



(12) 发明专利申请



(10) 申请公布号 CN 118043455 A

(43) 申请公布日 2024. 05. 14

(21) 申请号 202280056037.9

(22) 申请日 2022.08.18

(30) 优先权数据

21191995.6 2021.08.18 EP

(85) PCT国际申请进入国家阶段日

2024.02.08

(86) PCT国际申请的申请数据

PCT/EP2022/073110 2022.08.18

(87) PCT国际申请的公布数据

W02023/021151 EN 2023.02.23

(71) 申请人 恩多罗基因有限公司

地址 丹麦哥本哈根

(72) 发明人 P·鲁格比约 K·B·法尔肯伯格

J·库恩 C·蒙克

L·塔普·洛吉

(74) 专利代理机构 北京市中伦律师事务所

11410

专利代理师 杨黎峰 刘烽

(51) Int.Cl.

C12N 15/09 (2006.01)

C12N 15/67 (2006.01)

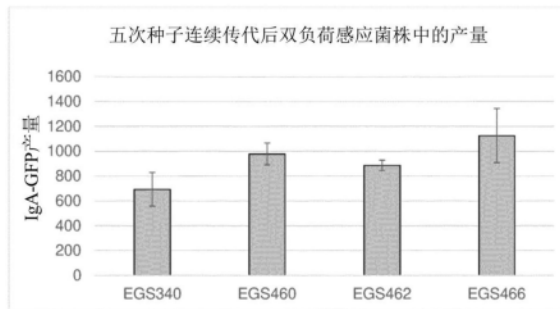
权利要求书3页 说明书223页 附图12页

(54) 发明名称

负荷成瘾生产菌株

(57) 摘要

本发明提供了一种用于合成产物的微生物生产细胞,其进一步包含负荷成瘾基因线路,所述基因线路表达赋予合成所述产物的那些细胞选择性生长和/或存活优势;同时限制低生产性或非生产性逃逸细胞的增殖。



1. 一种经基因工程化以合成产物的微生物生产细胞,所述微生物细胞进一步包含:

- a. 可操作地连接到第一负荷感应启动子的第一必需基因,和
- b. 可操作地连接到第二负荷感应启动子的第二必需基因,

其中所述第一负荷感应启动子相对于所述第一必需基因是异源的,并且所述第二负荷感应启动子相对于所述第二必需基因是异源的;

其中所述产物的合成赋予所述细胞负荷,并且

其中当所述第一负荷感应启动子被所述负荷诱导时,所述第一必需基因的表达相对于所述第一负荷感应启动子未被诱导时所述第一必需基因的基础水平表达被上调,并且当所述第二负荷感应启动子被所述负荷诱导时,所述第二必需基因的表达相对于所述第二负荷感应启动子未被诱导时所述第二必需基因的基础水平表达被上调。

2. 根据权利要求1所述的微生物生产细胞,其中当所述细胞中的所述第一必需基因和所述第二必需基因可操作地连接到其天然启动子时,由所述产物的合成赋予的所述负荷具有适合度代价,其测量为相对于对应非生产性微生物细胞测量的所述微生物生产细胞的最大指数期生长速率的选自 $\geq 5\%$ 、 $\geq 10\%$ 、 $\geq 15\%$ 、 $\geq 20\%$ 、 $\geq 25\%$ 、 $\geq 35\%$ 和 $\geq 45\%$ 的百分比降低。

3. 根据权利要求1或2中任一项所述的微生物生产细胞,其中所述细胞是:

- i. 细菌,其属于选自埃希氏杆菌属、乳杆菌属、乳球菌属、棒状杆菌属、芽孢杆菌属、醋酸杆菌属、不动杆菌属、假单胞菌属;丙酸杆菌属、拟杆菌属和双歧杆菌属的属;或
- ii. 酵母,其属于选自酵母属、克鲁维酵母菌属、假丝酵母属、毕赤酵母属、驹形酵母属、隐球菌属、德巴利酵母属、汉逊酵母属、耶氏酵母属、接合酵母属和裂殖酵母属的属;或
- iii. 丝状真菌,其选自青霉属、根霉菌属、镰孢菌属、梭链孢属、赤霉菌属、毛霉菌属、被孢霉属、木霉菌属、嗜热真菌属、链霉菌属和曲霉菌属。

4. 根据权利要求1-3中任一项所述的微生物生产细胞,其中所述第一必需基因和所述第二必需基因是相同的。

5. 根据权利要求1-4中任一项所述的微生物生产细胞,其中所述第一负荷感应启动子和所述第二负荷感应启动子是不同的。

6. 根据权利要求1-5中任一项所述的微生物生产细胞,其中所述第一必需基因和/或所述第二必需基因是非条件必需基因。

7. 根据权利要求1-6中任一项所述的微生物生产细胞,其中所述第一必需基因和/或所述第二必需基因可操作地连接到合成RBS,所述合成RBS的序列经选择以分别独立于所述第一负荷感应启动子和/或所述第二负荷感应启动子的诱导来修饰所述第一必需基因和/或所述第二必需基因的翻译强度。

8. 根据权利要求1-7中任一项所述的微生物生产细胞,其中所述第一负荷感应启动子和所述第二负荷感应启动子选自:

- i. 核糖体RNA启动子,
- ii. 由oxyR或其同源物上调的启动子,
- iii. 包含由HAC1或其同源物上调的UPR元件的启动子,和
- iv. DNA损伤感应启动子。

9. 根据权利要求1-8中任一项所述的微生物生产细胞,其中所述细胞的特征在于,与缺

乏分别可操作地连接到所述第一负荷感应启动子和所述第二负荷感应启动子的所述第一必需基因和所述第二必需基因的亲代微生物生产细胞相比,在从单个细胞进行至少25、30、35、40、45、50、60、70、80、90或100代次细胞分裂后,产物收率增加。

10. 根据权利要求1-9中任一项所述的微生物生产细胞,其中所述细胞的特征在于,与缺乏分别可操作地连接到所述第一负荷感应启动子和所述第二负荷感应启动子的所述第一必需基因和所述第二必需基因的亲代微生物生产细胞相比,在从单个细胞进行至少50代次细胞分裂后,产物收率增加至少10%、25%、50%或80%。

11. 一种产物生物合成的方法,其包括以下步骤:

- i. 提供至少一种根据权利要求1-10中任一项所述的微生物生产细胞,
- ii. 将所述至少一种细胞引入包含用于生产所述产物的底物的培养基中,
- iii. 回收所述产物。

12. 根据权利要求11所述的方法,其进一步包括分离所述产物和/或将所述产物配制成组合物,如营养组合物、药物组合物、化妆品组合物、洗涤剂组合物、润滑剂组合物或燃料组合物的步骤。

13. 根据权利要求11或12所述的产物生物合成方法,其中所述产物选自:有机酸、萜类化合物、类异戊二烯、聚酮化合物、醇、糖、维生素、醛、羧酸、脂肪酸、氨基酸、肽、酶(如淀粉酶、脂肪酶、蛋白酶、barnase、 β -半乳糖苷酶、晶体蛋白、角质酶、PET酶、漆酶和碳水化合物活性酶(如木聚糖酶、地衣聚糖酶、纤维素酶、裂解多糖单氧化酶和果胶酶))、治疗性蛋白质及其前体(如人生长激素、胰岛素、胰高血糖素样肽-1、单克隆抗体和多克隆抗体、单片段抗体和纳米抗体)、鸡蛋中天然存在的蛋白质(如卵清蛋白)、乳蛋白(如酪蛋白、乳凝集素、乳铁蛋白)、分泌型免疫球蛋白A和G、分泌成分。

14. 分别可操作地连接到第一负荷感应启动子和第二负荷感应启动子的第一必需基因和第二必需基因用于提高来源于单个细胞的至少50代后的微生物生产细胞培养群体的产物收率的用途;

其中所述第一负荷感应启动子相对于所述第一必需基因是异源的,并且所述第二负荷感应启动子相对于所述第二必需基因是异源的;

其中所述产物的合成赋予所述细胞负荷;并且

其中当所述第一负荷感应启动子被所述负荷诱导时,所述第一必需基因的表达相对于所述第一负荷感应启动子未被诱导时所述第一必需基因的基础水平表达被上调,并且当所述第二负荷感应启动子被所述负荷诱导时,所述第二必需基因的表达相对于所述第二负荷感应启动子未被诱导时所述第二必需基因的基础水平表达被上调。

15. 根据权利要求14所述的分别可操作地连接到所述第一负荷感应启动子和所述第二负荷感应启动子的所述第一必需基因和所述第二必需基因的用途,其中与缺乏分别可操作地连接到所述第一负荷感应启动子和所述第二负荷感应启动子的所述第一必需基因和所述第二必需基因的亲代微生物生产细胞相比,在从单个细胞进行至少50代次细胞分裂后,产物收率增加至少10%、25%、50%或80%。

16. 根据权利要求1-10中任一项所述的微生物生产细胞用于生产生物合成产物的用途,所述生物合成产物如有机酸、萜类化合物、类异戊二烯、聚酮化合物、醇、糖、维生素、醛、羧酸、脂肪酸、氨基酸、肽、酶(如淀粉酶、脂肪酶、蛋白酶、barnase、 β -半乳糖苷酶、晶体蛋

白、角质酶、PET酶、漆酶和碳水化合物活性酶(如木聚糖酶、地衣聚糖酶、纤维素酶、裂解多糖单氧化酶和果胶酶))、治疗性蛋白质及其前体(如人生长激素、胰岛素、胰高血糖素样肽-1、单克隆抗体和多克隆抗体、单片段抗体)、鸡蛋中天然存在的蛋白质(如卵清蛋白)、乳蛋白(如酪蛋白、乳凝集素、乳铁蛋白)、分泌型免疫球蛋白A和G、分泌成分、纳米抗体。

负荷成瘾生产菌株

技术领域

[0001] 本发明提供了一种用于合成产物的微生物生产细胞,其进一步包含负荷成瘾基因线路,其表达赋予合成产物的那些细胞选择性生长和/或存活优势;同时限制非/低生产性逃逸细胞的增殖。

背景技术

[0002] 越来越多的世界化学品生产依赖于微生物,所述微生物被基因工程化以发挥细胞工厂的功能,并为给定分子的生物合成量身定制。采用这些细胞工厂的生产过程通常始于生产生物的少量细胞的起始培养物,其在大发酵罐(高达1000000L体积)中经历细胞数量的生长和扩增期。在一些设置中,给定分子的生产在生长期和随后的时期(分批和补料分批培养)都进行;并且可替代地,生产可以是连续的。恒化发酵罐允许生产生物在不断稀释的发酵液中生长,由此从培养物中取出产物和细胞,同时补充新鲜的营养培养基。在工业规模上,这种生产过程可以持续操作1-2个月,然后在干净的罐中开始新的培养。在此工业中使用的发酵方法和设备非常相似,既用于生产多种商品小分子,也用于生产治疗性蛋白质,因此这些方法遇到相似的问题。在其他应用中,微生物细胞可被工程化以产生非天然分子,所述非天然分子在培养细胞的环境中赋予商业优势。

[0003] 为给定分子的生物合成分配有限代谢资源而工程化的细胞工厂承受非自然负荷;通常体现在生长较慢。这种负荷对不能合成产物分子(非生产或低生产)的细胞施加了选择压力,当生产运行持续较长时间时(例如在恒化器中),这个问题尤其突出。工业发酵中的这种非生产细胞是非常不希望的,因为它们消耗营养、氧气和空间。由于低生产和非生产细胞生长更快,因此它们比生产细胞具有更强的选择优势。在生长的细胞培养物中,适合度的此类改善可能导致生产细胞随时间推移的显著退出竞争。从最佳生产状态的这种偏离是丢弃发酵液和将资源花费在清洁、灭菌上的最终原因,更不用说营养物,以新的生产微生物补充发酵罐。此类非生产细胞来源于经历多次生长分裂的原始生产生物的细胞中出现的基因突变。

[0004] 由于无法避免在产生生物的细胞中发生基因突变和非基因适应,从而导致生产运行过程中细胞产物形成的下降,因此需要用于消除或减缓生产中非生产细胞生长的方法。优选地,此类消除方法足够有效,它们防止观察到的从生产状态的偏离,从而延长工业发酵的寿命。

[0005] 本发明解决了如何剥夺细胞工厂的非生产成员的适合度优势的问题;如随时间推移延缓它们在生产性细胞工厂成员中的增殖,从而提高细胞工厂的生产率。

发明内容

[0006] 本发明的第一方面提供了一种经基因工程化以合成产物的微生物生产细胞,其中所述细胞进一步包含:

[0007] -可操作地连接到第一负荷感应启动子的第一必需基因,和

[0008] -可操作地连接到第二负荷感应启动子的第二必需基因,

[0009] 其中所述第一负荷感应启动子相对于所述第一必需基因是异源的,并且所述第二负荷感应启动子相对于所述第二必需基因是异源的;其中产物的合成赋予所述细胞负荷和/或适合度代价;并且其中当所述第一负荷感应启动子被所述负荷和/或适合度代价诱导时,所述第一必需基因的表达相对于所述第一负荷感应启动子未被诱导时所述必需基因的基础水平表达被上调,并且当所述第二负荷感应启动子被所述负荷和/或适合度代价诱导时,所述第二必需基因的表达相对于所述第二负荷感应启动子未被诱导时所述第二必需基因的基础水平表达被上调。

[0010] 本发明的第二方面提供了一种产物生物合成的方法,其包括以下步骤:

[0011] -提供本发明的至少一种微生物生产细胞,

[0012] -将所述至少一种细胞引入包含用于生产所述产物的底物的培养基中,

[0013] -回收所述产物。

[0014] 本发明的第三方面提供了分别可操作地连接到第一和第二负荷感应启动子的第一和第二必需基因用于提高来源于单个细胞在至少50代后的微生物生产细胞培养群体的产物收率的用途;其中所述第一负荷感应启动子相对于所述第一必需基因是异源的,并且所述第二负荷感应启动子相对于所述第二必需基因是异源的;其中产物的生产赋予所述细胞负荷;并且其中当所述第一负荷感应启动子被所述负荷诱导时,所述第一必需基因的表达相对于所述第一负荷感应启动子未被诱导时所述第一必需基因的基础水平表达被上调,并且当所述第二负荷感应启动子被所述负荷诱导时,所述第二必需基因的表达相对于所述第二负荷感应启动子未被诱导时所述第二必需基因的基础水平表达被上调。

[0015] 本发明的第四方面提供了本发明的微生物生产细胞用于生产生物合成产物的用途。

[0016] 本发明的另一方面提供了一种经基因工程化以合成产物的微生物生产细胞,其中所述细胞进一步包含:

[0017] -可操作地连接到负荷感应启动子的必需基因,

[0018] 其中所述负荷感应启动子相对于所述必需基因是异源的;其中产物的合成赋予所述细胞负荷和/或适合度代价;其中当所述负荷感应启动子被所述负荷和/或适合度代价诱导时,所述必需基因的表达相对于所述负荷感应启动子未被诱导时所述必需基因的基础水平表达被上调;并且其中所述微生物生产细胞属于选自芽孢杆菌属 (*Bacillus*)、棒状杆菌属 (*Corynebacterium*) 和曲霉属 (*Aspergillus*) 的属。

附图说明

[0019] 图1:示出亲代菌株大肠杆菌 (*E. coli*) BL21 (DE3) (灰线) 的细胞和用hGH-GFP生产质粒pEG34转化的亲代菌株 (黑线; pEG34) 的细胞随时间推移生长 (如光密度所测量: OD_{630nm}) 的图。在hGH-GFP表达诱导条件下测量每个菌株的四个重复培养物的生长;误差条描绘平均值的标准误差 (n=4)。

[0020] 图2:示出亲代菌株大肠杆菌BL21 (DE3) (黑线) 的细胞和来源于亲代菌株的非生产负荷成瘾大肠杆菌菌株s5.0#3、s7.0#8和s9.0#8 (如所指出的那样;其中必需基因fo1P-gl1mM分别由p_{ycv}、p_{ycjX}和p_{mutM}控制) 的细胞随时间推移生长 (如光密度所测量: OD_{630nm}) 的

图。测得的生长速率基于四次重复培养 ($n=4$) 的平均值。

[0021] 图3: 示出含有hGH-GFP生产质粒pEG34 (黑线) 或对照质粒pEG0 (灰线) 的负荷成瘾菌株s9.0#8的细胞随时间推移生长 (如光密度所测量: OD_{630nm}) 的图。在hGH-GFP表达诱导条件下测量每个菌株的四个重复培养物的生长; 误差条描绘平均值的标准误差 ($n=4$)。

[0022] 图4: 示出亲代菌株大肠杆菌BL21 (DE3) (黑线) 的细胞和含有并表达hGH-GFP生产质粒pEG34的三种负载成瘾菌株: s5.0#3、s7.0#8和s9.0#8 (如所指出的那样) 的细胞随时间推移生长 (如光密度所测量: OD_{630nm}) 的图。在hGH-GFP表达诱导条件下测得的生长速率基于四次重复培养 ($n=4$) 的平均值。

[0023] 图5: 在以下hGH-GFP生产菌株的培养物的第一次和第六次培养转移 (种子1和6) 期间测量的GFP/OD值 (Y轴) 的直方图: s3.6 (p_{ibpA})、s6.6 (p_{grpE})、s7.6 (p_{ycjX}) 和s10.6 (p_{ybbN}), 其包含各自的负荷感应启动子和驱动必需基因fo1P-glmm表达的四种变体RBS编码序列之一。包括非负荷成瘾大肠杆菌菌株BEG34和亲代菌株大肠杆菌BL21 (DE3)、BEG0的培养物作为对照。对在hGH-GFP表达诱导条件下生长的培养物进行测量。误差条描绘平均值的标准误差 ($n=4$)。

[0024] 图6: 在以下hGH-GFP生产菌株的培养物的系列培养转移 (种子1至12) 后测量的GFP/OD值 (Y轴) 的直方图: s3.6#2 (p_{ibpA})、s6.6#6 (p_{grpE})、s7.6#8 (p_{ycjX}) 和s10.6#7 (p_{ybbN}), 其包含各自的负荷感应启动子和驱动必需基因fo1P-glmm表达的四种变体RBS编码序列之一。包括非负荷成瘾大肠杆菌菌株BEG34和亲代菌株大肠杆菌BL21 (DE3) 的培养物作为对照。对在hGH-GFP表达诱导条件下生长的培养物进行测量。误差条描绘平均值的标准误差 ($n=3$)。

[0025] 图7: OD/min的直方图, 作为负荷成瘾hGH-GFP生产菌株s3.6#2、s6.6#6、s7.6#8和s10.6#7与非负荷成瘾hGH-GFP生产大肠杆菌菌株BEG34和亲代菌株大肠杆菌BL21-DE3 (BEG0) 相比的生长速率的量度。各个菌株的生长速率基于在hGH-GFP表达诱导条件下在种子0测量的7个重复培养物; 误差条表示平均值的标准误差 ($n=7$)。

[0026] 图8: 示出在hGH-GFP生产菌株的培养物的系列培养转移 (种子1、4至9) 后测量的GFP/OD值 (Y轴) 减去空质粒对照BEG0-空菌株的测量值的图。菌株显示为: (A) s7.6#8 (p_{ycjX})、(B) s13.6#2 (p_{fsxA}) 和 (C) s13.6#2evo (p_{fsxA}), 各自包含驱动必需基因fo1P-glmm表达的各自的负荷感应启动子 (灰线), 和非负荷成瘾hGH-GFP生产大肠杆菌菌株BEG34 (黑线)。对在hGH-GFP表达诱导条件下生长的培养物进行测量。误差条描绘平均值的标准误差 ($n=5$)。

[0027] 图9: 示出在hGH-GFP生产菌株的培养物的系列培养转移 (种子1、4至9) 后测量的GFP/OD值 (Y轴) 减去空质粒对照BEG0-空菌株的测量值的图。菌株显示为: (A) s15.6.7 (p_{rrnB}) 和 (B) s16.6.6 (p_{rrnE}), 各自包含驱动必需基因fo1P-glmm表达的各自的负荷感应启动子 (灰线), 和非负荷成瘾hGH-GFP生产大肠杆菌菌株BEG34 (黑线)。对在hGH-GFP表达诱导条件下生长的培养物进行测量。误差条描绘平均值的标准误差 ($n=5$)。

[0028] 图10: 示出在hGH-GFP生产菌株s29.6#3的培养物的系列培养转移 (种子2至5, 以及7) 后测量的相对hGH-GFP生产 (Y轴) 减去平行空质粒对照BEG0菌株的测量值的图。与非负荷成瘾hGH-GFP生产大肠杆菌菌株BEG34 (灰点) 相比, 菌株s29.6#3包含驱动必需基因fo1P-glmm表达的渗透应力感应启动子p_{poxB} (黑色方块)。在72小时后分析连续传代7次。对在hGH-GFP表达诱导条件下生长的培养物进行测量。误差条描绘平均值的标准误差 (对于s29.6#3,

n=5并且对于BEG34,n=8)。

[0029] 图11:示出每分钟金黄色葡萄球菌(*S.aureus*)裂解速率的直方图,归一化为大肠杆菌培养物OD,作为与非负荷成瘾溶葡萄球菌酶生产大肠杆菌菌株BENDU5cam相比,包含负荷感应启动子 p_{grpE} 和驱动必需基因 $folP$ - $glmM$ 表达的四种变体RBS编码序列之一的溶葡萄球菌酶生产菌株s6.4#5的培养物的分泌溶葡萄球菌酶生产速率的量度。在系列培养转移(种子1和7)后,在溶葡萄球菌酶表达诱导条件下生长的培养物上测量所述速率。误差条描绘平均值的标准误差(n=3)。

[0030] 图12:示出与亲代非负荷成瘾甲羟戊酸生产大肠杆菌菌株BL21(DE3)- p_{MevT} 相比,在系列培养转移(种子2、5和6)后测量的由负荷成瘾甲羟戊酸生产菌株s19.1#1(p_{cspD})和s9.1.4(p_{mutM})的培养物合成的甲羟戊酸滴度(g/L)的图。负荷感应启动子 p_{cspD} 和 p_{mutM} 各自与驱动必需基因 $folP$ - $glmM$ 表达的四种变体RBS编码序列之一组合。对在 $MevT$ 表达诱导条件下生长的培养物进行测量。误差条描绘平均值的标准误差(n=3)。

[0031] 图13:示出毕赤酵母(*Pichia pastoris*)菌株EGS31的培养物的人血清白蛋白(hSA)生产滴度的柱状图,所述菌株EGS31包含在AOX1启动子控制下的hSA编码ALB1基因的cDNA版本的基因组整合拷贝,和可操作地连接到 $kanMX$ 选择基因的PDI1启动子。培养物在选择条件(由750 μ g/mL G418激活成瘾性)或非选择条件(成瘾性未激活)下生长。“成瘾性未激活”培养物的hSA滴度设定为100%,而“成瘾性激活”培养物的滴度显示为其百分比。

[0032] 图14A和图14B:示出毕赤酵母菌株EGS31的负荷成瘾版本的培养物的人血清白蛋白(hSA)生产滴度的柱状图,每个菌株包含在AOX1启动子控制下的hSA编码ALB1基因的cDNA版本的基因组整合拷贝,和可操作地融合到负荷响应启动子FPR2或RPL3或RPL6A的 $kanMX$ 选择基因。培养物在非选择条件(成瘾性未激活)下或在选择条件(用150或750 μ g/mL G418激活成瘾性)下生长约30个细胞分裂期([A]种子1),或在进一步10个细胞分裂([B]种子2)后。平均hSA滴度以种子1“EGS31”培养物中滴度的百分比显示。误差条描绘平均值的标准误差(n=4)。

[0033] 图15:长期培养中负荷成瘾EGS340和亲代EGS84的IgA-GFP产量的比较。使用IgA生产负荷感应 $perR$ 启动子调节必需 $iscU$ 基因操纵子($sufC$ - $sufD$ - $sufS$ - $sufU$ - $sufB$)的转录来证明芽孢杆菌属中的负荷成瘾以维持生产。在30℃下连续传代实验期间EGS340和EGS84的比较,其中每24小时将500倍稀释的种子转移到新的培养基中。点是平均值,并且误差条显示s.e.m。(n=7)。

[0034] 图16:EGS343和EGS84在长期培养中的比较。使用 $perR$ 启动子调节必需 $accC$ 基因操纵子($accB$ - $accC$ - $yhqY$)的转录来证明枯草芽孢杆菌中的负荷成瘾以维持生产。在30℃下连续传代实验期间EGS343和EGS84中的IgA-GFP产量的比较,其中每24小时将500倍稀释的种子转移到新的培养基中。点是平均值,并且误差条显示s.e.m。(n=7)。

[0035] 图17:使用几个负荷感应器同时调节不同的必需基因,从而维持枯草芽孢杆菌中异源蛋白的生产。在对应于大约75代的五次连续传代(500倍稀释种子)后比较IgA-GFP生产菌株。单个负荷感应器菌株EGS340仅携带控制 $iscU$ 必需基因操纵子(EGS340)转录的基于 p_{perR} 的负荷感应器。将EGS340与衍生菌株进行比较,其中 $accC$ 必需基因操纵子的转录也分别由基于 p_{ctsR} 的负荷感应器(EGS460)、基于 p_{dnaK} 的负荷感应器(EGS462)和基于 p_{hrcA} 的负荷感应器(EGS466)调节。条形图显示平均值,并且误差条显示s.e.m。(n=4)。

[0036] 图18:在非负荷成瘾对照菌株EGS084和双负荷成瘾衍生菌株EGS466(基于控制iscU必需基因操纵子的 p_{perR} 和控制accC必需基因操纵子的 p_{hrcA})中,在大约75代细胞培养后,通过IgA特异性ELISA测量的相对IgA-GFP水平的比较。

[0037] 图19:使用必需基因上的TIS变异来调节生长速度。在YPD中生长的亲代菌株(EGS621)和具有必需基因CIA1的TIS序列变异的菌株(EGS1100、EGS1101、EGS1102、EGS1104)的两个生物复制品。毕赤酵母属菌株表现出与四个TIS序列的相对强度相关的生长。

具体实施方式

[0038] 缩写、术语和定义:

[0039] 负荷(Burden)是指在生产条件下产物生产期间微生物细胞(如经基因工程化以合成产物的细胞)的细胞状态,并且其导致归因于所述产物合成的适合度代价,特别是在表现出产物高水平有效合成的细胞中。适合度代价是所述负荷的量度,其可以通过测量在生产条件下合成产物期间细胞的最大指数生长速率(沿生长曲线测量)与在可比生产条件下生长时不能进行所述合成的亲代微生物细胞相比的减少百分比来量化。

[0040] 负荷感应启动子(Burden-sensing promoter)是指由所述“负荷”诱导的启动子,并且当被诱导时,与所述微生物生产细胞的基本上不合成生产细胞的预期产物或少合成其至少50%的突变体衍生物相比,当在生产条件下培养时,所述负荷感应启动子可以上调微生物生产细胞中与其可操作地连接的基因的表达式。这种突变体衍生物包括在长期培养来源于所述生产细胞的细胞群体后分离的非生产性逃逸突变体。负荷感应启动子的非限制性示例(详情参见表2)包括大肠杆菌基因htpG、ibpA、clpB、yccV、grpE、ycjX、ldhA、mutM、ybbN、prlC、groES、fxsA、htpX、rrnB、rrnE、cspD、katE、xthA、uspE、gadB、ahpC、katG、grxA、oxyS、poxB、trxC的启动子和它们在其他革兰氏阳性和革兰氏阴性细菌中的同源版本;对于酿酒酵母(*S.cerevisiae*),负荷感应启动子的非限制性示例包括基因KAR2、PDI1、SAA1、FPR2、RPL3、RPL6A、RPL28、OGG1、RAD51、RAD54的启动子;对于芽孢杆菌属,负荷感应启动子的非限制性示例包括基因groES、ctsR、dnaK、perR、hrcA、spx、sigB、yflT的启动子和它们的同源版本;对于棒状杆菌属,负荷感应启动子的非限制性示例包括基因groES、kata、cplX、mutM的启动子和它们的同源版本;对于曲霉属,负荷感应启动子的非限制性示例包括基因bipA、c1xA的启动子和它们的同源版本。负荷感应启动子还可以是此类天然启动子的杂合、混杂或截短版本,只要此类启动子仍保持对负荷的反应。

[0041] 负荷成瘾微生物细胞(Burden-addicted microbial cell)是指被工程化以包含基因线路(genetic circuit)的微生物细胞,所述基因线路包含可操作地连接到所述细胞的必需基因的至少一种“负荷感应启动子”。优选地,所述细胞具有包含两个必需基因、即可操作地连接到第一负荷感应启动子的第一必需基因和可操作地连接到第二负荷感应启动子的第二必需基因的基因线路,其中所述第一负荷感应启动子相对于所述第一必需基因是异源的,并且所述第二负荷感应启动子相对于所述第二必需基因是异源的。

[0042] 负荷成瘾微生物生产细胞(Burden-addicted microbial production cell)是指经基因工程化以合成期望产物的微生物细胞,所述细胞经进一步基因工程化以包含基因线路,所述基因线路包含可操作地连接到细胞的“必需基因”的所述至少一种、优选两种“负荷

感应启动子”，如上所定义。所述期望产物可以是由所述工程化微生物细胞中的核酸分子编码的蛋白质，或由待由所述细胞表达的一种或多种核酸分子编码的异源途径的产物。

[0043] 非负荷成瘾微生物生产细胞 (Non-burden addicted microbial production cell) 被认为是负荷成瘾微生物生产细胞的“亲代”细胞。它是一种微生物生产细胞，其包含并表达编码产物或编码产物生物合成代谢途径的一种或多种基因，但缺乏包含可操作地连接到如本发明中所公开的必需基因的所述至少一种、优选两种负荷感应启动子的基因线路。这种非负荷成瘾微生物生产细胞中的所述至少一种、优选两种必需基因，其中每个基因可操作地连接到其天然启动子。

[0044] 本发明的期望产物包括(但不限于)当由基因工程化微生物生产细胞合成时，给生产细胞带来适合度代价(负荷)的产物。这种产物的非限制性示例包括有机酸、萜类化合物、类异戊二烯、聚酮化合物、醇、糖、维生素、醛、羧酸、脂肪酸、氨基酸、肽、酶、治疗性蛋白质及其前体(如人生长激素)、胰岛素、胰高血糖素样肽-1、单克隆或多克隆抗体、单片段抗体和纳米抗体。示例进一步包括鸡蛋中天然存在的蛋白质如卵清蛋白，或牛奶中天然存在的蛋白质如酪蛋白、乳凝集素、 α -lact乳铁蛋白、分泌型免疫球蛋白A和G、分泌成分。酶的示例包括淀粉酶、脂肪酶、蛋白酶、barnase、 β -半乳糖苷酶、晶体蛋白、角质酶、PET酶和漆酶以及碳水化合物活性酶如木聚糖酶、地衣聚糖酶、纤维素酶、裂解多糖单氧化酶和果胶酶。

[0045] 必需基因(或必需基因操纵子)是指微生物生产细胞中的一个或多个基因，如果所述基因被下调，将导致微生物生产细胞在生产条件下的生长速率降低。优选地，无论这些生产条件的营养组成如何，必需基因对于生长都是必需的，由此支持细胞生长的此类必需基因的充分表达将不依赖于生产条件下提供的特定抑制剂或营养物的存在与否。标准实验室条件下大肠杆菌必需基因的非限制性示例包括folP-glmM、glmM、murI、asd、thyA、rpoD、nusG、rpsU、accD、degS、fldA、ftsN、hflB、lolA、mraY、mreD、murA、murB、murF、nadD、rplV、rpsG、及其同源物。芽孢杆菌属的非限制性示例包括iscU操纵子(sufC-sufD-sufS-sufU-sufB)、accC操纵子(accB-accC-yhqY)、glmM、ylaN、infA、dapA及其同源物。酵母属的非限制性示例包括FOL1、MED7、RRP40、NOP8、PGI1、NEP1、URA3 LEU2、TRP1、HIS3及其同源物。曲霉属的非限制性示例包括ARG3、adeA及其同源物。棒状杆菌属的非限制性示例包括alr、glmM及其同源物。例如，合适的基因编码负责合成细胞壁组分的酶。此外，本发明上下文中使用的必需基因优选地既不编码待由生产菌株表达的期望产物；也不编码有助于在待由生产菌株表达的异源途径中合成期望产物或其中间体的蛋白质。

[0046] 条件必需基因是允许微生物生产细胞中的负荷成瘾系统通过在必需条件下(如添加抗生素或去除营养物)培养细胞而被“激活”的基因。

[0047] 必需基因基础表达水平是当“负荷感应启动子”未被诱导时，可操作地连接到负荷成瘾微生物生产细胞中的“负荷感应启动子”的每个“必需基因”(如上定义)的转录水平。必需基因可操作地连接的负荷感应启动子相对于所述必需基因是异源的，在某种意义上，所述启动子不是所述必需基因的天然启动子(即使所述启动子可能存在于同一基因组中)，并且因此在自然界中没有发现与所述必需基因可操作地连接。必需基因的所述基础(即未诱导的)水平表达是足以支持细胞在生产条件下(或在指数生长期)生长的水平，所述水平等于或小于其中每个所述必需基因可操作地连接到其天然启动子的对应细胞生长速率的10%、20%、50%、90%或95%。由于在其中负荷感应启动子未被激活的那些细胞中基础水

平必需基因表达的降低的生长速率构成了非生产性细胞、例如在产物生产过程中出现的非生产性突变体的选择性缺点。

[0048] 适合度代价 (Fitness cost) 通过比较非负荷成瘾微生物生产细胞或其细胞培养物 (即缺乏可操作地连接到所述至少一种、优选两种必需基因的负荷感应启动子, 但包含并表达编码产物或编码产物生物合成代谢途径的一种或多种基因) 的最大指数生长速率 (沿生长曲线测量时) 相对于在可比生产条件下生长时的以下的最大指数生长速率来量化: (i) 亲代微生物细胞 (或其细胞培养物), 其缺乏或不能表达编码产物或用于生物合成所述产物的代谢途径的一种或多种基因, 微生物生产细胞从所述亲代微生物细胞衍生 (或 (ii) 衍生自无负荷成瘾微生物生产细胞的逃逸细胞, 其中与无负荷成瘾微生物生产细胞相比, 所述逃逸细胞产生少于 50% 的产物)。在将各细胞或其细胞培养物引入包含用于生产所述产物的底物的培养基后, 在生产开始时测量各细胞或其细胞培养物的最大指数生长速率。

[0049] 生产条件是指待在生产过程中使用的特定培养条件, 并且可以涉及用于通过微生物生产细胞生产期望产物的培养基和/或其他培养条件 (例如温度、pH、搅拌、通气等)。

[0050] 核糖体结合位点 (RBS)、翻译起始区或翻译强度元件是指控制特定信使 RNA 翻译强度的 5' 非翻译区的基因区。

[0051] I. 负荷成瘾微生物生产细胞

[0052] 本发明的第一方面提供了一种经基因工程化以合成产物的微生物生产细胞, 其中所述细胞进一步包含:

[0053] a. 可操作地连接到负荷感应启动子的必需基因,

[0054] 其中所述启动子相对于所述必需基因是异源的; 其中产物的产生赋予所述细胞负荷和/或适合度代价, 并且其中当所述负荷感应启动子被所述负荷和/或适合度代价诱导时, 所述必需基因的表达相对于所述负荷感应启动子未被诱导时所述必需基因的基础水平表达被上调。

[0055] 在本发明的另一方面中, 经基因工程化以合成产物的微生物生产细胞包含:

[0056] a. 可操作地连接到第一负荷感应启动子的第一必需基因, 和

[0057] b. 可操作地连接到第二负荷感应启动子的第二必需基因,

[0058] 其中所述第一负荷感应启动子相对于所述第一必需基因是异源的, 并且所述第二负荷感应启动子相对于所述第二必需基因是异源的;

[0059] 其中产物的合成赋予所述细胞负荷, 并且

[0060] 其中当所述第一负荷感应启动子被所述负荷诱导时, 所述第一必需基因的表达相对于所述第一负荷感应启动子未被诱导时所述第一必需基因的基础水平表达被上调, 并且当所述第二负荷感应启动子被所述负荷诱导时, 所述第二必需基因的表达相对于所述第二负荷感应启动子未被诱导时所述第二必需基因的基础水平表达被上调。

