr Cys Leu Asn Ile Thr Gln Pro Gly Ser Asn Pro Trp
85 90 95

gac gtg ctg ttc ggg cag ggc ggc gtg ggc ctg aac gag ggc ggc acg 336
Asp Val Leu Phe Gly Gln Gly Gly Val Gly Leu Asn Glu Gly Gly Thr
100 105 110

tac acc ctg agc ttc acg gcg cgc gcc gcg cag ccc acg tcg ttc agg 384
Tyr Thr Leu Ser Phe Thr Ala Arg Ala Ala Gln Pro Thr Ser Phe Arg
115 120 125

acg ctg ctg cag ttc gac ggc gcg ccg tac acc aac tac ttc gtg cag 432
Thr Leu Leu Gln Phe Asp Gly Ala Pro Tyr Thr Asn Tyr Phe Val Gln
130 135 140

gac gcg gac gtg acc agc cag ccg aag acc ttc acg tcc acg ttc acg 480
Asp Ala Asp Val Thr Ser Gln Pro Lys Thr Phe Thr Ser Thr Phe Thr
145 150 155 160

atg gcg cag ccc agt gac gcg aag gcc gcg ttc cag ttc cag ctg ggc 528
Met Ala Gln Pro Ser Asp Ala Lys Ala Ala Phe Gln Phe Gln Leu Gly
165 170 175

gcc agg gcc gcc acg acc gtg tgc ttc agc cgc att tca ctg act ggc 576
Ala Arg Ala Ala Thr Thr Val Cys Phe Ser Arg Ile Ser Leu Thr Gly
180 185 190

cct gcc ttc ggc agc gcc gtg ccc gcc tcg ggt gcg gac gac ctg aag 624
Pro Ala Phe Gly Ser Ala Val Pro Ala Ser Gly Ala Asp Asp Leu Lys
195 200 205

ctg gtg cgg ctc aac cag acc ggg tac ctg ccg gac cgg ccg aaa ctg 672
Leu Val Arg Leu Asn Gln Thr Gly Tyr Leu Pro Asp Arg Pro Lys Leu
210 215 220

gcg gcc ctg ccg ttc gac tcg gac cgg ccg ctg ccg tgg act ctg ctg 720
Ala Ala Leu Pro Phe Asp Ser Asp Arg Pro Leu Pro Trp Thr Leu Leu
225 230 235 240

gac ggc acg cgc acg gtc gcc agt ggc gtg acc cgc gtg ttc ggc gcg 768
Asp Gly Thr Arg Thr Val Ala Ser Gly Val Thr Arg Val Phe Gly Ala
245 250 255

gac gcc gcg tcc ggc gag cac gtg cat cag gtg gac ttc agt gcc gtg 816
Asp Ala Ala Ser Gly Glu His Val His Gln Val Asp Phe Ser Ala Val
260 265 270

acc gcc ccg gcc gac ggg ctg gtg ctg gac gtc gcg ggt ttc cgc agc 864
Thr Ala Pro Ala Asp Gly Leu Val Leu Asp Val Ala Gly Phe Arg Ser
275 280 285

cac ccg ttc cgg atc ggg cgc gtg tac gac ggc ctg aaa cgc gac gcg 912
His Pro Phe Arg Ile Gly Arg Val Tyr Asp Gly Leu Lys Arg Asp Ala

[0029]

```
      290                 295                 300

ctg gcg tac ttc tac cac aac cgc agc ggc acg ccc atc aag gcg aag      960
Leu Ala Tyr Phe Tyr His Asn Arg Ser Gly Thr Pro Ile Lys Ala Lys
305                 310                 315                 320

tac gtc ggg gac gcc tgg gcg cgc ccg gcc ggt cac gcc ggg acc agc     1008
Tyr Val Gly Asp Ala Trp Ala Arg Pro Ala Gly His Ala Gly Thr Ser
                325                 330                 335

ccg aac cag ggg gac acg cgc gtc agc tgc ttc aag ggc acc gat cag     1056
Pro Asn Gln Gly Asp Thr Arg Val Ser Cys Phe Lys Gly Thr Asp Gln
            340                 345                 350

gcc ggg aac gtc tgg ccc ggc tgc ggg tac gaa ctg gac gcc agc ggc     1104
Ala Gly Asn Val Trp Pro Gly Cys Gly Tyr Glu Leu Asp Ala Ser Gly
        355                 360                 365

ggc tgg tac gac gcc ggg gat cac ggg aag tac gtc gtg aac ggc ggc     1152
Gly Trp Tyr Asp Ala Gly Asp His Gly Lys Tyr Val Val Asn Gly Gly
    370                 375                 380

gtg agc gtc tgg acg ctg ctg aac ctc gcc gag cgg ggc gcg cgc ctg     1200
Val Ser Val Trp Thr Leu Leu Asn Leu Ala Glu Arg Gly Ala Arg Leu
385                 390                 395                 400

aac ctc ccg gac gcg gac ggc agc ctg aat atc ccg gaa agc ggg aac     1248
Asn Leu Pro Asp Ala Asp Gly Ser Leu Asn Ile Pro Glu Ser Gly Asn
                405                 410                 415

ggc cgc agc gac ctg ctg gac gag gtc cgc tgg gag ctg gac ttc atg     1296
Gly Arg Ser Asp Leu Leu Asp Glu Val Arg Trp Glu Leu Asp Phe Met
            420                 425                 430

ctg cgg atg cag gtc ccg gat ggg cag acc ctg gtc ctg ccg cgc ggc     1344
Leu Arg Met Gln Val Pro Asp Gly Gln Thr Leu Val Leu Pro Arg Gly
        435                 440                 445

gac cag cgc ggc gcc ccg ctg acc ctg acg ccc acc ccg gcg ggc ggg     1392
Asp Gln Arg Gly Ala Pro Leu Thr Leu Thr Pro Thr Pro Ala Gly Gly
    450                 455                 460

ctg gtg cac cag aaa ctc acg gac gtc gcc tgg acc ggc ctg ccg ctg     1440
Leu Val His Gln Lys Leu Thr Asp Val Ala Trp Thr Gly Leu Pro Leu
465                 470                 475                 480

cgc ccc gat cag gat ccg cag ccg cgc gcg ctg tac tac ccc acg acc     1488
Arg Pro Asp Gln Asp Pro Gln Pro Arg Ala Leu Tyr Tyr Pro Thr Thr
                485                 490                 495

gcc gcg acc ctg aac ctc gcg ggc gtg gcg gcg cag tgc gcc cgc gtg     1536
Ala Ala Thr Leu Asn Leu Ala Gly Val Ala Ala Gln Cys Ala Arg Val
            500                 505                 510

tac cgc gcc tcg gac ccg gcc ttc gcg gtc cgc tgc ctg agc gcc gcg     1584
Tyr Arg Ala Ser Asp Pro Ala Phe Ala Val Arg Cys Leu Ser Ala Ala
```

[0030]

```
          515                 520                 525

cgc cgc gcg tgg cag gcc gcg cgg gcc gcg ccg gac gtg tac gcg tac      1632
Arg Arg Ala Trp Gln Ala Ala Arg Ala Ala Pro Asp Val Tyr Ala Tyr
    530                 535                 540

gac ctg ttc gtg ggc ggc ggc ccg tac gac gat acg gac gtc agc gac      1680
Asp Leu Phe Val Gly Gly Gly Pro Tyr Asp Asp Thr Asp Val Ser Asp
545                 550                 555                 560

gag ttc tac tgg gcg gcg gcg gaa ctg tac gcc acg acc ggc gag gcg      1728
Glu Phe Tyr Trp Ala Ala Ala Glu Leu Tyr Ala Thr Thr Gly Glu Ala
                565                 570                 575

gcg ttc ctg gag gcg ctg cgg gcc agt ccg ctg ttc ctg cag atg ccc      1776
Ala Phe Leu Glu Ala Leu Arg Ala Ser Pro Leu Phe Leu Gln Met Pro
            580                 585                 590

gag ggg cgt gaa ctg ggc tgg agt gac ctg acg gca gcg ggc acc ctg      1824
Glu Gly Arg Glu Leu Gly Trp Ser Asp Leu Thr Ala Ala Gly Thr Leu
        595                 600                 605

acg ctc gcc agc gtg ccc acg gcg ctg ccc ggc gcg gac gtg cag cag      1872
Thr Leu Ala Ser Val Pro Thr Ala Leu Pro Gly Ala Asp Val Gln Gln
    610                 615                 620

gcc cgc gcg aac gtc gtg gcg gcg gcg cgg gcg ttc cgg gac gcg gca      1920
Ala Arg Ala Asn Val Val Ala Ala Ala Arg Ala Phe Arg Asp Ala Ala
625                 630                 635                 640

ggc acg cag ggg tac cgc ctg ccg atg acc ggc gcc gag gcc acc tgg      1968
Gly Thr Gln Gly Tyr Arg Leu Pro Met Thr Gly Ala Glu Ala Thr Trp
                645                 650                 655

ggc tcg aac agt ggc gtg ctg aac cgc tcg gtc gtg atg ggc gcc gcg      2016
Gly Ser Asn Ser Gly Val Leu Asn Arg Ser Val Val Met Gly Ala Ala
            660                 665                 670

tgg gac ttc acg ggt gac gat tcg ttc gtg aat gtc gtg ctg gag ggc      2064
Trp Asp Phe Thr Gly Asp Asp Ser Phe Val Asn Val Val Leu Glu Gly
        675                 680                 685

ctg aac tac ctg ctg gga cgc aac ccg atg gac aag tcg tac gtg tcc      2112
Leu Asn Tyr Leu Leu Gly Arg Asn Pro Met Asp Lys Ser Tyr Val Ser
    690                 695                 700

ggg tac ggc gag cgt ccg ctg ctg aac ccg cac cac cgc ttc tgg gcg      2160
Gly Tyr Gly Glu Arg Pro Leu Leu Asn Pro His His Arg Phe Trp Ala
705                 710                 715                 720

cgg tcg ctg gac gcg gcg ctg ccc ggc ccg ccg cgt ggg gtg gtg tcg      2208
Arg Ser Leu Asp Ala Ala Leu Pro Gly Pro Pro Arg Gly Val Val Ser
                725                 730                 735

ggc ggc ccg aac agc gtg aac ttc agt gat ccg gtc gcg gcg aaa ctc      2256
Gly Gly Pro Asn Ser Val Asn Phe Ser Asp Pro Val Ala Ala Lys Leu
```