[0061] 在一个实施方案中, 第一和第二必需基因是不同的。在另一个实施方案中, 第一和第二必需基因是相同的。在特定实施方案中, 所述第一和第二必需基因是相同的, 而所述第一和第二负荷感应启动子是不同的。

[0062] 在一个实施方案中, 所述第一和第二负荷感应启动子是不同的。在另一个实施方案中, 所述第一和第二负荷感应启动子是相同的。在特定实施方案中, 所述第一和第二负荷感应启动子是相同的, 而所述第一和第二必需基因是不同的。

[0063] 在一个实施方案中,微生物生产细胞包含多于两种的必需基因,其各自可操作地连接到负荷感应启动子。在一个实施方案中,本发明提供了一种经基因工程化以合成产物的微生物生产细胞,所述微生物细胞包含两种、三种、四种、五种、六种或更多种负荷感应启动子以控制细胞中的必需基因。

[0064] 在一个实施方案中,微生物细胞包含三种必需基因,其各自可操作地连接到负荷感应启动子。在一个实施方案中,微生物细胞包含四种必需基因,其各自可操作地连接到负荷感应启动子。在一个实施方案中,所有的负荷感应启动子是不同的。在另一个实施方案中,负荷感应启动子中的一些是相同的,但它们与不同的必需基因相连。在一个实施方案中,所有的必需基因是不同的。在另一个实施方案中,必需基因中的一些是相同的,但它们与不同的负荷感应启动子相连。

[0065] 经基因工程化以合成产物的微生物生产细胞是包含编码产物或编码产物合成代谢途径的一种或多种基因的细胞,其中任选地,所述一种或多种基因可操作地连接到组成型或诱导型启动子。在一个实施方案中,编码产物或编码产物合成代谢途径的所述一种或多种基因可以是重组的。

[0066] 根据一个实施方案,本发明的所述微生物生产细胞的所述负荷和/或适合度代价通过相对于缺乏或不能表达编码产物或编码产物合成代谢途径的一种或多种基因并且所述微生物生产细胞从其衍生的亲代微生物细胞的最大指数生长速率,比较包含编码产物或编码产物合成代谢途径的一种或多种基因(但缺乏可操作地连接到所述至少一种、优选两种必需基因的负荷感应启动子)的微生物生产细胞的最大指数生长速率来量化,其中各个细胞在基本上相同的生产条件下培养。所述微生物生产细胞的负荷和/或适合度代价优选地对应于选自 $\geq 5\%$ 、 $\geq 10\%$ 、 $\geq 15\%$ 、 $\geq 20\%$ 、 $\geq 25\%$ 、 $\geq 35\%$ 和 $\geq 45\%$ 的量化最大指数生长速率的百分比降低。

[0067] 根据本发明的负荷成瘾微生物生产细胞可以是任何原核或真核微生物,如细菌、酵母和丝状真菌。

[0068] 在一个实施方案中,本发明的微生物生产细胞是原核生物。合适的细菌的非详尽列表给出如下:属于选自埃希氏杆菌属(*Escherichia*)、乳杆菌属(*Lactobacillus*)、乳球菌属(*Lactococcus*)、棒状杆菌属、芽孢杆菌属、醋酸杆菌属(*Acetobacter*)、不动杆菌属(*Acinetobacter*)、假单胞菌属(*Pseudomonas*);丙酸杆菌属(*Propionibacterium*)、拟杆菌属(*Bacteroides*)和双歧杆菌属(*Bifidobacterium*)的属的物种。

[0069] 在一个优选的实施方案中,本发明提供了经基因工程化以合成产物的微生物生产细胞,所述细胞包含可操作地连接到负荷感应启动子的必需基因;其中所述微生物生产细胞属于芽孢杆菌属。在另一个实施方案中,本发明的微生物生产细胞是真核生物,如酵母或真菌。在某些实施方案中,真核生物可以是酵母属(*Saccharomyces*)、*Komagataella*属或曲霉属的成员。

[0070] 合适的酵母的非详尽列表给出如下:属于酵母属的酵母,例如酿酒酵母、克鲁弗酵母(*S.kluyveri*)、贝酵母(*S.bayanus*)、少孢酵母(*S.exiguus*)、*S.sevazzi*、葡萄汁酵母(*S.uvarum*);属于克鲁维酵母属(*Kluyveromyces*)的酵母,例如马克斯克鲁维酵母乳酸变种(*K.lactis K.marxianus var.marxianus*)、耐热克鲁维酵母(*K.thermotolerans*);属于假丝酵母属(*Candida*)的酵母,例如产朊假丝酵母(*C.utilis*)、热带假丝酵母

(*C.tropicalis*)、白色假丝酵母(*C.albicans*)、解脂假丝酵母(*C.lipolytica*)、易变假丝酵母(*C.versatilis*)；属于毕赤酵母属(*Pichia*)的酵母,例如树干毕赤酵母(*P.stipidis*)、巴斯德毕赤酵母(*P.pastoris*)、嗜山梨醇毕赤氏酵母(*P.sorbitophila*),或其他酵母属,例如隐球菌属(*Cryptococcus*)、德巴利酵母属(*Debaromyces*)、汉逊酵母属(*Hansenula*)、耶氏酵母属(*Yarrowia*)、接合酵母属(*Zygosaccharomyces*)或裂殖酵母属(*Schizosaccharomyces*)。

[0071] 合适的丝状真菌的非详尽列表给出如下:属于青霉属(*Penicillium*)、根霉菌属(*Rhizopus*)、镰孢菌属(*Fusarium*)、梭链孢属(*Fusidium*)、赤霉菌属(*Gibberella*)、毛霉属(*Mucor*)、被孢霉属(*Mortierella*)、木霉属(*Trichoderma*)、嗜热真菌属(*Thermomyces*)、链霉菌属(*Streptomyces*)和曲霉属的丝状真菌。更具体地,丝状真菌可以选自尖孢镰刀菌(*Fusarium oxysporum*)、黑曲霉(*A.niger*)、泡盛曲霉(*A.awamori*)、米曲霉(*A.oryzae*)和构巢曲霉(*A.nidulans*)。

[0072] 由本发明的微生物生产细胞合成的一种或多种产物是带来细胞负荷和适合度代价(反映为生长速率的降低)的产物,所述一种或多种产物包括氨基酸、有机酸、萜类化合物、类异戊二烯、聚酯化合物、醇、糖、维生素、醛、羧酸、脂肪酸、肽、酶、治疗性蛋白质及其前体(如人生长激素)、胰岛素、胰高血糖素样肽-1、单克隆和多克隆抗体、单片段抗体和纳米抗体。示例进一步包括鸡蛋中天然存在的蛋白质如卵清蛋白,或牛奶中天然存在的蛋白质如酪蛋白、乳凝集素、 α -乳白蛋白、 β -乳球蛋白、骨桥蛋白、乳铁蛋白、分泌型免疫球蛋白A和G、分泌成分。酶的进一步示例是淀粉酶、脂肪酶、蛋白酶、barnase、 β -半乳糖苷酶、晶体蛋白、角质酶、PET酶和漆酶以及碳水化合物-活性酶如木聚糖酶、地衣聚糖酶、纤维素酶、裂解多糖单氧化酶和果胶酶。在一个实施方案中,所述产物不是本发明的微生物生产细胞的天然产物,因此不是由衍生所述微生物生产细胞的亲代细胞产生的。在一个实施方案中,编码产物或编码产物合成代谢途径的一种或多种基因对于本发明的微生物生产细胞是异源的;其中所述一种或多种基因可以是转基因。

[0073] II. 本发明的微生物生产细胞的特征

[0074] II.i 适合度代价-生产负荷

[0075] 微生物细胞中期望产物的异源表达或产生期望产物的途径给细胞带来了负荷。负荷还可以来自天然产物的过量生产,例如在诱变后选择的微生物菌株的细胞中。所述负荷可以被描述为导致细胞生长较慢的适合度代价(实施例1;图1)。这可能是由于微生物细胞在生产条件下在本发明产物的生物合成期间所经历的一般和/或产物特异性代谢毒性、潜在有限细胞资源和代谢物的使用和/或消耗增加。特别是在工业规模的发酵过程中,适合度代价和较慢生长的组合对发酵罐中的细胞群体施加了选择压力,随着时间的推移,由于遗传突变或非遗传变异导致的非生产性细胞的积累,导致产物收率下降(实施例2,图6中的BEG34)。

[0076] 在本发明中,编码产物或编码产物生物合成代谢途径的一种或多种基因在微生物生产细胞中的存在和表达赋予了适合度代价(负荷)。

[0077] 微生物生产细胞的生产产物的适合度代价通过比较非负荷成癮微生物生产细胞(即缺乏可操作地连接到必需基因的至少一种、优选两种负荷感应启动子,但包含并表达编码产物或编码产物生物合成代谢途径的一种或多种基因)的生长速率相对于以下的生长速

率来量化：(i) 亲代微生物细胞，非负荷成瘾微生物生产细胞从所述亲代微生物细胞衍生，其中所述亲代微生物细胞缺乏编码产物或用于生物合成所述产物的代谢途径的一种或多种基因；或(ii) 衍生自无负荷成瘾微生物生产细胞的逃逸细胞，其中与无负荷成瘾微生物生产细胞相比，所述逃逸细胞产生少于50%的产物。所述微生物生产细胞的负荷和/或适合度代价优选地对应于选自 $\geq 5\%$ 、 $\geq 10\%$ 、 $\geq 15\%$ 、 $\geq 20\%$ 、 $\geq 25\%$ 、 $\geq 35\%$ 和 $\geq 45\%$ ，更优选至少5%的量化生长速率的百分比降低。

[0078] 优选地，当在基本上相同的条件(选择这些条件以接近地模拟最终将发生大规模发酵的条件)下培养时，通过测量它们的最大指数生长速率，对相应的微生物细胞进行相对生长速率的量化。生长速率(growth rate)的所有使用都是指术语特定生长速率(specific growth rate)。

[0079] II.ii必需基因

[0080] 本发明的微生物生产细胞包含分别编码其表达是细胞生长和/或存活所需的至少一种、优选两种蛋白质的至少一种、优选两种必需基因。

[0081] 在一个实施方案中，微生物生产细胞包含分别编码第一和第二蛋白质(其中所述第一和第二蛋白质可以相同或不同)的第一和第二必需基因，其中所述第一和所述第二蛋白质两者的表达是细胞生长和/或存活所需的。

[0082] 优选地，所述至少一种、优选两种必需基因及其表达不会间接导致生产细胞期望产物的产量下降。此外，本发明上下文中使用的一种或多种必需基因不编码或导致待在微生物生产细胞中表达的异源途径的期望产物或中间体的合成。

[0083] 所述至少一种、优选两种必需基因优选地是一种或多种非条件必需基因，使得无论生长培养基的组成或培养细胞的条件如何，所述一种或多种基因的表达对于细胞生长和/或存活都是必需的。

[0084] 在本发明的一个实施方案中，当生产细胞是原核生物，如大肠杆菌时，所述至少一种、优选两种非条件必需基因选自folP-glmM、glmM、murI、asd、thyA、usA、rpoD、nusG、rpsU、accD、degS、fldA、ftsN、hflB、lolA、mraY、mreD、murA、murB、murF、nadD、rplV和rpsG、及其同源物。

[0085] 在本发明的另一个实施方案中，当生产细胞是原核生物如芽孢杆菌属菌株时，所述至少一种、优选两种非条件必需基因选自iscU操纵子、accC操纵子、glmM、ylaN、infA和dapA、及其同源物。

[0086] 在本发明的一个实施方案中，当生产细胞是原核生物如棒状杆菌属菌株时，所述至少一种、优选两种非条件必需基因选自alr、glmM、及其同源物。

[0087] 在本发明的一个实施方案中，当生产细胞是真核生物时，必需基因在非条件时选自酿酒酵母FOL1、MED7、RRP40、NOP8、PGI1、NEP1及其同源物；并且在有条件时选自酿酒酵母URA3、LEU2、TRP1、HIS3及其同源物。

[0088] 在本发明的一个实施方案中，当生产细胞是丝状真菌如曲霉属时，必需基因选自ARG3、adeA、ERG10、PFS2和TUB1及其同源物。

[0089] 在本发明的一个实施方案中，所述至少一种、优选两种必需基因是导致合成营养缺陷型生长所需的产物或对生长抑制剂如抗生素、特异性毒素、原毒素等具有抗性所需的蛋白质产物的条件性必需基因。

[0090] 当所述至少一种、优选两种必需基因是条件性必需时,则必须调整生产细胞的生长培养基/生长条件的组成,如使用缺乏特定营养物或补充有生长抑制剂的生长培养基。

[0091] 在一个进一步的实施方案中,生产细胞可以包含非条件和条件必需基因的组合。

[0092]

表 1 必需基因的非详尽列表			
SEQ ID NO.	基因 ^{#1}	一种或多种基因编码产物	
	大肠杆菌		Ecocyc/ Biocyc 登录 ID^{#2}
1	folP-glmM	编码二氢蝶酸合酶和磷酸葡萄糖胺变位酶的基因的操纵子	EG50011
3	glmM	磷酸葡萄糖胺变位酶	EG11553
5	murI	谷氨酸盐消旋酶	EG11204
7	asd	天冬氨酸半醛脱氢酶	EG10088
9	thyA	胸苷酸合酶	EG11002
11	rpoD	RNA 聚合酶, $\sigma 70$	EG10896
13	nusG	转录终止因子 NusG	EG10667
15	rpsU	30S 核糖体亚单位蛋白 S21	EG10920
17	accD	乙酰辅酶 A 羧基转移酶亚单位 β	EG10217
19	degS	丝氨酸内切蛋白酶	EG11652
21	fldA	黄素氧化还原蛋白 1	EG10318
23	ftsN	细胞分裂蛋白 FtsN	EG11529
25	hflB	ATP 依赖性锌金属蛋白酶 FtsH	EG11506
27	lolA	外膜脂蛋白载体蛋白	G6465
29	mraY	磷-N-乙酰胞壁酰基-五肽转移酶	EG10604
31	mreD	细胞形状决定蛋白 MreD	EG10610
33	murA	UDP-N-乙酰氨基葡萄糖 1-羧乙烯基转移酶	EG11358
35	murB	UDP-N-乙酰烯醇丙酮酰葡萄糖胺还原酶	G11205
37	murF	D-丙氨酰-D-丙氨酸添加酶	EG10622

[0093]

39	<i>nadD</i>	烟酸-单核苷酸腺苷酰转移酶	G6350
41	<i>rplV</i>	50S 核糖体亚单位蛋白 L22	EG10882
43	<i>rpsG</i>	30S 核糖体亚单位蛋白 S7	EG10906
	酿酒酵母		SGD ID^{#3}
45	<i>FOL1</i>	二氢新蝶呤醛缩酶	SGD:S000005200
47	<i>MED7</i>	介体复合物亚单位 MED7	SGD:S000005495
49	<i>RRP40</i>	外来体非催化核心亚单位 RRP40	SGD:S000005502
51	<i>NOP8</i>	60S 核糖体亚单位生物发生所需的核仁蛋白	SGD:S000005504
53	<i>PGI1</i>	葡萄糖-6-磷酸异构酶	SGD:S000000400
55	<i>NEP1</i>	18S rRNA 假尿苷甲基转移酶	SGD:S000004176
57	<i>URA3</i>	乳清酸核苷-5'-磷酸脱羧酶	SGD:S000000747
59	<i>LEU2</i>	3-异丙基苹果酸脱氢酶	SGD:S000000523
61	<i>TRP1</i>	磷酸核糖邻氨基苯甲酸异构酶	SGD:S000002414
63	<i>HIS3</i>	咪唑甘油磷酸脱水酶	SGD:S000005728
	芽孢杆菌属		BL 百脉根属 (lotus) 标记^{#4}
147	<i>glmM</i>	磷酸葡萄糖胺变位酶	BL02703
149	<i>ylaN</i>	控制细胞形状的蛋白质	BL02974
151	<i>infA</i>	翻译起始因子	BL01028
153	<i>dapA</i>	二氢吡啶二羧酸合酶	BL01208
165	<i>sufC</i> (在 <i>iscU</i> 操纵子中)		
167	<i>sufD</i> (在 <i>iscU</i> 操纵子中)		
169	<i>sufS</i> (在 <i>iscU</i> 操纵子中)		
171	<i>sufU</i> (在 <i>iscU</i> 操纵子中)		
173	<i>sufB</i> (在 <i>iscU</i> 操纵子中)		
175	<i>accB</i> (在 <i>accC</i> 操纵子中)		
177	<i>accC</i> (在 <i>accC</i> 操纵子中)		
179	<i>yghY</i> (在 <i>accC</i> 操纵子中)		
	棒状杆菌属^{#3}		
181	<i>glmM</i>	磷酸葡萄糖胺变位酶	G18NG-10145
183	<i>alr</i>	丙氨酸消旋酶	G18NG-10150
	曲霉属^{#5}		
185	<i>ARG3 / argB</i>	鸟氨酸氨甲酰转移酶	AO090023000856
187	<i>adeA</i>	磷酸核糖基氨基咪唑-琥珀酰氨基合成酶	AO090020000395
240	<i>ERG10</i>	乙酰辅酶 A 乙酰转移酶	AO090103000012
242	<i>PFS2</i>	葡萄糖胺 6-磷酸合成酶	AO090003000003
244	<i>TUB1</i>	α -1 微管蛋白	AO090005000840

[0094] ^{#1}根据本发明的必需基因是编码与由表中列出的相应必需基因编码的蛋白质具有至少70%、80%、90%、95%或甚至100%氨基酸同源性的蛋白质的基因。

[0095] ^{#2}BioCyc/EcoCyc数据库集合是途径/基因组数据库的大型在线集合-通过<https://biocyc.org/>访问

[0096] ^{#3}SGD是“酵母基因组数据库”-通过<https://yeastgenome.org/>访问

[0097] ^{#4}BL:通过https://www.genome.jp/kegg-bin/show_organism?org=bli访问

[0098] ^{#5}Aspergillusgenome.org

[0099] 基因的必要性可以通过创建细胞来确定,在所述细胞中相应的基因被敲除,其中非条件性基因敲除将导致细胞生存力的丧失,或生长失败,而与生长条件无关;而条件性基

因敲除将导致生长失败或细胞生存力的丧失,这取决于培养的组成培养基/条件。

[0100] 合适的必需基因可以通过进行测试测定来鉴定,所述测试测定将经测定的必需基因的天然启动子替换为诱导型启动子,如L-阿拉伯糖诱导型 p_{BAD} (在大肠杆菌中)、木糖诱导型 p_{XYL} (在枯草芽孢杆菌或地衣芽孢杆菌中)或半乳糖诱导型 p_{GAL} (在酿酒酵母或毕赤酵母中)或 p_{lac} (在大肠杆菌或乳杆菌中)或 p_{thiA} (在曲霉属中)。在这种必需基因测定中,使用标准方法将包含诱导型启动子的整合DNA构建体染色体整合到宿主微生物细胞中。在原核生物的情况下,整合DNA优选地含有RBS目录(例如,表3),以指导必需基因以不同的速率翻译,从而说明不同启动子的不同基线表达水平和必需基因所需的不同基线所需表达水平,以支持以对应的野生型特异性生长速率生长。在整合DNA的靶向整合后,将整合体细胞铺在补充有许可条件(诱导物的温度或浓度: $>0.9\%$ L-阿拉伯糖、 $>0.8\%$ 木糖、 $>1\%$ IPTG、 $>2\%$ 半乳糖)的板上。随后在标准生长曲线测定中测试细胞对诱导物条件的合成成瘾,并且合适的必需基因可以被鉴定为诱导物依赖性生长速率可以被鉴定的基因。在不存在诱导条件的情况下,可以观察到诱导物依赖性生长,生长速率降低 $>5\%$ 。合适的必需基因的特征在于,在与生产相关的培养基中,在最佳诱导物条件下,赋予指数期特定生长速率的降低小于 $<5\%$ 。

[0101] 鉴定必需基因的方法在“基因必需性-方法和方案”2015年1月[DOI 10.1007/978-1-4939-2398-4]中详细描述,这使得技术人员能够确定基因在特定生长条件下对于微生物细胞生长是否是必需的;包括用于绘制和鉴定以下各项的必需基因的方法:空肠弯曲杆菌(*Campylobacter jejuni*);血链球菌(*Streptococcus sanguinis*);牙龈卟啉单胞菌(*Porphyromonas gingivalis*);大肠杆菌;钩端螺旋体病(*Leptospirosis*);结核分枝杆菌(*Mycobacterium tuberculosis*);铜绿假单胞菌(*Pseudomonas aeruginosa*);和白色念珠菌(*Candida albicans*);主要通过筛选转座子标记的文库。另外,其中描述的各种“计算工具”使得技术人员能够直接预测和鉴定微生物基因组中编码必需蛋白质的基因,这得益于全基因组微生物序列的广泛可得性;和许多已知必需基因的结构特征。因此,向技术人员提供了必需基因的两个数据库;以及各种可自由访问的在线工具和算法,用于在没有过度负荷的情况下成功且可重复地鉴定微生物细胞基因组中的大量必需基因。应进一步注意的是,酿酒酵母中82%的必需基因编码与另一种生物中的蛋白质相似的必需蛋白质[Giaever G等人,2002];证明大多数必需基因,无论是已知的还是在一种微生物(例如酵母)中鉴定的,将在其他微生物中具有对应的基因,并且因此可以通过简单的BLAST搜索在选择微生物中进行鉴定。

[0102] II.iii 负荷感应启动子

[0103] 如上所述,本发明的微生物生产细胞包含分别可操作地连接到编码细胞生长和/或存活所需的至少一种、优选两种蛋白质的所述至少一种、优选两种必需基因的至少一种、优选两种负荷感应启动子。负荷感应启动子是在微生物生产细胞中被诱导的启动子,所述细胞由于其产物的生产或过量生产而经受负荷和/或适合度代价。通常,当微生物细胞的产物生产导致细胞生长速率降低(这又归因于生产的适合度代价)时,将发生负荷感应启动子的诱导。重要的是,根据本发明,所述至少一种、优选两种负荷感应启动子是由产物生产导致的生产细胞中的“负荷状态”诱导的,而不是由产物本身诱导的。由于使用所述至少一种、优选两种负荷感应启动子和必需基因来提高微生物生产细胞的生产率在很大程度上不依赖于其被工程化制造的产物,因此由本发明赋予的负荷成瘾支持广泛的微生物生产细胞应

用。

[0104] 由于必需基因可操作地连接到负荷感应启动子,因此其在负荷感应启动子被诱导时的转录和表达水平仅增加到高于基础水平。合适的负荷感应启动子是这样一种启动子,当未被诱导时,其允许可操作地连接的必需基因的基础水平表达,所述表达显著限制或阻止细胞生长和/或存活。必需基因的基础水平表达是足以支持细胞在生产条件下(或在指数生长期)生长的水平,所述水平等于或小于其中所述必需基因可操作地连接到其天然启动子的对应细胞生长速率的10%、20%、50%、90%或95%。当负荷感应启动子未被有效产物合成的负荷激活时,必需基因的这种基础表达水平将在负荷成瘾细胞中发挥作用,因此,在缺乏产物合成时,由此导致的生长速率降低将构成选择性劣势。

[0105] 与必需基因可操作地连接到其天然启动子的细胞相比,包含处于未诱导状态的负荷感应启动子的细胞将明显较慢地生长(实施例1,图2)。相比之下,与不生产产物的细胞相比,当细胞合成期望产物时,包含与必需基因相连的负荷感应启动子的微生物生产细胞表现出显著增加的生长速率(实施例1,图3)。此外,尽管存在生产产物的固有负荷,但本发明的微生物生产细胞可以令人惊讶地实现与亲代微生物细胞的生长速率没有区别的初始生长速率,所述亲代微生物细胞缺乏编码产物或用于生物合成所述产物的代谢途径的一种或多种基因并且微生物生产细胞从其衍生(实施例1,图4)。因此,在生产条件下,本发明的微生物生产细胞将比停止合成产物的生产细胞的逃逸变体具有选择优势,从而降低逃逸率。

[0106] 为了增加期望产物的产量并进一步延长生产细胞具备生产性的时段,多于一种的负荷感应启动子可能适用于促进有益细胞变体的生长。通过使用两种或更多种不同的负荷感应启动子来使生产细胞成瘾,可以使高产细胞变体更独特地与随着时间的推移可能来自同一祖先起始细胞的低生产性(遗传/非遗传)细胞变体区分开来。例如,环境因素(例如由生产培养中的外部压力导致的环境因素)可能在某种程度上激活甚至更低生产性的细胞,从而允许生长。

[0107] 通过使用多于一个的负荷感应器来控制必需基因,可以进一步控制指导所施加的基于成瘾的选择方案的转录空间。这对限制潜在的细胞逃逸模式可能很重要,在这种模式下,单个负荷生物感应器的负荷成瘾反应可能允许足够的生长以允许产量下降。

[0108] 表2提供了潜在负荷感应启动子的非详尽列表。

[0109] 在一个实施方案中,负荷感应启动子就微生物生产细胞而言可以是天然启动子,或经修饰的天然启动子,同时就必需基因而言是异源的。

[0110] 在一个实施方案中,启动子含有天然TF/ σ 因子结合位点。合适的负荷感应启动子是大肠杆菌 σ^{32} 调节子的启动子,因为它们通常响应于蛋白质的过量表达及其细胞内聚集而被上调。大肠杆菌 σ^{32} 启动子的非详尽列表包括在本文的表2中,而另外的合适的启动子由Nonaka等人,2006鉴定。

[0111] 在一个实施方案中,负荷感应启动子是非植物性 σ 因子调节的启动子。在优选的实施方案中,负荷感应启动子经由 σ^{32} 调节子激活,并且可以选自大肠杆菌基因htpG、ibpA、clpB、yccV、grpE、ycjX、ldhA、mutM、ybbN、prlC、groES、fxsA和htpX的启动子。

[0112] 在一个实施方案中,负荷感应启动子是mutM基因的启动子。mutM编码一种功能上保守的DNA糖基化酶,其负责启动一种最常见的氧化应激诱导的DNA损伤的修复,即鸟嘌呤氧化成7,8-二氢-8-氧鸟嘌呤(8-oxoG)(Jain等人,2007)。

[0113] 其他合适的负荷感应启动子可以选自大肠杆菌基因rrnB、rrnE、cspD、katE、xthA、uspE、gadB、ahpC、katG、grxA、oxyS、poxB、trxC的启动子。

[0114] 负荷感应启动子还包括与细胞碳营养状态和生长期相关的启动子。因此,在一个实施方案中,负荷感应启动子是毒素基因cspD的启动子,其被降低的生长速率和碳饥饿激活(Uppal等人,2014;Yamanaka和Inouye,1997),并因此预测在由于代谢产物的产生而经受负荷或适合度代价的细胞中差异上调。

[0115] 在一个实施方案中,负荷感应启动子是核糖体RNA启动子,特别是rrnB和rrnE基因的启动子,其响应于营养供给,被Fis蛋白诱导,并被ppGpp警报素抑制。Rrn启动子活性在具有高蛋白质产量的细胞中被诱导,如在其早期指数生长期所见(Nonaka等人,2006)。

[0116] 在一个实施方案中,负荷感应启动子是响应于与代谢负荷和微生物异源生产相关的氧化应激的启动子(Dragosits和Mattanovich,2013)。氧化应激反应启动子包括由大肠杆菌OxyR或其同源物(如芽孢杆菌PerR)调节的启动子,其激活与改善和保护细胞对氧化损伤的反应相关的几个基因;和属于rpoS(σ^S)的基因的启动子,其在蛋白质的过表达过程中,主要是在稳定生长期被上调。

[0117] 对于芽孢杆菌属,合适的负荷感应启动子可以选自例如HrcA调节的启动子,包括groES的启动子。其他合适的负荷感应启动子可以选自枯草芽孢杆菌基因ctsR、dnaK、perR、hrcA、spx、sigB、yflT、mutM的启动子。

[0118] 对于棒状杆菌属,此类合适的负荷感应启动子可以选自katA、cplX、mutM和groES。

[0119] 对于酵母菌属(Saccharomyes),合适的负荷感应启动子可以选自基因KAR2、PDI1、SAA1、FPR2、RPL3、RPL6A、RPL28、OGG1、RAD51、RAD54的启动子。

[0120] 对于曲霉属,此类合适的负荷感应启动子可以选自bipA、PDI、clxA以及酿酒酵母OGG1、RAD51和RAD54的同源物。

表 2 负荷感应启动子的非详尽列表

编号 *	启动子	由启动子调节的基因的经编码蛋白质	Ecocyc/ Biocyc 登录 ID**/AspGD ID***
2	p(<i>htpG</i>) (SEQ ID NO.: 65)	热休克蛋白 G(HtpG)是 Hsp90 家族的成员, 是一种分子伴侣, 其有助于细胞中蛋白质聚集体的重折叠。 <i>htpG</i> 转录经由 σ_{32} 调节子被激活。(酵母同源物: Hsp90)	EG10461-MONOMER
3	p(<i>ibpA</i>) (SEQ ID NO.: 66)	小热休克蛋白 LbpA 和 LbpB 是与异源蛋白表达和包涵体形成相关的伴侣蛋白, 其有助于聚集蛋白的重折叠。 <i>ibpA</i> 转录经由 σ_{32} 调节子被激活。	EG11534
4	p(<i>clpB</i>) (SEQ ID NO.: 67)	ClpB 是一种酪蛋白水解蛋白酶, 其催化蛋白质解聚。 <i>clpB</i> 转录经由 σ_{32} 和/或 σ_{70} 调节子被激活。(酵母同源物: Hsp104)。	EG10157
5	p(<i>yccV</i>) (SEQ ID NO.: 68)	YccV 也称为 HspQ, 其在热休克过程中参与变性蛋白的降解。 <i>yccV</i> 转录经由 σ_{32} 调节子被激活。	G6500
6	p(<i>grpE</i>) (SEQ ID NO.: 69)	GrpE 是 DnaK-DnaJ-GrpE 伴侣系统中的核苷酸交换因子。 <i>grpE</i> 启动子经由 σ_{32} 调节子被激活。	EG10416
7	p(<i>ycjX</i>) (SEQ ID NO.: 70)	<i>ycjX</i> 经由 σ_{32} 调节子被激活。	G6659
8	p(<i>ldhA</i>) (SEQ ID NO.: 71)	D-乳酸脱氢酶和 <i>ldhA</i> 启动子经由 σ_{32} 调节子被激活。	G592
9	p(<i>mutM</i>) (SEQ ID NO.: 72)	DNA-甲酰胺嘧啶糖基化酶是修复某些类型 DNA 损伤的碱基切除修复途径的组分。 <i>mutM</i> 启动子经由 σ_{32} 调节子被激活, 但在氨基酸饥饿时被 DksA-ppGpp 结合抑制。	EG10329
10	p(<i>ybbN</i>) (SEQ ID NO.: 73)	YbbN 也称为 CnoX, 是一种伴侣蛋白 (chaperedoxin)。 <i>ybbN</i> 启动子经由 σ_{32} 调节子被激活。	G6268
11	p(<i>prlC</i>) (SEQ ID NO.: 74)	PrlC 也称为寡肽酶 A; 并且属于 σ_{32} 调节子。	EG11441
12	p(<i>groES</i>) (SEQ ID NO.: 75)	GroES 是一种伴侣蛋白, 并且属于 σ_{32} 和 σ_{70} 调节子。	EG10600
13	p(<i>fxsA</i>) (SEQ ID NO.: 76)	FxsA 是属于 σ_{32} 调节子的多胞膜蛋白。	G7832
14	p(<i>htpX</i>) (SEQ ID NO.: 77)	HtpX 是一种锌依赖性内切蛋白酶和热休克蛋白, 其属于 σ_{32} 调节子。	EG10462
15	p(<i>rrnB</i>)	16S rRNA 基因, 其启动子响应于营养供给, 并在	EG30085

[0121]

[0122]

	(SEQ ID NO.: 78)	具有高蛋白质产量的细胞中被诱导。	
16	<i>p(rrnE)</i> (SEQ ID NO.: 79)	16S rRNA 基因, 其启动子响应于营养供给, 并在具有高蛋白质产量的细胞中被诱导。	EG30088
19	<i>p(cspD)</i> (SEQ ID NO.: 80)	CspD 是一种毒素。 <i>cspD</i> 启动子反应与生长速率成反比。 <i>cspD</i> 的转录由环-AMP-CRP、H-NS 和 ppGpp 激活。	EG11111
21	<i>p(katE)</i> (SEQ ID NO.: 81)	单功能过氧化氢酶(HPII), 其负责在细胞中将有毒的过氧化氢分解成水和氧气。 <i>katE</i> 转录依赖于 σ^S 并被 Fur-Fe ²⁺ 和 ppGpp 激活。	EG10509
22	<i>p(xthA)</i> (SEQ ID NO.: 82)	多功能核酸内切酶, 脱氧核糖核酸外切酶 III, 其例如去除由于过氧化氢而损伤的 DNA。 <i>xthA</i> 属于 σ^{32} 调节子。	EG11073
23	<i>p(uspE)</i> (SEQ ID NO.: 83)	通用应激蛋白 E; 其在稳定期被诱导, 作为各种应激条件如氧化应激的结果。 <i>uspE</i> 转录被 GadX 和 ppGpp 激活。	EG11246
24	<i>p(gadB)</i> (SEQ ID NO.: 84)	谷氨酸脱羧酶 B 通过将谷氨酸转化为 γ -氨基丁酸, 在反应中接受一个质子, 帮助细胞在暴露于酸性条件时保持接近中性的细胞内 pH。 <i>gadB</i> 是 σ^{70} 和 σ^S 调节子的一部分, 并且由 AdiY、GadE、GadX、Pasp ⁵⁶ -RcsB 和 ppGpp 激活。	EG11490
25	<i>p(ahpC)</i> (SEQ ID NO.: 85)	烷基氢过氧化物还原酶 C 通过催化内源性合成的氢过氧化物的还原, 在保护细胞免受氧化应激方面发挥作用。 <i>ahpC</i> 是 σ^{70} 的一部分, 并且转录被 OxyR 激活。	EG11384
26	<i>p(katG)</i> (SEQ ID NO.: 86)	双功能氢过氧化物酶 I(HPI), 其既作为过氧化氢酶又作为过氧化物酶。	EG10511
27	<i>p(grxA)</i> (SEQ ID NO.: 87)	谷氧还蛋白主要负责催化二硫化物的还原。 <i>grxA</i> 的转录起始被 OxyR 激活。	EG10417
28	<i>p(oxyS)</i> (SEQ ID NO.: 88)	<i>oxyS</i> 是由氧化应激经由 OxyR 激活的小 RNA 基因。	EG31116
29	<i>p(poxB)</i> (SEQ ID NO.: 89)	丙酮酸氧化酶, 其经由 OxyR, 但也通过 Cra、MarA、fnr、ppGpp 和 σ^S 被氧化和渗透应力上调	EG10754
30	<i>p(trxC)</i> (SEQ ID NO.: 90)	硫氧还蛋白, 其在氧化应激过程中经由 OxyR 被诱导	EG11887
酵母属			
	<i>p(KAR2)</i> (SEQ ID NO.: 91)	包含未折叠蛋白质反应元件的 HAC 上调的启动子, 积极响应于高产量	
	<i>p(PDII)</i> (SEQ ID NO.: 92)	包含未折叠蛋白质反应元件的 HAC 上调的启动子, 积极响应于高产量	
	<i>p(SAA1)</i> (SEQ ID NO.: 93)	包含未折叠蛋白质反应元件的 HAC 上调的启动子, 积极响应于高产量	
	<i>p(FPR2)</i> (SEQ ID NO.: 94)	包含未折叠蛋白质反应元件的 HAC 上调的启动子	
	<i>p(RPL3)</i> (SEQ ID NO.: 95)	RNA 聚合酶 I 上调启动子, 积极响应于高产量	
	<i>p(RPL6A)</i> (SEQ ID NO.: 96)	RNA 聚合酶 I 上调启动子, 积极响应于高产量	

[0123]