[0031]

740 745 750

agg ggc cgc tgc gtg ggc ctg cgc tgc tac acc gac gac atc ggc gcg 2304
Arg Gly Arg Cys Val Gly Leu Arg Cys Tyr Thr Asp Asp Ile Gly Ala
755 760 765

tac acc atg aac gag gtg acc atc aac tgg aac gcg ccg ctg gcg tgg 2352
Tyr Thr Met Asn Glu Val Thr Ile Asn Trp Asn Ala Pro Leu Ala Trp
770 775 780

gtg gcg gcg ttc gtg gag cac agc acc cga cgc taa 2388
Val Ala Ala Phe Val Glu His Ser Thr Arg Arg
785 790 795

<210> 10
<211> 795
<212> PRT
<213> Deinococcus sp.

<400> 10

Val Lys Leu Gly Glu Asp Thr Thr Ala Pro Tyr Glu Phe Thr Val Asn
1 5 10 15

Ala Asp Pro Gly Leu Asn Gly Thr His Val Tyr Ser Ala Gln Ala Val
20 25 30

Ala Gly Asp Ala Ala Gly Ile Ser Ala Pro Val Ser Val Gln Ile Arg
35 40 45

Ile Ala Asp Thr Arg Thr Thr Glu Leu Leu Ser Asn Gly Asp Phe Ser
50 55 60

Gln Gly Leu Asn Pro Trp Trp Thr Ala Gly Thr Ala Ala Ser Thr Thr
65 70 75 80

Gly Gly Glu Thr Cys Leu Asn Ile Thr Gln Pro Gly Ser Asn Pro Trp
85 90 95

Asp Val Leu Phe Gly Gln Gly Gly Val Gly Leu Asn Glu Gly Gly Thr
100 105 110

Tyr Thr Leu Ser Phe Thr Ala Arg Ala Ala Gln Pro Thr Ser Phe Arg
115 120 125

Thr Leu Leu Gln Phe Asp Gly Ala Pro Tyr Thr Asn Tyr Phe Val Gln
130 135 140

[0032]

Asp Ala Asp Val Thr Ser Gln Pro Lys Thr Phe Thr Ser Thr Phe Thr
145 150 155 160

Met Ala Gln Pro Ser Asp Ala Lys Ala Ala Phe Gln Phe Gln Leu Gly
165 170 175

Ala Arg Ala Ala Thr Thr Val Cys Phe Ser Arg Ile Ser Leu Thr Gly
180 185 190

Pro Ala Phe Gly Ser Ala Val Pro Ala Ser Gly Ala Asp Asp Leu Lys
195 200 205

Leu Val Arg Leu Asn Gln Thr Gly Tyr Leu Pro Asp Arg Pro Lys Leu
210 215 220

Ala Ala Leu Pro Phe Asp Ser Asp Arg Pro Leu Pro Trp Thr Leu Leu
225 230 235 240

Asp Gly Thr Arg Thr Val Ala Ser Gly Val Thr Arg Val Phe Gly Ala
245 250 255

Asp Ala Ala Ser Gly Glu His Val His Gln Val Asp Phe Ser Ala Val
260 265 270

Thr Ala Pro Ala Asp Gly Leu Val Leu Asp Val Ala Gly Phe Arg Ser
275 280 285

His Pro Phe Arg Ile Gly Arg Val Tyr Asp Gly Leu Lys Arg Asp Ala
290 295 300

Leu Ala Tyr Phe Tyr His Asn Arg Ser Gly Thr Pro Ile Lys Ala Lys
305 310 315 320

Tyr Val Gly Asp Ala Trp Ala Arg Pro Ala Gly His Ala Gly Thr Ser
325 330 335

Pro Asn Gln Gly Asp Thr Arg Val Ser Cys Phe Lys Gly Thr Asp Gln
340 345 350

Ala Gly Asn Val Trp Pro Gly Cys Gly Tyr Glu Leu Asp Ala Ser Gly
355 360 365

[0033]

Gly Trp Tyr Asp Ala Gly Asp His Gly Lys Tyr Val Val Asn Gly Gly
370 375 380

Val Ser Val Trp Thr Leu Leu Asn Leu Ala Glu Arg Gly Ala Arg Leu
385 390 395 400

Asn Leu Pro Asp Ala Asp Gly Ser Leu Asn Ile Pro Glu Ser Gly Asn
405 410 415

Gly Arg Ser Asp Leu Leu Asp Glu Val Arg Trp Glu Leu Asp Phe Met
420 425 430

Leu Arg Met Gln Val Pro Asp Gly Gln Thr Leu Val Leu Pro Arg Gly
435 440 445

Asp Gln Arg Gly Ala Pro Leu Thr Leu Thr Pro Thr Pro Ala Gly Gly
450 455 460

Leu Val His Gln Lys Leu Thr Asp Val Ala Trp Thr Gly Leu Pro Leu
465 470 475 480

Arg Pro Asp Gln Asp Pro Gln Pro Arg Ala Leu Tyr Tyr Pro Thr Thr
485 490 495

Ala Ala Thr Leu Asn Leu Ala Gly Val Ala Ala Gln Cys Ala Arg Val
500 505 510

Tyr Arg Ala Ser Asp Pro Ala Phe Ala Val Arg Cys Leu Ser Ala Ala
515 520 525

Arg Arg Ala Trp Gln Ala Ala Arg Ala Ala Pro Asp Val Tyr Ala Tyr
530 535 540

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

Glu Phe Tyr Trp Ala Ala Ala Glu Leu Tyr Ala Thr Thr Gly Glu Ala
565 570 575

Ala Phe Leu Glu Ala Leu Arg Ala Ser Pro Leu Phe Leu Gln Met Pro
580 585 590

[0034]

Glu Gly Arg Glu Leu Gly Trp Ser Asp Leu Thr Ala Ala Gly Thr Leu
595 600 605

Thr Leu Ala Ser Val Pro Thr Ala Leu Pro Gly Ala Asp Val Gln Gln
610 615 620

Ala Arg Ala Asn Val Val Ala Ala Ala Arg Ala Phe Arg Asp Ala Ala
625 630 635 640

Gly Thr Gln Gly Tyr Arg Leu Pro Met Thr Gly Ala Glu Ala Thr Trp
645 650 655

Gly Ser Asn Ser Gly Val Leu Asn Arg Ser Val Val Met Gly Ala Ala
660 665 670

Trp Asp Phe Thr Gly Asp Asp Ser Phe Val Asn Val Val Leu Glu Gly
675 680 685

Leu Asn Tyr Leu Leu Gly Arg Asn Pro Met Asp Lys Ser Tyr Val Ser
690 695 700

Gly Tyr Gly Glu Arg Pro Leu Leu Asn Pro His His Arg Phe Trp Ala
705 710 715 720

Arg Ser Leu Asp Ala Ala Leu Pro Gly Pro Pro Arg Gly Val Val Ser
725 730 735

Gly Gly Pro Asn Ser Val Asn Phe Ser Asp Pro Val Ala Ala Lys Leu
740 745 750

Arg Gly Arg Cys Val Gly Leu Arg Cys Tyr Thr Asp Asp Ile Gly Ala
755 760 765

Tyr Thr Met Asn Glu Val Thr Ile Asn Trp Asn Ala Pro Leu Ala Trp
770 775 780

Val Ala Ala Phe Val Glu His Ser Thr Arg Arg
785 790 795

<210> 11
<211> 1320

[0035]

<212> DNA
<213> Deinococcus sp.