<i>p(RPL28)</i> (SEQ ID NO.: 97)	RNA 聚合酶 I 上调启动子, 积极响应于高产量	
<i>p(OGGI)</i> (SEQ ID NO.: 98)	由异源表达导致的 DNA 损伤反应广泛诱导 DNA 修复系统的转录。	
<i>P(RAD51)</i> (SEQ ID NO.: 99)	由异源表达导致的 DNA 损伤反应广泛诱导 DNA 修复系统的转录。	
<i>P(RAD54)</i> (SEQ ID NO.: 100)	由异源表达导致的 DNA 损伤反应广泛诱导 DNA 修复系统的转录。	
芽孢杆菌属		
<i>P_{groES}</i> (SEQ ID NO.: 212)	GroES 是一种伴侣蛋白, 并且由 SigA 和 HrcA 调节。	BSUB : BSU06020-MONOMER
<i>P_{ctsR}</i> (SEQ ID NO.: 213)	CtsR 是 III 类热休克基因的转录阻遏物, 并且由 SigM、SigA、SigB、SigF、Spx 调节, 也是自动调节的。	BSUB : BSU00830-MONOMER
<i>P_{dnaK}</i> (SEQ ID NO.: 214)	DnaK 是一种分子伴侣, 并且由 HrcA 和 SigA 调节	BSUB : BSU25470-MONOMER
<i>P_{perR}</i> (SEQ ID NO.: 215)	PerR 是过氧化物调节子的转录阻遏物和细胞内 Fe/Mn 比率的感应器, 并且由 SigA 调节, 也是自动调节的。	BSUB : BSU08730-MONOMER
<i>P_{hrcA}</i> (SEQ ID NO.: 216)	HrcA 是 I 类热休克基因的转录阻遏物, 并且由 SigA 调节, 也是自动调节的	BSUB : BSU25490-MONOMER
<i>P_{spx}</i> (SEQ ID NO.: 217)	参与防止蛋白质聚集的转录调节因子, 并且由 PerR、SigW、SigX、SigM 和 SigA 调节	BSUB : BSU11500-MONOMER
<i>P_{sigB}</i> (SEQ ID NO.: 218)	RNA 聚合酶 σ 因子 σ^B 参与一般应激反应, 并且由 SigA 和 CcpA 调节, 也是自动调节的	BSUB : BSU04730-MONOMER
<i>P_{yflT}</i> (SEQ ID NO.: 219)	由 SigB 调节的一般应激蛋白 17M	BSUB : BSU07550-MONOMER
曲霉属***		
<i>P_{bipA}</i> (SEQ ID NO.: 189)	一种参与蛋白质分泌的 ER 定位伴侣蛋白 (BiP)	AO090003000257
<i>P_{clxA}</i> (SEQ ID NO.: 190)	一种参与蛋白质分泌的钙联蛋白同源物	AO090009000313
<i>P_{agsA}</i> (SEQ ID NO.: 239)	1,3- α -D-葡聚糖合酶	AO090010000106
棒状杆菌属**		
<i>p_{groES}</i> (SEQ ID NO.: 191)	伴侣蛋白	G18NG-10159
<i>p_{katA}</i> (SEQ ID NO.: 192)	过氧化氢酶	G18NG-9810
<i>p_{cplX}</i> (SEQ ID NO.: 193)		
<i>p_{mutM}</i> (SEQ ID NO.: 194)	DNA 损伤修复	G18NG-11665

[0124] *编号对应于分配给在本申请的实施例部分中测试的不同菌株的编号。

[0125] **BioCyc/EcoCyc 数据库集合是途径/基因组数据库的大型在线集合 - 通过 <https://biocyc.org/> 访问

[0126] ***对于曲霉属, Aspergillusgenome.org 是一个以相关曲霉属物种基因组为特征的数据库 (AspGD)。

[0127] 不同负荷感应启动子 (例如在表 2 中) 诱导/激活时的基础表达水平和特异性反应曲线是不同的, 当选择与特定微生物生产细胞及其产物具有最佳匹配 (就强度而言) 的负荷感应启动子时, 可以利用这一点。

[0128] 在一些情况下,感兴趣的基因是操纵子的一部分,在这种情况下,操纵子的第一个基因优选地用于定义相关的启动子序列。

[0129] 在生产过程中由微生物生产细胞中的负荷和/或适合度代价诱导的合适的负荷感应启动子可以选自技术人员已知的公共数据库,例如用于大肠杆菌启动子的<https://ecocyc.org>;用于芽孢杆菌属启动子(也适用于棒状杆菌属启动子),它们各自的原核同源物的<https://bsubcyc.org>;用于酿酒酵母启动子及其真菌同源物的<https://yeastgenome.org>;并且<https://aspergillusgenome.org>是一个以相关曲霉属物种基因组为特征的数据库(AspGD)。

[0130] 还可以在测试测定中验证合适的负荷感应启动子,其中候选负荷感应启动子可操作地连接到编码荧光蛋白(例如绿色荧光蛋白)的基因,并整合到经基因工程化以合成产物的微生物生产细胞中,并整合到对应的非生产细胞(对照)中,所述细胞在生产培养基中在50-150代细胞的连续稀释培养实验后或在大规模发酵培养结束时被分离。与在非生产细胞(对照)中相比,负荷感应启动子在生产细胞中诱导可操作地连接到荧光蛋白表达的候选负荷感应启动子的转录,与分离的相应非生产细胞中的活性相比,在生产细胞中诱导至少5%、7.5%、10%、15%、25%、60%、150%、300%。

[0131] II.iv将负荷感应启动子与微生物生产细胞匹配

[0132] 响应于生产产物的负荷或适合度代价,微生物生产细胞表现出以表达某些基因为特征的转录状态,在生产发酵期间,所述基因的表达在对应的非生产性亲代细胞或衍生自微生物生产细胞的非生产性逃逸细胞中不被诱导。负荷或适合度代价的生理性质将取决于由给定微生物生产细胞合成的产物,并将反映在其表达被诱导的基因类型中。在生产给定产物的细胞中诱导的基因启动子的身份可以通过本领域熟知的技术(例如转录组学,参见实施例7)来确定。由于许多经诱导的基因启动子将与表2中列出的那些启动子相关或对应,因此这些为给定微生物生产细胞寻找匹配的负荷感应启动子提供了起点。通过鉴定在感兴趣的生产菌株中被发现特异性上调的基因的5至15个启动子,可以扩展启动子的选择。如本文实施例中所示,用于以产物特异性方式将可操作地连接到必需基因的负荷感应启动子与微生物生产细胞相匹配的方法在实验上是快速的,并且可以在简单的实验室设置中进行。

[0133] II.v翻译控制元件

[0134] 通过必需基因表达调节本发明的微生物生产细胞的生长,要求负荷感应启动子的反应阈值和曲线与必需基因的表达水平平衡;使得基础水平的必需基因表达支持有限的生长或不支持生长,而在高产细胞中,经诱导的负荷感应启动子驱动足够的必需基因表达以支持显著增加的生长速率,优选地,当在指数生长期测量时,生长速率类似于缺乏本发明的负荷感应启动子的生产性细胞的生长速率或低5%。

[0135] 用于平衡负荷感应启动子对必需基因表达水平的负荷反应的一种合适方法是修饰必需基因的翻译强度。在细菌中,翻译强度由起始密码子直接上游的Shine-Dalgarno/核糖体结合位点(RBS)序列定义,而在真核细胞中,翻译起始区、翻译起始位点(TIS)或Kozak元件可以用于修饰翻译强度。大肠杆菌中赋予宽范围翻译强度的RBS在表3.1中提供,而进一步的示例可以在文献(例如Bonde等人,2016)中找到。本领域技术人员可以通过为每个负荷感应启动子构建核糖体结合位点的四个变体(Rugbjerg等人,2018,PNAS[美国国家科学院院刊]),并测试哪个变体能够使所选择必需基因有效调节细胞的生长速率,来平衡经调

节的必需基因的翻译强度。用于地衣芽孢杆菌或枯草芽孢杆菌的示例性RBS在表3.2中提供。表3.3中提供了用于毕赤酵母属的示例性TIS。

[0136]	表 3.1 通常覆盖如通过表达红色荧光蛋白的测定中所定义的广泛表达范围的大肠杆菌中合适的 RBS 的列表 (Bonde 等人, 2016)。	
	RBS	对数 (表达水平)
	ACTTGA	2.80244
	ACTTGC	2.73094
	ACTTGG	3.14847
	ACTTGT	2.77523

[0137]	表 3.2 通常覆盖广泛表达范围的枯草芽孢杆菌或地衣芽孢杆菌中合适的 RBS 的列表。	
	RBS	
	ACGAGA	
	ACGAGC	
	ACGAGT	
	ACGAGA	

[0138]	表 3.3: 巴斯德毕赤酵母中调节必需基因表达的合适 TIS 序列变异的列表。	
	TIS	序列:
	TIS1	TGATAT
	TIS3	ACGTTC
	TIS5	TAGGTT
	TIS9	AAAACA

[0139] II.vi 工程化负荷感应启动子和翻译控制元件

[0140] 选择的天然负荷感应启动子通常被原核生物中天然调节的ORF的-1至-300bp区域(上游)和真核生物中的-1至-500bp区域包含。必须包含的核心启动子及其调节位点(例如转录因子结合位点)的序列边界是普通常识,如对于 σ^{32} (Nonaka等人, 2006)。翻译控制元件/RBS可以被添加至启动子的下游,以避免所选择启动子序列的调节特性的改变。

[0141] II.vii 提高生产水平

[0142] 除了与非生产性细胞相比具有增加的初始生长速率之外,本发明的微生物生产细胞的一个令人惊讶的优势是与缺乏工程化负荷成癮的微生物生产细胞相比,经过多次细胞分裂,所述细胞在大规模发酵期间仍保持显著提高的生产率(在实施例2中模拟;图6)。无论合成的产物类型(如实施例2-5中由微生物生产细胞合成的产物范围所示)如何,都获得了由本发明的微生物生产细胞中负荷成癮赋予的令人惊讶的优势。

[0143] 根据一个实施方案,本发明的微生物生产细胞的特征在于,在从单个细胞进行至少20、25、30、35、40、45、50、55、60、70、100、150、250或400代次细胞分裂后,与相同代次细胞分裂后的非负荷成癮生产细胞相比,产物收率提高。将产物收率测量为每单位底物生产的产物的摩尔数或克数。

[0144] 在一个实施方案中,在从单个细胞进行至少50代次细胞分裂后,与相同代次细胞分裂后的参考非负荷成癮生产细胞相比,生产水平增加了至少10%、25%、50%或80%。

[0145] 微生物细胞工厂工程化中的普遍实践旨在最小化已知由于异源途径的维持和表达而产生的负荷。本发明提供了一种提高细胞工厂收率的反直觉解决方案,因为它使微生物细胞生长和存活依赖于生产其产物的细胞,同时经受恒定的负荷或适合度代价。虽然不

希望被理论束缚,但推测在本发明的微生物生产细胞及其后代上施加的选择压力(其中仅处于负荷状态的生产性细胞可以生长/存活)可以用于在初始等基因细胞的起始群体中的长期培养期间逐步选择高负荷细胞亚群。

[0146] III. 用于制备和鉴定负荷成瘾微生物生产菌株的细胞的方法

[0147] 在进行本发明的优选方法中,通过以下步骤制备和鉴定负荷成瘾微生物生产菌株。为了获得负荷感应启动子和必需基因之间的最佳工作关系,制备和测试具有不同RBS的不同候选负荷感应启动子以调节所选择必需基因可能是相关的。

[0148] III.i 特定生产菌株中一种或多种负荷感应启动子候选物的鉴定

[0149] 可以测试表2中列出的候选负荷感应启动子。可替代地,可以鉴定在感兴趣的特定生产菌株中检测到的独特的正分化基因转录物;它们各自的启动子提供了候选负荷感应启动子的来源。优选地,通过将微生物生产菌株(在生产过程中)的转录物图谱与来源于对应生产菌株的经分离的一种或多种基因或非基因逃逸突变变体进行比较来鉴定此类分化基因转录物,其中这种逃逸突变变体的特征在于至多低50%更低的生产速率。正分化基因转录物可以通过使用转录组学来鉴定;例如通过与非/低产逃逸突变变体相比,对从一种或多种生产性微生物生产细胞中提取的经转录RNA进行RNA测序。经鉴定的启动子可以与微生物生产细胞中的给定必需基因组合进行测试。

[0150] 用于鉴定至少一种、优选两种负荷感应启动子的方法在实施例7中示出;然后将所述一种或多种启动子可操作地连接到用于测试的所选微生物生产细胞中的必需基因。

[0151] III.ii 在生长调节过程中引入负荷感应启动子

[0152] 通过将选自一般候选负荷感应启动子列表(表2)或选自经鉴定的一种或多种特定负荷感应启动子(参见IIIi)的所述至少一种、优选两种启动子可操作地连接到微生物生产细胞中的必需基因来测试所述启动子。必需基因的不同翻译强度可以通过提供同源必需基因的替代RBS序列来同时测试;以及测试不同的必需基因(如第Iiii节所述)。就微生物生产细胞而言,必需基因可以是天然的或异源的。

[0153] 当必需基因是天然基因时,其同源天然启动子可以被例如突变或缺失事件破坏(即变得无功能),并且然后通过靶向引入将负荷感应启动子(其相对于必需基因是异源的)可操作地连接到天然必需基因。在另一个实施方案中,天然必需基因启动子通过靶向引入被负荷感应启动子替代。

[0154] 在微生物中靶向引入基因序列的合适的方法包括细菌中的重组工程,例如大肠杆菌中的Lambda Red重组工程;而同源重组可以用于酵母和丝状真菌。这两个概念都可以任选地与CRISPR组合使用,以提高基因替换效率。

[0155] III.iii 经平衡的负荷感应和生长调节的筛选

[0156] 然后筛选包含可操作地连接到必需基因的插入的负荷感应启动子的微生物生产细胞克隆,以鉴定负荷感应启动子以生长控制方式调节必需基因的克隆。例如,在控制产物生产的条件下(例如使用生产基因可诱导的微生物生产菌株),将许多此类克隆(例如8-96个克隆)的生长与对应的非负荷成瘾微生物生产菌株的细胞进行比较。

[0157] 当在两种菌株的产物合成低/不存在的条件下测量细胞生长时;合适的负荷成瘾生产克隆是表现出比非负荷成瘾生产菌株显著且优选至少5%更低生长速率的克隆(实施例1,图2)。相反,如果产量高/存在产量,合适的负荷成瘾生产克隆中必需基因表达的上调

意味着与相应的非负荷成瘾生产菌株相比时,生长速率的差异显著降低或消除。

[0158] III.iv未受干扰的中枢和生产代谢的验证

[0159] 必需基因的转录调节或表达的变化可能导致生产基因和中枢碳或氮代谢的不必要的间接干扰,从而损害产物形成,并反过来降低负荷成瘾微生物生产细胞的负荷。为了排除此类克隆,在两种菌株都合成产物的条件下测量细胞生长;其中合适的负荷成瘾生产克隆是表现出生长速率等于或低于非负荷成瘾生产菌株的克隆(如实施例2,图7所示)。

[0160] III.v筛选生长速率稳定性

[0161] 通过培养至少20代次细胞分裂,例如通过连续传代和测量生长速率,或通过从放大生产过程的不同步骤中取样,来测试满足上述标准的克隆在模拟生产的条件下的生长速率稳定性。随着时间的推移,合适的负荷成瘾微生物生产细胞保持低于非负荷成瘾生产菌株的生长速率(例如实施例2,图7)。

[0162] IV.使用负荷成瘾微生物生产菌株的细胞生产期望产物的方法

[0163] 本发明的第二方面涉及一种生产期望产物的方法,其包括以下步骤:

[0164] a.提供经基因工程化以合成产物的至少一种微生物生产细胞,其中所述至少一种细胞进一步包含分别可操作地连接到至少一种、优选两种负荷感应启动子的至少一种、优选两种必需基因,并且其中所述一种或多种启动子相对于所述必需基因是异源的,

[0165] b.将所述至少一种经基因修饰的微生物细胞引入包含用于生产所述产物的底物的培养基中,

[0166] c.回收由所述培养物合成的所述产物,

[0167] 其中产物的合成赋予所述至少一种细胞负荷和/或适合度代价,并且其中当所述至少一种、优选两种负荷感应启动子被所述负荷和/或适合度代价诱导时,所述至少一种、优选两种必需基因的表达相对于所述至少一种、优选两种必需基因的基础水平表达被上调;并且其中在所述至少一种生产细胞或其后代细胞中缺乏所述产物合成降低了所述至少一种细胞的生长速率。

[0168] 用于生产期望产物的方法可以包括以下步骤:提供所述至少一种经基因工程化以合成产物的微生物生产细胞的细胞培养物;以及将所述至少一种经基因修饰的微生物细胞的细胞培养物引入包含用于生产所述产物的底物的培养基中。所述细胞在培养基中培养,以允许细胞产生期望的产物并支持细胞的生长和增殖。培养的时间可以根据期望产物进行优化。

[0169] 用于生产期望产物的方法可以进一步包括分离产物和/或将产物配制成组合物如营养组合物、药物组合物、化妆品组合物、洗涤剂组合物、润滑剂组合物或燃料组合物的步骤。

[0170] V.负荷成瘾微生物生产菌株的细胞用于生产期望产物的用途

[0171] 本发明的第三方面涉及本发明的负荷成瘾微生物生产细胞用于生产期望产物的用途,其中在所述负荷成瘾生产菌株或其后代细胞中缺乏产物合成降低了所述菌株的生长速率,即使产物存在于细胞内或细胞外环境中也是如此。

[0172] 实施例

[0173] 实施例1:工程化生产重组人生长激素的负荷成瘾大肠杆菌菌株

[0174] 1.1细胞工厂对异源生产基因的维持和表达对其细胞构成不自然的负荷

[0175] 将亲代大肠杆菌BL21 (DE3) 细胞的生长速率与用质粒pEG34转化的工程化衍生物进行比较,所述质粒pEG34包含编码与绿色荧光蛋白 (GFP) 融合的重组人生长激素 (hGH) 的基因。

[0176] 1.1.1材料和方法

[0177] 将宿主菌株大肠杆菌BL21 (DE3) (耶鲁大学大肠杆菌遗传保藏中心) 的细胞制成电感受态,并通过标准电穿孔方法 (1800V, 25 μ F, 200欧姆, 1mm比色皿宽度) 用质粒pEG34转化, 并且铺在含有氯霉素的LB琼脂板上。

表 4 质粒		
质粒	基因型	表达的蛋白质
pEG34 [SEQ ID No.: 101]	hGH-GFP 融合 camR 基因* [SEQ ID No.: 103]	与 GFP 融合的 hGH [SEQ ID No.: 102] 氯霉素抗性蛋白 [SEQ ID No.: 104]
[0178] pEG0 [SEQ ID No.: 105]	camR 基因* [SEQ ID No.: 103]	氯霉素抗性蛋白 [SEQ ID No.: 104]
pENDU5CAM [SEQ ID No.: 106]	lss 基因 camR 基因* [SEQ ID No.: 103]	溶葡萄球菌酶 [SEQ ID No.: 107] 氯霉素抗性蛋白 [SEQ ID No.: 104]
pMevT [SEQ ID No.: 108]	mevT 操纵子	乙酰乙酰辅酶 A 合酶 [SEQ ID No.: 109] 羟甲基戊二酰辅酶 A 合酶
[0179]	camR 基因* [SEQ ID No.: 103]	[SEQ ID No.: 110] 羟甲基戊二酰辅酶 A 还原酶 [SEQ ID No.: 111] 氯霉素抗性蛋白 [SEQ ID No.: 104]
	pSIM5-tet [Koskiniemi 等 人 2011]	

[0180] *编码氯霉素抗性蛋白 [SEQ ID NO:104] 的 camR 基因 [SEQ ID NO:103] 在 pEG34 [SEQ ID NO:101] 中,位于互补链上的核苷酸位置 3768-4427 处;在 pEG0 [SEQ ID NO:105] 中,位于互补链上的核苷酸位置 2339-2995 处;在 pENDU5CAM [SEQ ID NO:106] 中,位于互补链上的核苷酸位置 1617-2276 处;并且在 pMevT [SEQ ID NO:108] 中,位于互补链上的核苷酸位置 6585-7244 处。

[0181] 在含有 500 μ M IPTG 和 30mg/L 氯霉素的 200 μ L 2xYT 培养基 (16g/L 胰蛋白胨、5g/L NaCl、10g/L 酵母提取物) 上的 96 孔微量滴定板中,在 37 $^{\circ}$ C 下培养单个菌落转化体的预培养物,同时在 E1x808 板阅读器 (伯腾公司 (Biotek)) 中快速水平摇动,每 10 分钟读取 OD630。使用第一次读取的 OD630 值背景减去 OD630 值。

[0182] 1.1.2结果

[0183] 包含 hGH-GFP 表达质粒 pEG34 的大肠杆菌细胞比衍生它的亲代大肠杆菌宿主的细胞生长更慢 (图 1);说明了由于将细胞资源分配给合成构建体的维持和表达而给细胞带来的负荷或负载。

[0184] 1.2工程化负荷成瘾细胞工厂

[0185] 亲代大肠杆菌宿主的负荷成瘾菌株被基因工程化,以整合基因线路,所述基因线路被设计为在细胞工厂中赋予生产性细胞选择性适合度优势。具体地,在大肠杆菌 BL21 (DE3) 染色体中,负荷感应启动子 (p_{mutM} 、 p_{yccV} 或 p_{yccX}) 取代了必需基因操纵子 fo1P-glmM 的天

然启动子。基因线路进一步包括启动子和必需基因之间的RBS,其中测试了四种不同的RBS(表3)来调节必需基因的表达水平。将包含基因线路的菌株的生长速率与亲代宿主菌株大肠杆菌BL21(DE3)进行比较。

[0186] 1.2.1材料和方法

[0187] 负荷感应启动子:启动子 p_{mutM} 、 p_{yccV} 或 p_{ycjX} 通过使用表5中指定的引物和来源于裂解的大肠杆菌BL21(DE3)细胞的基因组DNA作为模板,PCR扩增紧接相应基因上游的0.3kb区域(参见表3)而产生。PCR混合物:10 μ l MQ水、2 μ l正向引物(10 μ M)、2 μ l反向引物(10 μ M)、1 μ l DNA模板、15 μ l Phusion U MasterMix(赛默科技(Thermo Scientific))。PCR反应方案:95C持续180秒(1x);95C持续20秒,68-58C(接地)持续30秒,72C持续60秒(35x);72 C持续300秒(1x);15 C离开。

表 5 引物			
大肠杆菌基因的启动子	引物名称	引物序列 (5'-3') *	SEQ ID No.:
[0188]	yccV		
	yccV for	ATTTCATATGUGGTTATAAAGACATCAACATGC	112
	yccV rev	ATCCCCATTACGTGAGGATGGCTAAGGTCCAGTGAAG TACCCTGGGCAAAGAGTTTCATacctt NCAAGT AGTCAC CTCCCGGAAATCT	113
	ycjX		
	ycjX for	ATTTCATATGUTTGGCTGATACTGTAATTCTTC	114
	ycjX rev	ATCCCCATTACGTGAGGATGGCTAAGGTCCAGTGAAG TACCCTGGGCAAAGAGTTTCATacctt NCAAGT TGCTGT CCTGTGCTGCTCTG	115
	mutM		
	mutM for	ATTTCATATGUGGCGAAAAGTACTAAGTACTTA	116
	mutM rev	ATCCCCATTACGTGAGGATGGCTAAGGTCCAGTGAAG TACCCTGGGCAAAGAGTTTCATacctt NCAAGT AGCATC	117
[0189]		TCCAGGAATGAACA	
	P492	TGCAACTCTTCTCAACGC	118
	P493	AAGAACCAGGCGCTTCTAA	119
	P525	ATTGATGACCTGATGGCAC	120
	P526	GATCCAGACTTCGAAGCG	121

[0190] *RBS互补编码序列以粗体显示;其中N是A、G、T或C。RBS编码在所示寡核苷酸序列的相反链上。

[0191] 使用USER克隆来产生整合序列,所述整合序列包含与线性1.5kb DNA片段融合的经扩增启动子区域,所述DNA片段含有用于选择正确重组工程产物的kanR基因和与fo1P基因直接上游221bp相同的221bp靶向序列。通过将经扩增启动子区域和kanR-fo1P片段以等摩尔量混合,并添加1 μ l 10x T4连接缓冲液(赛默科技)和0.75 μ l USER酶(新英格兰生物实验室(New England Biolabs)),以总的10 μ l反应体积使其融合。将USER反应置于37℃下30分钟。然后将反应物在室温下放置15分钟,随后添加0.75 μ l T4 DNA连接酶(赛默科技)并在室温下孵育30分钟。包含启动子、KanR抗性基因和fo1P靶向序列的这种整合序列的示例可以在序列表SEQ ID NO.140(s9_pmutM_fo1P)中找到。然后使用引物“rev”(根据表5中的特异性启动子)和P493(表5),使用以下PCR反应方案将连接的产物扩增至大约250ng/ μ l:98C持续180秒(1x);85C持续20秒,72-68C(接地)持续30秒,72C持续60秒(35x);72C持续5分钟(1x);15C离开。引物突出端提供了50bp fo1P相同的靶向序列,以指导重组工程进入fo1P基因座。

[0192] 负荷成瘾启动子的染色体整合:如下将启动子整合到大肠杆菌BL21(DE3)细胞基因组中fo1P基因的上游:用600 μ l用pSIM5-tet预转化的BL21(DE3)过夜培养物接种100ml含

有四环素的2xYT培养基 (Koskiniemi等人, 2011)。将细胞在30℃下培养, 并当达到OD600=0.20时, 将培养物转移到42℃的振荡浴中15分钟, 以允许位于pSIM5-tet上的重组工程酶的表达。然后将培养物转移到2x冷的50ml离心管中; 以4000g离心10分钟; 弃去上清液; 并用20ml冰冷的10%甘油洗涤残留的细胞沉淀。然后将部分重悬浮的细胞以4000g离心6分钟; 弃去上清液; 并用20ml冰冷的10%甘油洗涤细胞沉淀。将部分重悬浮的细胞以4000g离心6分钟; 将每个细胞沉淀小心地重悬浮在495μl冰冷的10%甘油中, 并合并。将90μl重悬浮的细胞添加到电穿孔比色皿中, 每个比色皿含有1μl (>250ng) 负荷感应启动子整合序列中的一种。用以下设置对细胞进行电穿孔: 1800V, 25μF, 200欧姆, 1mm比色皿宽度。在电穿孔后立即向比色皿中添加900ml 2xYT培养基, 并在1.5ml Eppendorf管中于37℃下使细胞恢复1.5小时。将电穿孔的细胞在室温下孵育过夜, 以便有时间进行重组; 并且随后铺在含有50mg/L卡那霉素的LB琼脂板上, 并在37℃下培养过夜, 以确保pSIM5-tet固化。使用引物P525和P526验证了对必需基因座的正确靶向。一般来说, 挑取单个kan^R菌落, 其大小平均或小于菌落群体。

[0193] 使用Taq DNA聚合酶和靶向folP启动子区域的引物, 通过菌落PCR来验证kan^R选择的菌落的基因组中的负荷成瘾启动子整合。另外, 通过Sanger测序确定所选择菌落中RBS的身份。

[0194] 随后, 通过在37℃下在水平摇动下在补充有500μM IPTG的200μl 2xYT中培养, 将包含基因线路的菌株的生长速率与亲代宿主菌株大肠杆菌BL21 (DE3) 进行比较。

[0195] 1.2.2结果

[0196] 与它们从其衍生的亲代大肠杆菌宿主相比, 所选择的负荷成瘾大肠杆菌菌株的生长速率在不同程度上较慢 (图2; 示出负荷成瘾菌株s5.0#3、s7.0#8和s9.0#8的生长, 其中必需基因folP-glmm分别由负荷感应启动子 p_{yccV} 、 p_{yccX} 和 p_{mutM} 控制)。由于细胞生长依赖于必需folP-glmm基因的表达水平, 因此可以得出结论, 在非生产性、负荷成瘾菌株中, 由相应负荷感应启动子驱动的必需基因表达的水平不足以支持以野生型亲代菌株的水平生长。

[0197] 在这些负荷成瘾菌株中观察到的生长速率的降低提供了可以施加在非生产细胞上的惩罚的量子度, 所述非生产细胞在包含负荷成瘾基因线路的细胞工厂群体的培养期间自发进化。惩罚的程度由负荷感应启动子的选择和所选RBS的强度决定。惩罚越大-则培养过程中自发产生的任何低或非生产变体 (例如由生产基因的突变产生) 的“负选择”窗口越宽。

[0198] 1.3生产增加了负荷成瘾细胞工厂的生长速率

[0199] 将负荷成瘾大肠杆菌菌株 (其对不生产赋予生长惩罚 (图2)) 用作宿主菌株, 用于证明具有生产能力的负荷成瘾细胞的选择性适合度优势。

[0200] 1.3.1材料和方法

[0201] 将产生的负荷成瘾大肠杆菌菌株 (s9.0#8; s5.0#3; 和s7.0#8) 的生长最慢的克隆分别用pEG34或pEG0 (表4) 通过标准电穿孔方法 (1800V, 25μF, 200欧姆, 1mm比色皿宽度) 制成电感受态并转化, 并且铺在含有氯霉素和卡那霉素的LB琼脂板上。将单个菌落转化体在含有氯霉素的2xYT培养基上培养过夜, 然后用于接种包含200μL 2xYT培养基 (500μM IPTG, 氯霉素) 的96孔微量滴定板, 并在37℃下培养, 同时在E1x808板阅读器 (伯腾公司) 中快速水平摇动, 每10分钟读取OD630。使用第一次读取的OD630值背景减去OD630值。

[0202] 将大肠杆菌菌株各自用质粒pEG34或pEG0 (表4) 转化, 并且当诱导细胞合成重组蛋

白(与GFP融合的hGH)时,测量它们的生长速率特性。

[0203] 1.3.2结果

[0204] 当与亲代大肠杆菌BL21 (DE3) 菌株相比时,包含控制folP-glM必需基因表达的负荷感应启动子 p_{mutM} 的负荷成瘾菌株大肠杆菌菌株s9.0#8,表现出经测试菌株中最慢的生长(图2)。相比于用空质粒BEG0转化的大肠杆菌菌株s9.0#8的细胞,在用pEG34质粒转化并诱导表达重组蛋白hGH-GFP的大肠杆菌菌株s9.0#8的细胞中,指数生长速率的这种降低被逆转(图3)。这表明大肠杆菌菌株s9.0#8中的负荷感应启动子 p_{mutM} 是由重组蛋白hGH-GFP的合成引起的对细胞的负荷或负载诱导的。这反过来导致folP-glM基因的上调表达和增强的指数生长。

[0205] hGH-GFP在每个负荷成瘾菌株(含有pEG34质粒的大肠杆菌菌株s5.0#3、s7.0#8和s9.0#8)中的合成,不仅导致指数生长速率下降的逆转;而且另外地,这些菌株的生长速率与亲代大肠杆菌BL21 (DE3) 的生长无法区分(图4)。

[0206] 总之,此实施例说明了负荷成瘾基因线路可以用于赋予细胞工厂中其蛋白质或代谢物合成足够高以构成负荷或负载的那些细胞选择性适合度优势;并且这种负荷或负载可通过可操作地连接到细胞中必需基因的负荷感应启动子来检测。相比之下,在负荷成瘾细胞工厂的培养过程中自发出现的非生产性变异细胞(例如由生产基因突变产生)将经受负选择压力,因为它们的生长将被减缓至由必需基因的基础表达所支持的速率。降低的生长速率本身足以减缓细胞工厂中此类变异细胞频率的增加,从而延缓细胞工厂生产率随时间推移的下降。

[0207] 实施例2:生产人生长激素的负荷成瘾大肠杆菌菌株显示出增强的长期生产稳定性

[0208] 负荷感应启动子显示为可以感应并被细胞中负荷诱导状态激活的启动子,所述负荷诱导状态由重组蛋白hGH-GFP的细胞合成产生。一旦被激活,则负荷感应启动子显示出将必需基因folP-glM的表达提高到足以赋予细胞在与非生产性细胞相比时的选择性生长优势的水平。为了最大化响应于其同源负荷感应启动子的必需基因表达的动态范围,将负荷感应启动子与赋予不同翻译强度的变体RBS编码序列(表3)随机组合。

[0209] 如以下在模拟大规模生产条件下培养的工程化生产菌株所示,具有适合用于负荷成瘾基因线路的负荷感应特性的启动子显示包括热休克启动子、DNA损伤启动子、氧化应激反应启动子和rRNA启动子。

[0210] 首先,为每种负荷感应启动子选择许多含有随机变异RBS编码序列的克隆,并在许多细胞分裂中跟踪它们的hGH-GFP合成,以便证明它们随时间推移提高/保持重组hGH-GFP合成的相对能力。

[0211] 2.1材料和方法

[0212] 负荷感应启动子:使用表6中指定的引物,如实施例1.2.1所述通过PCR和USER克隆产生启动子 p_{yccV} 、 p_{ycjX} 、 p_{ibpA} 、 p_{grpE} 、 p_{ldhA} 和 p_{ybbN} 及其各自的整合序列;同时使用引物“rev”(根据表6中的特异性启动子)和P493(表5)扩增连接的产物。

表 6 引物			
大肠杆菌基因的启动子	引物名称	引物序列 (5'-3') *	SEQ ID No.:
ibpA	ibpA fwd	ATCCCCATTACGTGAGGATGGCTAAGGTCCAGTGAAGT ACCCTGGGCAAAGAGTTTCATacett NCAAGTAATCAATA GTCCTGAAATC	122
	ibpA rev	ATTCATATGTTTGGCCTGATGAGTTATAGCG	123
grpE	grpE fwd	ATTCATATGUTCTCCGCGAGCGTGCCAGTTTT	124
	grpE rev	ATCCCCATTACGTGAGGATGGCTAAGGTCCAGTGAAGT ACCCTGGGCAAAGAGTTTCATacett NCAAGTGAATTTCT CCGCGTTTTTTT	125
ldhA	ldhA fwd	ATTCATATGUGGGAACCCACAGCCCCGAGCGTC	126
	ldhA rev	ATCCCCATTACGTGAGGATGGCTAAGGTCCAGTGAAGT ACCCTGGGCAAAGAGTTTCATacett NCAAGTAAGACTT TCTCCAGTGATGT	127
ybbN	ybbN fwd	ATTCATATGUCCCATTCGTA CTGCTTACCGATCCCCA TTACGTGAGGATGGCTAAGGTCCAGTGAAGTACCCTG GGCAAAGAGTTTCATacett NCAAGTGGAGTCGCTCTCT GTTGTCG	128
	ybbN rev	ATCCCCATTACGTGAGGATGGCTAAGGTCCAGTGAAGT ACCCTGGGCAAAGAGTTTCATacett NCAAGTGGAGTCG CTCTCTGTTGTCG	129
fxsA	fxsA fwd	ATTCATATGUTAAACACACAGAATATATGTGG	130
	fxsA rev	ATCCCCATTACGTGAGGATGGCTAAGGTCCAGTGAAGT ACCCTGGGCAAAGAGTTTCATacett NCAAGTAGGTTTCT CCTGTAATAGCA	131
rrnB	rrnB fwd	ATTCATATGUGCAAACACGCCGCCGGGTCAG	132
	rrnB rev	ATCCCCATTACGTGAGGATGGCTAAGGTCCAGTGAAGT ACCCTGGGCAAAGAGTTTCATacett NCAAGTAAAAGTT TGACGCTCAAAGA	133
rrnE	rrnE fwd	ATTCATATGUAATGCGCCTCCGTTGAGACGACAAC	134
	rrnE rev	ATCCCCATTACGTGAGGATGGCTAAGGTCCAGTGAAGT ACCCTGGGCAAAGAGTTTCATacett NCAAGTAAAAGTT TGATGCTCAAAGA	135
poxB	poxB fwd	ATTCATATGUTAGGTTGTCGCTGCCTGCCGTG	136
	poxB rev	ATCCCCATTACGTGAGGATGGCTAAGGTCCAGTGAAGT ACCCTGGGCAAAGAGTTTCATacett NCAAGTGGTTCTCC ATCTCCTGAATG	137

[0214] *RBS互补编码序列以粗体显示;其中N是A、G、T或C。RBS编码在所示寡核苷酸序列的相反链上。

[0215] 负荷成瘾启动子的染色体整合:如下将每个启动子整合到大肠杆菌菌株BEG34(对应于含有质粒pEG34的大肠杆菌BL21(DE3))基因组中fo1P基因的上游。如实施例1.2.1所述,制备用重组工程质粒pSIM5-tet预转化的大肠杆菌菌株BEG34的细胞,并用每种启动子整合序列通过电穿孔进行转化。在重组工程和pSIM5-tet固化的所述步骤后,从每个板中挑取八个菌落,对应于具有相同启动子整合但具有随机变异RBS序列的八个克隆(参见表6)。然后将每个克隆转移到含有补充有用于维持质粒pEG34的氯霉素的200μl 2xYT的96孔板的一个孔中。如实施例1.2.1所述,将2μl每个培养的克隆用于经由菌落PCR验证启动子整合;在冷冻96孔板之前。

[0216] 短期hGH-GFP生产筛选测定:将冷冻的96孔板解冻,并使用针式复制器将细胞转移到含有补充有氯霉素的200μL 2xYT的新96孔板中。用Breathe-Easy密封膜(西格玛奥德里奇公司(Sigma-Aldrich))密封此板,并将其置于SynergyH1板读数器(伯腾公司)中,在37℃和754rpm下线性摇动20小时过夜;并且在20小时内每10分钟测量每个孔的OD(600nm)和GFP荧光(ex/em 485nm/528nm)。然后在室温下将板置于常规的桌面板振荡器中4小时,随后将

来自每个孔的2 μ L培养物转移到新的96孔板,所述板含有补充有氯霉素和0.5mM IPTG的200 μ L 2xYT。将新的96孔板同样用Breathe-Easy密封膜(西格玛奥德里奇公司)密封,并将其置于SynergyH1板读数器中,在37 $^{\circ}$ C和754rpm下线性摇动20小时过夜;并且在20小时内每10分钟在600nm和ex/em 485nm/528nm下测量每个孔的OD和GFP荧光。第二天,将2 μ L培养物转移到新的96孔板中,所述板含有补充有氯霉素和0.5mM IPTG的200 μ L 2xYT,并且每天重复所述过程持续总共6天。

[0217] 长期hGH-GFP生产测定(图6):将来自每个负荷成瘾hGH-GFP生产菌株的单个所选择菌落用于接种15mL Greiner培养管,所述管含有补充有氯霉素和0.5mM IPTG的4mL 2xYT。将培养物在250rpm的摇床上于37 $^{\circ}$ C下生长23小时。然后在相同条件下,将2 μ L每种培养物以2000倍稀释(对应于大约11代次细胞分裂)接种到新的Greiner培养管中,其中对总共12个种子重复这些方法步骤。每次传代后,取200 μ L培养物样品,以通过在SynergyH1板读数器(伯腾公司)中测量OD600和GFP荧光(ex/em 485nm/528nm)来确定hGH-GFP合成。

[0218] 长期hGH-GFP生产测定(图8):将来自每个负荷成瘾hGH-GFP生产菌株的单个所选择菌落用于接种24深孔板,所述板携带补充有氯霉素和0.5mM IPTG的1.8mL 2xYT。将培养物在200rpm的摇床上于30 $^{\circ}$ C下生长23小时。然后在相同条件下,将2 μ L每种培养物以1000倍稀释(对应于大约10代次细胞分裂)接种到新的深孔板中,其中对总共9个种子重复这些方法步骤。每次传代后,取200 μ L培养物样品,以通过在SynergyH1板读数器(伯腾公司)中测量OD600和GFP荧光(ex/em 485nm/528nm)来确定hGH-GFP合成。

[0219] 所选菌株的生长速率测量:将所选择的高hGH生产菌株s3.6#2、s6.6#6、s7.6#8和s10.6#7在含有氯霉素和卡那霉素的LB琼脂板上划线。将BEG34和BEG0菌株在含有氯霉素的LB琼脂板上划线。将来自每个板的7个菌落用于接种含有补充有氯霉素的200 μ L 2xYT的96孔板;将其用Breathe-Easy密封膜(西格玛奥德里奇公司)密封。在37 $^{\circ}$ C和754rpm线性摇动下,在SynergyH1板读数器(伯腾公司)中过夜测量它们的生长。将来自每个孔的2 μ L样品转移到新的96孔板中的孔中,所述板含有补充有氯霉素和0.5mM IPTG的200 μ L 2xYT。用透气膜密封96孔板,并将其在37 $^{\circ}$ C和754rpm线性摇动下培养20小时,并且使用Synergy H1板读数器每10分钟测量生长(OD600nm)。

[0220] 菌株目录:

[0221] BEG34=携带pEG34的大肠杆菌BL21(DE3);

[0222] BEG0=携带空质粒pEG0的大肠杆菌BL21(DE3)

[0223] sX.Z#Y=携带pEG34的大肠杆菌BL21(DE3),将wt fo1P启动子替换为负荷成瘾启动子X,其中X是指表2中的启动子ID。#后面的数字Y表示由简并引物引入的四个RBS变体库中选择的克隆(表7)。Z分别是指hGH-GFP(0或6)、甲羟戊酸(1)、溶葡萄球菌酶(4)的生产基因。

表 7: 相应的大肠杆菌菌株的 RBS 的“N”核苷酸的标识

菌株标识符	RBS “N”标识	菌株标识符	RBS “N”标识	菌株标识符	RBS “N”标识
s5.0 #3	C	s16.6#6	A	s19.6#2	A
s7.0 #8	A	s18.6#8	T	s21.6#8	C
s9.0#8	NGS	s9.1#1	C	s23.6#4	T
s6.4#5	C	s15.1#1	A	s24.6#2	T
s3.6#2	A	s19.1#1	T	s25.6#1	C
s6.6 #6	C	s14.1#5	T	s29.6#3	T
s7.6#8	T	s15.1#5	A	s22.6#8	T
s10.6#7	A	s19.1#7	A	s26.6#3	A
s13.6#2	C	s15.6 #1	A	s29.6#4	T
s15.6#7	A	s16.6#7	T	s30.6#6	T

[0224]

[0225] 2.2结果

[0226] 2.2.1短期hGH-GFP生产率筛选

[0227] 如下,在模拟的大规模生产条件下测试四种hGH-GFP生产菌株的生产率,所述菌株包含选自一组热休克启动子 (p_{ibpA} 、 p_{grpE} 、 p_{ycjX} 和 p_{ybbN}) 的负荷感应启动子,并具有四种RBS编码序列变体中的一种。菌株每天通过100倍的反向稀释连续传代6天,相当于每个种子大约6代。到第6天(种子6),在种子1和种子6期间,与非负荷成瘾生产菌株BEG34相比,几个菌株显示出升高的hGH-GFP合成。相对于菌株BEG34具有高短期hGH-GFP生产率的菌株是s3.6#2 (p_{ibpA})、s6.6#6 (p_{grpE})、s7.6#8 (p_{ycjX}) 和s10.6#7 (p_{ybbN}) -参见图5。

[0228] 2.2.2增强的长期hGH-GFP生产率

[0229] 通过连续转移,在模拟的大规模发酵下监测负荷成瘾菌株s3.6#2、s6.6#6、s7.6#8、s10.6#7和非负荷成瘾菌株BEG34和非生产菌株BEG0的hGH-GFP生产率。在种子2之后,负荷成瘾菌株s3.6#2、s6.6#6、s7.6#8、s10.6#7和非负荷成瘾对照菌株BEG34在hGH-GFP合成方面表现同样良好(图6)。然而,在另外的大约20次传代(种子4)后,对照菌株BEG34的hGH-GFP生产率显著下降;并且到种子6时,所述菌株基本上停止了生产。令人惊讶的是,负荷成瘾菌株s7.6#8(具有控制fo1P-glmM的 p_{ycjX}) 仍保持约50%种子2水平的生产率。另外,与非负荷成瘾菌株BEG34相比,负荷成瘾菌株s3.6#2(具有控制fo1P-glmM的 p_{ibpA})、s10.6#7(具有控制fo1P-glmM的 p_{ybbN}) 和s6.6#6(具有控制fo1P-glmM的 p_{grpE}) 也显示出显著更好的长期生产率。

[0230] 2.2.3负荷成瘾菌株的生长速率

[0231] 生产hGH-GFP的负荷成瘾菌株s3.6#2、s6.6#6、s7.6#8和s10.6#7的生长速率不高于非负荷成瘾hGH-GFP生产菌株BEG34;并且因此,随着时间的推移,它们的hGH-GFP生产水平的提高并不简单地源于初始生产水平的降低,以及因此固有的较低负荷(图7)。

[0232] 2.2.4负荷成瘾基因线路对所选择必需基因的依赖性

[0233] 将包含控制必需基因fo1P-glmM的热休克启动子 p_{fxsA} 的生产hGH-GFP的负荷成瘾菌株(s13.6#2 ($fxsA$))与包含消除fo1P表达的移码fo1P的突变衍生物菌株s13.6#2evo ($fxsA$) (其生长仅依赖于glmM表达)进行比较。尽管fo1P表达的缺失导致在复杂的2xYT培养基中的生长速率降低(数据未示出),但当与非负荷成瘾hGH-GFP生产菌株BEG34相比时,单独的必需基因glmM被证明足以在负荷成瘾菌株s13.6#2evo ($fxsA$) 中以与菌株s7.6#8 ($pycJX$) 可比的水平赋予增强的长期hGH-GFP生产稳定性(图8)。

[0234] 2.2.5核糖体RNA启动子在负荷成瘾基因线路中的用途

[0235] 除了热休克启动子之外,核糖体RNA启动子如 p_{rrnB} 和 p_{rrnE} 显示出能够感应细胞中的

负荷成瘾状态,并作为响应,控制必需基因表达,如提高hGH-GFP生产菌株的长期生产率。如图9中所见,在对应于约110代次细胞分裂的模拟的大规模生产中,当与非负荷成瘾hGH-GFP生产大肠杆菌菌株BEG34相比时,菌株s15.6.7(p_{rrnB})和s16.6.6(p_{rrnE})中hGH-GFP生产的长期稳定性显著提高。

[0236] 2.2.6氧化应激感应启动子在负荷成瘾基因线路中的用途

[0237] 氧化应激感应启动子如 p_{poxB} 显示出能够感应细胞中的负荷成瘾状态,并作为响应,控制必需基因表达,如提高hGH-GFP生产菌株的长期生产率。如图10中所见,在对应于大约90代的7个连续传代的模拟的大规模生产和最终种子的时间延长培养下,当与非负荷成瘾hGH-GFP生产大肠杆菌菌株BEG34相比时,菌株s29.6#3(p_{poxB})中hGH-GFP生产的生产水平和长期稳定性二者都显著增加。如图10中所见,使用基于 p_{poxB} 的负荷成瘾基因线路的生产令人惊讶地导致随时间推移而提高的产量,这表明群体中非基因高性能变体的富集。

[0238] 总之,本实施例说明,通过将必需基因folP-glmm的转录偶联到大肠杆菌热休克、氧化应激和DNA损伤响应启动子以及rRNA启动子中的任一个,负荷成瘾增强了被工程化以合成与GFP融合的人生长激素的大肠杆菌菌株的长期稳定性和产量。

[0239] 实施例3:生产溶葡萄球菌酶的负荷成瘾大肠杆菌菌株显示出增强的长期生产稳定性

[0240] 溶葡萄球菌酶是一种27kDa的肽链内切酶,其切割金黄色葡萄球菌(*Staphylococcus aureus*)细胞壁肽聚糖中的交联五甘氨酸桥,从而导致细胞裂解。此抗菌剂可以在大肠杆菌中重组合成。如由以下在模拟的大规模生产条件下生产溶葡萄球菌酶的负荷成瘾大肠杆菌菌株所说明,包含具有负荷感应特性的启动子的负荷成瘾基因线路被证明增强了被工程化以合成溶葡萄球菌酶的此类大肠杆菌菌株的生产稳定性。负荷成瘾生产菌株的生产率进一步被证明通过选择其中负荷感应启动子与赋予同源必需基因最佳翻译强度的RBS序列(表6)组合的菌株而得到优化。这些优势由负荷成瘾菌株s6.4#5(p_{grpE})举例说明,如下所述,所述菌株被证明对产量损失更具弹性,因为对由溶葡萄球菌酶表达导致的负荷或适合度代价成瘾。

[0241] 3.1材料和方法

[0242] 负荷感应启动子:使用表6中指定的 p_{grpE} 特异性引物,通过由PCR扩增 p_{grpE} 启动子,如实施例1.2.1所述产生包含启动子 p_{grpE} 的启动子整合序列。

[0243] 负荷成瘾启动子的染色体整合:如下将 $grpE$ 启动子整合到大肠杆菌菌株BENDU5cam(对应于含有生产溶葡萄球菌酶的质粒pENDU5cam的大肠杆菌BL21(DE3))基因组中folP-glmm基因的上游。如实施例1.2.1所述,制备用重组工程质粒pSIM5-tet预转化的大肠杆菌菌株BENDU5cam的细胞,并用 p_{grpE} 启动子整合序列通过电穿孔进行转化。在重组工程和pSIM5-tet固化的所述步骤后,挑取5个菌落,对应于具有相同启动子整合但具有随机变异RBS序列的克隆(参见表6)。然后将每个克隆转移到含有补充有用于维持质粒pENDU5cam的氯霉素的4mL 2xYT的15mL Greiner培养管中,并进行培养。如实施例1.2.1所述,将2 μ l每个培养的克隆用于验证启动子整合;在冷冻之前。

[0244] 溶葡萄球菌酶生产筛选:使用筛选测定鉴定具有最高溶葡萄球菌酶生产率的大肠杆菌菌株s6.4#5(p_{grpE}),其中与非负荷成瘾生产菌株BENDU5cam相比,将RBS和 p_{grpE} 的三种其他潜在组合在为96孔形式的补充有氯霉素的2xYT中培养1000倍稀释度的六个连续传代(对

应于60代)。将菌株在37℃下在15mL培养管中于含有30mg/L氯霉素的3mL 2xYT培养基中生长21小时。将150μL培养物与50μL 50%甘油混合,并储存在-80℃下的96孔板中。另外,将30μL培养物转移到补充有氯霉素的3mL新鲜2xYT培养基中,并在相同条件下再生长21小时;并且重复总共5次转移,其中每次转移都将过夜培养物冷冻保存。

[0245] 溶葡萄球菌酶表达和检测测定:将含有来自5次转移的所有冷冻保存物的96孔板在冰上解冻。将来自每个孔的20μL转移到含有补充有氯霉素的180μL 2xYT的96孔板中,并在37℃下在板阅读器中培养生长,直到大多数孔达到 $OD_{630}=0.35$;并且然后向每个孔中添加补充了30mg/L氯霉素和20mM IPTG的10μL 2xYT至最终浓度为1mM IPTG,足以诱导溶葡萄球菌酶基因表达。将诱导的培养物在37℃下生长4小时,然后将板以4000RPM离心10分钟。将100μL上清液转移到含有100μL过夜金黄色葡萄球菌培养物的新96孔板中。每10分钟,在37℃下,在板阅读器上以 OD_{630} 监测每个孔中的金黄色葡萄球菌裂解持续2小时。通过将金黄色葡萄球菌OD的变化除以发生变化的60分钟时间范围来量化金黄色葡萄球菌裂解速率。以下可以看到这个等式:

$$[0246] \quad \text{速率}_{裂解} = \frac{OD_{t_0} - OD_{t_1}}{t_1 - t_0}$$

[0247] 通过将测得的速率归一化为相应生产大肠杆菌培养物的最终 OD_{630} 测量值来确定特异性溶葡萄球菌酶合成速率。

[0248] 3.2结果

[0249] 在通过连续传代实现的模拟的大规模生产条件下,培养具有 p_{grpE} 启动子和控制必需基因 $folP$ - $g1mM$ 表达的四种RBS编码序列变体中的一种的负荷成瘾溶葡萄球菌酶生产菌株。如图11中所见,与在非负荷成瘾溶葡萄球菌酶生产BENDU5cam菌株中看到的产量急剧下降相比,菌株s6.4#5(p_{grpE})保持了显著更高水平的溶葡萄球菌酶产量。因为s6.4#5(p_{grpE})和BENDU5cam菌株在种子0时的初始溶葡萄球菌酶生产速率相同;因此s6.4#5(p_{grpE})的生产速率下降较小,并不是因为溶葡萄球菌酶生产的固有初始负荷较低。

[0250] 总之:通过将必需基因 $folP$ - $g1mM$ 转录偶联到大肠杆菌热休克启动子 p_{grpE} ,证明负荷成瘾增强了被工程化以合成经分泌的溶葡萄球菌酶的大肠杆菌的长期稳定性和产量。

[0251] 实施例4:生产甲羟戊酸的负荷成瘾大肠杆菌菌株显示出增强的长期生产稳定性

[0252] 通过引入表达异源三步酶促途径的质粒 p_{MevT} 来工程化大肠杆菌BL21(DE3)细胞以合成甲羟戊酸(Martin等人,2003),所述异源三步酶促途径经由乙酰辅酶A库将葡萄糖转化为甲羟戊酸。包含具有负荷感应特性并控制必需基因转录的启动子的负荷成瘾基因线路被证明在模拟的大规模生产条件下增强了此类甲羟戊酸生产大肠杆菌菌株的生产稳定性。负荷成瘾生产菌株的生产率被证明通过选择其中负荷感应启动子与赋予同源必需基因最佳翻译强度的RBS编码序列(表5)组合的菌株而得到进一步优化。具体地,包含控制 $folP$ - $g1mM$ 转录的负荷成瘾启动子 p_{cspD} (一种氧化应激和葡萄糖饥饿感应启动子)或 p_{mutM} (热休克/DNA损伤感应启动子)的甲羟戊酸生产菌株二者都被证明比非负荷成瘾菌株对产量损失更具弹性,因为它们对由甲羟戊酸生产产生的转录信号成瘾。

[0253] 4.1材料和方法

[0254] 负荷感应启动子:使用表8中指定的 p_{mutM} 和 p_{cspD} 特异性引物,通过由PCR扩增 p_{mutM} 和 p_{cspD} 启动子,如实施例1.2.1所述产生包含启动子 p_{mutM} 和 p_{cspD} 的启动子整合序列。

[0255]

表 8 引物			
大肠杆菌基因的启动子	引物名称	引物序列 (5'-3') *	SEQ ID No.:
<i>mutM</i>	mutM for	ATTCATATGUGGCGAAAAGTACTAAGTAC TTA	116
	mutM rev	ATCCCCATTACGTGAGGATGGCTAAGGTC CAGTGAAGTACCCTGGGCAAAGAGTTTCA Tacctt NCAAGT AGCATCTCCAGGAATGAAC A	117
<i>cspD</i>	cspD fwd	ATTCATATGUCGGTTATCGGCAGAACGCC CTG	138
	cspD rev	ATCCCCATTACGTGAGGATGGCTAAGGTC CAGTGAAGTACCCTGGGCAAAGAGTTTCA Tacctt NCAAGT GCTTCGACATCCTTCGCAA A	139

[0256] *RBS互补编码序列以粗体显示;其中N是A、G、T或C。RBS编码在所示寡核苷酸序列的相反链上。

[0257] 负荷成瘾启动子的染色体整合:如下将 p_{mutM} 和 p_{cspD} 分别整合到含有溶葡萄球菌酶生产质粒pMevT的大肠杆菌菌株BL21 (DE3) pMevT的细胞基因组中fo1P-glmM基因的上游。如实施例1.2.1所述,制备用重组工程质粒pSIM5-tet预转化的大肠杆菌菌株BL21 (DE3) pMevT的细胞,并用每种启动子整合序列通过电穿孔进行转化。在重组工程和pSIM5-tet固化的所述步骤后,挑取5个菌落,对应于具有相同启动子整合但具有随机变异RBS序列的克隆(参见表8)。然后将每个克隆转移到含有补充有用于维持质粒pMevT的氯霉素的200 μ l 2xYT的96孔板的一个孔中,并进行培养。如实施例1.2.1所述,将2 μ l每个培养的克隆用于验证启动子整合;在冷冻96孔板之前。

[0258] 甲羟戊酸生产筛选:如下使用筛选测定鉴定具有最高甲羟戊酸生产率的大肠杆菌菌株s19.1.1(p_{cspD})和s9.1.4(p_{mutM}),其中与非负荷成瘾生产大肠杆菌菌株BL21 (DE3) pMevT平行,将包含RBS和启动子的四种其他潜在组合的菌株在96孔板中培养1000倍稀释度的六个连续传代(对应于60代)。使菌株在用透气密封件密封的微量滴定板中含有30mg/L氯霉素和0.5mM IPTG的200 μ L 2xYT培养基中,于37 $^{\circ}$ C下在水平摇动下生长21小时。将150 μ L培养物与50 μ l 50%甘油混合,并储存在-80 $^{\circ}$ C下的96孔板中。另外,将2 μ l 10倍稀释的培养物转移到补充有氯霉素和0.5mM IPTG的200 μ L新鲜2xYT培养基中,并在相同条件下再生长21小时。总共进行了5次与上述相同的转移,并且每次转移都将过夜培养物冷冻保存。

[0259] 甲羟戊酸合成和检测测定:将含有来自第二、第五和第六次转移的冷冻保存物的96孔板在冰上解冻,并且用于接种含有0.5mM IPTG和30mg/L氯霉素的10mL 2xYT,并在37 $^{\circ}$ C下在水平摇动(250rpm)下培养54小时。用23 μ L 20%硫酸处理来自每个培养物的300 μ L等分试样;剧烈摇动,然后以13000 $\times$ g旋转2分钟。将上清液(培养基)样品注射到Ultimate 3000高效液相色谱仪中,在50 $^{\circ}$ C下,在Aminex HPX-87H离子排斥柱(300mm $\times$ 7.8mm, Bio-Rad实验室)上运行5mM硫酸流动相(0.6mL/分钟)。将折光率检测器用于检测。甲羟戊酸的标准曲线是用溶于在相同条件下孵育的非生产大肠杆菌菌株的2xYT培养基上清液中的甲羟戊酸内酯(西格玛奥德里奇公司)生成的。

[0260] 4.2结果

[0261] 在通过连续传代实现的模拟的大规模生产条件下,培养具有 p_{cspD} 或 p_{mutM} 启动子和控制必需基因fo1P-glmM表达的四种RBS编码序列变体中的一种的负荷成瘾甲羟戊酸生产

菌株。如图12中所见,与在非负荷成瘾甲羟戊酸生产菌株BL21 (DE3) pMevT中看到的产量急剧下降相比,所选择的负荷成瘾菌株s19.1.1(p_{cspD})和s9.1.4(p_{mutM})二者都保持了显著更高水平的甲羟戊酸产量。因为每个负荷成瘾菌株和对照BL21 (DE3) pMevT菌株在种子0时的初始甲羟戊酸生产速率相同;菌株s19.1.1(p_{cspD})和s9.1.4(p_{mutM})的生产速率下降较小并不是因为甲羟戊酸生产的固有初始负荷较低。

[0262] 总之:通过将必需基因fo1P-glM转录偶联到大肠杆菌氧化应激和葡萄糖饥饿感应启动子p_{cspD}和热休克/DNA损伤感应启动子p_{mutM},证明负荷成瘾增强了被工程化以合成甲羟戊酸的大肠杆菌的长期稳定性和产量。

[0263] 实施例5:评估启动子作为负荷感应器以使酵母细胞对工程化重组蛋白生产的负荷成瘾

[0264] 能够感应由蛋白质或生物合成途径的重组表达引起的对细胞的负荷或负载,然后诱导必需基因表达的启动子,可以用于创建适合在酵母中使用的负荷成瘾基因线路。一种用于评估候选启动子的方法在经基因工程化以合成任选地与绿色荧光蛋白 (GFP) 翻译融合的重组人血清白蛋白 (hSA) 或胰岛素前体 (IP) 的酵母细胞中进行了说明。

[0265] 为了使生长响应于候选负荷感应启动子的活性,使用标准可选择标记,使用转化到酵母细胞中的线性DNA构建体的同源重组,用候选启动子基因替代酿酒酵母中必需基因的天然启动子。举例来说,候选启动子可以选自核糖体RNA基因的上调启动子,如由RNA聚合酶I转录的启动子(Laferté等人,2006)、DNA损伤感应启动子(例如pOGG1),和由HAC1转录因子上调的未折叠蛋白反应 (UPR) 启动子(Kimata等人,2006)。为了确保负荷感应启动子一旦被激活,则与非感应启动子相比,赋予细胞选择性生长优势,这可能有必要微调必需基因的表达水平。可以通过改变与负荷感应启动子一起引入的翻译起始区来工程化一系列翻译强度。

[0266] 选择具有候选负荷感应启动子的潜在不同组合的克隆,并评估在30-100代次细胞分裂中维持的蛋白质生产。

[0267] 5.1材料和方法

[0268] 生长培养基:YPD培养基包含1%酵母提取物、2%蛋白胨、2%葡萄糖。SC培养基包含6.7g/L酵母氮源,不含氨基酸,含硫酸铵,但缺乏尿嘧啶。

[0269] 启动子构建体的染色体整合和验证:用于染色体整合的构建体将含有天然调节基因上游的300-600bp。使用酿酒酵母的标准电穿孔程序,通过转化和同源重组进行启动子构建体的染色体整合。使用菌落PCR验证启动子的正确染色体整合。

[0270] 长期培养和生产:将来自每个菌株的单菌落转移到24深孔板中,并在诱导重组蛋白产生的条件下,以250rpm和30C在1.8mL YPD培养基中培养。在培养48小时后,通过1000倍反向稀释,在相同条件下将细胞传代至新的深孔板。使用重组蛋白特异性测定(例如GFP检测)分析样品的生产,并通过OD600监测细胞密度。

[0271] 所选择菌株的生长速率测量:将单个菌株的生长速率与非负荷成瘾生产菌株进行比较。用透气膜密封96孔板,并在37°C和754rpm线性摇动下,在Synergy H1板读数器中测量生长20小时。每10分钟测量OD600。

[0272] 5.2结果

[0273] 5.2.1HAC1上调启动子

[0274] 在生产人胰岛素前体或人血清白蛋白的重组蛋白生产酵母菌株中,当在天然生长调节基因(例如编码嘧啶生物合成所必需的乳清酸核苷5'-磷酸脱羧酶的条件性必需基因URA3)前面引入时,包含未折叠蛋白反应(UPR)元件的HAC1上调启动子(例如KAR2(SEQ ID NO.:91)、PDI1(SEQ ID NO.:92)、SSA1(SEQ ID NO.:93)或FPR2(SEQ ID NO.:94))被证明可用于调节生长,任选地与GFP偶联。经由对应于60-80代次细胞分裂的连续传代,在模拟的长期生产下跟踪生产的稳定性。在包含控制生长调节基因(URA3)转录的HAC1上调启动子的酵母菌株中,预期生产比对应的亲代重组蛋白生产酵母菌株更稳定。

[0275] 5.3.2RNA聚合酶I上调启动子

[0276] RNA聚合酶I在酵母中转录核糖体RNA基因。RNA聚合酶I上调启动子,如基因:RPL3(SEQ ID NO.:95)、RPL6A(SEQ ID NO.:96)和RPL28(SEQ ID NO.:97)的启动子可用于调节酵母必需基因的生长。在生产人胰岛素前体或人血清白蛋白的重组蛋白过量生产菌株中,此类上调启动子在天然生长调节基因(例如条件性必需基因URA3)前面被引入,可能与GFP偶联。经由对应于60-80代次细胞分裂的连续传代,通过实验模拟的长期生产跟踪生产的稳定性。在包含控制条件性必需基因URA3转录的RNA聚合酶I上调启动子的酵母菌株中,生产将比对应的亲代重组蛋白生产酵母菌株更稳定。

[0277] 5.3.3DNA损伤响应启动子

[0278] 由异源表达导致的酵母中的DNA损伤反应广泛地诱导DNA修复系统的转录,包括OGG1(SEQ ID NO.:98)、RAD51(SEQ ID NO.:99)和RAD54(SEQ ID NO.:100)。编码OGG1、RAD51或RAD54的基因的启动子,当可操作地连接到必需基因时,可用于调节本发明的酵母生产细胞的生长。在生产人胰岛素前体或人血清白蛋白的酵母蛋白生产菌株的细胞中,此类启动子在天然必需基因(例如编码URA3的生长调节基因)前面被引入,任选地与GFP融合。

[0279] 经由对应于60-80代次细胞分裂的连续传代,在模拟的长期生产下跟踪生产的稳定性。在具有 p_{RAD51} 和/或 p_{RAD54} 上调必需基因的菌株中,生产将更加稳定。

[0280] 总之:通过将必需基因转录偶联到选自重组蛋白生产负荷期间激活的启动子或与核糖体RNA启动子相关的负荷感应启动子,负荷成瘾增强了被工程化以合成人血清白蛋白生产或胰岛素前体的出芽酵母细胞的长期稳定性和生产(表2)。

[0281] 实施例6生产人血清白蛋白的负荷成瘾酵母菌株的示例显示出增强的长期生产稳定性

[0282] 用于酵母菌株的负荷成瘾系统包括来源于编码以下各项的基因的启动子:1) 蛋白异构酶PDI1,其编码属于酵母(如酿酒酵母和巴斯德毕赤酵母)的未折叠蛋白反应的伴侣蛋白,并且其丰度响应于重组蛋白的过量表达而经常上调;2) 由RPL6A和RPL3基因编码的核糖体亚单位,和3) 编码肽基-脯氨酰基顺反异构酶的FPR2,已知所述异构酶在DNA复制应激时被激活。将基于这些启动子的负荷成瘾系统引入被工程化以表达和分泌人血清白蛋白的毕赤酵母菌株中,以便确定它们对长期人血清白蛋白(hSA)生产稳定性的影响。

[0283] 6.1材料和方法:

[0284] 毕赤酵母(甲基营养型毕赤酵母(*Komagataella phaffii*))菌株EGS31是CBS7435菌株(NRRL-Y11430或ATCC 76273)的衍生物,其通过在AOX1启动子控制下的hSA编码ALB1基因的基因组整合的cDNA版本被工程化以分泌hSA。

[0285] 菌株构建:EGS31菌株的负荷成瘾版本是通过基因整合构建体产生的,所述构建体

包含可操作地连接到以下负荷响应启动子中的一种的kanMX条件性可选择G418抗性基因：PDI1、pFPR2、pRPL3和pRPL6A。通过用侧接>750bp同源臂的线性整合DNA(分别为序列N1-N3)转化EGS31细胞,将这些负荷成瘾构建体整合到巴斯德毕赤酵母菌株的KU70基因组基因座中。使用标准电穿孔程序对用乙酸锂和二硫苏糖醇预处理的指数生长细胞进行转化(Wu和Letchworth,2004)。

表 9: 引入巴斯德毕赤酵母中的重组基因和负荷成瘾构建体		
基因/结构名称	构建核苷酸 SEQ ID NO:	经编码的蛋白质 SEQ ID NO.:
pAOX1-ALB 构建体	SEQ ID NO: 155	hSA 蛋白 SEQ ID NO: 156
[0286] KU70 (上) -pPDI1-kanMX-终止子-KU70 (下)	SEQ ID NO: 157	kanMX 蛋白 SEQ ID NO: 158
KU70 (上) -pFPR2-kanMX-终止子-KU70 (下)	SEQ ID NO: 159	kanMX 蛋白 SEQ ID NO: 160
KU70 (上) -pRPL3-kanMX-终止子-KU70 (下)	SEQ ID NO: 161	kanMX 蛋白 SEQ ID NO: 162
KU70 (上) -pRPL6A-kanMX-终止子-KU70 (下)	SEQ ID NO: 163	kanMX 蛋白 SEQ ID NO: 164

[0287] 生长培养基:如下制备BMGY和BMMY液体培养基(1L):将10g酵母提取物和20g蛋白胨添加到700mL H₂O中;用磁力搅拌器混合,然后高压灭菌。在冷却至室温后;将以下物质添加到溶液中:100ml 1M磷酸钾缓冲液(pH 6.0);100ml 13.4% (w/v) 酵母氮源,含硫酸铵,不含氨基酸;2ml 0.02% (w/v) 生物素,并且在BMGY的情况下,添加100ml 10% (v/v) 甘油;并且在BMMY的情况下,添加100mL 5% (v/v) 甲醇。

[0288] 培养:将EGS31和负荷成瘾EGS31菌株在YPD(1%酵母提取物、2%蛋白胨、2% D-葡萄糖)琼脂板上划线,并在30℃下孵育过夜。挑取单个菌落并在2mL BMGY中预培养,其中对负荷成瘾菌株的培养物补充50ug/mL G418,并在30℃下以300rpm水平摇动过夜进行培养。使用1uL预培养物将表达培养物接种到具有通气盖的96孔深孔板中含有不同浓度(0μg/mL、750μg/mL)G418的500μL BMMY培养基中,以产生“种子1”。

[0289] 将培养物在30℃下孵育72小时(300rpm水平摇动)。为了产生下一个种子,另外四次,将生长的培养物连续传代(500倍稀释)至新的具有通气盖的96孔深孔板中含有不同浓度(0μg/mL、750μg/mL)G418的500μL BMMY培养基。在每次连续传代中,将来自生长培养物的甘油菌(20%甘油)储存在-80℃下。

[0290] 为了定量分泌的hSA的生产,使来自甘油菌的定量培养物在具有通气盖的96孔深孔板中含有不同浓度(0μg/mL和750μg/mL)G418的500μL BMMY培养基中重新生长72小时。

[0291] 将培养物以3000g离心15分钟,并按照制造商的说明使用hSA特异性ELISA试剂盒(Abcam目录号:ab179887:人白蛋白SimpleStepELISA®试剂盒)对50μL上清液中的浓度进行定量。

[0292] 6.2结果

[0293] 当菌株的负荷成瘾系统被添加G418激活时,具有基因组整合的负荷成瘾启动子PDI1(可操作地连接到可选择的kanMX基因)的菌株展现出更高的经分泌的hSA产量(图13)。此外,在大约30次细胞分裂后,对于每种其他经测试的负荷成瘾启动子,随着更高的选择(750μg/mL G418)观察到hSA产量提高(图14A-种子1)。当进一步培养约另外10次细胞分裂

时(图14B-种子2),与非负荷成瘾对照和未激活负荷成瘾的菌株(0 μ g/mL G418)相比,当负荷成瘾被激活时(150-750 μ g/mL G418),所述菌株显示出hAS生产水平的提高。

[0294] 总之,在激活其各自的负荷成瘾系统的条件下,培养包含本发明的示例性负荷成瘾系统的酵母细胞被认为在经培养群体中富集了高产酵母变体。虽然在相对短的培养时间(30次细胞分裂)后可以检测到hSA产量的增加,但当传统培养物通常表现出生产率的显著下降时,高产酵母变体的富集在更长的培养期内得到维持和进一步增强。

[0295] 实施例7合适的负荷感应启动子候选物的鉴定

[0296] 不同的工程化生产基因和途径引发指示生产过程的不同转录反应。为了鉴定用作负荷感应器的合适的候选启动子,进行了以下实验。

[0297] 方法:典型的基因逃逸细胞是从在预期的发酵培养基中与感兴趣的基因工程化微生物生产细胞的长期培养中分离的。合适的基因逃逸细胞的特征在于具有比原始基因工程化生产细胞高至少5%的指数期生长速率和低至少30%的生产速率或产物收率。

[0298] 在预期的发酵条件、按比例缩小的条件或模拟预期发酵条件的摇瓶条件下培养生产细胞和对应的逃逸细胞。在对应于生产细胞中最高生产率的时间点,取样进行RNA测序。使用Purelink RNA Mini试剂盒(赛默飞世尔公司)纯化总RNA,并按照试剂盒制造商的说明使用TruSeq Stranded mRNA试剂盒(依诺米那公司(Illumina))制备总RNA。将读数作图并分析菌株的参考基因组,然后分析生产细胞和对应逃逸细胞之间的差异表达。

[0299] 结果:候选合适启动子被鉴定为驱动基因表达的那些启动子,所述基因在生产生物中相对于至少一种经分离的基因逃逸菌株显示出>3倍更高表达的差异表达。

[0300] 实施例8:生产IgA片段的负荷成瘾枯草芽孢杆菌菌株显示出增强的长期生产稳定性

[0301] 负荷感应启动子显示为可以感应并被细胞中负荷诱导状态激活的启动子,所述负荷诱导状态由重组蛋白(一种C-末端融合到GFP的IgA片段(IgA-GFP))的细胞合成产生。一旦被激活,则负荷感应启动子显示出将必需基因(操纵子)iscU或accC的表达提高到足以赋予细胞在与非生产性细胞相比时的选择性生长优势的水平。为了最大化响应于其同源负荷感应启动子的必需基因表达的动态范围,将负荷感应启动子与赋予不同翻译强度的变体RBS编码序列(表3.2)随机组合。