<220>
<221> CDS
<222> (1)..(1320)

<400> 11

```
atg aac cct gtt tca ggc ttt ctt caa agg gtt ctt gct ttt cca gag      48
Met Asn Pro Val Ser Gly Phe Leu Gln Arg Val Leu Ala Phe Pro Glu
1               5                  10                  15

caa aaa agc aca gcc cct ttc cca gag cag aaa gaa agg cat gag aaa      96
Gln Lys Ser Thr Ala Pro Phe Pro Glu Gln Lys Glu Arg His Glu Lys
            20                  25                  30

tgc cat tca tct gaa ccc ggt aca gcc aca gaa tgc atg aca gaa aaa     144
Cys His Ser Ser Glu Pro Gly Thr Ala Thr Glu Cys Met Thr Glu Lys
        35                  40                  45

cat tcc atc gtg aaa aaa acg tca tac aat aaa agc atg cat ttc caa     192
His Ser Ile Val Lys Lys Thr Ser Tyr Asn Lys Ser Met His Phe Gln
    50                  55                  60

ccc agc gac gac agc caa ctg atc ctt tcc cat gtg cgc gaa ttc gca     240
Pro Ser Asp Asp Ser Gln Leu Ile Leu Ser His Val Arg Glu Phe Ala
65                  70                  75                  80

cag cat gaa att gcc cct ctg gcc gcc cag tac gac cgc tct gga gaa     288
Gln His Glu Ile Ala Pro Leu Ala Ala Gln Tyr Asp Arg Ser Gly Glu
                85                  90                  95

tac ccc tgg gag caa ctg cgc aaa ctg gcc gac atg ggc ctt ctg ggg     336
Tyr Pro Trp Glu Gln Leu Arg Lys Leu Ala Asp Met Gly Leu Leu Gly
            100                 105                 110

gcc acc atc ccc gag gaa tat gga ggt gca ggt ctg gat tcg gtg act     384
Ala Thr Ile Pro Glu Glu Tyr Gly Gly Ala Gly Leu Asp Ser Val Thr
        115                 120                 125

tac gcc ctc tgt ctg gaa gaa atc gct gct gcc gat gcc agt gtg gcc     432
Tyr Ala Leu Cys Leu Glu Glu Ile Ala Ala Ala Asp Ala Ser Val Ala
    130                 135                 140

gtg atc atg agc gtc cag aac gga ctt cca gag caa atg atc ctg aag     480
Val Ile Met Ser Val Gln Asn Gly Leu Pro Glu Gln Met Ile Leu Lys
145                 150                 155                 160

tac ggc tct gaa gca caa aaa cag cag tac ctc gct cct ctg gcg cag     528
Tyr Gly Ser Glu Ala Gln Lys Gln Gln Tyr Leu Ala Pro Leu Ala Gln
                165                 170                 175

ggc aaa aaa att ggg gcc ttc tgt ctc act gag ccg aat gca ggt tcg     576
Gly Lys Lys Ile Gly Ala Phe Cys Leu Thr Glu Pro Asn Ala Gly Ser
            180                 185                 190
```

[0036]

```
gat gca gcc agt ttg cgt ctg agt gct caa aaa gat ggt gac agt tgg      624
Asp Ala Ala Ser Leu Arg Leu Ser Ala Gln Lys Asp Gly Asp Ser Trp
        195                 200                 205

gtt ctg aac ggc cag aaa gca tgg atc acc tct gga gga cag gcg gac      672
Val Leu Asn Gly Gln Lys Ala Trp Ile Thr Ser Gly Gly Gln Ala Asp
    210                 215                 220

acc tat ctg gtg atg gcc cgc acc tct ggc tct ggg gca aaa ggg gtc      720
Thr Tyr Leu Val Met Ala Arg Thr Ser Gly Ser Gly Ala Lys Gly Val
225                 230                 235                 240

acc tgt ttt gtg gtg gaa aaa gac acc ccc ggc ctc tct ttt ggc aaa      768
Thr Cys Phe Val Val Glu Lys Asp Thr Pro Gly Leu Ser Phe Gly Lys
                245                 250                 255

ccc gag gag aaa atg ggc ctt cat gct gcc cac acc acc aca gtc aat      816
Pro Glu Glu Lys Met Gly Leu His Ala Ala His Thr Thr Thr Val Asn
            260                 265                 270

ttt gaa gag ctg cgg att cca gat gca caa cgc att ggc gat gaa ggg      864
Phe Glu Glu Leu Arg Ile Pro Asp Ala Gln Arg Ile Gly Asp Glu Gly
        275                 280                 285

caa ggt ctg atc atc gcc ctg agc agt ctg gat tcc ggg cgc att ggc      912
Gln Gly Leu Ile Ile Ala Leu Ser Ser Leu Asp Ser Gly Arg Ile Gly
    290                 295                 300

att gcc atg cag tcc atc ggg att gcc aga gcg gct ctg gaa gca gca      960
Ile Ala Met Gln Ser Ile Gly Ile Ala Arg Ala Ala Leu Glu Ala Ala
305                 310                 315                 320

gca aag tac gcc ctg acc cgc gag caa ttt ggc aag aaa att gca gag     1008
Ala Lys Tyr Ala Leu Thr Arg Glu Gln Phe Gly Lys Lys Ile Ala Glu
                325                 330                 335

ttt gag ggg gtc tct ttc aag att gcg gaa atg gct tcc cgc att gag     1056
Phe Glu Gly Val Ser Phe Lys Ile Ala Glu Met Ala Ser Arg Ile Glu
            340                 345                 350

gct gcc cgt ctg gtg gcc ctc aaa gct gca tgg ctg cgc gat cag ggg     1104
Ala Ala Arg Leu Val Ala Leu Lys Ala Ala Trp Leu Arg Asp Gln Gly
        355                 360                 365

aag aag tac acc aaa gag gct tcc atg gcc aag ttt ctg gct tcg gat     1152
Lys Lys Tyr Thr Lys Glu Ala Ser Met Ala Lys Phe Leu Ala Ser Asp
    370                 375                 380

gct gcg gtg tac gtg acc cgt gaa gcg gtg cag att ttt ggg ggc aac     1200
Ala Ala Val Tyr Val Thr Arg Glu Ala Val Gln Ile Phe Gly Gly Asn
385                 390                 395                 400

ggg tac agc cgg gaa tac ccg gtg gaa cgg tat tac cgg gac gcc aaa     1248
Gly Tyr Ser Arg Glu Tyr Pro Val Glu Arg Tyr Tyr Arg Asp Ala Lys
                405                 410                 415
```

[0037]

atc acc gaa atc tat gaa ggc acc aac gag atc cag aaa ctg gtg att 1296
Ile Thr Glu Ile Tyr Glu Gly Thr Asn Glu Ile Gln Lys Leu Val Ile
420 425 430

tcc cgt cag gtg ttc tca gag tat 1320
Ser Arg Gln Val Phe Ser Glu Tyr
435 440

<210> 12
<211> 440
<212> PRT
<213> Deinococcus sp.

<400> 12

Met Asn Pro Val Ser Gly Phe Leu Gln Arg Val Leu Ala Phe Pro Glu
1 5 10 15

Gln Lys Ser Thr Ala Pro Phe Pro Glu Gln Lys Glu Arg His Glu Lys
20 25 30

Cys His Ser Ser Glu Pro Gly Thr Ala Thr Glu Cys Met Thr Glu Lys
35 40 45

His Ser Ile Val Lys Lys Thr Ser Tyr Asn Lys Ser Met His Phe Gln
50 55 60

Pro Ser Asp Asp Ser Gln Leu Ile Leu Ser His Val Arg Glu Phe Ala
65 70 75 80

Gln His Glu Ile Ala Pro Leu Ala Ala Gln Tyr Asp Arg Ser Gly Glu
85 90 95

Tyr Pro Trp Glu Gln Leu Arg Lys Leu Ala Asp Met Gly Leu Leu Gly
100 105 110

Ala Thr Ile Pro Glu Glu Tyr Gly Gly Ala Gly Leu Asp Ser Val Thr
115 120 125

Tyr Ala Leu Cys Leu Glu Glu Ile Ala Ala Ala Asp Ala Ser Val Ala
130 135 140

Val Ile Met Ser Val Gln Asn Gly Leu Pro Glu Gln Met Ile Leu Lys
145 150 155 160

[0038]

Tyr Gly Ser Glu Ala Gln Lys Gln Gln Tyr Leu Ala Pro Leu Ala Gln
165 170 175

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

Asp Ala Ala Ser Leu Arg Leu Ser Ala Gln Lys Asp Gly Asp Ser Trp
195 200 205

Val Leu Asn Gly Gln Lys Ala Trp Ile Thr Ser Gly Gly Gln Ala Asp
210 215 220

Thr Tyr Leu Val Met Ala Arg Thr Ser Gly Ser Gly Ala Lys Gly Val
225 230 235 240

Thr Cys Phe Val Val Glu Lys Asp Thr Pro Gly Leu Ser Phe Gly Lys
245 250 255

Pro Glu Glu Lys Met Gly Leu His Ala Ala His Thr Thr Thr Val Asn
260 265 270

Phe Glu Glu Leu Arg Ile Pro Asp Ala Gln Arg Ile Gly Asp Glu Gly
275 280 285

Gln Gly Leu Ile Ile Ala Leu Ser Ser Leu Asp Ser Gly Arg Ile Gly
290 295 300

Ile Ala Met Gln Ser Ile Gly Ile Ala Arg Ala Ala Leu Glu Ala Ala
305 310 315 320

Ala Lys Tyr Ala Leu Thr Arg Glu Gln Phe Gly Lys Lys Ile Ala Glu
325 330 335

Phe Glu Gly Val Ser Phe Lys Ile Ala Glu Met Ala Ser Arg Ile Glu
340 345 350

Ala Ala Arg Leu Val Ala Leu Lys Ala Ala Trp Leu Arg Asp Gln Gly
355 360 365

Lys Lys Tyr Thr Lys Glu Ala Ser Met Ala Lys Phe Leu Ala Ser Asp
370 375 380

[0039]

Ala Ala Val Tyr Val Thr Arg Glu Ala Val Gln Ile Phe Gly Gly Asn
385 390 395 400

Gly Tyr Ser Arg Glu Tyr Pro Val Glu Arg Tyr Tyr Arg Asp Ala Lys
405 410 415

Ile Thr Glu Ile Tyr Glu Gly Thr Asn Glu Ile Gln Lys Leu Val Ile
420 425 430

Ser Arg Gln Val Phe Ser Glu Tyr
435 440

<210> 13
<211> 1026
<212> DNA
<213> Deinococcus sp.