[0302] 如以下在模拟大规模生产条件下培养的工程化生产菌株所示,具有适合用于负荷成瘾基因线路的负荷感应特性的启动子显示包括热休克启动子、DNA损伤启动子和氧化应激反应启动子。

[0303] 8.1材料和方法

[0304] 负荷感应启动子:通过PCR和USER克隆产生启动子 P_{hrcA} P_{perR} 及其各自的靶向必需基因的整合序列,从而产生整合载体(表10)。

[0305]

表 10: 用于插入负荷感应启动子以控制必需基因的芽孢杆菌属染色体整合构建体。

质粒	启动子必需基因	用于选择染色体整合的抗性基因
pEG151 (SEQ ID NO.195)	<i>P_{groES-iscU}</i>	specR (SEQ ID No. 211)
pEG152 (SEQ ID NO.196)	<i>P_{ctsR-iscU}</i>	specR (SEQ ID No. 211)
pEG153 (SEQ ID NO.197)	<i>P_{dnaK-iscU}</i>	specR (SEQ ID No. 211)
pEG154 (SEQ ID NO.198)	<i>P_{perR-iscU}</i>	specR (SEQ ID No. 211)
pEG155 (SEQ ID NO.199)	<i>P_{hrcA-iscU}</i>	specR (SEQ ID No. 211)
pEG156 (SEQ ID NO.200)	<i>P_{spx-iscU}</i>	specR (SEQ ID No. 211)
pEG157 (SEQ ID NO.201)	<i>P_{sigB-iscU}</i>	specR (SEQ ID No. 211)
pEG158 (SEQ ID NO.202)	<i>P_{yflT-iscU}</i>	specR (SEQ ID No. 211)
pEG159 (SEQ ID NO.203)	<i>P_{groES-accC}</i>	specR (SEQ ID No. 211)
pEG160 (SEQ ID NO.204)	<i>P_{ctsR-accC}</i>	specR (SEQ ID No. 211)
pEG161 (SEQ ID NO.205)	<i>P_{dnaK-accC}</i>	specR (SEQ ID No. 211)
pEG162 (SEQ ID NO.206)	<i>P_{perR-accC}</i>	specR (SEQ ID No. 211)

[0306]

pEG163 (SEQ ID NO.207)	<i>P_{hrcA-accC}</i>	specR (SEQ ID No. 211)
pEG164 (SEQ ID NO.208)	<i>P_{spx-accC}</i>	specR (SEQ ID No. 211)
pEG165 (SEQ ID NO.209)	<i>P_{sigB-accC}</i>	specR (SEQ ID No. 211)
pEG166 (SEQ ID NO.210)	<i>P_{yflT-accC}</i>	specR (SEQ ID No. 211)
pEG204	<i>P_{groES-accC}</i>	camR (SEQ ID No. 220)
pEG205	<i>P_{ctsR-accC}</i>	camR (SEQ ID No. 220)
pEG206	<i>P_{dnaK-accC}</i>	camR (SEQ ID No. 220)
pEG207	<i>P_{perR-accC}</i>	camR (SEQ ID No. 220)
pEG208	<i>P_{hrcA-accC}</i>	camR (SEQ ID No. 220)
pEG209	<i>P_{spx-accC}</i>	camR (SEQ ID No. 220)
pEG210	<i>P_{sigB-accC}</i>	camR (SEQ ID No. 220)
pEG211	<i>P_{yflT-accC}</i>	camR (SEQ ID No. 220)

[0307]

表 11 研究的芽孢杆菌属菌株列表

菌株名称	新特征	用以下转化	新的被靶向的必需基因的 RBS 中的最后一个核苷酸	亲代菌株
枯草芽孢杆菌 KO7	-	-	-	-
EGS084	IgA-GFP 表达	pEG062 (SEQ ID NO.222)	-	枯草芽孢杆菌 KO7
EGS340	<i>P_{perR-iscU}</i>	pEG154	T	EGS084
EGS343	<i>P_{perR-accC}</i>	pEG162	T	EGS084
EGS460	<i>P_{ctsR-accC}</i>	pEG205	C	EGS340
EGS462	<i>P_{dnaK-accC}</i>	pEG206	G	EGS340
EGS466	<i>P_{hrcA-accC}</i>	pEG208	T	EGS340

[0308] 负荷成瘾启动子的染色体整合:为了构建候选负荷成瘾菌株(表11),如下通过同源重组将特征为融合到RBS变体(表3.2)的每个启动子的整合载体(表10)整合到产生IgA-GFP的EGS084菌株(对应于枯草芽孢杆菌KO7,含有插入amyE基因座中的pEG062表达盒)基因组中含有iscU或accC的操纵子上游。按照前述的方案(dx.doi.org/10.17504/protocols.io.bdmti46n)根据标准芽孢杆菌属转化方法制备感受态细胞,并用pEG151转化为pEG166。在补充有壮观霉素(200μg/mL)的LB琼脂板上选择转化体,并使用引物E257/E258(iscU)或E261/E262(accC)通过菌落PCR进行确认。

[0309]

表 12 引物		
SEQ ID NO.	寡聚 ID	序列 (5'-)
223	E257	CTGAGTCAATTTTCATCAAGG
224	E258	CAGAAGAAAACATTCTCGAAGC
225	E261	GAGGTCGCAACTGAAAAAGC
226	E262	ATGTTTCAGTCTCCTTACTCC

[0310] 长期IgA-GFP生产测定:将来自每个负荷成瘾IgA-GFP生产菌株的单个所选择菌落用于接种96深孔板,所述板含有补充有卡那霉素(10 μ g/mL)的500 μ L/孔Ca118-2培养基(dx.doi.org/10.17504/protocols.io.bdmui46w)。使培养物在30 $^{\circ}$ C/250RPM水平摇动培养箱(Innova公司,2英寸振幅)中生长24小时。然后在相同的培养条件下,将499 μ L新鲜Ca118-2培养基中的1 μ L每种培养物转移(作为连续传代)(对应于每次转移大约10代)到新的深孔板中。这些方法步骤总共重复多达13次。在每次连续传代中,将100 μ L每种菌株存在等量的50%甘油中用于储存和后续检查。为了在每次传代后定量IgA-GFP表达,使培养物在前述生长条件下再生长24小时。将培养物以2000g离心5分钟,并在1体积PBS中洗涤两次。将20 μ L培养物样品添加到96孔板中的180 μ L PBS(10倍稀释)中,并且通过在协同H4板读数器(Biotek公司)中测量OD600和GFP荧光($\lambda_{ex}/\lambda_{em}$ =485nm/528nm)来定量细胞密度和IgA-GFP产量。将每个培养物的特定生产水平量化为GFP信号,将所述信号归一化为OD600值,随后减去在不产生GFP的枯草芽孢杆菌K07菌株中测量的相同值。

[0311] 8.1结果

[0312] 在五次500倍稀释连续传代的长期培养(大约65代细胞)后,IgA-GFP产量被定量并被证实负荷成瘾菌株中得到改善(图15和16)。

[0313] 基于所述方法,可以使用可能与其他生产菌株的转录负荷反应相匹配的其他候选负荷感应器来构建类似的负荷成瘾菌株。可以使用设计的长期生产测定或例如在连续小规模(例如400mL)生产培养中容易地评估它们的适用性。

[0314] 实施例9:控制多种必需基因的负荷成瘾提高产量

[0315] 这里证明了如何在单个生产细胞中引入两个(并以此类推)控制不同必需基因(操纵子)的负荷感应启动子,以增加期望产物的产量并进一步延长生产细胞具备生产性的时段。另外,通过使用多于一个的负荷感应器来控制必需基因,可以进一步控制指导所施加的基于成瘾的选择方案的转录空间。

[0316] 对负荷成瘾的IgA-GFP生产菌株EGS340(控制iscU必需基因操纵子的 P_{hrcA})进行遗传转化,以将accC必需基因操纵子的天然启动子分别更换为候选负荷感应器 P_{ctsR} 、 P_{dnaK} 和 P_{hrcA} 。

[0317] 9.1材料和方法

[0318] 负荷成瘾启动子的染色体整合:使用引物E372/E373(用于载体骨架)和E374/E375(用于pDG1662中的氯霉素抗性标记)将壮观霉素抗性标记更换为pEG159-pEG166中的氯霉素抗性标记(Guerout-Fleury等人1996)以产生pEG204至pEG211(表10)。如下将每个启动子整合到EGS340菌株(对应于含有调节iscU必需基因操纵子的 P_{perR} 的EGS084)基因组中含有accC的必需基因操纵子的上游。根据前述的方法(dx.doi.org/10.17504/protocols.io.bdmti46n)制备感受态细胞,并用pEG204转化为pEG211。在补充有氯霉素(5 μ g/mL)的LB琼脂板上选择转化体,并使用引物E261/E262通过菌落PCR进行确认。

表 13 引物			
SEQ ID NO	寡聚 ID	序列 (5'-)	
[0319]	227	E372	ATTATAACAUGTATTCACGAACGAAAATCG
	228	E373	ATCAGATTUAGAAAACAATAAACCCCTTGCATAGG
	229	E374	AAAATCTGAUACGGGGCAGG
	230	E375	ATGTTATAAUACCCATTAGTTCAACAAACGAAAATTG

[0320] 长期IgA-GFP生产测定:将来自每个负荷成瘾IgA-GFP生产菌株的单个所选择菌落用于接种96深孔板,所述板含有补充有卡那霉素(10 μ g/mL)的500 μ L/孔Ca118-2培养基(dx.doi.org/10.17504/protocols.io.bdmui46w)。使培养物在30 $^{\circ}$ C/250RPM水平摇动培养箱(Innova公司,2英寸振幅)中生长24小时。然后在相同的培养条件下,将499 μ L新鲜Ca118-2培养基中的1 μ L每种培养物转移(作为连续传代)(对应于每次转移大约10代)到新的深孔板中。这些方法步骤总共重复多达13次。在每次连续传代中,将100 μ L每种菌株存在等量的50%甘油中用于储存和后续检查。为了在每次传代后定量IgA-GFP表达,使培养物在前述生长条件下再生长24小时。将培养物以2000g离心5分钟,并在1体积PBS中洗涤两次。将20 μ L培养物样品添加到96孔板中的180 μ L PBS(10倍稀释)中,并且通过在协同H4板读数器(Biotek公司)中测量OD600和GFP荧光($\lambda_{ex}/\lambda_{em}$ =485nm/528nm)来定量细胞密度和IgA-GFP产量。将每个培养物的特定生产水平量化为GFP信号,将所述信号归一化为OD600值,随后减去在不产生GFP的枯草芽孢杆菌K07菌株中测量的相同值。

[0321] 细胞破碎和ELISA:在PBS中洗涤后,将培养物以2000g离心5分钟。将沉淀重悬于1体积的裂解缓冲液(10mM Tris,pH 7.5,150mM NaCl,500 μ M EDTA)中,再次以6000g离心5分钟,并重悬于1体积的补充有溶菌酶(10mg/mL)的裂解缓冲液中。通过在37 $^{\circ}$ C下将反应孵育30分钟来破碎细胞。之后,将反应物短暂涡旋并在4 $^{\circ}$ C下以12000g离心30分钟。除了在添加生色底物后,在10分钟内每40秒在600nm处读取每个孔的吸光度,根据制造商的方案,使用来自Abcam(ab190536)的猪IgA ELISA试剂盒对上清液进行ELISA。将相对IgA浓度计算为原始培养物每OD600值的斜率。

[0322] 9.2结果

[0323] 单个负荷感应器菌株EGS340仅携带控制iscU必需基因操纵子(EGS340)转录的基于 p_{perR} 的负荷感应器。将EGS340与衍生菌株进行比较,其中accC必需基因操纵子的转录也分别由基于 p_{ctsR} 的负荷感应器(EGS460)、基于 p_{dnaK} 的负荷感应器(EGS462)和基于 p_{hrcA} 的负荷感应器(EGS466)调节。在大约75代细胞的长期培养后,IgA-GFP产量被定量并被证实负在负荷成瘾菌株中得到改善(图17和18)。

[0324] 基于所述方法,可以使用可能与其他生产菌株的转录负荷反应相匹配的其他候选负荷感应器来容易地构建类似的负荷成瘾菌株。可以使用设计的长期生产测定或在连续小规模(例如400mL)生产培养中容易地评估它们的适用性。

[0325] 实施例10:鉴定合适的负荷感应启动子的另外的方法

[0326] 在这个实施例中,鉴定了合适的候选负荷感应启动子用于负荷成瘾以维持特定生产菌株中的长期异源生产。

[0327] 所述方法分为用于从头发现(RNA测序)的任选途径1和用于确认推定的负荷感应启动子(q-PCR)的途径2。

[0328] 在途径1中,使用RNA测序将典型生产条件(就例如温度、搅拌、产物/副产物/底物浓度/生长阶段而言)期间生产生物中编码基因的转录活性与相同典型生产条件期间分离的非/低产生物的转录活性进行比较。可以任选地对分离的非/低产生物的培养物补加产物至在给定时间产生生物的相应培养物中大致发现的浓度。非/低产分离物通常可以在培养>50-100个分裂(优选)后分离,但合适的分离菌株也可以通过基因灭活一个或几个关键异源生产基因进行基因工程化。

[0329] 进行RNA测序,优选地重复至少三次,优选地根据本领域已知的标准方法使用Illumina短读测序。

[0330] 使用用于差异基因表达的标准生物信息学分析,例如结合edgeR工作流程(DOI:10.18129/B9.bioc.edgeR),在相同的典型生产条件下,从高产培养物中与非/低产培养物相比上调至少5-10倍的基因中鉴定合适的候选负荷感应启动子。通过使用这种方法鉴定候选的负荷感应启动子,可以发现优先地被产物形成的负荷激活的启动子(与仅仅是生产压力相比)。

[0331] 接下来,在途径2中,使用q-PCR在样品上确认推定的负荷感应启动子(来自途径1或其他列表),类似地在相同的典型生产条件(就例如温度、搅拌、产物/副产物/底物浓度/生长阶段而言)期间将一个非/低产分离物与高产量分离物进行比较。在相同的典型生产条件下,从高产培养物中与非/低产培养物相比上调至少5-10倍的基因中鉴定合适的候选负荷感应启动子。

[0332] 实施例11:控制多种必需基因的负荷成瘾提高重组人生长激素在大肠杆菌中的产量

[0333] 通过整合一个另外的不同负荷感应启动子来改善负荷成瘾大肠杆菌菌株s7.6#8的长期生产稳定性,所述启动子控制选自例如所述列表并按照建议进行筛选的第二必需基因的转录。

[0334] 11.1材料和方法

[0335] 使用标准电穿孔用重组工程质粒如pKD46转化s7.6#8单个菌落。使用重组工程,将融合到可变RBS(表3.1)的候选负荷感应启动子(表2)直接染色体整合到必需基因murI或类似必需基因的上游,并选择使用整合构建体中存在的壮观霉素抗性基因。

[0336] 在重组工程和重组工程质粒固化的所述步骤后,挑取5个菌落,对应于具有相同启动子整合但具有随机变异RBS序列的克隆。然后将每个克隆转移到含有补充有助于维持质粒pMevT的氯霉素的200 μ l 2xYT的96孔板的孔中,并进行培养。在冷冻96孔板之前,如实施例1.2.1所述,将2 μ l每个培养的克隆用于验证启动子整合。

[0337] 11.2结果

[0338] 通过200倍稀释度的连续传代来评估不同的所得克隆的长期生产稳定性。在大约75代细胞的长期培养后,IgA-GFP产量被定量并被证实与在特征为仅一个负荷调节必需基因的单一负荷成瘾菌株中相比,在新菌株中得到改善。

[0339] 实施例12:具有TIS序列变异的必需基因CIA1的表达调节在巴斯德毕赤酵母菌株中显示出响应性生长

[0340] 此实施例证明了必需基因的TIS(翻译起始位点)序列的变化对生长的调节。必需基因的翻译强度被TIS序列中的此类变异修饰,所述变异可以在真核生物中用于滴定负荷

感应启动子对必需基因表达水平的负荷反应,类似于在原核生物中使用核糖体结合位点。

[0341] 12.1材料和方法

[0342] 具有TIS(翻译起始位点)序列变异的负荷响应启动子:通过PCR和USER克隆产生启动子(P_{gsh2} (巴斯德毕赤酵母同源物CDS:chr1-4_0496的启动子),其具有不同的强度增加的TIS序列和*cia1*(巴斯德毕赤酵母同源物CDS:chr1-3_0207)必需基因整合序列)的变异,从而产生整合片段(表14)。

表 14: 用于插入具有 TIS 变异的启动子 P_{gsh2} 以控制必需基因 <i>cia1</i> 的毕赤酵母属染色体整合构建体			
整合片段	启动子必需基因	TIS 变异 (参见表 3.3)	用于选择染色体整合的抗性基因
[0343] INT1(SEQ ID NO 232)	P_{gsh2} - <i>cia1</i>	TIS1	<i>Sh ble</i> (SEQ ID NO 233)
INT3(SEQ ID NO 234)	P_{gsh2} - <i>cia1</i>	TIS3	<i>Sh ble</i> (SEQ ID NO 233)
INT5(SEQ ID NO 235)	P_{gsh2} - <i>cia1</i>	TIS5	<i>Sh ble</i> (SEQ ID NO 233)
INT9(SEQ ID NO 236)	P_{gsh2} - <i>cia1</i>	TIS9	<i>Sh ble</i> (SEQ ID NO 233)

[0344] 具有TIS序列变异的负荷响应启动子的染色体整合:为了构建必需基因*cia1*表达受调节的菌株,如下通过同源重组将特征为具有TIS序列变异的启动子 P_{gsh2} 的整合片段(表14)整合到产生IgAL-NanoLuc的EGS621菌株的*cia1* CDS上游。根据标准毕赤酵母属转化方法(Wu an Letchworth2018)制备电感受态细胞,并将其用整合片段INT1、INT3、INT5和INT9转化。在补充有Zeocin® (50μg/mL)的YPD琼脂板上选择转化体,并使用引物E521(SEQ ID NO.237)和E627(SEQ ID NO.238)通过菌落PCR进行确认。

表 15: 研究的毕赤酵母属菌株列表				
菌株名称	新特征	用以下转化	靶向必需基因 <i>cia1</i> 的启动子的 TIS 变异	亲代菌株
[0345] EGS621	IgAL-NanoLuc 表达	-	-	EGS75
EGS1100	P_{gsh2} - <i>cia1</i>	INT1	TIS1	EGS621
EGS1101	P_{gsh2} - <i>cia1</i>	INT3	TIS3	EGS621
[0346] EGS1102	P_{gsh2} - <i>cia1</i>	INT5	TIS5	EGS621
EGS1104	P_{gsh2} - <i>cia1</i>	INT9	TIS9	EGS621

[0347] 毕赤酵母属TIS序列生长反应测定:将亲代菌株EGS621和TIS序列变体EGS1100、EGS1101、EGS1102和EGS1104的单个菌落分别稀释在30μL MilliQ H₂O中。将5μL稀释的菌落接种在95μL YPD培养基中以进行预培养。将EGS621菌落用G418 (50μg/mL) 补充,并且将其其余菌落用Zeocin® (50μg/mL) 补充。将预培养物在30℃和300RPM摇动下生长过夜。然后在补充有适当抗生素的YPD培养基中将过夜预培养物稀释1000倍。将1μL稀释的培养物接种在96孔培养板中补充有适当抗生素的99μL YPD中,并用Breathe-Easy®密封膜(德国达姆施塔特的西格玛奥德里奇公司(Sigma-Aldrich,Darmstadt,Germany))密封。将96孔板放置在Synergy 4板读数器(美国佛蒙特州的BioTek公司)中,并且使用以下设置使培养物生长:30℃,高摇动速度,48小时内每10分钟测量一次OD₆₀₀。

[0348] 12.2结果

[0349] 展现了所有菌株之间的生长差异-参见图19。具有天然启动子 P_{cia1} 的亲代菌株(EGS621)表现出最快的生长。具有正好整合在*cia1* CDS上游的 P_{gsh2} 启动子的TIS序列变异的菌株(EGS1100、EGS1101、EGS1102和EGS1104)根据其相对TIS强度展示生长。具有最强TIS序列的菌株TIS9(EGS1104)表现出比亲代菌株(EGS621)略慢的生长,而TIS菌株(EGS1100、EGS1101、EGS1102和EGS1104)的生长最快。具有最弱TIS序列TIS1的菌株(EGS1100)表现出

在所有菌株中的最慢生长。TIS3 (EGS1101) 和TIS5 (EGS1102) 菌株表现出与四个TIS序列之间的相对强度一致的生长。

[0350] 总之：证明毕赤酵母属菌株的生长响应于 P_{gsh2} 的TIS序列强度，所述 P_{gsh2} 正好整合在cia1CDS的上游，证实了cia1是毕赤酵母属菌株生长的必需基因。

[0351] 实施例13：在米曲霉生产菌株中实施负荷成瘾

[0352] 本发明的负荷成瘾系统还可以在其他真核生物如真菌中实施。本文提供了在米曲霉生产菌株中实施负荷成瘾的实施例。

[0353] 13.1 负荷感应启动子元件

[0354] 负荷感应启动子的示例包括来自bipA (例如SEQ ID NO.189)、clxA (例如SEQ ID NO.190) 和agsA (例如SEQ ID NO.239) 的基因同源物的启动子序列 (起始密码子上游750bp)。为了构建负荷成瘾的米曲霉生产菌株，将负荷启动子整合到必需基因如ERG10 (例如SEQ ID NO.240)、PFS2 (例如SEQ ID NO.242) 或TUB1 (例如SEQ ID NO.244) 同源物的前面。为了调整必需基因的表达，将负荷感应启动子与4个不同变体的翻译起始序列 (TIS) 文库整合，如表3.3所呈现。这些变体替代了启动子序列的最后6个核苷酸，即相对于起始密码子的-6至-1。负荷成瘾整合片段的示例在表16中呈现。

表 16：米曲霉负荷成瘾整合构建体		
整合片段	负荷感应启动子	必需基因
[0355] agsA_ERG10_pyrG_NC_036442 (SEQ ID NO.246)	agsA	ERG10
	agsA	PFS2
	agsA	TUB1
	bipA	ERG10
[0356] bipA_PFS2_pyrG_NC_036436 (SEQ ID NO.250)	bipA	PFS2
	bipA	TUB1
	clxA	ERG10
	clxA	PFS2
	clxA	TUB1

[0357] 13.2 负荷感应整合构建体的构建

[0358] 为了将天然必需基因启动子替换为负荷感应启动子，通过标准分子克隆制备了三个必需基因靶向DNA片段。从米曲霉RIB40基因组DNA中扩增每个必需基因起始密码子上游和下游大约1.5kbp的区域 (分别为ERG10、PFS和TUB1)。接下来，同样从米曲霉基因组DNA中扩增紧接bipA、clxA和agsA各自起始密码子上游的750bp启动子区域。最后，获得含有选择标记例如amdS或pyrG的合成基因片段。使用例如Gibson组装技术组装所有片段。使用PCR从组装反应中扩增全长敲除构建体。

[0359] 根据标准方案，如Christensen等人1988制备了米曲霉原生质体，并且将其用负荷感应整合构建体转化并在适当的培养基上选择。

[0360] 使用含有10ml YPM培养基(2g/1酵母提取物、2g/1蛋白胨和2%麦芽糖)的摇瓶测定表达,所述YPM培养基接种了来自转化体菌株和不含负荷感应启动子元件的参照生产菌株的孢子,并在30℃、200rpm下孵育4天,随后对产物进行取样。

[0361] 还在补料分批发酵中进行了进一步测定:罐培养基(蔗糖24g/L,酵母提取物10g/L, $(\text{NH}_4)_2\text{SO}_4$ 5g/L, $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 2g/L, K_2SO_4 2g/L,柠檬酸1g/L, KH_2PO_4 2g/L痕量金属溶液0.5ml/L),温度34℃,通气量1vvm,并使用10% NH_4OH 将pH控制在6.0。对培养基接种包含负荷感应整合构建体和参照菌株的转化体菌株的种子培养物(将孢子在摇瓶(甘油20g/L,酵母提取物18g/L)中在30℃和250rpm下预生长1天)。当pH>6.4时,以3.33g/L/h的速率开始补料(400g/L麦芽糖糖浆、1g/L柠檬酸)。控制搅拌器速度以避免氧张力过低(<20%)。

[0362] 参考文献

[0363] Bonde, M.T., Pedersen, M., Klausen, M.S., Jensen, S.I., Wulff, T., Harrison, S., 等人(2016). Predictable tuning of protein expression in bacteria. *Nat. Methods* 13. doi:10.1038/nmeth.3727.

[0364] Christensen, T., Woeldike, H., Boel, E. 等人(1988) High Level Expression of Recombinant Genes in *Aspergillus Oryzae*. *Nat Biotechnol* 6, 1419-1422. doi.org/10.1038/nbt1288-1419.

[0365] Falkenberg et al (2022). Protein expression in *Bacillus subtilis*. Oct 27, 2022. Website: protocols.io. doi.org/10.17504/protocols.io.bdmti46n.

[0366] Giaever G 等人, *Nature*. 2002 Jul 25; 418 (6896) : 387-91. DOI: 10.1038/nature00935.

[0367] Jain, R., Kumar, P. & Varshney, U. A distinct role of formamidopyrimidine DNA glycosylase (MutM) in down-regulation of accumulation of G,C mutations and protection against oxidative stress in mycobacteria. *DNA Repair (Amst)*. 6, 1774-1785 (2007).

[0368] Guerout-Fleury AM, Frandson N, Stragier P. 1996. Plasmids for ectopic integration in *Bacillus subtilis*. *Gene* 180:57-61.

[0369] Koskineniemi, S., Pr nting, M., Gullberg, E., N svall, J., & Andersson, D. I. (2011). Activation of cryptic aminoglycoside resistance in *Salmonella enterica*. *Molecular Microbiology*, Vol. 80, pp. 1464-1478. https://doi.org/10.1111/j.1365-2958.2011.07657.x.

[0370] Kimata, Y., Ishiwata-Kimata, Y., Yamada, S., and Kohno, K. (2006). Yeast unfolded protein response pathway regulates expression of genes for anti-oxidative stress and for cell surface proteins. *Genes to Cells* 11, 59-69. doi: 10.1111/j.1365-2443.2005.00921.x..

[0371] Lafert , A., Favry, E., Sentenac, A., Riva, M., Carles, C., and Ch din, S. (2006). The transcriptional activity of RNA polymerase I is a key determinant for the level of all ribosome components. *Genes Dev.* 20, 2030-2040. doi: 10.1101/gad.386106.

[0372] Maeda, M., Shimada, T., and Ishihama, A. (2015). Strength and Regulation of

Seven rRNA Promoters in *Escherichia coli*. PLoS One 10,1-19.doi:10.1371/journal.pone.0144697.

[0373] Nonaka,G.,Blankschien,M.,Herman,C.,Gross,C.a,and Rhodius,V.a(2006).Regulon and promoter analysis of the *E.coli* heat-shock factor,sigma32, reveals a multifaceted cellular response to heat stress.Genes Dev.20,1776-89.doi:10.1101/gad.1428206.

[0374] Pitera,D.J.,Paddon,C.J.,Newman,J.D.,and Keasling,J.D.(2007).Balancing a heterologous mevalonate pathway for improved isoprenoid production in *Escherichia coli*.Metab.Eng.9,193-207.doi:10.1016/j.ymben.2006.11.002.

[0375] Rugbjerg,P.,Sarup-Lytzen,K.,Nagy,M.,and Sommer,M.O.A.(2018).Synthetic addiction extends the productive life time of engineered *Escherichia coli* populations.Proc.Natl.Acad.Sci.115,2347-2352.doi:10.1073/pnas.1718622115.

[0376] Wu,S.and Letchworth,G.J.(2004)High efficiency transformation by electroporation of *Pichia pastoris* pretreated with lithium acetate and dithiothreitol.Biotechniques 36,152-154.

[0377] Wu an Letchworth(2018)High efficiency transformation by electroporation of *Pichia pastoris* pretreated with lithium acetate and dithiothreitol.BIOTECHNIQUESVOL.36,NO.1DRUG DISCOVERY AND GENOMIC TECHNOLOGIES.doi.org/10.2144/04361DD02.

[0378] Yoon,S.H.,Han,M.J.,Lee,S.Y.,Jeong,K.J.,and Yoo,J.S.(2003).Combined transcriptome and proteome analysis of *Escherichia coli* during high cell density culture.Biotechnol.Bioeng.81,753-767.doi:10.1002/bit.10626.

[0379] Uppal,S.,Shetty,D.M.,&Jawali,N.(2014).Cyclic AMP receptor protein regulates cspD,a bacterial toxin gene,in *Escherichia coli*.Journal of Bacteriology,196(8),1569-1577.https://doi.org/10.1128/JB.01476-13.

[0380] Yamanaka,K.,&Inouye,M.(1997).Growth-phase-dependent expression of cspD,encoding a member of the CspA family in *Escherichia coli*.Journal of Bacteriology,179(16),5126-5130.https://doi.org/10.1128/jb.179.16.5126-5130.1997.

序列表	
<110>	丹麦技术大学
<120>	负荷成瘾生产菌株
<130>	P3061EP00
<140>	21191995.6
<141>	2021-08-18
<160>	254
<170>	PatentIn 3.5 版
<210>	1
<211>	2179
<212>	DNA
<213>	大肠杆菌 (E. coli)
<220>	
<221>	基因
<222>	(1)..(2179)
<223>	大肠杆菌必需基因 folP-glmM
<220>	
<221>	CDS
<222>	(1)..(849)
<223>	folP
<400>	1
atg aaa ctc ttt gcc cag ggt act tca ctg gac ctt agc cat cct cac	48
Met Lys Leu Phe Ala Gln Gly Thr Ser Leu Asp Leu Ser His Pro His	
1 5 10 15	
gta atg ggg atc ctc aac gtc acg cct gat tcc ttt tcg gat ggt ggc	96
Val Met Gly Ile Leu Asn Val Thr Pro Asp Ser Phe Ser Asp Gly Gly	
20 25 30	
acg cat aac tcg ctg ata gat gcg gtg aaa cat gcg aat ctg atg atc	144
Thr His Asn Ser Leu Ile Asp Ala Val Lys His Ala Asn Leu Met Ile	
35 40 45	
aat gct ggc gcg acg atc att gac gtt ggt ggc gag tcc acg cgc cca	192
Asn Ala Gly Ala Thr Ile Ile Asp Val Gly Gly Glu Ser Thr Arg Pro	
50 55 60	
ggg gcg gcg gaa gtt agc gtt gaa gaa gag ttg caa cgt gtt att cct	240
Gly Ala Ala Glu Val Ser Val Glu Glu Glu Leu Gln Arg Val Ile Pro	
65 70 75 80	
gtg gtt gag gca att gct caa cgc ttc gaa gtc tgg atc tcg gtc gat	288
Val Val Glu Ala Ile Ala Gln Arg Phe Glu Val Trp Ile Ser Val Asp	
85 90 95	
aca tcc aaa cca gaa gtc atc cgt gag tca gcg aaa gtt ggc gct cac	336
Thr Ser Lys Pro Glu Val Ile Arg Glu Ser Ala Lys Val Gly Ala His	
100 105 110	
att att aat gat atc cgc tcc ctt tcc gaa cct ggc gct ctg gag gcg	384
Ile Ile Asn Asp Ile Arg Ser Leu Ser Glu Pro Gly Ala Leu Glu Ala	
115 120 125	
gct gca gaa acc ggt tta ccg gtt tgt ctg atg cat atg cag gga aat	432
Ala Ala Glu Thr Gly Leu Pro Val Cys Leu Met His Met Gln Gly Asn	
130 135 140	
cca aaa acc atg cag gaa gct ccg aag tat gac gat gtc ttt gca gaa	480
Pro Lys Thr Met Gln Glu Ala Pro Lys Tyr Asp Asp Val Phe Ala Glu	
145 150 155 160	
gtg aat cgc tac ttt att gag caa ata gca cgt tgc gag cag gcg ggt	528
Val Asn Arg Tyr Phe Ile Glu Gln Ile Ala Arg Cys Glu Gln Ala Gly	
165 170 175	
atc gca aaa gag aaa ttg ttg ctc gac gcc gga ttc ggt ttc ggt aaa	576
Ile Ala Lys Glu Lys Leu Leu Leu Asp Pro Gly Phe Gly Phe Gly Lys	
180 185 190	
aat ctc tcc cat aac tat tca tta ctg gcg cgc ctg gct gaa ttt cac	624
Asn Leu Ser His Asn Tyr Ser Leu Leu Ala Arg Leu Ala Glu Phe His	
195 200 205	
cat ttc aac ctg ccg ctg ttg gtg ggt atg tca cga aaa tcg atg att	672
His Phe Asn Leu Pro Leu Leu Val Gly Met Ser Arg Lys Ser Met Ile	
210 215 220	
ggg cag ctg ctg aac gtg ggg ccg tcc gag cgc ctg agc ggt agt ctg	720
Gly Gln Leu Leu Asn Val Gly Pro Ser Glu Arg Leu Ser Gly Ser Leu	
225 230 235 240	
gcc tgt gcg gtc att gcc gca atg caa ggc gcg cac atc att cgt gtt	768
Ala Cys Ala Val Ile Ala Ala Met Gln Gly Ala His Ile Ile Arg Val	
245 250 255	
cat gac gtc aaa gaa acc gta gaa gcg atg cgg gtg gtg gaa gcc act	816
His Asp Val Lys Glu Thr Val Glu Ala Met Arg Val Val Glu Ala Thr	
260 265 270	
ctg tct gca aag gaa aac aaa cgc tat gag taa tcgtaaatat ttcggtaccg	869
Leu Ser Ala Lys Glu Asn Lys Arg Tyr Glu	
275 280	
atgggattcg tggctcgtgta ggggatgcgc cgatcacacc tgattttgtg cttaagctgg	929
gttggggccgc gggtaaagtgc ctggcgcgcc acgggtcccc taagatcatt attggtaaag	989