<220>
<221> CDS
<222> (1)..(1026)

<400> 13

```
atg caa gct gtt gca ctc acc aga cgg ggc aac att gac gcc ctt gaa   48
Met Gln Ala Val Ala Leu Thr Arg Arg Gly Asn Ile Asp Ala Leu Glu
1               5                  10                  15

ccc atc cgc ctg ccc att tca gag cct caa gct gga gaa gtc ctg gtg   96
Pro Ile Arg Leu Pro Ile Ser Glu Pro Gln Ala Gly Glu Val Leu Val
            20                  25                  30

cgc atc cgt gca gtg gcc ctc aac cac ctg gat gtg tgg gtg cgc aag  144
Arg Ile Arg Ala Val Ala Leu Asn His Leu Asp Val Trp Val Arg Lys
        35                  40                  45

ggg gtc gcc agc ccg aaa ctt ccc ctg cca cac ctg ctc ggc tca gac  192
Gly Val Ala Ser Pro Lys Leu Pro Leu Pro His Leu Leu Gly Ser Asp
    50                  55                  60

att gcg gga gag gtg gct gca atg ggt cca ggg gtt gaa ggt ttg tca  240
Ile Ala Gly Glu Val Ala Ala Met Gly Pro Gly Val Glu Gly Leu Ser
65                  70                  75                  80

gag ggc aca aaa gtg atg ctg aac ccg ggc gta tcc tgc ggc cac tgc  288
Glu Gly Thr Lys Val Met Leu Asn Pro Gly Val Ser Cys Gly His Cys
                85                  90                  95

gaa cgc tgc ctc agt gga cac gac aac ctg tgc cgc cac tac caa att  336
Glu Arg Cys Leu Ser Gly His Asp Asn Leu Cys Arg His Tyr Gln Ile
            100                 105                 110
```

[0040]

```
ctc ggg gaa cac cgc tgg gga ggg tat gcc cag tac atc agc atc ccg      384
Leu Gly Glu His Arg Trp Gly Gly Tyr Ala Gln Tyr Ile Ser Ile Pro
        115                 120                 125

aga acc aat gtg ctg ccc atg cca gag ggc ctt gac ttt gta gag gcc      432
Arg Thr Asn Val Leu Pro Met Pro Glu Gly Leu Asp Phe Val Glu Ala
    130                 135                 140

gcc tct gtc ccc ctg tcg gcc ctc acc gct tac cag atg gtg ttc gac      480
Ala Ser Val Pro Leu Ser Ala Leu Thr Ala Tyr Gln Met Val Phe Asp
145                 150                 155                 160

cgt gca cag ctg aaa ccc tgg gaa acc gtc ctg atc ctg gcg gcg gca      528
Arg Ala Gln Leu Lys Pro Trp Glu Thr Val Leu Ile Leu Ala Ala Ala
                165                 170                 175

agt ggg gtc agc gtc aac ctg atc cag ctc tgc aaa ctg gtg ggt gca      576
Ser Gly Val Ser Val Asn Leu Ile Gln Leu Cys Lys Leu Val Gly Ala
            180                 185                 190

aaa gtc atc gct gtg gcc agc acc cct gaa aag cag gcc ctg gca ctg      624
Lys Val Ile Ala Val Ala Ser Thr Pro Glu Lys Gln Ala Leu Ala Leu
        195                 200                 205

aaa ctt ggt gca gat cac gtg atc ggt tca cat gaa gac cag gct cag      672
Lys Leu Gly Ala Asp His Val Ile Gly Ser His Glu Asp Gln Ala Gln
    210                 215                 220

gcc gtc aaa gcc ctg act gca gga gaa ggg gca gac gtg gtg ttt gac      720
Ala Val Lys Ala Leu Thr Ala Gly Glu Gly Ala Asp Val Val Phe Asp
225                 230                 235                 240

cac acc gga gcg gac aac tgg caa cgc agc ctg aaa agc ctg aag tgg      768
His Thr Gly Ala Asp Asn Trp Gln Arg Ser Leu Lys Ser Leu Lys Trp
                245                 250                 255

gga ggc cgt ctg gtc acc tgt ggg gca acc agt gga cat gaa gcc gtg      816
Gly Gly Arg Leu Val Thr Cys Gly Ala Thr Ser Gly His Glu Ala Val
            260                 265                 270

acc ccc ctc aac tgg gtg ttt ttc aag cag ctc agc atc ctg ggt tcc      864
Thr Pro Leu Asn Trp Val Phe Phe Lys Gln Leu Ser Ile Leu Gly Ser
        275                 280                 285

acc atg ggc tcc aaa gca gac ctc cac aaa att cag aag ttt gtg caa      912
Thr Met Gly Ser Lys Ala Asp Leu His Lys Ile Gln Lys Phe Val Gln
    290                 295                 300

gaa gga aaa ctc agg cct gtg gtg ggc cat gtt ctg gac ttt gct cag      960
Glu Gly Lys Leu Arg Pro Val Val Gly His Val Leu Asp Phe Ala Gln
305                 310                 315                 320

gca aga gag gcc cat ggg ctt ctg gaa tcc agg cag gct ctg ggc aag     1008
Ala Arg Glu Ala His Gly Leu Leu Glu Ser Arg Gln Ala Leu Gly Lys
                325                 330                 335
```

[0041]

gtg gtt ttg agg gtc ccg 1026
Val Val Leu Arg Val Pro
340

<210> 14
<211> 342
<212> PRT
<213> Deinococcus sp.

<400> 14

Met Gln Ala Val Ala Leu Thr Arg Arg Gly Asn Ile Asp Ala Leu Glu
1 5 10 15

Pro Ile Arg Leu Pro Ile Ser Glu Pro Gln Ala Gly Glu Val Leu Val
20 25 30

Arg Ile Arg Ala Val Ala Leu Asn His Leu Asp Val Trp Val Arg Lys
35 40 45

Gly Val Ala Ser Pro Lys Leu Pro Leu Pro His Leu Leu Gly Ser Asp
50 55 60

Ile Ala Gly Glu Val Ala Ala Met Gly Pro Gly Val Glu Gly Leu Ser
65 70 75 80

Glu Gly Thr Lys Val Met Leu Asn Pro Gly Val Ser Cys Gly His Cys
85 90 95

Glu Arg Cys Leu Ser Gly His Asp Asn Leu Cys Arg His Tyr Gln Ile
100 105 110

Leu Gly Glu His Arg Trp Gly Gly Tyr Ala Gln Tyr Ile Ser Ile Pro
115 120 125

Arg Thr Asn Val Leu Pro Met Pro Glu Gly Leu Asp Phe Val Glu Ala
130 135 140

Ala Ser Val Pro Leu Ser Ala Leu Thr Ala Tyr Gln Met Val Phe Asp
145 150 155 160

Arg Ala Gln Leu Lys Pro Trp Glu Thr Val Leu Ile Leu Ala Ala Ala
165 170 175

[0042]

Ser Gly Val Ser Val Asn Leu Ile Gln Leu Cys Lys Leu Val Gly Ala
180 185 190

Lys Val Ile Ala Val Ala Ser Thr Pro Glu Lys Gln Ala Leu Ala Leu
195 200 205

Lys Leu Gly Ala Asp His Val Ile Gly Ser His Glu Asp Gln Ala Gln
210 215 220

Ala Val Lys Ala Leu Thr Ala Gly Glu Gly Ala Asp Val Val Phe Asp
225 230 235 240

His Thr Gly Ala Asp Asn Trp Gln Arg Ser Leu Lys Ser Leu Lys Trp
245 250 255

Gly Gly Arg Leu Val Thr Cys Gly Ala Thr Ser Gly His Glu Ala Val
260 265 270

Thr Pro Leu Asn Trp Val Phe Phe Lys Gln Leu Ser Ile Leu Gly Ser
275 280 285

Thr Met Gly Ser Lys Ala Asp Leu His Lys Ile Gln Lys Phe Val Gln
290 295 300

Glu Gly Lys Leu Arg Pro Val Val Gly His Val Leu Asp Phe Ala Gln
305 310 315 320

Ala Arg Glu Ala His Gly Leu Leu Glu Ser Arg Gln Ala Leu Gly Lys
325 330 335

Val Val Leu Arg Val Pro
340

<210> 15
<211> 1710
<212> DNA
<213> Zymomonas mobilis

<400> 15
atgagttata ctgtcggtac ctatttagcg gagcggcttg tccagattgg tctcaagcat 60

cacttcgcag tcgcgggcga ctacaacctc gtccttcttg acaacctgct tttgaacaaa 120

aacatggagc aggtttattg ctgtaacgaa ctgaactgcg gtttcagtgc agaaggttat 180

[0043]

gctcgtgcca aaggcgcagc agcagccgtc gttacctaca gcgtcggtgc gctttccgca 240

tttgatgcta tcggtggcgc ctatgcagaa aaccttccgg ttatcctgat ctccggtgct 300

ccgaacaaca atgatcacgc tgctggtcac gtgttgcatc acgctcttgg caaaaccgac 360

tatcactatc agttggaaat ggccaagaac atcacggccg ccgctgaagc gatttacacc 420

ccggaagaag ctccggctaa aatcgatcac gtgattaaaa ctgctcttcg tgagaagaag 480

ccggtttatc tcgaaatcgc ttgcaacatt gcttccatgc cctgcgccgc tcctggaccg 540

gcaagcgcat tgttcaatga cgaagccagc gacgaagctt ctttgaatgc agcggttgaa 600

gaaaccctga aattcatcgc caaccgcgac aaagttgccg tcctcgtcgg cagcaagctg 660

cgcgcagctg gtgctgaaga agctgctgtc aaatttgctg atgctctcgg tggcgcagtt 720

gctaccatgg ctgctgcaaa aagcttcttc ccagaagaaa acccgcatta catcggcacc 780

tcatggggtg aagtcagcta tccgggcgtt gaaaagacga tgaaagaagc cgatgcggtt 840

atcgctctgg ctcctgtctt caacgactac tccaccactg gttggacgga tattcctgat 900

cctaagaaac tggttctcgc tgaaccgcgt tctgtcgtcg ttaacggcat tcgcttcccc 960

agcgtccatc tgaaagacta tctgacccgt ttggctcaga aagtttccaa gaaaaccggt 1020

gcattggact tcttcaaatc cctcaatgca ggtgaactga agaaagccgc tccggctgat 1080

ccgagtgctc cgttggtcaa cgcagaaatc gcccgtcagg tcgaagctct tctgaccccg 1140

aacacgacgg ttattgctga aaccggtgac tcttggttca atgctcagcg catgaagctc 1200

ccgaacggtg ctcgcgttga atatgaaatg cagtggggtc acattggttg gtccgttcct 1260

gccgccttcg gttatgccgt cggtgctccg gaacgtcgca acatcctcat ggttggtgat 1320

ggttccttcc agctgacggc tcaggaagtc gctcagatgg ttcgcctgaa actgccggtt 1380

atcatcttct tgatcaataa ctatggttac accatcgaag ttatgatcca tgatggtccg 1440

tacaacaaca tcaagaactg ggattatgcc ggtctgatgg aagtgttcaa cggtaacggt 1500

ggttatgaca gcggtgctgg taaaggcctg aaggctaaaa ccggtggcga actggcagaa 1560

gctatcaagg ttgctctggc aaacaccgac ggcccaaccc tgatcgaatg cttcatcggt 1620

cgtgaagact gcactgaaga attggtcaaa tggggtaagc gcgttgctgc cgccaacagc 1680

cgtaagcctg ttaacaagct cctctagtag 1710

<210> 16
<211> 1707
<212> DNA

[0044]

<213> 人工

<220>
<223> Deinococcus codon optimized PDC gene derived from Zm

<400> 16
atgagctaca cggtgggcac gtacctggcg gagcggctgg tgcagatcgg cctgaagcac 60

cactttgcgg tggcgggcga ctacaacctg gtgctgctgg acaacctgct gctgaacaag 120

aacatggagc aggtgtactg ctgcaacgag ctgaactgcg gctttagcgc cgagggctac 180

gcccgtgcca agggcgctgc cgccgccgtg gtgacctaca gcgtgggcgc cctgagcgcg 240

ttcgacgcca tcggcggcgc ctacgcggag aacctgccgg tgatcctgat cagcggtgcc 300

cccaacaaca acgaccacgc cgcgggccac gtgctgcacc acgccctggg caagaccgac 360

taccactacc agctggagat ggcgaagaac atcaccgctg ccgccgaggc catctacacg 420

ccggaggagg cccccgcgaa gatcgaccac gtgatcaaga ccgccctgcg cgagaagaag 480

ccggtgtacc tggagatcgc gtgcaacatc gccagcatgc cctgcgcggc ccccggcccg 540

gccagcgcgc tgttcaacga cgaggcgagc gacgaggcca gcctgaacgc ggccgtggag 600

gagaccctga agtttatcgc gaaccgcgac aaggtggccg tgctggtggg cagcaagctg 660

cgggcggccg gcgccgagga ggcggccgtg aagttcgccg acgcgctggg cggcgccgtg 720

gcgacgatgg ccgccgccaa gagcttcttt ccggaggaga acccccacta catcggcacc 780

agctggggcg aagtgagcta cccgggcgtg gagaagacga tgaaggaggc cgacgcggtg 840

atcgccctgg cgcccgtgtt taacgactac agcaccacgg gctggaccga catccccgac 900

ccgaagaagc tggtgctggc cgagccgcgc agcgtggtgg tgaacggcgt gcggttcccc 960

agcgtgcacc tgaaggacta cctgacccgc ctggcgcaga aggtgagcaa gaagacgggc 1020

gccctggact tctttaagag cctgaacgcg ggcgagctga agaaggccgc gcccgcggac 1080

ccgagcgccc ccctggtgaa cgcggagatc gcccggcagg tggaggccct gctgaccccg 1140

aacaccacgg tgatcgccga gacgggcgac agctggttca acgcgcagcg catgaagctg 1200

cccaacggcg cccgggtgga gtacgagatg cagtggggcc acatcggctg gagcgtgccg 1260

gccgcgttcg gctacgcggt gggtgccccc gagcgccgga acatcctgat ggtgggcgac 1320

ggcagctttc agctgaccgc gcaggaggtg gcccagatgg tgcgcctgaa gctgccggtg 1380

atcatcttcc tgatcaacaa ctacggctac acgatcgagg tgatgatcca cgacggcccc 1440

tacaacaaca tcaagaactg ggactacgcc ggcctgatgg aggtgtttaa cggcaacggc 1500

[0045]

ggctacgaca gcggcgcggg caagggcctg aaggccaaga ccggcggcga gctggccgag 1560

gcgatcaagg tggccctggc gaacaccgac ggccccacgc tgatcgagtg cttcatcggc 1620

cgggaggact gcacggagga gctggtgaag tggggcaagc gcgtggccgc cgccaacagc 1680

cgcaagccgg tgaacaagct gctgtga 1707

<210> 17
<211> 1152
<212> DNA
<213> Zymomonas mobilis

<400> 17
atggcttctt caacttttta tattcctttc gtcaacgaaa tgggcgaagg ttcgcttgaa 60

aaagcaatca aggatcttaa cggcagcggc tttaaaaatg cgctgatcgt ttctgatgct 120

ttcatgaaca aatccggtgt tgtgaagcag gttgctgacc tgttgaaagc acagggtatt 180

aattctgctg tttatgatgg cgttatgccg aacccgactg ttaccgcagt tctggaaggc 240

cttaagatcc tgaaggataa caattcagac ttcgtcatct ccctcggtgg tggttctccc 300

catgactgcg ccaaagccat cgctctggtc gcaaccaatg gtggtgaagt caaagactac 360

gaaggtatcg acaaatctaa gaaacctgcc ctgcctttga tgtcaatcaa cacgacggct 420

ggtacggctt ctgaaatgac gcgtttctgc atcatcactg atgaagtccg tcacgttaag 480

atggccattg ttgaccgtca cgttaccccg atggtttccg tcaacgatcc tctgttgatg 540

gttggtatgc caaaaggcct gaccgccgcc accggtatgg atgctctgac ccacgcattt 600

gaagcttatt cttcaacggc agctactccg atcaccgatg cttgcgcttt gaaagcagct 660

tccatgatcg ctaagaatct gaagaccgct tgcgacaacg gtaaggatat gccggctcgt 720

gaagctatgg cttatgccca attcctcgct ggtatggcct tcaacaacgc ttcgcttggt 780

tatgtccatg ctatggctca ccagttgggc ggttactaca acctgccgca tggtgtctgc 840

aacgctgttc tgcttccgca tgttctggct tataacgcct ctgtcgttgc tggtcgtctg 900

aaagacgttg gtgttgctat gggtctcgat atcgccaatc tcggtgataa agaaggcgca 960

gaagccacca ttcaggctgt tcgcgatctg gctgcttcca ttggtattcc agcaaacctg 1020

accgagctgg gtgctaagaa agaagatgtg ccgcttcttg ctgaccacgc tctgaaagat 1080

gcttgtgctc tgaccaaccc gcgtcagggt gatcagaaag aagttgaaga actcttcctg 1140

agcgctttct aa 1152

[0046]