[0381]

```

acacgcgtat ttctggctat atgctggagt cagcactgga agcgggtctg gcggcagcgg 1049
gcctttccgc actcttcact ggcccgatgc caacaccggc cgtggcttat ctgacgcgta 1109
ccttcgcgcg agaggccgga attgtgatat ctgcatcgca taacccgttc tacgataatg 1169
gcattaaatt cttctctatc gacggcacca aactgcccga tgcggtagaa gaggccatcg 1229
aagcggaaat ggaaaaggag atcagctgcg ttgattcggc agaactgggt aaagccagcc 1289
gtatcggtga tgccgcgggt cgctatatcg agttttgcaa agccacgttc ccgaacgaac 1349
ttagcctcag tgaactgaag attgtggtgg attgtgcaaa cgggtgcgact tatcacatcg 1409
cgccgaacgt gctgcgcgaa ctggggggcga acgttatcgc tatcggttgt gagccaaacg 1469
gtgtaaacaat caatgccgaa gtggggggcta ccgacgttcg cgcgctccag gctcgtgtgc 1529
tggctgaaaa agcggatctc ggtattgcct tcgacggcga tggcgatcgc gtgattatgg 1589
ttgacctga aggcaataaa gtcgatggcg atcagatcat gtatatcatc gcgcgtgaag 1649
gtcttcgtca gggccagctg cgtggtggcg ctgtgggtac attgatgagc aacatggggc 1709
ttgaactggc gctgaaacag ttaggaattc catttgccgc cgcgaaagtg ggtgaccgct 1769
acgtactgga aaaaatgcag gagaaaaggc ggcgatcggg tgcaagaat tccggtcag 1829
tgatcctgct ggataaaaact actaccgggt acggcatcgt tgctggcttg caggtgctgg 1889
cggctatggc acgtaaccat atgagcctgc acgacctttg cagcggcatg aaaatgttcc 1949
cgcagattct ggttaacgta cgttacaccg caggtagcgg cgatccactt gagcatgagt 2009
cagttaaagc cgtgaccgca gaggttgaag ctgcgctggg caaccgtgga cgcgtgttgc 2069
tgcgtaaatc cggcaccgaa ccgttaattc gcgtgatggt ggaaggcgaa gacgaagcgc 2129
aggtgactga atttgcacac cgcacgcgcg atgcagtaaa agccgtttaa 2179
<210> 2
<211> 282
<212> PRT
<213> 大肠杆菌
<400> 2
Met Lys Leu Phe Ala Gln Gly Thr Ser Leu Asp Leu Ser His Pro His
1 5 10 15
Val Met Gly Ile Leu Asn Val Thr Pro Asp Ser Phe Ser Asp Gly Gly
20 25 30
Thr His Asn Ser Leu Ile Asp Ala Val Lys His Ala Asn Leu Met Ile
35 40 45
Asn Ala Gly Ala Thr Ile Ile Asp Val Gly Gly Glu Ser Thr Arg Pro
50 55 60
Gly Ala Ala Glu Val Ser Val Glu Glu Glu Leu Gln Arg Val Ile Pro
65 70 75 80
Val Val Glu Ala Ile Ala Gln Arg Phe Glu Val Trp Ile Ser Val Asp
85 90 95
Thr Ser Lys Pro Glu Val Ile Arg Glu Ser Ala Lys Val Gly Ala His
100 105 110
Ile Ile Asn Asp Ile Arg Ser Leu Ser Glu Pro Gly Ala Leu Glu Ala
115 120 125
Ala Ala Glu Thr Gly Leu Pro Val Cys Leu Met His Met Gln Gly Asn
130 135 140
Pro Lys Thr Met Gln Glu Ala Pro Lys Tyr Asp Asp Val Phe Ala Glu
145 150 155 160
Val Asn Arg Tyr Phe Ile Glu Gln Ile Ala Arg Cys Glu Gln Ala Gly
165 170 175
Ile Ala Lys Glu Lys Leu Leu Leu Asp Pro Gly Phe Gly Phe Gly Lys
180 185 190
Asn Leu Ser His Asn Tyr Ser Leu Ala Arg Leu Ala Glu Phe His
195 200 205
His Phe Asn Leu Pro Leu Leu Val Gly Met Ser Arg Lys Ser Met Ile
210 215 220
Gly Gln Leu Leu Asn Val Gly Pro Ser Glu Arg Leu Ser Gly Ser Leu
225 230 235 240
Ala Cys Ala Val Ile Ala Ala Met Gln Gly Ala His Ile Ile Arg Val
245 250 255
His Asp Val Lys Glu Thr Val Glu Ala Met Arg Val Val Glu Ala Thr
260 265 270
Leu Ser Ala Lys Glu Asn Lys Arg Tyr Glu
275 280
<210> 3
<211> 1338
<212> DNA
<213> 大肠杆菌
<220>
<221> CDS
<222> (1)..(1338)
<223> 大肠杆菌必需基因 glmM
<400> 3
atg agt aat cgt aaa tat ttc ggt acc gat ggg att cgt ggt cgt gta 48
1 5 10 15
Met Ser Asn Arg Lys Tyr Phe Gly Thr Asp Gly Ile Arg Gly Arg Val
20 25 30
ggg gat gcg ccg atc aca cct gat ttt gtg ctt aag ctg ggt tgg gcc 96
Gly Asp Ala Pro Ile Thr Pro Asp Phe Val Leu Lys Leu Gly Trp Ala
35 40 45
gcg ggt aaa gtg ctg gcg cgc cac ggc tcc cgt aag att att att ggt 144
Ala Gly Lys Val Leu Ala Arg His Gly Ser Arg Lys Ile Ile Ile Gly

```


51

<212> PRT
 <213> 大肠杆菌
 <400> 4
 Met Ser Asn Arg Lys Tyr Phe Gly Thr Asp Gly Ile Arg Gly Arg Val
 1 5 10 15
 Gly Asp Ala Pro Ile Thr Pro Asp Phe Val Leu Lys Leu Gly Trp Ala
 20 25 30
 Ala Gly Lys Val Leu Ala Arg His Gly Ser Arg Lys Ile Ile Ile Gly
 35 40 45
 Lys Asp Thr Arg Ile Ser Gly Tyr Met Leu Glu Ser Ala Leu Glu Ala
 50 55 60
 Gly Leu Ala Ala Ala Gly Leu Ser Ala Leu Phe Thr Gly Pro Met Pro
 65 70 75 80
 Thr Pro Ala Val Ala Tyr Leu Thr Arg Thr Phe Arg Ala Glu Ala Gly
 85 90 95
 Ile Val Ile Ser Ala Ser His Asn Pro Phe Tyr Asp Asn Gly Ile Lys
 100 105 110
 Phe Phe Ser Ile Asp Gly Thr Lys Leu Pro Asp Ala Val Glu Glu Ala
 115 120 125
 Ile Glu Ala Glu Met Glu Lys Glu Ile Ser Cys Val Asp Ser Ala Glu
 130 135 140
 Leu Gly Lys Ala Ser Arg Ile Val Asp Ala Ala Gly Arg Tyr Ile Glu
 145 150 155 160
 Phe Cys Lys Ala Thr Phe Pro Asn Glu Leu Ser Leu Ser Glu Leu Lys
 165 170 175
 Ile Val Val Asp Cys Ala Asn Gly Ala Thr Tyr His Ile Ala Pro Asn
 180 185 190
 Val Leu Arg Glu Leu Gly Ala Asn Val Ile Ala Ile Gly Cys Glu Pro
 195 200 205
 Asn Gly Val Asn Ile Asn Ala Glu Val Gly Ala Thr Asp Val Arg Ala
 210 215 220
 Leu Gln Ala Arg Val Leu Ala Glu Lys Ala Asp Leu Gly Ile Ala Phe
 225 230 235 240
 Asp Gly Asp Gly Asp Arg Val Ile Met Val Asp His Glu Gly Asn Lys
 245 250 255
 Val Asp Gly Asp Gln Ile Met Tyr Ile Ile Ala Arg Glu Gly Leu Arg
 260 265 270
 Gln Gly Gln Leu Arg Gly Gly Ala Val Gly Thr Leu Met Ser Asn Met
 275 280 285
 Gly Leu Glu Leu Ala Leu Lys Gln Leu Gly Ile Pro Phe Ala Arg Ala
 290 295 300
 Lys Val Gly Asp Arg Tyr Val Leu Glu Lys Met Gln Glu Lys Gly Trp
 305 310 315 320
 Arg Ile Gly Ala Glu Asn Ser Gly His Val Ile Leu Leu Asp Lys Thr
 325 330 335
 Thr Thr Gly Asp Gly Ile Val Ala Gly Leu Gln Val Leu Ala Ala Met
 340 345 350
 Ala Arg Asn His Met Ser Leu His Asp Leu Cys Ser Gly Met Lys Met
 355 360 365
 Phe Pro Gln Ile Leu Val Asn Val Arg Tyr Thr Ala Gly Ser Gly Asp
 370 375 380
 Pro Leu Glu His Glu Ser Val Lys Ala Val Thr Ala Glu Val Glu Ala
 385 390 395 400
 Ala Leu Gly Asn Arg Gly Arg Val Leu Leu Arg Lys Ser Gly Thr Glu
 405 410 415
 Pro Leu Ile Arg Val Met Val Glu Gly Glu Asp Glu Ala Gln Val Thr
 420 425 430
 Glu Phe Ala His Arg Ile Ala Asp Ala Val Lys Ala Val
 435 440 445
 <210> 5
 <211> 858
 <212> DNA
 <213> 大肠杆菌
 <220>
 <221> CDS
 <222> (1) .. (858)
 <223> 大肠杆菌必需基因 murI
 <400> 5
 atg gct acc aaa ctg cag gac ggg aat aca cct tgt ctg gca gct aca 48
 Met Ala Thr Lys Leu Gln Asp Gly Asn Thr Pro Cys Leu Ala Ala Thr
 1 5 10 15
 cct tct gaa cca cgt ccc acc gtg ctg gtg ttt gac tcc ggc gtc ggt 96
 Pro Ser Glu Pro Arg Pro Thr Val Leu Val Phe Asp Ser Gly Val Gly
 20 25 30
 ggg ttg tcg gtc tat gac gag atc cgg cat ctc tta ccg gat ctc cat 144
 Gly Leu Ser Val Tyr Asp Glu Ile Arg His Leu Leu Pro Asp Leu His
 35 40 45
 tac att tat gct ttc gat aac gtc gct ttc ccg tat ggc gaa aaa agc 192

	Tyr	Ile	Tyr	Ala	Phe	Asp	Asn	Val	Ala	Phe	Pro	Tyr	Gly	Glu	Lys	Ser	
	50					55					60						
	gaa	gcg	ttt	att	gtt	gag	cga	gtg	gtg	gca	att	gtc	acc	gcg	gtg	caa	240
	Glu	Ala	Phe	Ile	Val	Glu	Arg	Val	Val	Ala	Ile	Val	Thr	Ala	Val	Gln	
	65					70					75					80	
	gaa	cgt	tat	ccc	ctt	gcg	ctg	gct	gtg	gtc	gct	tgc	aac	act	gcc	agt	288
	Glu	Arg	Tyr	Pro	Leu	Ala	Leu	Ala	Val	Val	Ala	Cys	Asn	Thr	Ala	Ser	
					85					90					95		
	acc	gtt	tca	ctt	cct	gca	tta	cgc	gaa	aag	ttc	gac	ttc	ccg	gtt	gtt	336
	Thr	Val	Ser		100	Pro	Ala	Leu	Arg	Glu	Lys	Phe	Asp	Phe	Pro	Val	Val
					105									110			
	ggc	gtc	gtg	ccg	gcg	att	aaa	cct	gct	gca	cgt	ctg	acg	gca	aat	ggc	384
	Gly	Val	Val	Pro	Ala	Ile	Lys	Pro	Ala	Ala	Arg	Leu	Thr	Ala	Asn	Gly	
					115									125			
	att	gtc	gga	tta	ctg	gca	acc	cgc	gga	aca	gtt	aaa	cgt	tct	tat	act	432
	Ile	Val	Gly	Leu	Leu	Ala	Thr	Arg	Gly	Thr	Val	Lys	Arg	Ser	Tyr	Thr	
					130									140			
	cat	gag	ctg	atc	gcg	cgt	ttc	gct	aat	gaa	tgc	cag	ata	gaa	atg	ctg	480
	His	Glu	Leu	Ile	Ala	Arg	Phe	Ala	Asn	Glu	Cys	Gln	Ile	Glu	Met	Leu	
	145					150					155				160		
	ggc	tcg	gca	gag	atg	gtt	gag	ttg	gct	gaa	gcg	aag	cta	cat	ggc	gaa	528
	Gly	Ser	Ala	Glu	Met	Val	Glu	Leu	Ala	Glu	Ala	Lys	Leu	His	Gly	Glu	
					165					170					175		
	gat	gtt	tct	ctg	gat	gca	cta	aaa	cgt	atc	cta	cgc	ccg	tgg	tta	aga	576
	Asp	Val	Ser	Leu	Asp	Ala	Leu	Lys	Arg	Ile	Leu	Arg	Pro	Trp	Leu	Arg	
					180					185				190			
	atg	aaa	gag	ccg	cca	gat	acc	gtt	gta	ttg	ggc	tgc	acc	cat	ttc	cct	624
	Met	Lys	Glu	Pro	Pro	Asp	Thr	Val	Val	Leu	Gly	Cys	Thr	His	Phe	Pro	
						200							205				
	cta	cta	caa	gaa	gaa	ctg	tta	caa	gtg	ctg	cca	gag	gga	acc	cgg	ctg	672
	Leu	Leu	Gln	Glu	Glu	Leu	Leu	Gln	Val	Leu	Pro	Glu	Gly	Thr	Arg	Leu	
						215							220				
	gtg	gat	tct	ggc	gca	gcg	att	gct	cgc	cga	acg	gcc	tgg	ttg	tta	gaa	720
	Val	Asp	Ser	Gly	Ala	Ala	Ile	Ala	Arg	Arg	Thr	Ala	Trp	Leu	Leu	Glu	
	225					230					235				240		
	cat	gaa	gcc	ccg	gat	gca	aaa	tct	gcc	gat	gcg	aat	att	gcc	ttt	tgt	768
	His	Glu	Ala	Pro	Asp	Ala	Lys	Ser	Ala	Asp	Ala	Asn	Ile	Ala	Phe	Cys	
					245					250					255		
[0385]	atg	gca	atg	acg	cca	gga	gct	gaa	caa	tta	ttg	ccc	gtt	tta	cag	cgt	816
	Met	Ala	Met	Thr	Pro	Gly	Ala	Glu	Gln	Leu	Leu	Pro	Val	Leu	Gln	Arg	
					260				265					270			
	tac	ggc	ttc	gaa	acg	ctc	gaa	aaa	ctg	gca	gtt	tta	ggc	tga			858
	Tyr	Gly	Phe	Glu	Thr	Leu	Glu	Lys	Leu	Ala	Val	Leu	Gly				
					275			280					285				
	<210>	6															
	<211>	285															
	<212>	PRT															
	<213>	大肠杆菌															
	<400>	6															
	Met	Ala	Thr	Lys	Leu	Gln	Asp	Gly	Asn	Thr	Pro	Cys	Leu	Ala	Ala	Thr	
	1				5					10					15		
	Pro	Ser	Glu	Pro	Arg	Pro	Thr	Val	Leu	Val	Phe	Asp	Ser	Gly	Val	Gly	
					20				25					30			
	Gly	Leu	Ser	Val	Tyr	Asp	Glu	Ile	Arg	His	Leu	Leu	Pro	Asp	Leu	His	
					35			40					45				
	Tyr	Ile	Tyr	Ala	Phe	Asp	Asn	Val	Ala	Phe	Pro	Tyr	Gly	Glu	Lys	Ser	
	50					55					60						
	Glu	Ala	Phe	Ile	Val	Glu	Arg	Val	Val	Ala	Ile	Val	Thr	Ala	Val	Gln	
	65					70					75				80		
	Glu	Arg	Tyr	Pro	Leu	Ala	Leu	Ala	Val	Val	Ala	Cys	Asn	Thr	Ala	Ser	
					85					90					95		
	Thr	Val	Ser	Leu	Pro	Ala	Leu	Arg	Glu	Lys	Phe	Asp	Phe	Pro	Val	Val	
					100				105					110			
	Gly	Val	Val	Pro	Ala	Ile	Lys	Pro	Ala	Ala	Arg	Leu	Thr	Ala	Asn	Gly	
					115				120					125			
	Ile	Val	Gly	Leu	Leu	Ala	Thr	Arg	Gly	Thr	Val	Lys	Arg	Ser	Tyr	Thr	
					130			135					140				
	His	Glu	Leu	Ile	Ala	Arg	Phe	Ala	Asn	Glu	Cys	Gln	Ile	Glu	Met	Leu	
	145					150					155				160		
	Gly	Ser	Ala	Glu	Met	Val	Glu	Leu	Ala	Glu	Ala	Lys	Leu	His	Gly	Glu	
					165					170					175		
	Asp	Val	Ser	Leu	Asp	Ala	Leu	Lys	Arg	Ile	Leu	Arg	Pro	Trp	Leu	Arg	
					180				185					190			
	Met	Lys	Glu	Pro	Pro	Asp	Thr	Val	Val	Leu	Gly	Cys	Thr	His	Phe	Pro	
					195			200					205				
	Leu	Leu	Gln	Glu	Glu	Leu	Leu	Gln	Val	Leu	Pro	Glu	Gly	Thr	Arg	Leu	
					210			215				220					
	Val	Asp	Ser	Gly	Ala	Ala	Ile	Ala	Arg	Arg	Thr	Ala	Trp	Leu	Leu	Glu	

225	His	Glu	Ala	Pro	Asp	230	Ala	Lys	Ser	Ala	Asp	235	Ala	Asn	Ile	Ala	Phe	240	Cys
Met	Ala	Met	Thr	245	Pro	Gly	Ala	Glu	Gln	250	Leu	Leu	Pro	Val	Leu	255	Gln	Arg	
Tyr	Gly	Phe	Glu	260	Thr	Leu	Glu	Lys	Leu	265	Ala	Val	Leu	Gly	270				
<210>	7																		
<211>	1104																		
<212>	DNA																		
<213>	大肠杆菌																		
<220>																			
<221>	CDS																		
<222>	(1)..(1104)																		
<223>	大肠杆菌必需基因 asd																		
<400>	7																		
atg	aaa	aat	ggt	ggt	ttt	atc	ggc	tgg	cgc	ggt	atg	gtc	ggc	tcc	ggt				48
Met	Lys	Asn	Val	Gly	Phe	Ile	Gly	Trp	Arg	Gly	Met	Val	Gly	Ser	Val				
1				5					10					15					
ctc	atg	caa	cgc	atg	ggt	gaa	gag	cgc	gac	ttc	gac	gcc	att	cgc	cct				96
Leu	Met	Gln	Arg	Met	Val	Glu	Glu	Arg	Asp	Phe	Asp	Ala	Ile	Arg	Pro				
				20				25					30						
gtc	tac	ttt	tct	act	tct	cag	ctt	ggt	cag	gct	gcg	ccg	tct	ttt	ggc				144
Val	Tyr	Phe	Ser	Thr	Ser	Gln	Leu	Gly	Gln	Ala	Ala	Pro	Ser	Phe	Gly				
				35			40					45							
gga	acc	act	ggc	aca	ctt	cag	gat	gcc	ttt	gat	ctg	gag	gcg	cta	aag				192
Gly	Thr	Thr	Gly	Thr	Leu	Gln	Asp	Ala	Phe	Asp	Leu	Glu	Ala	Leu	Lys				
				50			55				60								
gcc	ctc	gat	atc	att	gtg	acc	tgt	cag	ggc	ggc	gat	tat	acc	aac	gaa				240
Ala	Leu	Asp	Ile	Ile	Val	Thr	Cys	Gln	Gly	Gly	Asp	Tyr	Thr	Asn	Glu				
					70				75						80				
atc	tat	cca	aag	ctt	cgt	gaa	agc	gga	tgg	cag	ggt	tac	tgg	att	gac				288
Ile	Tyr	Pro	Lys	Leu	Arg	Glu	Ser	Gly	Trp	Gln	Gly	Tyr	Trp	Ile	Asp				
				85				90						95					
gca	gca	tcg	tct	ctg	cgc	atg	aaa	gat	gac	gcc	atc	atc	att	ctt	gac				336
Ala	Ala	Ser	Ser	Leu	Arg	Met	Lys	Asp	Asp	Ala	Ile	Ile	Ile	Leu	Asp				
				100				105					110						
ccc	gtc	aat	cag	gac	gtc	att	acc	gac	gga	tta	aat	aat	ggc	atc	agg				384
Pro	Val	Asn	Gln	Asp	Val	Ile	Thr	Asp	Gly	Leu	Asn	Asn	Gly	Ile	Arg				
				115			120					125							
act	ttt	ggt	ggc	ggt	aac	tgt	acc	gta	agc	ctg	atg	ttg	atg	tcg	ttg				432
Thr	Phe	Val	Gly	Gly	Asn	Cys	Thr	Val	Ser	Leu	Met	Leu	Met	Ser	Leu				
				130			135				140								
ggt	ggt	tta	ttc	gcc	aat	gat	ctt	ggt	gat	tgg	gtg	tcc	ggt	gca	acc				480
Gly	Gly	Leu	Phe	Ala	Asn	Asp	Leu	Val	Asp	Trp	Val	Ser	Val	Ala	Thr				
					150				155					160					
tac	cag	gcc	gct	tcc	ggc	ggt	ggt	gcg	cga	cat	atg	cgt	gag	tta	tta				528
Tyr	Gln	Ala	Ala	Ser	Gly	Gly	Gly	Ala	Arg	His	Met	Arg	Glu	Leu	Leu				
				165				170					175						
acc	cag	atg	ggc	cat	ctg	tat	ggc	cat	gta	gca	gat	gaa	ctc	gcg	acc				576
Thr	Gln	Met	Gly	His	Leu	Tyr	Gly	His	Val	Ala	Asp	Glu	Leu	Ala	Thr				
				180				185					190						
ccg	tcc	tct	gct	att	ctc	gat	atc	gaa	cgc	aaa	gtc	aca	acc	tta	acc				624
Pro	Ser	Ser	Ala	Ile	Leu	Asp	Ile	Glu	Arg	Lys	Val	Thr	Thr	Leu	Thr				
				195			200					205							
cgt	agc	ggt	gag	ctg	ccg	gtg	gat	aac	ttt	ggc	gtg	ccg	ctg	gcg	ggt				672
Arg	Ser	Gly	Glu	Leu	Pro	Val	Asp	Asn	Phe	Gly	Val	Pro	Leu	Ala	Gly				
				210			215				220								
agc	ctg	att	ccg	tgg	atc	gac	aaa	cag	ctc	gat	aac	ggt	cag	agc	cgc				720
Ser	Leu	Ile	Pro	Trp	Ile	Asp	Lys	Gln	Leu	Asp	Asn	Gly	Gln	Ser	Arg				
				225			230				235								
gaa	gag	tgg	aaa	ggg	cag	gcg	gaa	acc	aac	aag	atc	ctc	aac	aca	tct				768
Glu	Glu	Trp	Lys	Gly	Gln	Ala	Glu	Thr	Asn	Lys	Ile	Leu	Asn	Thr	Ser				
				245					250					255					
tcc	gta	att	ccg	gta	gat	ggt	tta	tgt	gtg	cgt	gtc	ggg	gca	ttg	cgc				816
Ser	Val	Ile	Pro	Val	Asp	Gly	Leu	Cys	Val	Arg	Val	Gly	Ala	Leu	Arg				
				260				265					270						
tgc	cac	agc	cag	gca	ttc	act	att	aaa	ttg	aaa	aaa	gat	gtg	tcg	att				864
Cys	His	Ser	Gln	Ala	Phe	Thr	Ile	Lys	Leu	Lys	Lys	Asp	Val	Ser	Ile				
				275			280					285							
ccg	acc	gtg	gaa	gaa	ctg	ctg	gct	gcg	cac	aat	ccg	tgg	gcg	aaa	gtc				912
Pro	Thr	Val	Glu	Glu	Leu	Leu	Ala	Ala	His	Asn	Pro	Trp	Ala	Lys	Val				
				290			295				300								
gtt	ccg	aac	gat	cgg	gaa	atc	act	atg	cgt	gag	cta	acc	cca	gct	gcc				960
Val	Pro	Asn	Asp	Arg	Glu	Ile	Thr	Met	Arg	Glu	Leu	Thr	Pro	Ala	Ala				
				305			310				315				320				
gtt	acc	ggc	acg	ctg	acc	acg	ccg	gta	ggc	cgc	ctg	cgt	aag	ctg	aat				1008
Val	Thr	Gly	Thr	Leu	Thr	Thr	Pro	Val	Gly	Arg	Leu	Arg	Lys	Leu	Asn				

[0387]

325

330

335

atg gga cca gag ttc ctg tca gcc ttt acc gtg ggc gac cag ctg ctg

Met Gly Pro Glu Phe Leu Ser Ala Phe Thr Val Gly Asp Gln Leu Leu

340 345 350

tgg ggg gcc gcg gag ccg ctg cgt cgg atg ctt cgt caa ctg gcg taa

Trp Gly Ala Ala Glu Pro Leu Arg Arg Met Leu Arg Gln Leu Ala

355 360 365

<210> 8

<211> 367

<212> PRT

<213> 大肠杆菌

<400> 8

Met Lys Asn Val Gly Phe Ile Gly Trp Arg Gly Met Val Gly Ser Val

1 5 10 15

Leu Met Gln Arg Met Val Glu Glu Arg Asp Phe Asp Ala Ile Arg Pro

20 25 30

Val Tyr Phe Ser Thr Ser Gln Leu Gly Gln Ala Ala Pro Ser Phe Gly

35 40 45

Gly Thr Thr Gly Thr Leu Gln Asp Ala Phe Asp Leu Glu Ala Leu Lys

50 55 60

Ala Leu Asp Ile Ile Val Thr Cys Gln Gly Gly Asp Tyr Thr Asn Glu

65 70 75 80

Ile Tyr Pro Lys Leu Arg Glu Ser Gly Trp Gln Gly Tyr Trp Ile Asp

85 90 95

Ala Ala Ser Ser Leu Arg Met Lys Asp Asp Ala Ile Ile Ile Leu Asp

100 105 110

Pro Val Asn Gln Asp Val Ile Thr Asp Gly Leu Asn Asn Gly Ile Arg

115 120 125

Thr Phe Val Gly Gly Asn Cys Thr Val Ser Leu Met Leu Met Ser Leu

130 135 140

Gly Gly Leu Phe Ala Asn Asp Leu Val Asp Trp Val Ser Val Ala Thr

145 150 155 160

Tyr Gln Ala Ala Ser Gly Gly Ala Arg His Met Arg Glu Leu Leu

165 170 175

Thr Gln Met Gly His Leu Tyr Gly His Val Ala Asp Glu Leu Ala Thr

180 185 190

Pro Ser Ser Ala Ile Leu Asp Ile Glu Arg Lys Val Thr Thr Leu Thr

195 200 205

Arg Ser Gly Glu Leu Pro Val Asp Asn Phe Gly Val Pro Leu Ala Gly

210 215 220

Ser Leu Ile Pro Trp Ile Asp Lys Gln Leu Asp Asn Gly Gln Ser Arg

225 230 235 240

Glu Glu Trp Lys Gly Gln Ala Glu Thr Asn Lys Ile Leu Asn Thr Ser

245 250 255

Ser Val Ile Pro Val Asp Gly Leu Cys Val Arg Val Gly Ala Leu Arg

260 265 270

Cys His Ser Gln Ala Phe Thr Ile Lys Leu Lys Lys Asp Val Ser Ile

275 280 285

Pro Thr Val Glu Glu Leu Leu Ala Ala His Asn Pro Trp Ala Lys Val

290 295 300

Val Pro Asn Asp Arg Glu Ile Thr Met Arg Glu Leu Thr Pro Ala Ala

305 310 315 320

Val Thr Gly Thr Leu Thr Thr Pro Val Gly Arg Leu Arg Lys Leu Asn

325 330 335 340

Met Gly Pro Glu Phe Leu Ser Ala Phe Thr Val Gly Asp Gln Leu Leu

340 345 350 355

Trp Gly Ala Ala Glu Pro Leu Arg Arg Met Leu Arg Gln Leu Ala

360 365

<210> 9

<211> 1104

<212> DNA

<213> 大肠杆菌

<220>

<221> CDS

<222> (1)..(1104)

<223> 大肠杆菌必需基因 thyA

<400> 9

atg aaa aat gtt ggt ttt atc gcc tgg cgc ggt atg gtc gcc tcc gtt

Met Lys Asn Val Gly Phe Ile Gly Trp Arg Gly Met Val Gly Ser Val

5 10 15

ctc atg caa cgc atg gtt gaa gag cgc gac ttc gac gcc att cgc cct

Leu Met Gln Arg Met Val Glu Glu Arg Asp Phe Asp Ala Ile Arg Pro

20 25 30

gtc tac ttt tct act tct cag ctt ggt cag gct gcg ccg tct ttt ggc

Val Tyr Phe Ser Thr Ser Gln Leu Gly Gln Ala Ala Pro Ser Phe Gly

35 40 45

gga acc act gcc aca ctt cag gat gcc ttt gat ctg gag gcg cta aag

Gly Thr Thr Gly Thr Leu Gln Asp Ala Phe Asp Leu Glu Ala Leu Lys

50 55 60

48

96

144

192

	50	gcc	ctc	gat	atc	att	gtg	acc	tgt	cag	ggc	ggc	gat	tat	acc	aac	gaa	240
	Ala	Leu	Asp	Ile	Ile	Val	Thr	Cys	Gln	Gly	Gly	Asp	Tyr	Thr	Asn	Glu	80	
	65	atc	tat	cca	aag	ctt	cgt	gaa	agc	gga	tgg	cag	ggg	tac	tgg	att	gac	288
	Ile	Tyr	Pro	Lys	Leu	Arg	Glu	Ser	Gly	Trp	Gln	Gly	Tyr	Trp	Ile	Asp	95	
	85	gca	gca	tcg	tct	ctg	cgc	atg	aaa	gat	gac	gcc	atc	atc	att	ctt	gac	336
	Ala	Ala	Ser	Ser	Leu	Arg	Met	Lys	Asp	Asp	Ala	Ile	Ile	Ile	Leu	Asp	110	
	100	ccc	gtc	aat	cag	gac	gtc	att	acc	gac	gga	tta	aat	aat	ggc	atc	agg	384
	Pro	Val	Asn	Gln	Asp	Val	Ile	Thr	Asp	Gly	Leu	Asn	Asn	Gly	Ile	Arg	125	
	115	act	ttt	gtt	ggc	ggg	aac	tgt	acc	gta	agc	ctg	atg	ttg	atg	tcg	ttg	432
	Thr	Phe	Val	Gly	Gly	Asn	Cys	Thr	Val	Ser	Leu	Met	Leu	Met	Ser	Leu	140	
	130	ggg	ggg	tta	ttc	gcc	aat	gat	ctt	gtt	gat	tgg	gtg	tcc	gtt	gca	acc	480
	Gly	Gly	Leu	Phe	Ala	Asn	Asp	Leu	Val	Asp	Trp	Val	Ser	Val	Ala	Thr	160	
	145	tac	cag	gcc	gct	tcc	ggc	ggg	ggg	gcg	cga	cat	atg	cgt	gag	tta	tta	528
	Tyr	Gln	Ala	Ala	Ser	Gly	Gly	Gly	Ala	Arg	His	Met	Arg	Glu	Leu	Leu	175	
	165	acc	cag	atg	ggc	cat	ctg	tat	ggc	cat	gta	gca	gat	gaa	ctc	gag	acc	576
	Thr	Gln	Met	Gly	His	Leu	Tyr	Gly	His	Val	Ala	Asp	Glu	Leu	Ala	Thr	190	
	180	ccg	tcc	tct	gct	att	ctc	gat	atc	gaa	cgc	aaa	gtc	aca	acc	tta	acc	624
	Pro	Ser	Ser	Ala	Ile	Leu	Asp	Ile	Glu	Arg	Lys	Val	Thr	Thr	Leu	Thr	205	
	195	cgt	agc	ggg	gag	ctg	ccg	gtg	gat	aac	ttt	ggc	gtg	ccg	ctg	gag	ggg	672
	Arg	Ser	Gly	Glu	Leu	Pro	Val	Asp	Asn	Phe	Gly	Val	Pro	Leu	Ala	Gly	220	
	210	agc	ctg	att	ccg	tgg	atc	gac	aaa	cag	ctc	gat	aac	ggg	cag	agc	cgc	720
	Ser	Leu	Ile	Pro	Trp	Ile	Asp	Lys	Gln	Leu	Asp	Asn	Gly	Gln	Ser	Arg	240	
	225	gaa	gag	tgg	aaa	ggg	cag	gcg	gaa	acc	aac	aag	atc	ctc	aac	aca	tct	768
	Glu	Glu	Trp	Lys	Gln	Ala	Glu	Thr	Asn	Lys	Ile	Leu	Asn	Thr	Ser	255		
	245	tcc	gta	att	ccg	gta	gat	ggg	tta	tgt	gtg	cgt	gtc	ggg	gca	ttg	cgc	816
	Ser	Val	Ile	Pro	Val	Asp	Gly	Leu	Cys	Val	Arg	Val	Gly	Ala	Leu	Arg	270	
	260	tgc	cac	agc	cag	gca	ttc	act	att	aaa	ttg	aaa	aaa	gat	gtg	tcg	att	864
	Cys	His	Ser	Gln	Ala	Phe	Thr	Ile	Lys	Leu	Lys	Lys	Asp	Val	Ser	Ile	285	
	275	ccg	acc	gtg	gaa	gaa	ctg	ctg	gct	gcg	cac	aat	ccg	tgg	gag	aaa	gtc	912
	Pro	Thr	Val	Glu	Glu	Leu	Leu	Ala	Ala	His	Asn	Pro	Trp	Ala	Lys	Val	300	
	290	gtt	ccg	aac	gat	cgg	gaa	atc	act	atg	cgt	gag	cta	acc	cca	gct	gcc	960
	Val	Pro	Asn	Asp	Arg	Glu	Ile	Thr	Met	Arg	Glu	Leu	Thr	Pro	Ala	Ala	320	
	305	gtt	acc	ggc	acg	ctg	acc	acg	ccg	gta	ggc	cgc	ctg	cgt	aag	ctg	aat	1008
	Val	Thr	Gly	Thr	Leu	Thr	Thr	Pro	Val	Gly	Arg	Leu	Arg	Lys	Leu	Asn	335	
	325	atg	gga	cca	gag	ttc	ctg	tca	gcc	ttt	acc	gtg	ggc	gac	cag	ctg	ctg	1056
	Met	Gly	Pro	Glu	Phe	Leu	Ser	Ala	Phe	Thr	Val	Gly	Asp	Gln	Leu	Leu	350	
	340	tgg	ggg	gcc	gag	gag	ccg	ctg	cgt	cgg	atg	ctt	cgt	caa	ctg	gag	taa	1104
	Trp	Gly	Ala	Ala	Glu	Pro	Leu	Arg	Arg	Met	Leu	Arg	Gln	Leu	Ala		360	
	355	<210>	10															
		<211>	367															
		<212>	PRT															
		<213>	大肠杆菌															
		<400>	10															
		Met	Lys	Asn	Val	Gly	Phe	Ile	Gly	Trp	Arg	Gly	Met	Val	Gly	Ser	Val	
	1	Leu	Met	Gln	Arg	Met	Val	Glu	Glu	Arg	Asp	Phe	Asp	Ala	Ile	Arg	Pro	
		Val	Tyr	Phe	Ser	Thr	Ser	Gln	Leu	Gly	Gln	Ala	Ala	Pro	Ser	Phe	Gly	
		Gly	Thr	Thr	Gly	Thr	Leu	Gln	Asp	Ala	Phe	Asp	Leu	Glu	Ala	Leu	Lys	
	50	Ala	Leu	Asp	Ile	Ile	Val	Thr	Cys	Gln	Gly	Gly	Asp	Tyr	Thr	Asn	Glu	
	65	Ile	Tyr	Pro	Lys	Leu	Arg	Glu	Ser	Gly	Trp	Gln	Gly	Tyr	Trp	Ile	Asp	
		Ala	Ala	Ser	Ser	Leu	Arg	Met	Lys	Asp	Asp	Ala	Ile	Ile	Ile	Leu	Asp	
		Pro	Val	Asn	Gln	Asp	Val	Ile	Thr	Asp	Gly	Leu	Asn	Asn	Gly	Ile	Arg	

Thr	Phe	Val	Gly	Gly	Asn	Cys	Thr	Val	Ser	Leu	Met	Leu	Met	Ser	Leu		
Gly	130	Leu	Phe	Ala	Asn	135	Leu	Val	Asp	Trp	140	Ser	Val	Ala	Thr		
145	Gln	Ala	Ala	Ser	150	Gly	Gly	Ala	Arg	155	His	Met	Arg	Glu	Leu	160	Leu
Tyr	Gln	Met	Gly	His	165	Leu	Tyr	Gly	His	170	Val	Ala	Asp	Glu	Leu	175	Ala
Thr	Gln	Met	180	Ile	Leu	Asp	Ile	Glu	Arg	185	Lys	Val	Thr	Leu	Thr		
Pro	Ser	Ser	195	Ala	Leu	Val	200	Asn	Phe	Gly	220	Pro	Leu	Ala	Gly		
Arg	Ser	Gly	Glu	Leu	Pro	215	Asp	Gln	Leu	235	Asn	Gly	Gln	Ser	Arg		
Ser	225	Leu	Ile	Pro	Trp	230	Ile	Asp	Lys	Gln	235	Asn	Gly	Gln	Ser	Arg	
Glu	Glu	Trp	Lys	Gly	Gln	245	Ala	Glu	Thr	250	Asn	Lys	Ile	Leu	Asn	Thr	Ser
Ser	Val	Ile	Pro	Val	Asp	Gly	Leu	Cys	Val	Arg	265	Val	Gly	Ala	Leu	Arg	
Cys	His	Ser	275	Gln	Ala	Phe	Thr	Ile	Lys	Leu	Lys	Lys	Asp	Val	Ser	Ile	
Pro	Thr	Val	Glu	Glu	Leu	Leu	280	Ala	His	Asn	Pro	285	Trp	Ala	Lys	Val	
Val	Pro	Asn	Asp	Arg	Glu	Ile	Thr	Met	Arg	Glu	Leu	Thr	Pro	Ala	Ala		
305	Thr	Gly	Thr	Leu	Thr	Thr	Pro	Val	Gly	315	Arg	Leu	Arg	Lys	Leu	Asn	
Val	Thr	Gly	Thr	Leu	Thr	Thr	Pro	Val	Gly	330	Arg	Leu	Arg	Lys	Leu	Asn	
Met	Gly	Pro	Glu	Phe	Leu	Ser	Ala	Phe	Thr	Val	Gly	Asp	Gln	Leu	Leu		
Trp	Gly	Ala	Ala	Glu	Pro	Leu	Arg	Arg	Met	Leu	Arg	Gln	Leu	Ala			
		355					360					365					
<210>	11																
<211>	1842																
<212>	DNA																
<213>	大肠杆菌																
<220>																	
<221>	CDS																
<222>	(1)..(1842)																
<223>	大肠杆菌必需基因 rpoD																
<400>	11																
atg	gag	caa	aac	ccg	cag	tca	cag	ctg	aaa	ctt	ctt	gtc	acc	cgt	ggt		48
Met	Glu	Gln	Asn	Pro	Gln	Ser	Gln	Leu	Lys	Leu	Leu	Val	Thr	Arg	Gly		
1			5						10					15			
aag	gag	caa	ggc	tat	ctg	acc	tat	gcc	gag	gtc	aat	gac	cat	ctg	ccg		96
Lys	Glu	Gln	Gly	Tyr	Leu	Thr	Tyr	Ala	Glu	Val	Asn	Asp	His	Leu	Pro		
			20					25					30				
gaa	gat	atc	gtc	gat	tca	gat	cag	atc	gaa	gac	atc	atc	caa	atg	atc		144
Glu	Asp	Ile	Val	Asp	Ser	Asp	Gln	Ile	Glu	Asp	Ile	Ile	Gln	Met	Ile		
			35				40					45					
aac	gac	atg	ggc	att	cag	gtg	atg	gaa	gaa	gca	ccg	gat	gcc	gat	gat		192