<210> 18
<211> 1152
<212> DNA
<213> 人工

<220>
<223> Deinococcus codon optimized ADH gene derived from Zm

<400> 18
atggcgagca gcacgtttta catcccgttt gtgaacgaga tgggcgaggg cagcctggag 60

aaggcgatca aggacctgaa cggcagcggc ttcaagaacg ccctgatcgt gagcgacgcg 120

tttatgaaca agagcggcgt ggtgaagcag gtggccgacc tgctgaaggc gcagggcatc 180

aacagcgccg tgtacgacgg cgtgatgccc aacccgaccg tgacggcggt gctggagggc 240

ctgaagatcc tgaaggacaa caacagcgac ttcgtgatca gcctgggcgg cggcagcccg 300

cacgactgcg ccaaggcgat cgccctggtg gccaccaacg gcggcgaggt gaaggactac 360

gagggcatcg acaagagcaa gaagcccgcc ctgccgctga tgagcatcaa caccacggcc 420

ggcacggcga gcgagatgac ccgcttctgc atcatcacgg acgaagtgcg gcacgtgaag 480

atggccatcg tggaccgcca cgtgaccccg atggtgtccg tgaacgaccc gctgctgatg 540

gtgggcatgc ccaagggcct gaccgccgcg acgggcatgg acgccctgac ccacgccttt 600

gaggcgtaca gcagcaccgc cgcgacgccg atcaccgacg cgtgcgccct gaaggccgcg 660

agcatgatcg ccaagaacct gaagaccgcg tgcgacaacg gcaaggacat gcccgcccgg 720

gaggcgatgg cctacgcgca gttcctggcc ggcatggcgt ttaacaacgc gagcctgggc 780

tacgtgcacg cgatggccca ccagctgggc ggctactaca acctgccgca cggcgtgtgc 840

aacgccgtgc tgctgcccca cgtgctggcc tacaacgcga gcgtggtggc gggccggctg 900

aaggacgtgg gcgtggcgat gggcctggac atcgcgaacc tgggcgacaa ggagggcgcc 960

gaggcgacga tccaggccgt gcgcgacctg gccgcgccca tcggcatccc ggccaacctg 1020

accgagctgg gcgcgaagaa ggaggacgtg cccctgctgg ccgaccacgc gctgaaggac 1080

gcctgcgcgc tgaccaaccc ccggcagggc gaccagaagg aggtggagga gctgtttctg 1140

agcgcgtttt ga 1152

<210> 19
<211> 1200
<212> DNA
<213> Moorella sp

<400> 19

[0047]

atgtgggaaa caaaaataaa catcaacgaa gtccgggaaa tccgggctaa aacaaccgtc 60

tactttggag ttggagctat taaaaagatt gacgacatag ccagggaatt taaggaaaag 120

ggatacgata ggatcatcgt aataaccggc aaggggggctt ataaagccac cggcgcgtgg 180

gaatatatag ttccggcctt aaataaaaac cagataacct atatccatta cgaccaggtg 240

acgcccaacc cgacggtaga ccaggttgac gaggcaacca aacaagcccg ggaattcggt 300

gcccgagccg tcctggccat cggcgggggt agccccattg atgccgctaa aagcgtagcc 360

gtcttgctct cctaccccga caaaaatgcc cgacagctct accagttaga atttacacct 420

gttaaggccg cacctatcat cgctattaat cttacccatg gtacggggac ggaagccgat 480

cgctttgccg ttgtcagcat ccctgaaaag gcatataaac ccgctattgc ctatgattgc 540

atttaccccct tatattcaat tgacgacccg gccctcatgg taaaactgcc gtccgaccag 600

acagcttatg tctctgttga tgccctcaac catgtcgtcg aagcagccac cagcaaagta 660

gccagcccct atactattat cctggccaag gaaacggtac ggctcatcgc ccgatacctg 720

ccccaggccc tgtcccatcc ggcggatttg acggccaggt attatctcct ctatgcttcc 780

ctgattgccg gaatagcctt tgacaacggt ttgctccact tcacccacgc cctggaacac 840

cccctgagcg ccgtcaaacc ggagctcgcc cacggtctgg ggctgggtat gctgctgccg 900

gccgtagtca agcagattta cccggcaacc ccggaggtac tggcggagat actggagccc 960

attgttcccg atctcaaagg cgttcccggt gaagcagaaa aggctgccag cggggtggca 1020

aaatggcttg ccggagccgg tattaccatg aagctaaaag atgcgggctt tcaagcggaa 1080

gatatcgcca ggttaactga cctggccttt accaccccga gtctcgagct tctcctgagt 1140

atggccccgg taacggccga cagggaaagg gttaaggcaa tttaccagga cgccttttaa 1200

<210> 20
<211> 1532
<212> DNA
<213> Zymomonas mobilis

<400> 20
atgaaaacaa aacatctttc tgacagtatt gtcgtaacat ttataagcac ttattctgat 60

taaacatacg aaaattgcta ttgtgagtca aatatcacac cttaagttac aaacttatag 120

ttgtacacgt ttggtctaaa aaaacctttt atgtggcttt tactagatac gaatgagcac 180

atttggaggg ttcgctatgg tgtgggaatc ccacgttccg ataaaccagg tctttgagct 240

gaggtgcaaa acaacagact acttcggcct atgcgccata cacaaattta acggcttcgt 300

[0048]

ccgggagctt aaggggaaag gcgtggacag ggttatcctc gttactggga gcagctccta 360

caagaagtgc ggtgcatggg acgttgtcag gcctgccctt gaggaaaacg gggtcgagta 420

tgtccattac gacaaggtgg gggccaaccc aacggttgac atgatcgatg aagctgccga 480

gatggggaga gagttcgggg cccaagcggt gataggcata ggtggtggca gccccatcga 540

tagtgccaag agcgttgcga ttctgctgga ataccctgac aagaccgcca gggacctcta 600

cgagttcagg tttacccccg taaaggccaa gccagtaata gcgataaaca caacccatgg 660

agccggaacc gaagttgaca ggttcgcagt ggcgacgatt cctgagaaag agtacaagcc 720

ggccatagct tacgactgca tctacccgct ttacgcaatc gatgacccct cactgacggt 780

aaagctttcc ccggagcaga ccctttacct gaccatcgat gcactcaatc acgttaccga 840

agccgccaca accaagctgg ccaatccata ttcgatactc ctcgccaaag aggccgcgag 900

gctgatattt gactacctcc cagaggccct caaacacccc gacaaccttc aggcaaggta 960

cgccctgctc tacgcctctg ccatagccgg aataagcttc gacaacggtc tgctccactt 1020

tacgcacgcc cttgagcacc cgctgagtgc ggtcaaaccg gacttccccc acggcattgg 1080

cctcgcaatg cttcttccgg cggtaatcag gcacatatac cctgccaccg caaagatact 1140

tgccgaggtc tacaggccgc tcgttcccga ggccaaaggt gttccgggag aggtggaact 1200

cgtcgccagg agggtagaag agtggctctt cagcatcggc ataactgaaa agcttgagga 1260

cgtcgggttc tctgagacgg atgtgaacag gctaacggaa cttgccatga aaacaccaag 1320

ccttaacctg ctcctctcca tggctccggg tgaggccact cgggagagaa tagcggccat 1380

ataccgtgat tcgctttatc ccatcagcag agggtgaggg gcgctttctt ctgccttttg 1440

attttctgtc cgctttgcga ggattttaaa cccttggttc aaaaaatctt tttattggct 1500

tacttttaga ccggatttga gaaagactct aa 1532

<210> 21
<211> 417
<212> DNA
<213> 人工

<220>
<223> pTufA promoter derived from Deinococcus radiodurans R1

<400> 21
gtcgggtgtc gagcatcgtg atcgctgtgc tgagcctgct gtttttttccc tttggaacgg 60

tgctcggtgt catcatgttg atcgggatct tcgacgaaca ggtgaccgct tacctgagcc 120

[0049]

gctgagcggg ttttcgggca agagttagtt cggggccagg aacctatacg gaacctatac 180

gggaactggc ccttttttgct ggattcccgg cccggcgctc ttgcccttct aggccttttc 240

atgtaagcta agcagtcggt gcgccgtggc ctctcagccc ggtgcaacgt gccagagagc 300

aggcgggaca ccgccctggc gagaatcaac ccaatgtggg tcacaccacc agaacctcgc 360

ggcatgcgcg ggggaagccc gatcaaaggg acgtttttac gtgtgaatcc accgctt 417

<210> 22
<211> 220
<212> DNA
<213> 人工

<220>
<223> pTufB promoter derived from Deinococcus radiodurans R1

<400> 22
gttcgagtct cgtttcccgc tttttacgct cctgtagctc agcggtagag cactcccttg 60

gtaagggaga ggtcatcggt tcaagtccga tcaggagctc caacgaagaa atgcggcccg 120

gataagaggg ccgcatttcg attttgaccg agcatggccc tacgccagcc cgcgcacttt 180

tgatatagtg ttcaggcttg cccgtttgcg ggcaaaaccc 220

<210> 23
<211> 651
<212> DNA
<213> 人工

<220>
<223> Deinococcus codon optimized CATgene derived from Staphilococcus aureus