[illegible]


```

        610
<210> 12
<211> 613
<212> PRT
<213> 大肠杆菌
<400> 12
Met Glu Gln Asn Pro Gln Ser Gln Leu Lys Leu Leu Val Thr Arg Gly
1      5      10      15
Lys Glu Gln Gly Tyr Leu Thr Tyr Ala Glu Val Asn Asp His Leu Pro
      20      25
Glu Asp Ile Val Asp Ser Asp Gln Ile Glu Asp Ile Ile Gln Met Ile
      35      40      45
Asn Asp Met Gly Ile Gln Val Met Glu Glu Ala Pro Asp Ala Asp Asp
      50      55      60
Leu Met Leu Ala Glu Asn Thr Ala Asp Glu Asp Ala Ala Glu Ala Ala
      65      70      75      80
Ala Gln Val Leu Ser Ser Val Glu Ser Glu Ile Gly Arg Thr Thr Asp
      85      90      95
Pro Val Arg Met Tyr Met Arg Glu Met Gly Thr Val Glu Leu Leu Thr
      100      105      110
Arg Glu Gly Glu Ile Asp Ile Ala Lys Arg Ile Glu Asp Gly Ile Asn
      115      120      125
Gln Val Gln Cys Ser Val Ala Glu Tyr Pro Glu Ala Ile Thr Tyr Leu
      130      135      140
Leu Glu Gln Tyr Asp Arg Val Glu Ala Glu Glu Ala Arg Leu Ser Asp
      145      150      155      160
Leu Ile Thr Gly Phe Val Asp Pro Asn Ala Glu Glu Asp Leu Ala Pro
      165      170      175
Thr Ala Thr His Val Gly Ser Glu Leu Ser Gln Glu Asp Leu Asp Asp
      180      185      190
Asp Glu Asp Glu Asp Glu Glu Asp Gly Asp Asp Asp Ser Ala Asp Asp
      195      200      205
Asp Asn Ser Ile Asp Pro Glu Leu Ala Arg Glu Lys Phe Ala Glu Leu
      210      215      220
Arg Ala Gln Tyr Val Val Thr Arg Asp Thr Ile Lys Ala Lys Gly Arg
      225      230      235      240
Ser His Ala Thr Ala Gln Glu Glu Ile Leu Lys Leu Ser Glu Val Phe
      245      250      255
Lys Gln Phe Arg Leu Val Pro Lys Gln Phe Asp Tyr Leu Val Asn Ser
      260      265      270
Met Arg Val Met Met Asp Arg Val Arg Thr Gln Glu Arg Leu Ile Met
      275      280      285
Lys Leu Cys Val Glu Gln Cys Lys Met Pro Lys Lys Asn Phe Ile Thr
      290      295      300
Leu Phe Thr Gly Asn Glu Thr Ser Asp Thr Trp Phe Asn Ala Ala Ile
      305      310      315      320
Ala Met Asn Lys Pro Trp Ser Glu Lys Leu His Asp Val Ser Glu Glu
      325      330      335
Val His Arg Ala Leu Gln Lys Leu Gln Gln Ile Glu Glu Thr Gly
      340      345      350
Leu Thr Ile Glu Gln Val Lys Asp Ile Asn Arg Arg Met Ser Ile Gly
      355      360      365
Glu Ala Lys Ala Arg Arg Ala Lys Lys Glu Met Val Glu Ala Asn Leu
      370      375      380
Arg Leu Val Ile Ser Ile Ala Lys Lys Tyr Thr Asn Arg Gly Leu Gln
      385      390      395      400
Phe Leu Asp Leu Ile Gln Glu Gly Asn Ile Gly Leu Met Lys Ala Val
      405      410      415
Asp Lys Phe Glu Tyr Arg Arg Gly Tyr Lys Phe Ser Thr Tyr Ala Thr
      420      425      430
Trp Trp Ile Arg Gln Ala Ile Thr Arg Ser Ile Ala Asp Gln Ala Arg
      435      440      445
Thr Ile Arg Ile Pro Val His Met Ile Glu Thr Ile Asn Lys Leu Asn
      450      455      460
Arg Ile Ser Arg Gln Met Leu Gln Glu Met Gly Arg Glu Pro Thr Pro
      465      470      475      480
Glu Glu Leu Ala Glu Arg Met Leu Met Pro Glu Asp Lys Ile Arg Lys
      485      490      495
Val Leu Lys Ile Ala Lys Glu Pro Ile Ser Met Glu Thr Pro Ile Gly
      500      505      510
Asp Asp Glu Asp Ser His Leu Gly Asp Phe Ile Glu Asp Thr Thr Leu
      515      520      525
Glu Leu Pro Leu Asp Ser Ala Thr Thr Glu Ser Leu Arg Ala Ala Thr
      530      535      540
His Asp Val Leu Ala Gly Leu Thr Ala Arg Glu Ala Lys Val Leu Arg
      545      550      555      560
Met Arg Phe Gly Ile Asp Met Asn Thr Asp His Thr Leu Glu Glu Val
      565      570      575

```

[0391]

	Gly	Lys	Gln	Phe	Asp	Val	Thr	Arg	Glu	Arg	Ile	Arg	Gln	Ile	Glu	Ala	
			580						585					590			
	Lys	Ala	Leu	Arg	Lys	Leu	Arg	His	Pro	Ser	Arg	Ser	Glu	Val	Leu	Arg	
			595					600					605				
	Ser	Phe	Leu	Asp	Asp												
		610															
	<210>	13															
	<211>	546															
	<212>	DNA															
	<213>	大肠杆菌															
	<220>																
	<221>	CDS															
	<222>	(1)..(546)															
	<223>	大肠杆菌必需基因 nusG															
	<400>	13															
	atg	tct	gaa	gct	cct	aaa	aag	cgc	tggt	tac	gtc	ggt	cag	gcg	ttt	tcc	48
	Met	Ser	Glu	Ala	Pro	Lys	Lys	Arg	Trp	Tyr	Val	Val	Gln	Ala	Phe	Ser	
	1			5					10					15			
	ggt	ttt	gaa	ggc	cgc	gta	gca	acg	tcg	ctg	cgt	gag	cat	atc	aaa	tta	96
	Gly	Phe	Glu	Gly	Arg	Val	Ala	Thr	Ser	Leu	Arg	Glu	His	Ile	Lys	Leu	
			20						25				30				
	cac	aac	atg	gaa	gat	ttg	ttt	ggg	gaa	gtc	atg	gta	cca	acc	gaa	gaa	144
	His	Asn	Met	Glu	Asp	Leu	Phe	Gly	Glu	Val	Met	Val	Pro	Thr	Glu	Glu	
			35					40					45				
	gtg	ggt	gaa	atc	cgt	ggc	ggg	cag	cgt	cgc	aaa	agc	gaa	cgt	aaa	ttc	192
	Val	Val	Glu	Ile	Arg	Gly	Gly	Gln	Arg	Arg	Lys	Ser	Glu	Arg	Lys	Phe	
		50					55					60					
	ttc	cct	ggc	tac	gtc	ctc	ggt	cag	atg	gtg	atg	aac	gac	gcg	agc	tggt	240
	Phe	Pro	Gly	Tyr	Val	Leu	Val	Gln	Met	Val	Met	Asn	Asp	Ala	Ser	Trp	
		65				70				75					80		
	cac	ctg	gtg	cgc	agc	gta	ccg	cgt	gtg	atg	ggc	ttc	atc	ggc	ggg	act	288
	His	Leu	Val	Arg	Ser	Val	Pro	Arg	Val	Met	Gly	Phe	Ile	Gly	Gly	Thr	
				85					90					95			
	tcc	gat	cgt	cct	gcg	cca	atc	agc	gat	aaa	gaa	gtc	gat	gcg	att	atg	336
	Ser	Asp	Arg	Pro	Ala	Pro	Ile	Ser	Asp	Lys	Glu	Val	Asp	Ala	Ile	Met	
				100					105					110			
	aac	cgc	ctg	cag	cag	ggt	ggg	gat	aag	ccg	cgt	ccg	aaa	acg	ctg	ttt	384
	Asn	Arg	Leu	Gln	Gln	Val	Gly	Asp	Lys	Pro	Arg	Pro	Lys	Thr	Leu	Phe	
			115					120					125				
[0392]	gaa	ccg	ggg	gaa	atg	gtc	cgt	ggt	aag	gat	ggg	ccg	ttc	gct	gac	ttc	432
	Glu	Pro	Gly	Glu	Met	Val	Arg	Val	Asn	Asp	Gly	Pro	Phe	Ala	Asp	Phe	
		130					135					140					
	aac	ggg	gtt	gtt	gaa	gaa	gtg	gat	tac	gag	aaa	tct	cgt	ctg	aaa	gtg	480
	Asn	Gly	Val	Val	Glu	Glu	Val	Asp	Tyr	Glu	Lys	Ser	Arg	Leu	Lys	Val	
		145				150				155					160		
	tct	gtt	tct	atc	ttc	ggg	cgt	gcg	acc	ccg	gta	gag	ctg	gac	ttc	agc	528
	Ser	Val	Ser	Ile	Phe	Gly	Arg	Ala	Thr	Pro	Val	Glu	Leu	Asp	Phe	Ser	
				165						170				175			
	cag	gtt	gaa	aaa	gcc	taa											546
	Gln	Val	Glu	Lys	Ala												
			180														
	<210>	14															
	<211>	181															
	<212>	PRT															
	<213>	大肠杆菌															
	<400>	14															
	Met	Ser	Glu	Ala	Pro	Lys	Lys	Arg	Trp	Tyr	Val	Val	Gln	Ala	Phe	Ser	
	1			5					10					15			
	Gly	Phe	Glu	Gly	Arg	Val	Ala	Thr	Ser	Leu	Arg	Glu	His	Ile	Lys	Leu	
			20						25				30				
	His	Asn	Met	Glu	Asp	Leu	Phe	Gly	Glu	Val	Met	Val	Pro	Thr	Glu	Glu	
			35					40					45				
	Val	Val	Glu	Ile	Arg	Gly	Gly	Gln	Arg	Arg	Lys	Ser	Glu	Arg	Lys	Phe	
		50					55					60					
	Phe	Pro	Gly	Tyr	Val	Leu	Val	Gln	Met	Val	Met	Asn	Asp	Ala	Ser	Trp	
		65				70				75					80		
	His	Leu	Val	Arg	Ser	Val	Pro	Arg	Val	Met	Gly	Phe	Ile	Gly	Gly	Thr	
				85					90					95			
	Ser	Asp	Arg	Pro	Ala	Pro	Ile	Ser	Asp	Lys	Glu	Val	Asp	Ala	Ile	Met	
				100					105					110			
	Asn	Arg	Leu	Gln	Gln	Val	Gly	Asp	Lys	Pro	Arg	Pro	Lys	Thr	Leu	Phe	
			115					120					125				
	Glu	Pro	Gly	Glu	Met	Val	Arg	Val	Asn	Asp	Gly	Pro	Phe	Ala	Asp	Phe	
		130					135					140					
	Asn	Gly	Val	Val	Glu	Glu	Val	Asp	Tyr	Glu	Lys	Ser	Arg	Leu	Lys	Val	
		145				150				155					160		
	Ser	Val	Ser	Ile	Phe	Gly	Arg	Ala	Thr	Pro	Val	Glu	Leu	Asp	Phe	Ser	
				165						170					175		

Gln Val Glu Lys Ala
180

<210> 15
<211> 216
<212> DNA
<213> 大肠杆菌
<220>
<221> CDS
<222> (1)..(216)
<223> 大肠杆菌必需基因 rpsU
<400> 15

atg ccg gta att aaa gta cgt gaa aac gag ccg ttc gac gta gct ctg 48
Met Pro Val Ile Lys Val Arg Glu Asn Glu Pro Phe Asp Val Ala Leu
1 5 10 15
cgt cgc ttc aag cgt tcc tgc gaa aaa gca ggt gtt ctg gcg gaa gtt 96
Arg Arg Phe Lys Arg Ser Cys Glu Lys Ala Gly Val Leu Ala Glu Val
20 25 30
cgt cgt cgt gag ttc tat gaa aaa ccg act acc gaa cgt aag cgc gct 144
Arg Arg Arg Glu Phe Tyr Glu Lys Pro Thr Thr Glu Arg Lys Arg Ala
35 40 45
aaa gct tct gca gtg aaa cgt cac gcg aag aaa ctg gct cgc gaa aac 192
Lys Ala Ser Ala Val Lys Arg His Ala Lys Lys Leu Ala Arg Glu Asn
50 55 60
gca cgc cgc act cgt ctg tac taa 216
Ala Arg Arg Thr Arg Leu Tyr
65 70

<210> 16
<211> 71
<212> PRT
<213> 大肠杆菌
<400> 16

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

[0393]

<210> 17
<211> 216
<212> DNA
<213> 大肠杆菌
<220>
<221> CDS
<222> (1)..(216)
<223> 大肠杆菌必需基因 accD
<400> 17

atg ccg gta att aaa gta cgt gaa aac gag ccg ttc gac gta gct ctg 48
Met Pro Val Ile Lys Val Arg Glu Asn Glu Pro Phe Asp Val Ala Leu
1 5 10 15
cgt cgc ttc aag cgt tcc tgc gaa aaa gca ggt gtt ctg gcg gaa gtt 96
Arg Arg Phe Lys Arg Ser Cys Glu Lys Ala Gly Val Leu Ala Glu Val
20 25 30
cgt cgt cgt gag ttc tat gaa aaa ccg act acc gaa cgt aag cgc gct 144
Arg Arg Arg Glu Phe Tyr Glu Lys Pro Thr Thr Glu Arg Lys Arg Ala
35 40 45
aaa gct tct gca gtg aaa cgt cac gcg aag aaa ctg gct cgc gaa aac 192
Lys Ala Ser Ala Val Lys Arg His Ala Lys Lys Leu Ala Arg Glu Asn
50 55 60
gca cgc cgc act cgt ctg tac taa 216
Ala Arg Arg Thr Arg Leu Tyr
65 70

<210> 18
<211> 71
<212> PRT
<213> 大肠杆菌
<400> 18

Met Pro Val Ile Lys Val Arg Glu Asn Glu Pro Phe Asp Val Ala Leu
1 5 10 15
Arg Arg Phe Lys Arg Ser Cys Glu Lys Ala Gly Val Leu Ala Glu Val
20 25 30
Arg Arg Arg Glu Phe Tyr Glu Lys Pro Thr Thr Glu Arg Lys Arg Ala
35 40 45
Lys Ala Ser Ala Val Lys Arg His Ala Lys Lys Leu Ala Arg Glu Asn

	50		55		60	
	Ala Arg Arg Thr Arg Leu Tyr					
	65		70			
	<210> 19					
	<211> 216					
	<212> DNA					
	<213> 大肠杆菌					
	<220>					
	<221> CDS					
	<222> (1)..(216)					
	<223> 大肠杆菌必需基因 degS					
	<400> 19					
	atg ccg gta att aaa gta cgt gaa aac gag ccg ttc gac gta gct ctg					48
	Met Pro Val Ile Lys Val Arg Glu Asn Glu Pro Phe Asp Val Ala Leu					
	1	5	10	15		
	cgt cgc ttc aag cgt tcc tgc gaa aaa gca ggt gtt ctg gcg gaa gtt					96
	Arg Arg Phe Lys Arg Ser Cys Glu Lys Ala Gly Val Leu Ala Glu Val					
		20	25	30		
	cgt cgt cgt gag ttc tat gaa aaa ccg act acc gaa cgt aag cgc gct					144
	Arg Arg Arg Glu Phe Tyr Glu Lys Pro Thr Thr Glu Arg Lys Arg Ala					
		35	40	45		
	aaa gct tct gca gtg aaa cgt cac gcg aag aaa ctg gct cgc gaa aac					192
	Lys Ala Ser Ala Val Lys Arg His Ala Lys Lys Leu Ala Arg Glu Asn					
		50	55	60		
	gca cgc cgc act cgt ctg tac taa					216
	Ala Arg Arg Thr Arg Leu Tyr					
	65	70				
	<210> 20					
	<211> 71					
	<212> PRT					
	<213> 大肠杆菌					
	<400> 20					
	Met Pro Val Ile Lys Val Arg Glu Asn Glu Pro Phe Asp Val Ala Leu					
	1	5	10	15		
	Arg Arg Phe Lys Arg Ser Cys Glu Lys Ala Gly Val Leu Ala Glu Val					
		20	25	30		
	Arg Arg Arg Glu Phe Tyr Glu Lys Pro Thr Thr Glu Arg Lys Arg Ala					
		35	40	45		
[0394]	Lys Ala Ser Ala Val Lys Arg His Ala Lys Lys Leu Ala Arg Glu Asn					
		50	55	60		
	Ala Arg Arg Thr Arg Leu Tyr					
	65	70				
	<210> 21					
	<211> 531					
	<212> DNA					
	<213> 大肠杆菌					
	<220>					
	<221> CDS					
	<222> (1)..(531)					
	<223> 大肠杆菌必需基因 fldA					
	<400> 21					
	atg gct atc act ggc atc ttt ttc ggc agc gac acc ggt aat acc gaa					48
	Met Ala Ile Thr Gly Ile Phe Phe Gly Ser Asp Thr Gly Asn Thr Glu					
	1	5	10	15		
	aat atc gca aaa atg att caa aaa cag ctt ggt aaa gac gtt gcc gat					96
	Asn Ile Ala Lys Met Ile Gln Lys Gln Leu Gly Lys Asp Val Ala Asp					
		20	25	30		
	gtc cat gac att gca aaa agc agc aaa gaa gat ctg gaa gct tat gac					144
	Val His Asp Ile Ala Lys Ser Ser Lys Glu Asp Leu Glu Ala Tyr Asp					
		35	40	45		
	att ctg ctg ctg ggc atc cca acc tgg tat tac ggc gaa gcg cag tgt					192
	Ile Leu Leu Leu Gly Ile Pro Thr Trp Tyr Tyr Gly Glu Ala Gln Cys					
		50	55	60		
	gac tgg gat gac ttc ttc ccg act ctc gaa gag att gat ttc aat ggc					240
	Asp Trp Asp Asp Phe Phe Pro Thr Leu Glu Glu Ile Asp Phe Asn Gly					
	65	70	75	80		
	aaa ctg gtt gcg ctg ttt ggt tgt ggt gac cag gaa gat tac gcc gaa					288
	Lys Leu Val Ala Leu Phe Gly Cys Gly Asp Gln Glu Asp Tyr Ala Glu					
		85	90	95		
	tat ttc tgc gac gca ttg ggc acc atc cgc gac atc att gaa ccg cgc					336
	Tyr Phe Cys Asp Ala Leu Gly Thr Ile Arg Asp Ile Ile Glu Pro Arg					
		100	105	110		
	ggt gca acc atc gtt ggt cac tgg cca act gcg ggc tat cat ttc gaa					384
	Gly Ala Thr Ile Val Gly His Trp Pro Thr Ala Gly Tyr His Phe Glu					
		115	120	125		
	gca tca aaa ggt ctg gca gat gac gac cac ttt gtc ggt ctg gct atc					432
	Ala Ser Lys Gly Leu Ala Asp Asp Asp His Phe Val Gly Leu Ala Ile					

63

	Gln	Ser	Trp	Gln	Gln	Gln	Thr	Arg	Thr	Ser	Gln	Ala	Ala	Pro	Val	Gln	
	gcg	caa	ccg	cgc	cag	tct	aaa	ccg	gct	tct	tct	cag	cag	ccg	tac	cag	624
	Ala	Gln	Pro	Arg	Gln	Ser	Lys	Pro	Ala	Ser	Ser	Gln	Gln	Pro	Tyr	Gln	
			195					200				205					
	gat	ctg	ctg	caa	act	cct	gcg	cac	acg	act	gcg	cag	tcg	aag	ccg	cag	672
	Asp	Leu	Leu	Gln	Thr	Pro	215	His	Thr	Thr	Ala	Gln	Ser	Lys	Pro	Gln	
		210									220						
	caa	gct	gcg	cca	gtt	gct	cgt	gcc	gct	gac	gcg	cca	aaa	ccg	acg	gcg	720
	Gln	Ala	Ala	Pro	Val	Ala	Arg	Ala	Ala	Asp	Ala	Pro	Lys	Pro	Thr	Ala	
	225					230					235					240	
	gag	aaa	aaa	gac	gaa	cgc	cgc	tggt	atg	gtg	cag	tcg	ggt	tcg	ttc	aga	768
	Glu	Lys	Lys	Asp	Glu	Arg	Arg	Trp	Met	Val	Gln	Cys	Gly	Ser	Phe	Arg	
					245					250					255		
	ggc	gcg	gaa	cag	gca	gag	acg	gta	cgt	gct	cag	ctg	gcg	ttc	gaa	ggc	816
	Gly	Ala	Glu	Gln	Ala	Glu	Thr	Val	Arg	Ala	Gln	Leu	Ala	Phe	Glu	Gly	
				260					265					270			
	ttt	gac	tcg	aaa	atc	acc	acc	aac	aat	ggc	tggt	aat	cgt	gtg	gtc	att	864
	Phe	Asp	Ser	Lys	Ile	Thr	Thr	Asn	Asn	Gly	Trp	Asn	Arg	Val	Val	Ile	
			275					280					285				
	ggc	ccg	gtg	aaa	ggc	aaa	gag	aac	gca	gac	agc	acc	ctc	aat	cgg	ttg	912
	Gly	Pro	Val	Lys	Gly	Lys	Glu	Asn	Ala	Asp	Ser	Thr	Leu	Asn	Arg	Leu	
		290					295					300					
	aag	atg	gcg	ggc	cat	aca	aac	tgc	att	cgg	ctc	gcc	gcc	ggg	ggt	tga	960
	Lys	Met	Ala	Gly	His	Thr	Asn	Cys	Ile	Arg	Leu	Ala	Ala	Gly	Gly		
	305					310					315						
	<210>	24															
	<211>	319															
	<212>	PRT															
	<213>	大肠杆菌															
	<400>	24															
	Val	Ala	Gln	Arg	Asp	Tyr	Val	Arg	Arg	Ser	Gln	Pro	Ala	Pro	Ser	Arg	
	1			5						10					15		
	Arg	Lys	Lys	Ser	Thr	Ser	Arg	Lys	Lys	Gln	Arg	Asn	Leu	Pro	Ala	Val	
			20						25				30				
	Ser	Pro	Ala	Met	Val	Ala	Ile	Ala	Ala	Ala	Val	Leu	Val	Thr	Phe	Ile	
			35				40					45					
	Gly	Gly	Leu	Tyr	Phe	Ile	Thr	His	His	Lys	Lys	Glu	Glu	Ser	Glu	Thr	
		50					55					60					
	Leu	Gln	Ser	Gln	Lys	Val	Thr	Gly	Asn	Gly	Leu	Pro	Pro	Lys	Pro	Glu	
	65					70					75					80	
	Glu	Arg	Trp	Arg	Tyr	Ile	Lys	Glu	Leu	Glu	Ser	Arg	Gln	Pro	Gly	Val	
				85						90				95			
	Arg	Ala	Pro	Thr	Glu	Pro	Ser	Ala	Gly	Gly	Glu	Val	Lys	Thr	Pro	Glu	
			100						105					110			
	Gln	Leu	Thr	Pro	Glu	Gln	Arg	Gln	Leu	Leu	Glu	Gln	Met	Gln	Ala	Asp	
			115					120					125				
	Met	Arg	Gln	Gln	Pro	Thr	Gln	Leu	Val	Glu	Val	Pro	Trp	Asn	Glu	Gln	
		130					135					140					
	Thr	Pro	Glu	Gln	Arg	Gln	Thr	Leu	Gln	Arg	Gln	Arg	Gln	Ala	Gln		
	145					150				155					160		
	Gln	Leu	Ala	Glu	Gln	Arg	Leu	Ala	Gln	Gln	Ser	Arg	Thr	Thr	Glu		
			165							170				175			
	Gln	Ser	Trp	Gln	Gln	Gln	Thr	Arg	Thr	Ser	Gln	Ala	Ala	Pro	Val	Gln	
			180						185					190			
	Ala	Gln	Pro	Arg	Gln	Ser	Lys	Pro	Ala	Ser	Ser	Gln	Gln	Pro	Tyr	Gln	
			195					200					205				
	Asp	Leu	Leu	Gln	Thr	Pro	Ala	His	Thr	Thr	Ala	Gln	Ser	Lys	Pro	Gln	
		210					215					220					
	Gln	Ala	Ala	Pro	Val	Ala	Arg	Ala	Ala	Asp	Ala	Pro	Lys	Pro	Thr	Ala	
	225					230					235				240		
	Glu	Lys	Lys	Asp	Glu	Arg	Arg	Trp	Met	Val	Gln	Cys	Gly	Ser	Phe	Arg	
					245					250					255		
	Gly	Ala	Glu	Gln	Ala	Glu	Thr	Val	Arg	Ala	Gln	Leu	Ala	Phe	Glu	Gly	
			260						265					270			
	Phe	Asp	Ser	Lys	Ile	Thr	Thr	Asn	Asn	Gly	Trp	Asn	Arg	Val	Val	Ile	
			275					280					285				
	Gly	Pro	Val	Lys	Gly	Lys	Glu	Asn	Ala	Asp	Ser	Thr	Leu	Asn	Arg	Leu	
		290					295					300					
	Lys	Met	Ala	Gly	His	Thr	Asn	Cys	Ile	Arg	Leu	Ala	Ala	Gly	Gly		
	305					310					315						
	<210>	25															
	<211>	1935															
	<212>	DNA															
	<213>	大肠杆菌															
	<220>																
	<221>	CDS															
	<222>	(1) .. (1935)															

[0396]

	<223> 大肠杆菌必需基因 hflB																
	<400> 25																
	atg	gcg	aaa	aac	cta	ata	ctc	tgg	ctg	gtc	att	gcc	gtt	gtg	ctg	atg	48
	Met	Ala	Lys	Asn	Leu	Ile	Leu	Trp	Leu	Val	Ile	Ala	Val	Val	Leu	Met	
	1				5					10					15		
	tca	gta	ttc	cag	agc	ttt	ggg	ccc	agc	gag	tct	aat	ggc	cgt	aag	gtg	96
	Ser	Val	Phe	Gln	Ser	Phe	Gly	Pro	Ser	Glu	Ser	Asn	Gly	Arg	Lys	Val	
				20					25					30			
	gat	tac	tct	acc	ttc	cta	caa	gag	gtc	aat	aac	gac	cag	gtt	cgt	gaa	144
	Asp	Tyr	Ser	Thr	Phe	Leu	Gln	Glu	Val	Asn	Asn	Asp	Gln	Val	Arg	Glu	
			35					40					45				
	gcg	cgt	atc	aac	gga	cgt	gaa	atc	aac	gtt	acc	aag	aaa	gat	agt	aac	192
	Ala	Arg	Ile	Asn	Gly	Arg	Glu	Ile	Asn	Val	Thr	Lys	Lys	Asp	Ser	Asn	
		50					55					60					
	cgt	tat	acc	act	tac	att	ccg	gtt	cag	gat	ccg	aaa	tta	ctg	gat	aac	240
	Arg	Tyr	Thr	Thr	Tyr	Ile	Pro	Val	Gln	Asp	Pro	Lys	Leu	Leu	Asp	Asn	
		65				70					75					80	
	ctg	ttg	acc	aag	aac	gtc	aag	gtt	gtc	ggg	gaa	ccg	cct	gaa	gaa	cca	288
	Leu	Leu	Thr	Lys	Asn	Val	Lys	Val	Val	Gly	Glu	Pro	Pro	Glu	Glu	Pro	
				85						90					95		
	agc	ctg	ctg	gct	tct	atc	ttc	atc	tcc	tgg	ttc	ccg	atg	ctg	ttg	ctg	336
	Ser	Leu	Leu	Ala	Ser	Ile	Phe	Ile	Ser	Trp	Phe	Pro	Met	Leu	Leu	Leu	
				100					105					110			
	att	ggg	gtc	tgg	atc	ttc	ttc	atg	cgt	caa	atg	cag	ggc	ggc	ggg	ggc	384
	Ile	Gly	Val	Trp	Ile	Phe	Phe	Met	Arg	Gln	Met	Gln	Gly	Gly	Gly	Gly	
			115					120					125				
	aaa	ggg	gcc	atg	tcg	ttt	ggg	aag	agc	aaa	gcg	cgc	atg	ctg	acg	gaa	432
	Lys	Gly	Ala	Met	Ser	Phe	Gly	Lys	Ser	Lys	Ala	Arg	Met	Leu	Thr	Glu	
		130					135					140					
	gat	cag	atc	aaa	acg	acc	ttt	gct	gac	gtt	gcg	ggc	tgc	gac	gaa	gca	480
	Asp	Gln	Ile	Lys	Thr	Thr	Phe	Ala	Asp	Val	Ala	Gly	Cys	Asp	Glu	Ala	
		145				150					155				160		
	aaa	gaa	gaa	gtt	gct	gaa	ctg	gtt	gag	tat	ctg	cgc	gag	ccg	agc	cgc	528
	Lys	Glu	Glu	Val	Ala	Glu	Leu	Val	Glu	Tyr	Leu	Arg	Glu	Pro	Ser	Arg	
				165						170					175		
	ttc	cag	aaa	ctc	ggc	ggg	aag	atc	ccg	aaa	ggc	gtc	ttg	atg	gtc	ggg	576
	Phe	Gln	Lys	Leu	Gly	Gly	Lys	Ile	Pro	Lys	Gly	Val	Leu	Met	Val	Gly	
				180					185					190			
[0397]	cct	ccg	ggg	acc	ggg	aaa	acg	ctg	ctg	gcg	aaa	gcg	att	gca	ggc	gaa	624
	Pro	Pro	Gly	Thr	Gly	Lys	Thr	Leu	Leu	Ala	Lys	Ala	Ile	Ala	Gly	Glu	
			195					200					205				
	gcg	aaa	gtt	ccg	ttc	ttt	act	atc	tcc	ggg	tct	gac	ttc	gta	gaa	atg	672
	Ala	Lys	Val	Pro	Phe	Phe	Thr	Ile	Ser	Gly	Ser	Asp	Phe	Val	Glu	Met	
		210					215					220					
	ttc	gtc	ggg	gtg	ggg	gca	tcc	cgt	gtt	cgt	gac	atg	ttc	gaa	cag	gcg	720
	Phe	Val	Gly	Val	Gly	Ala	Ser	Arg	Val	Arg	Asp	Met	Phe	Glu	Gln	Ala	
		225				230					235				240		
	aag	aaa	gcg	gca	ccg	tgc	atc	atc	ttt	atc	gat	gaa	atc	gac	gcc	gta	768
	Lys	Lys	Ala	Ala	Pro	Cys	Ile	Ile	Phe	Ile	Asp	Glu	Ile	Asp	Ala	Val	
				245						250					255		
	ggc	cgc	cag	cgt	ggc	gct	ggg	ctg	ggc	ggg	ggg	cac	gat	gaa	cgt	gaa	816
	Gly	Arg	Gln	Arg	Gly	Ala	Gly	Leu	Gly	Gly	Gly	His	Asp	Glu	Arg	Glu	
				260				265						270			
	cag	act	ctg	aac	cag	atg	ctg	gtt	gag	atg	gat	ggc	ttc	gaa	ggg	aac	864
	Gln	Thr	Leu	Asn	Gln	Met	Leu	Val	Glu	Met	Asp	Gly	Phe	Glu	Gly	Asn	
			275					280					285				
	gaa	ggg	atc	atc	gtt	atc	gcc	gcg	act	aac	cgt	ccg	gac	gtt	ctc	gac	912
	Glu	Gly	Ile	Ile	Val	Ile	Ala	Ala	Thr	Asn	Arg	Pro	Asp	Val	Leu	Asp	
		290					295					300					
	ccg	gcc	ctg	ctg	cgt	cct	ggc	cgt	ttc	gac	cgt	cag	gtt	gtg	gtc	ggc	960
	Pro	Ala	Leu	Leu	Arg	Pro	Gly	Arg	Phe	Asp	Arg	Gln	Val	Val	Val	Gly	
		305				310					315					320	
	ttg	cca	gat	gtt	cgc	ggg	cgt	gag	cag	atc	ctg	aaa	gtt	cac	atg	cgt	1008
	Leu	Pro	Asp	Val	Arg	Gly	Arg	Glu	Gln	Ile	Leu	Lys	Val	His	Met	Arg	
				325						330					335		
	cgc	gta	cca	ttg	gca	ccc	gat	atc	gac	gcg	gca	atc	att	gcc	cgt	ggg	1056
	Arg	Val	Pro	Leu	Ala	Pro	Asp	Ile	Asp	Ala	Ala	Ile	Ile	Ala	Arg	Gly	
				340					345					350			
	act	cct	ggg	ttc	tcc	ggg	gct	gac	ctg	gcg	aac	ctg	gtg	aac	gaa	gcg	1104
	Thr	Pro	Gly	Phe	Ser	Gly	Ala	Asp	Leu	Ala	Asn	Leu	Val	Asn	Glu	Ala	
			355				360						365				
	gca	ctg	ttc	gct	gct	cgt	ggc	aac	aaa	cgc	gtt	gtg	tcg	atg	gtt	gag	1152
	Ala	Leu	Phe	Ala	Ala	Arg	Gly	Asn	Lys	Arg	Val	Val	Ser	Met	Val	Glu	
		370					375					380					
	ttc	gag	aaa	gcg	aaa	gac	aaa	atc	atg	atg	ggg	gcg	gaa	cgt	cgc	tcc	1200
	Phe	Glu	Lys	Ala	Lys	Asp	Lys	Ile	Met	Met	Gly	Ala	Glu	Arg	Arg	Ser	
		385				390					395					400	
	atg	gtg	atg	acg	gaa	gcg	cag	aaa	gaa	tcg	acg	gct	tac	cac	gaa	gcg	1248