<400> 23
atgaacttta acaagatcga cctggacaac tggaagcgga aggagatctt taaccactac 60

ctgaaccagc agacgacgtt tagcatcacc acggagatcg acatcagcgt gctgtaccgc 120

aacatcaagc aggaaggcta caagttctac cccgccttca tctttctggt gacccgcgtg 180

atcaacagca acaccgcgtt tcggacgggc tacaacagcg acggcgagct gggctactgg 240

gacaagctgg agcccctgta cacgatcttc gacggcgtga gcaagacctt tagcggcatc 300

tggacgccgg tgaagaacga cttcaaggag ttttacgacc tgtacctgag cgacgtggag 360

aagtacaacg gcagcggcaa gctgttcccc aagaccccca tcccggagaa cgcctttagc 420

ctgagcatca tcccgtggac cagcttcacg ggctttaacc tgaacatcaa caacaacagc 480

[0050]

aactacctgc tgcccatcat caccgccggc aagttcatca acaagggcaa cagcatctac 540

ctgccgctga gcctgcaggt gcaccacagc gtgtgcgacg gctaccacgc gggcctgttt 600

atgaacagca tccaggagct gagcgaccgc ccgaacgact ggctgctgtg a 651

<210> 24
<211> 375
<212> DNA
<213> 人工

<220>
<223> Deinococcus codon optimized BLEO gene derived from Streptoalloteichus hindustanus

<400> 24
atggcgaagc tgacgagcgc ggtgcccgtg ctgacggcgc gggacgtggc ggaggcggtg 60

gagttctgga cggaccggct gggctttagc cgcgtgttcg tggaggacga ctttgccggc 120

gtggtgcggg acgacgtgac cctgttcatc agcgcggtgc aggaccaggt ggtgcccgac 180

aacacccagg cgtgggtgtg ggtgcgcggc ctggacgagc tgtacgcgga gtggagcgag 240

gtcgtgagca ccaactttcg cgacgccagc ggcccggcca tgaccgagat cgtggagcag 300

ccgtggggcc gggagttcgc gctgcgggac ccggcgggca actgcgtgca ctttgtggcg 360

gaggagcagg actga 375

<210> 25
<211> 23
<212> DNA
<213> 人工

<220>
<223> 引物

<400> 25
ggtatctacg cattccaccg cta 23

<210> 26
<211> 23
<212> DNA
<213> 人工

<220>
<223> 引物

<400> 26
gttacccgga atcactgggc gta 23

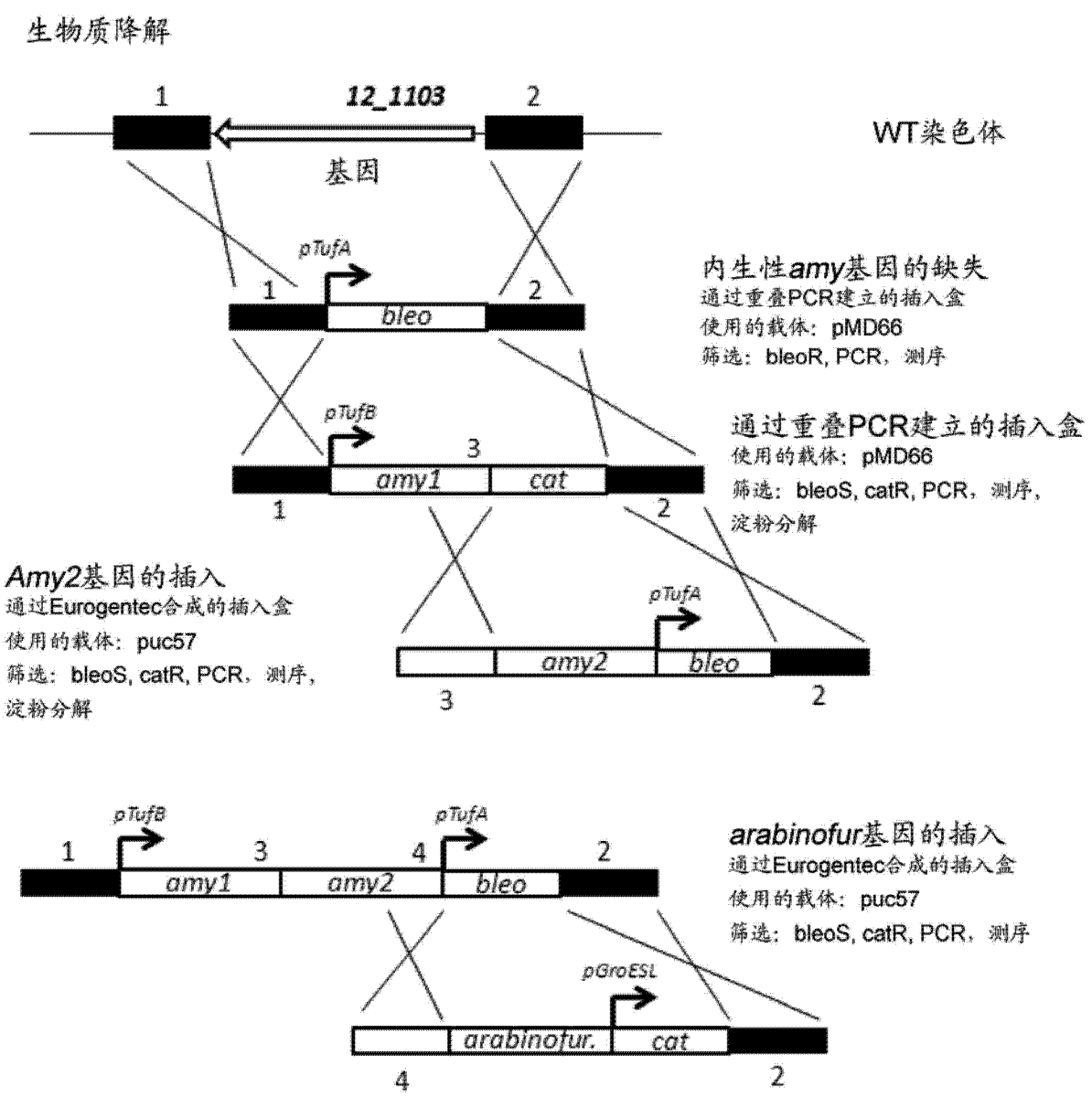

生物质降解
1
12_1103
2
基因
WT染色体
pTufA
1
2
bleo
内生性amy基因的缺失
通过重叠PCR建立的插入盒
使用的载体：pMD66
筛选：bleoR, PCR，测序
pTufB
3
amy1
cat
1
2
通过重叠PCR建立的插入盒
使用的载体：pMD66
筛选：bleoS, catR, PCR，测序，
淀粉分解
Amy2基因的插入
通过Eurogentec合成的插入盒
使用的载体：puc57
筛选：bleoS, catR, PCR，测序，
淀粉分解
pTufA
amy2
bleo
3
2
pTufB
pTufA
1
3
4
2
amy1
amy2
bleo
arabinofur基因的插入
通过Eurogentec合成的插入盒
使用的载体：puc57
筛选：bleoS, catR, PCR，测序
pGroESL
arabinofur.
cat
4
2

图 1

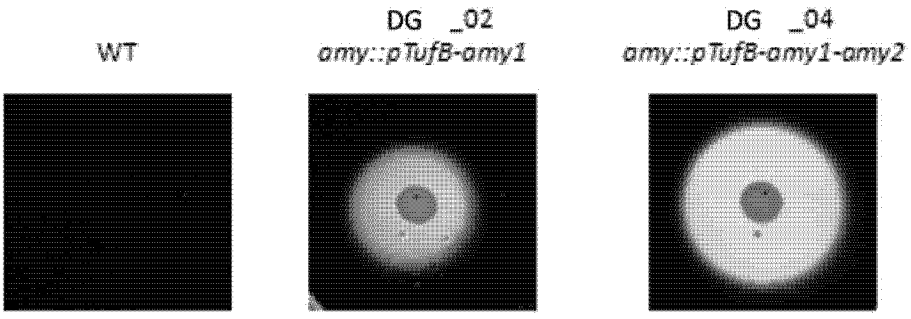

WT
DG _02
amy::pTufB-amy1
DG _04
amy::pTufB-amy1-amy2

图 2

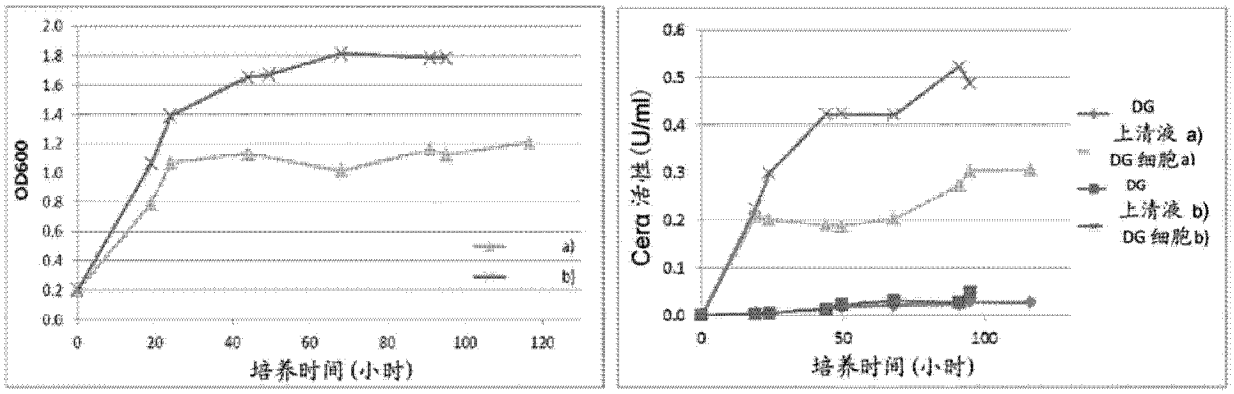

OD600
培养时间(小时)
a)
b)
Cera 活性(U/ml)
DG 上清液 a)
DG 细胞a)
DG 上清液 b)
DG 细胞b)
培养时间(小时)

图 3

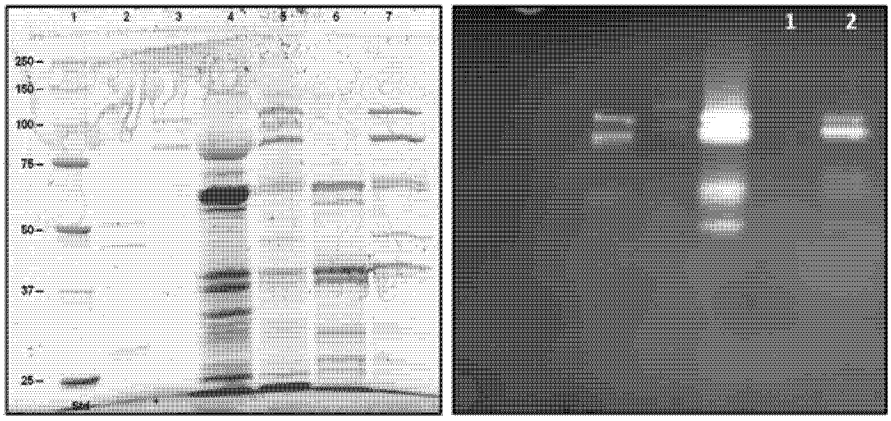

1
2
3
4
5
6
7
250
150
100
75
50
37
25
1
2

图 4

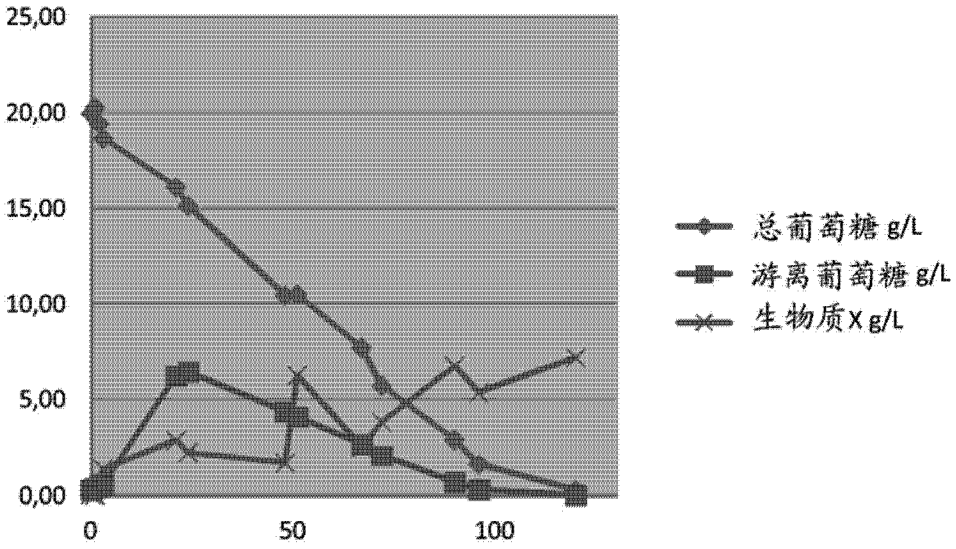

25,00
20,00
15,00
10,00
5,00
0,00
0
50
100
总葡萄糖 g/L
游离葡萄糖 g/L
生物质X g/L

图 5

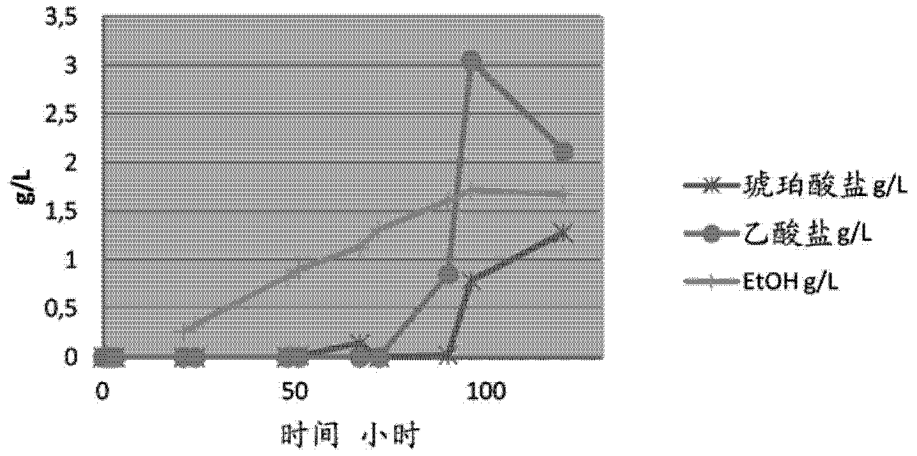

3,5
3
2,5
2
1,5
1
0,5
0
g/L
0
50
100
时间 小时
琥珀酸盐 g/L
乙酸盐 g/L
EtOH g/L

图 6

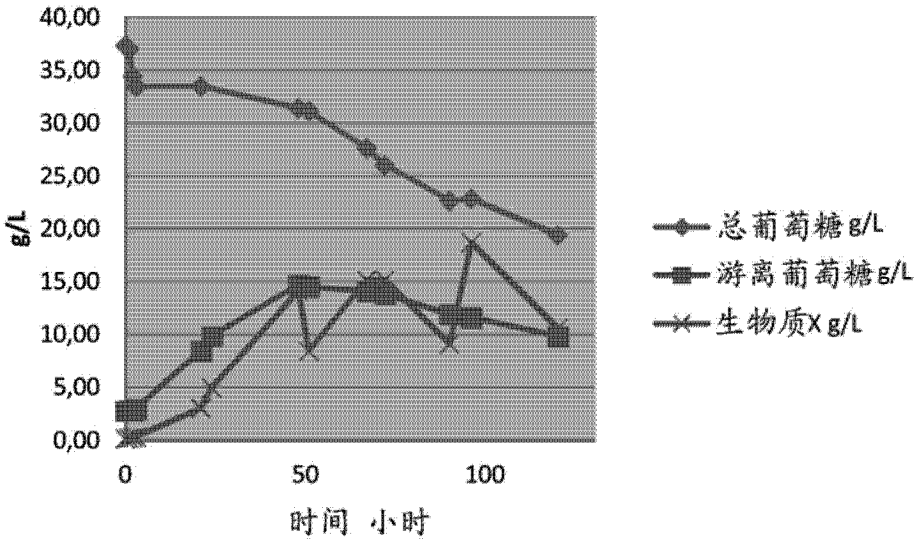

40,00
35,00
30,00
25,00
20,00
15,00
10,00
5,00
0,00
g/L
0
50
100
时间 小时
总葡萄糖 g/L
游离葡萄糖 g/L
生物质X g/L

图 7

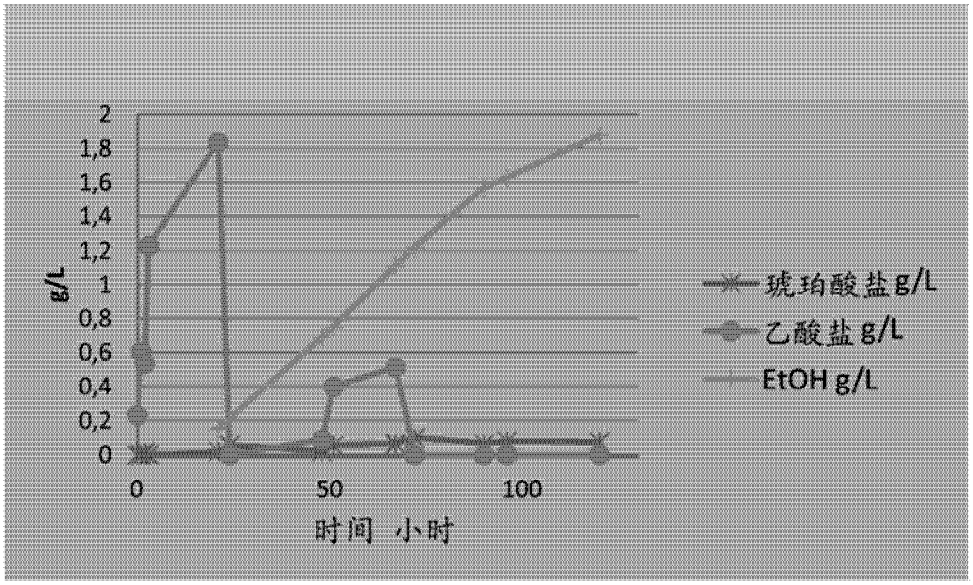



图 8