	Met	Val	Met	Thr	Glu	Ala	Gln	Lys	Glu	Ser	Thr	Ala	Tyr	His	Glu	Ala	
	ggt	cat	gcg	att	atc	ggt	cg	ctg	gtg	ccg	gaa	cac	gat	ccg	gtg	cac	1296
	Gly	His	Ala	Ile	Ile	Gly	Arg	Leu	Val	Pro	Glu	His	Asp	Pro	Val	His	
				420					425					430			
	aaa	gtg	acg	att	atc	cca	cg	ggt	cg	cg	ctg	ggt	gtg	act	ttc	ttc	1344
	Lys	Val	Thr	Ile	Ile	Pro	Arg	Gly	Arg	Ala	Leu	Gly	Val	Thr	Phe	Phe	
				435					440					445			
	ttg	cct	gag	ggc	gac	gca	atc	agc	gcc	agc	cg	cag	aaa	ctg	gaa	agc	1392
	Leu	Pro	Glu	Gly	Asp	Ala	Ile	Ser	Ala	Ser	Arg	Gln	Lys	Leu	Glu	Ser	
		450							455					460			
	cag	att	tct	acg	ctg	tac	ggt	ggt	cg	ctg	gca	gaa	gag	atc	atc	tac	1440
	Gln	Ile	Ser	Thr	Leu	Tyr	Gly	Gly	Arg	Leu	Ala	Glu	Glu	Ile	Ile	Tyr	
		465														480	
	ggg	ccg	gaa	cat	gta	tct	acc	ggt	gcg	tcc	aac	gat	att	aaa	ggt	gcg	1488
	Gly	Pro	Glu	His	Val	Ser	Thr	Gly	Ala	Ser	Asn	Asp	Ile	Lys	Val	Ala	
					485											495	
	acc	aac	ctg	gca	cg	aac	atg	gtg	act	cag	tgg	ggc	ttc	tct	gag	aaa	1536
	Thr	Asn	Leu	Ala	Arg	Asn	Met	Val	Thr	Gln	Trp	Gly	Phe	Ser	Glu	Lys	
				500										510			
	ttg	ggt	cca	ctg	ctg	tac	gcg	gaa	gaa	gaa	ggt	gaa	gtg	ttc	ctc	ggc	1584
	Leu	Gly	Pro	Leu	Leu	Tyr	Ala	Glu	Glu	Glu	Gly	Glu	Val	Phe	Leu	Gly	
			515						520					525			
	cg	agc	gta	gcg	aaa	gcg	aaa	cat	atg	tcc	gat	gaa	act	gca	cg	atc	1632
	Arg	Ser	Val	Ala	Lys	Ala	Lys	His	Met	Ser	Asp	Glu	Thr	Ala	Arg	Ile	
		530							535					540			
	atc	gac	cag	gaa	gtg	aaa	gca	ctg	att	gag	cg	aac	tat	aat	cg	gcg	1680
	Ile	Asp	Gln	Glu	Val	Lys	Ala	Leu	Ile	Glu	Arg	Asn	Tyr	Asn	Arg	Ala	
		545														560	
	cg	cag	ctt	ctg	acc	gac	aat	atg	gat	att	ctg	cat	gcg	atg	aaa	gat	1728
	Arg	Gln	Leu	Leu	Thr	Asp	Asn	Met	Asp	Ile	Leu	His	Ala	Met	Lys	Asp	
					565											575	
	gct	ctc	atg	aaa	tat	gag	act	atc	gac	gca	ccg	cag	att	gat	gac	ctg	1776
	Ala	Leu	Met	Lys	Tyr	Glu	Thr	Ile	Asp	Ala	Pro	Gln	Ile	Asp	Asp	Leu	
				580										590			
	atg	gca	cg	cg	gat	gta	cg	ccg	cca	gcg	ggc	tgg	gaa	gaa	cca	ggc	1824
	Met	Ala	Arg	Arg	Asp	Val	Arg	Pro	Pro	Ala	Gly	Trp	Glu	Glu	Pro	Gly	
			595						600					605			
[0398]	gct	tct	aac	aat	tct	ggc	gac	aat	ggt	agt	cca	aag	gct	cct	cg	ccg	1872
	Ala	Ser	Asn	Asn	Ser	Gly	Asn	Asn	Gly	Ser	Pro	Lys	Ala	Pro	Arg	Pro	
		610							615					620			
	ggt	gat	gaa	ccg	cg	acg	ccg	aac	ccg	ggt	aac	acc	atg	tca	gag	cag	1920
	Val	Asp	Glu	Pro	Arg	Thr	Pro	Asn	Pro	Gly	Asn	Thr	Met	Ser	Glu	Gln	
		625				630					635					640	
	tta	ggc	gac	aag	taa												1935
	Leu	Gly	Asp	Lys													
	<210>	26															
	<211>	644															
	<212>	PRT															
	<213>	大肠杆菌															
	<400>	26															
	Met	Ala	Lys	Asn	Leu	Ile	Leu	Trp	Leu	Val	Ile	Ala	Val	Val	Leu	Met	
	1				5					10					15		
	Ser	Val	Phe	Gln	Ser	Phe	Gly	Pro	Ser	Glu	Ser	Asn	Gly	Arg	Lys	Val	
				20					25					30			
	Asp	Tyr	Ser	Thr	Phe	Leu	Gln	Glu	Val	Asn	Asn	Asp	Gln	Val	Arg	Glu	
			35					40					45				
	Ala	Arg	Ile	Asn	Gly	Arg	Glu	Ile	Asn	Val	Thr	Lys	Lys	Asp	Ser	Asn	
		50					55					60					
	Arg	Tyr	Thr	Thr	Tyr	Ile	Pro	Val	Gln	Asp	Pro	Lys	Leu	Leu	Asp	Asn	
		65			70					75					80		
	Leu	Leu	Thr	Lys	Asn	Val	Lys	Val	Val	Gly	Glu	Pro	Pro	Glu	Glu	Pro	
				85						90					95		
	Ser	Leu	Leu	Ala	Ser	Ile	Phe	Ile	Ser	Trp	Phe	Pro	Met	Leu	Leu	Leu	
			100						105					110			
	Ile	Gly	Val	Trp	Ile	Phe	Phe	Met	Arg	Gln	Met	Gln	Gly	Gly	Gly	Gly	
			115					120					125				
	Lys	Gly	Ala	Met	Ser	Phe	Gly	Lys	Ser	Lys	Ala	Arg	Met	Leu	Thr	Glu	
		130					135					140					
	Asp	Gln	Ile	Lys	Thr	Thr	Phe	Ala	Asp	Val	Ala	Gly	Cys	Asp	Glu	Ala	
		145				150					155				160		
	Lys	Glu	Glu	Val	Ala	Glu	Leu	Val	Glu	Tyr	Leu	Arg	Glu	Pro	Ser	Arg	
				165						170					175		
	Phe	Gln	Lys	Leu	Gly	Gly	Lys	Ile	Pro	Lys	Gly	Val	Leu	Met	Val	Gly	
				180					185					190			
	Pro	Pro	Gly	Thr	Gly	Lys	Thr	Leu	Leu	Ala	Lys	Ala	Ile	Ala	Gly	Glu	
			195					200					205				

[0399]

Ala Lys Val Pro Phe Phe Thr Ile Ser Gly Ser Asp Phe Val Glu Met
 210 215 220
 Phe Val Gly Val Gly Ala Ser Arg Val Arg Asp Met Phe Glu Gln Ala
 225 230 235 240
 Lys Lys Ala Ala Pro Cys Ile Ile Phe Ile Asp Glu Ile Asp Ala Val
 245 250 255
 Gly Arg Gln Arg Gly Ala Gly Leu Gly Gly Gly His Asp Glu Arg Glu
 260 265 270
 Gln Thr Leu Asn Gln Met Leu Val Glu Met Asp Gly Phe Glu Gly Asn
 275 280 285
 Glu Gly Ile Ile Val Ile Ala Ala Thr Asn Arg Pro Asp Val Leu Asp
 290 295 300
 Pro Ala Leu Leu Arg Pro Gly Arg Phe Asp Arg Gln Val Val Val Gly
 305 310 315 320
 Leu Pro Asp Val Arg Gly Arg Glu Gln Ile Leu Lys Val His Met Arg
 325 330 335
 Arg Val Pro Leu Ala Pro Asp Ile Asp Ala Ala Ile Ile Ala Arg Gly
 340 345 350
 Thr Pro Gly Phe Ser Gly Ala Asp Leu Ala Asn Leu Val Asn Glu Ala
 355 360 365
 Ala Leu Phe Ala Ala Arg Gly Asn Lys Arg Val Val Ser Met Val Glu
 370 375 380
 Phe Glu Lys Ala Lys Asp Lys Ile Met Met Gly Ala Glu Arg Arg Ser
 385 390 395 400
 Met Val Met Thr Glu Ala Gln Lys Glu Ser Thr Ala Tyr His Glu Ala
 405 410 415
 Gly His Ala Ile Ile Gly Arg Leu Val Pro Glu His Asp Pro Val His
 420 425 430
 Lys Val Thr Ile Ile Pro Arg Gly Arg Ala Leu Gly Val Thr Phe Phe
 435 440 445
 Leu Pro Glu Gly Asp Ala Ile Ser Ala Ser Arg Gln Lys Leu Glu Ser
 450 455 460
 Gln Ile Ser Thr Leu Tyr Gly Gly Arg Leu Ala Glu Glu Ile Ile Tyr
 465 470 475 480
 Gly Pro Glu His Val Ser Thr Gly Ala Ser Asn Asp Ile Lys Val Ala
 485 490 495
 Thr Asn Leu Ala Arg Asn Met Val Thr Gln Trp Gly Phe Ser Glu Lys
 500 505 510
 Leu Gly Pro Leu Leu Tyr Ala Glu Glu Gly Glu Val Phe Leu Gly
 515 520 525
 Arg Ser Val Ala Lys Ala Lys His Met Ser Asp Glu Thr Ala Arg Ile
 530 535 540
 Ile Asp Gln Glu Val Lys Ala Leu Ile Glu Arg Asn Tyr Asn Arg Ala
 545 550 555 560
 Arg Gln Leu Leu Thr Asp Asn Met Asp Ile Leu His Ala Met Lys Asp
 565 570 575
 Ala Leu Met Lys Tyr Glu Thr Ile Asp Ala Pro Gln Ile Asp Asp Leu
 580 585 590
 Met Ala Arg Arg Asp Val Arg Pro Pro Ala Gly Trp Glu Glu Pro Gly
 595 600 605
 Ala Ser Asn Asn Ser Gly Asp Asn Gly Ser Pro Lys Ala Pro Arg Pro
 610 615 620
 Val Asp Glu Pro Arg Thr Pro Asn Pro Gly Asn Thr Met Ser Glu Gln
 625 630 635 640
 Leu Gly Asp Lys

<210> 27
 <211> 612
 <212> DNA
 <213> 大肠杆菌
 <220>
 <221> CDS
 <222> (1)..(612)
 <223> 大肠杆菌必需基因 lolA
 <400> 27

atg aaa aaa att gcc atc acc tgt gca tta ctc tca agc tta gta gca	48
Met Lys Lys Ile Ala Ile Thr Cys Ala Leu Leu Ser Ser Leu Val Ala	
1 5 10 15	
agc agc gtt tgg gct gat gcc gca agc gat ctg aaa agc cgc ctg gat	96
Ser Ser Val Trp Ala Asp Ala Ala Ser Asp Leu Lys Ser Arg Leu Asp	
20 25 30	
aaa gtc agc agc ttc cac gcc agc ttc aca caa aaa gtg act gac ggt	144
Lys Val Ser Ser Phe His Ala Ser Phe Thr Gln Lys Val Thr Asp Gly	
35 40 45	
agc ggc gcg gcg gtg cag gaa ggt cag ggc gat ctg tgg gtg aaa cgt	192
Ser Gly Ala Ala Val Gln Gly Gln Gly Asp Leu Trp Val Lys Arg	
50 55 60	
cca aac tta ttc aac tgg cat atg aca caa cct gat gaa agc att ctg	240

	Pro	Asn	Leu	Phe	Asn	Trp	His	Met	Thr	Gln	Pro	Asp	Glu	Ser	Ile	Leu	
	65					70					75					80	
	ggt	tct	gac	ggt	aaa	aca	ctg	tgg	ttc	tat	aac	ccg	ttc	ggt	gag	caa	288
	Val	Ser	Asp	Gly	Lys	Thr	Leu	Trp	Phe	Tyr	Asn	Pro	Phe	Val	Glu	Gln	
					85					90					95		
	gct	acg	gca	acc	tgg	ctg	aaa	gat	gcc	acc	ggt	aat	acg	ccg	ttt	atg	336
	Ala	Thr	Ala	Thr	Trp	Leu	Lys	Asp	Ala	Thr	Gly	Asn	Thr	Pro	Phe	Met	
					100				105					110			
	ctg	att	gcc	cgc	aac	cag	tcc	agc	gac	tgg	cag	cag	tac	aat	atc	aaa	384
	Leu	Ile	Ala	Arg	Asn	Gln	Ser	Ser	Asp	Trp	Gln	Gln	Tyr	Asn	Ile	Lys	
					115				120					125			
	cag	aat	ggc	gat	gac	ttt	gtc	ctg	acg	ccg	aaa	gcc	agc	aat	ggc	aat	432
	Gln	Asn	Gly	Asp	Asp	Phe	Val	Leu	Thr	Pro	Lys	Ala	Ser	Asn	Gly	Asn	
					130				135					140			
	ctg	aag	cag	ttc	acc	att	aac	gtg	gga	cgt	gat	ggc	aca	atc	cat	cag	480
	Leu	Lys	Gln	Phe	Thr	Ile	Asn	Val	Gly	Arg	Asp	Gly	Thr	Ile	His	Gln	
																160	
	ttt	agc	gcg	gtg	gag	cag	gac	gat	cag	cgc	agc	agt	tat	caa	ctg	aaa	528
	Phe	Ser	Ala	Val	Glu	Gln	Asp	Asp	Gln	Arg	Ser	Ser	Tyr	Gln	Leu	Lys	
																175	
	tcc	cag	caa	aat	ggg	gct	gtg	gat	gca	gcg	aaa	ttt	acc	ttc	acc	ccg	576
	Ser	Gln	Gln	Asn	Gly	Ala	Val	Asp	Ala	Ala	Lys	Phe	Thr	Phe	Thr	Pro	
																190	
	ccg	caa	ggc	gtc	acg	gta	gat	gat	caa	cgt	aag	tag					612
	Pro	Gln	Gly	Val	Thr	Val	Asp	Asp	Gln	Arg	Lys						
	<210>	28															
	<211>	203															
	<212>	PRT															
	<213>	大肠杆菌															
	<400>	28															
	Met	Lys	Lys	Ile	Ala	Ile	Thr	Cys	Ala	Leu	Leu	Ser	Ser	Leu	Val	Ala	
	1				5					10					15		
	Ser	Ser	Val	Trp	Ala	Asp	Ala	Ala	Ser	Asp	Leu	Lys	Ser	Arg	Leu	Asp	
	Lys	Val	Ser	Ser	Phe	His	Ala	Ser	Phe	Thr	Gln	Lys	Val	Thr	Asp	Gly	
	Ser	Gly	Ala	Ala	Val	Gln	Glu	Gly	Gln	Gly	Asp	Leu	Trp	Val	Lys	Arg	
[0400]	Pro	Asn	Leu	Phe	Asn	Trp	His	Met	Thr	Gln	Pro	Asp	Glu	Ser	Ile	Leu	
	65					70					75					80	
	Val	Ser	Asp	Gly	Lys	Thr	Leu	Trp	Phe	Tyr	Asn	Pro	Phe	Val	Glu	Gln	
					85					90					95		
	Ala	Thr	Ala	Thr	Trp	Leu	Lys	Asp	Ala	Thr	Gly	Asn	Thr	Pro	Phe	Met	
					100				105					110			
	Leu	Ile	Ala	Arg	Asn	Gln	Ser	Ser	Asp	Trp	Gln	Gln	Tyr	Asn	Ile	Lys	
					115				120					125			
	Gln	Asn	Gly	Asp	Asp	Phe	Val	Leu	Thr	Pro	Lys	Ala	Ser	Asn	Gly	Asn	
	Leu	Lys	Gln	Phe	Thr	Ile	Asn	Val	Gly	Arg	Asp	Gly	Thr	Ile	His	Gln	
																160	
	Phe	Ser	Ala	Val	Glu	Gln	Asp	Asp	Gln	Arg	Ser	Ser	Tyr	Gln	Leu	Lys	
																175	
	Ser	Gln	Gln	Asn	Gly	Ala	Val	Asp	Ala	Ala	Lys	Phe	Thr	Phe	Thr	Pro	
																190	
	Pro	Gln	Gly	Val	Thr	Val	Asp	Asp	Gln	Arg	Lys						
	<210>	29															
	<211>	1083															
	<212>	DNA															
	<213>	大肠杆菌															
	<220>																
	<221>	CDS															
	<222>	(1)..(1083)															
	<223>	大肠杆菌必需基因 mraY															
	<400>	29															
	atg	tta	gtt	tgg	ctg	gcc	gaa	cat	ttg	gtc	aaa	tat	tat	tcc	ggc	ttt	48
	Met	Leu	Val	Trp	Leu	Ala	Glu	His	Leu	Val	Lys	Tyr	Tyr	Ser	Gly	Phe	
	1				5					10					15		
	aac	gtc	ttt	tcc	tat	ctg	acg	ttt	cgc	gcc	atc	gtc	agc	ctg	ctg	acc	96
	Asn	Val	Phe	Ser	Tyr	Leu	Thr	Phe	Arg	Ala	Ile	Val	Ser	Leu	Leu	Thr	
	gcg	ctg	ttc	atc	tca	ttg	tgg	atg	ggc	ccg	cgt	atg	att	gct	cat	ttg	144
	Ala	Leu	Phe	Ile	Ser	Leu	Trp	Met	Gly	Pro	Arg	Met	Ile	Ala	His	Leu	
	caa	aaa	ctt	tcc	ttt	ggt	cag	gtg	gtg	cgt	aac	gac	ggg	cct	gaa	tca	192
	Gln	Lys	Leu	Ser	Phe	Gly	Gln	Val	Val	Arg	Asn	Asp	Gly	Pro	Glu	Ser	

	cac	ttc	agc	aag	cgc	ggt	acg	ccg	acc	atg	ggc	ggg	att	atg	atc	ctg	240
	His	Phe	Ser	Lys	Arg	Gly	Thr	Pro	Thr	Met	Gly	Gly	Ile	Met	Ile	Leu	
	65					70					75					80	
	acg	gcg	att	gtg	atc	tcc	gta	ctg	ctg	tgg	gct	tac	ccg	tcc	aat	ccg	288
	Thr	Ala	Ile	Val	Ile	Ser	Val	Leu	Leu	Trp	Ala	Tyr	Pro	Ser	Asn	Pro	
						85				90					95		
	tac	gtc	tgg	tgc	gtg	ttg	gtg	gtg	ctg	gta	ggg	tac	ggg	gtt	att	ggc	336
	Tyr	Val	Trp	Cys	Val	Leu	Val	Val	Leu	Val	Gly	Tyr	Gly	Val	Ile	Gly	
				100					105					110			
	ttt	gtt	gat	gat	tat	cgc	aaa	gtg	gtg	cgt	aaa	gac	acc	aaa	ggg	ttg	384
	Phe	Val	Asp	Asp	Tyr	Arg	Lys	Val	Val	Arg	Lys	Asp	Thr	Lys	Gly	Leu	
			115					120					125				
	atc	gct	cgt	tgg	aag	tat	ttc	tgg	atg	tcg	gtc	att	gcg	ctg	ggg	gtc	432
	Ile	Ala	Arg	Trp	Lys	Tyr	Phe	Trp	Met	Ser	Val	Ile	Ala	Leu	Gly	Val	
			130					135					140				
	gcc	ttc	gcc	ctg	tac	ctt	gcc	ggc	aaa	gac	acg	ccc	gca	acg	cag	ctg	480
	Ala	Phe	Ala	Leu	Tyr	Leu	Ala	Gly	Lys	Asp	Thr	Pro	Ala	Thr	Gln	Leu	
	145					150					155					160	
	gtg	gtc	cca	ttc	ttt	aaa	gat	gtg	atg	ccg	cag	ctg	ggg	ctg	ttc	tac	528
	Val	Val	Pro	Phe	Phe	Lys	Asp	Val	Met	Pro	Gln	Leu	Gly	Leu	Phe	Tyr	
						165				170					175		
	att	ctg	ctg	gct	tac	ttc	gtc	att	gtg	ggg	act	ggc	aac	gcg	gta	aac	576
	Ile	Leu	Leu	Ala	Tyr	Phe	Val	Ile	Val	Gly	Thr	Gly	Asn	Ala	Val	Asn	
						180				185				190			
	ctg	acc	gat	ggg	ctc	gac	ggc	ctg	gca	att	atg	ccg	acc	gta	ttt	gtc	624
	Leu	Thr	Asp	Gly	Leu	Asp	Gly	Leu	Ala	Ile	Met	Pro	Thr	Val	Phe	Val	
			195					200					205				
	gcc	ggg	ggg	ttt	gcg	ctg	gtg	gcg	tgg	gcg	acc	ggc	aat	atg	aac	ttt	672
	Ala	Gly	Gly	Phe	Ala	Leu	Val	Ala	Trp	Ala	Thr	Gly	Asn	Met	Asn	Phe	
		210					215					220					
	gcc	agc	tac	ttg	cat	ata	ccg	tat	ctg	cga	cac	gcc	ggg	gaa	ctg	gtt	720
	Ala	Ser	Tyr	Leu	His	Ile	Pro	Tyr	Leu	Arg	His	Ala	Gly	Glu	Leu	Val	
	225					230					235					240	
	att	gtc	tgt	acc	gcg	ata	gtc	ggg	gca	gga	ctg	ggc	ttc	ctg	tgg	ttt	768
	Ile	Val	Cys	Thr	Ala	Ile	Val	Gly	Ala	Gly	Leu	Gly	Phe	Leu	Trp	Phe	
						245				250					255		
	aac	acc	tat	ccg	gcg	cag	gtc	ttt	atg	ggc	gat	gta	ggg	tcg	ctg	gcg	816
	Asn	Thr	Tyr	Pro	Ala	Gln	Val	Phe	Met	Gly	Asp	Val	Gly	Ser	Leu	Ala	
				260					265					270			
[0401]	tta	ggg	ggg	gcg	tta	ggc	att	atc	gcc	gta	ctg	cta	cgt	cag	gaa	ttc	864
	Leu	Gly	Gly	Ala	Leu	Gly	Ile	Ile	Ala	Val	Leu	Leu	Arg	Gln	Glu	Phe	
			275					280					285				
	ctg	ctg	gtg	att	atg	ggg	ggc	gtg	ttc	gtg	gta	gaa	acg	ctt	tct	gtc	912
	Leu	Leu	Val	Ile	Met	Gly	Gly	Val	Phe	Val	Val	Glu	Thr	Leu	Ser	Val	
			290				295					300					
	atc	ctg	cag	gtc	ggc	tcc	ttt	aaa	ctg	cgc	gga	caa	cgt	att	ttc	cgc	960
	Ile	Leu	Gln	Val	Gly	Ser	Phe	Lys	Leu	Arg	Gly	Gln	Arg	Ile	Phe	Arg	
						310				315					320		
	atg	gca	ccg	att	cat	cac	cac	tat	gaa	ctg	aaa	ggc	tgg	ccg	gaa	ccg	1008
	Met	Ala	Pro	Ile	His	His	His	Tyr	Glu	Leu	Lys	Gly	Trp	Pro	Glu	Pro	
						325				330					335		
	cgc	gtc	att	gtg	cgt	ttc	tgg	att	att	tcg	ctg	atg	ctg	gtt	ctg	att	1056
	Arg	Val	Ile	Val	Arg	Phe	Trp	Ile	Ile	Ser	Leu	Met	Leu	Val	Leu	Ile	
						340			345					350			
	ggg	ctg	gca	acg	ctg	aag	gta	cgt	taa								1083
	Gly	Leu	Ala	Thr	Leu	Lys	Val	Arg									
			355					360									
	<210>	30															
	<211>	360															
	<212>	PRT															
	<213>	大肠杆菌															
	<400>	30															
	Met	Leu	Val	Trp	Leu	Ala	Glu	His	Leu	Val	Lys	Tyr	Tyr	Ser	Gly	Phe	
	1				5					10					15		
	Asn	Val	Phe	Ser	Tyr	Leu	Thr	Phe	Arg	Ala	Ile	Val	Ser	Leu	Leu	Thr	
				20					25					30			
	Ala	Leu	Phe	Ile	Ser	Leu	Trp	Met	Gly	Pro	Arg	Met	Ile	Ala	His	Leu	
				35				40					45				
	Gln	Lys	Leu	Ser	Phe	Gly	Gln	Val	Val	Arg	Asn	Asp	Gly	Pro	Glu	Ser	
		50					55				60						
	His	Phe	Ser	Lys	Arg	Gly	Thr	Pro	Thr	Met	Gly	Gly	Ile	Met	Ile	Leu	
	65					70					75				80		
	Thr	Ala	Ile	Val	Ile	Ser	Val	Leu	Leu	Trp	Ala	Tyr	Pro	Ser	Asn	Pro	
						85				90					95		
	Tyr	Val	Trp	Cys	Val	Leu	Val	Val	Leu	Val	Gly	Tyr	Gly	Val	Ile	Gly	
				100					105					110			
	Phe	Val	Asp	Asp	Tyr	Arg	Lys	Val	Val	Arg	Lys	Asp	Thr	Lys	Gly	Leu	
			115					120					125				

[0402]

```

Ile Ala Arg Trp Lys Tyr Phe Trp Met Ser Val Ile Ala Leu Gly Val
130 135 140
Ala Phe Ala Leu Tyr Leu Ala Gly Lys Asp Thr Pro Ala Thr Gln Leu
145 150 155
Val Val Pro Phe Phe Lys Asp Val Met Pro Gln Leu Gly Leu Phe Tyr
165 170 175
Ile Leu Leu Ala Tyr Phe Val Ile Val Gly Thr Gly Asn Ala Val Asn
180 185 190
Leu Thr Asp Gly Leu Asp Gly Leu Ala Ile Met Pro Thr Val Phe Val
195 200 205
Ala Gly Gly Phe Ala Leu Val Ala Trp Ala Thr Gly Asn Met Asn Phe
210 215 220
Ala Ser Tyr Leu His Ile Pro Tyr Leu Arg His Ala Gly Glu Leu Val
225 230 235
Ile Val Cys Thr Ala Ile Val Gly Ala Gly Leu Gly Phe Leu Trp Phe
245 250 255
Asn Thr Tyr Pro Ala Gln Val Phe Met Gly Asp Val Gly Ser Leu Ala
260 265 270
Leu Gly Gly Ala Leu Gly Ile Ile Ala Val Leu Leu Arg Gln Glu Phe
275 280 285
Leu Leu Val Ile Met Gly Gly Val Phe Val Val Glu Thr Leu Ser Val
290 295 300
Ile Leu Gln Val Gly Ser Phe Lys Leu Arg Gly Gln Arg Ile Phe Arg
305 310 315
Met Ala Pro Ile His His His Tyr Glu Leu Lys Gly Trp Pro Glu Pro
325 330 335
Arg Val Ile Val Arg Phe Trp Ile Ile Ser Leu Met Leu Val Leu Ile
340 345 350
Gly Leu Ala Thr Leu Lys Val Arg
355 360
<210> 31
<211> 489
<212> DNA
<213> 大肠杆菌
<220>
<221> CDS
<222> (1)..(489)
<223> 大肠杆菌必需基因 mreD
<400> 31
gtg gcg agc tat cgt agc cag gga cgc tgg gta atc tgg ctc tct ttc 48
Val Ala Ser Tyr Arg Ser Gln Gly Arg Trp Val Ile Trp Leu Ser Phe
1 5 10 15
ctc att gcg ctg ttg ctg caa atc atg ccc tgg ccg gat aac ctg att 96
Leu Ile Ala Leu Leu Leu Gln Ile Met Pro Trp Pro Asp Asn Leu Ile
20 25 30
gtt ttc cgg cca aac tgg gtg tta ctc atc ttg ttg tat tgg atc ctg 144
Val Phe Arg Pro Asn Trp Val Leu Leu Ile Leu Leu Tyr Trp Ile Leu
35 40 45
gcc ttg cct cat cgc gta aat gtg ggc aca ggt ttt gtg atg ggt gcc 192
Ala Leu Pro His Arg Val Asn Val Gly Thr Gly Phe Val Met Gly Ala
50 55 60
ata ctg gat ctg atc agc ggc tcg acg ctt ggc gta cgc gta ttg gcg 240
Ile Leu Asp Leu Ile Ser Gly Ser Thr Leu Gly Val Arg Val Leu Ala
65 70 75 80
atg agc atc att gct tac ctg gtg gcg ctg aaa tac cag ctt ttc cgc 288
Met Ser Ile Ile Ala Tyr Leu Val Ala Leu Lys Tyr Gln Leu Phe Arg
85 90 95
aac ctc gca tta tgg cag cag gcg ctg gtc gtc atg ttg ctt tcg ctg 336
Asn Leu Ala Leu Trp Gln Gln Ala Leu Val Val Met Leu Leu Ser Leu
100 105 110
gtg gtg gat att att gtt ttc tgg gca gag ttt tta gtg att aac gtc 384
Val Val Asp Ile Ile Val Phe Trp Ala Glu Phe Leu Val Ile Asn Val
115 120 125
tct ttc aga ccg gaa gtg ttc tgg agt agt gta gtc aat ggg gtg ctc 432
Ser Phe Arg Pro Glu Val Phe Trp Ser Ser Val Val Asn Gly Val Leu
130 135 140
tgg ccg tgg att ttc ttg ctg atg cgc aaa gtc cgt cag cag ttt gca 480
Trp Pro Trp Ile Phe Leu Leu Met Arg Lys Val Arg Gln Gln Phe Ala
145 150 155 160
gtg caa taa
Val Gln 489

<210> 32
<211> 162
<212> PRT
<213> 大肠杆菌
<400> 32
Val Ala Ser Tyr Arg Ser Gln Gly Arg Trp Val Ile Trp Leu Ser Phe

```

71

	gat	act	ctc	gac	gcc	gtg	ctg	gcg	aaa	ctg	cgt	gac	gct	gga	gcg	gac	816
	Asp	Thr	Leu	Asp	Ala	Val	Leu	Ala	Lys	Leu	Arg	Asp	Ala	Gly	Ala	Asp	
				260					265					270			
	atc	gaa	gtc	ggc	gaa	gac	tgg	att	agc	ctg	gat	atg	cat	ggc	aaa	cgt	864
	Ile	Glu	Val	Gly	Glu	Asp	Trp	Ile	Ser	Leu	Asp	Met	His	Gly	Lys	Arg	
				275					280					285			
	ccg	aag	gct	gtt	aac	gta	cgt	acc	gcg	ccg	cat	ccg	gca	ttc	ccg	acc	912
	Pro	Lys	Ala	Val	Asn	Val	Arg	Thr	Ala	Pro	His	Pro	Ala	Phe	Pro	Thr	
				290					295					300			
	gat	atg	cag	gcc	cag	ttc	acg	ctg	ttg	aac	ctg	gtg	gca	gaa	ggg	act	960
	Asp	Met	Gln	Ala	Gln	Phe	Thr	Leu	Leu	Asn	Leu	Val	Ala	Glu	Gly	Thr	
				305												320	
	gga	ttc	atc	acc	gaa	acg	gtc	ttt	gaa	aac	cgc	ttt	atg	cat	gtg	cca	1008
	Gly	Phe	Ile	Thr	Glu	Thr	Val	Phe	Glu	Asn	Arg	Phe	Met	His	Val	Pro	
				325												335	
	gag	ctg	agc	cgt	atg	ggc	gcg	cac	gcc	gaa	atc	gaa	agc	aat	acc	gtt	1056
	Glu	Leu	Ser	Arg	Met	Gly	Ala	His	Ala	Glu	Ile	Glu	Ser	Asn	Thr	Val	
				340					345							350	
	att	tgt	cac	ggt	gtt	gaa	aaa	ctt	tct	ggc	gca	cag	gtt	atg	gca	acc	1104
	Ile	Cys	His	Gly	Val	Glu	Lys	Leu	Ser	Gly	Ala	Gln	Val	Met	Ala	Thr	
				355					360							365	
	gat	ctg	cgt	gca	tca	gca	agc	ctg	gtg	ctg	gct	ggc	tgt	att	gcg	gaa	1152
	Asp	Leu	Arg	Ala	Ser	Ala	Ser	Leu	Val	Leu	Ala	Gly	Cys	Ile	Ala	Glu	
				370					375							380	
	ggg	acg	acg	gtg	gtt	gat	cgt	att	tat	cac	atc	gat	cgt	ggc	tac	gaa	1200
	Gly	Thr	Thr	Val	Val	Asp	Arg	Ile	Tyr	His	Ile	Asp	Arg	Gly	Tyr	Glu	
				385												400	
	cgc	att	gaa	gac	aaa	ctg	cgc	gct	tta	ggc	gca	aat	att	gag	cgt	gtg	1248
	Arg	Ile	Glu	Asp	Lys	Leu	Arg	Ala	Leu	Gly	Ala	Asn	Ile	Glu	Arg	Val	
					405					410						415	
	aaa	ggc	gag	taa													1260
	Lys	Gly	Glu														
	<210>	34															
	<211>	419															
	<212>	PRT															
	<213>	大肠杆菌															
	<400>	34															
[0404]	Met	Asp	Lys	Phe	Arg	Val	Gln	Gly	Pro	Thr	Lys	Leu	Gln	Gly	Glu	Val	
	1				5					10					15		
	Thr	Ile	Ser	Gly	Ala	Lys	Asn	Ala	Ala	Leu	Pro	Ile	Leu	Phe	Ala	Ala	
					20				25					30			
	Leu	Leu	Ala	Glu	Glu	Pro	Val	Glu	Ile	Gln	Asn	Val	Pro	Lys	Leu	Lys	
					35				40					45			
	Asp	Val	Asp	Thr	Ser	Met	Lys	Leu	Leu	Ser	Gln	Leu	Gly	Ala	Lys	Val	
							55					60					
	Glu	Arg	Asn	Gly	Ser	Val	His	Ile	Asp	Ala	Arg	Asp	Val	Asn	Val	Phe	
							70				75					80	
	Cys	Ala	Pro	Tyr	Asp	Leu	Val	Lys	Thr	Met	Arg	Ala	Ser	Ile	Trp	Ala	
							85			90					95		
	Leu	Gly	Pro	Leu	Val	Ala	Arg	Phe	Gly	Gln	Gly	Gln	Val	Ser	Leu	Pro	
									105					110			
	Gly	Gly	Cys	Thr	Ile	Gly	Ala	Arg	Pro	Val	Asp	Leu	His	Ile	Ser	Gly	
									120					125			
	Leu	Glu	Gln	Leu	Gly	Ala	Thr	Ile	Lys	Leu	Glu	Glu	Gly	Tyr	Val	Lys	
							135						140				
	Ala	Ser	Val	Asp	Gly	Arg	Leu	Lys	Gly	Ala	His	Ile	Val	Met	Asp	Lys	
							150				155					160	
	Val	Ser	Val	Gly	Ala	Thr	Val	Thr	Ile	Met	Cys	Ala	Ala	Thr	Leu	Ala	
							165				170					175	
	Glu	Gly	Thr	Thr	Ile	Ile	Glu	Asn	Ala	Ala	Arg	Glu	Pro	Glu	Ile	Val	
									185					190			
	Asp	Thr	Ala	Asn	Phe	Leu	Ile	Thr	Leu	Gly	Ala	Lys	Ile	Ser	Gly	Gln	
									200					205			
	Gly	Thr	Asp	Arg	Ile	Val	Ile	Glu	Gly	Val	Glu	Arg	Leu	Gly	Gly	Gly	
									215					220			
	Val	Tyr	Arg	Val	Leu	Pro	Asp	Arg	Ile	Glu	Thr	Gly	Thr	Phe	Leu	Val	
							230					235				240	
	Ala	Ala	Ala	Ile	Ser	Arg	Gly	Lys	Ile	Ile	Cys	Arg	Asn	Ala	Gln	Pro	
															255		
	Asp	Thr	Leu	Asp	Ala	Val	Leu	Ala	Lys	Leu	Arg	Asp	Ala	Gly	Ala	Asp	
															270		
	Ile	Glu	Val	Gly	Glu	Asp	Trp	Ile	Ser	Leu	Asp	Met	His	Gly	Lys	Arg	
									280					285			
	Pro	Lys	Ala	Val	Asn	Val	Arg	Thr	Ala	Pro	His	Pro	Ala	Phe	Pro	Thr	
																300	
	Asp	Met	Gln	Ala	Gln	Phe	Thr	Leu	Leu	Asn	Leu	Val	Ala	Glu	Gly	Thr	
							310					315				320	

	Gly	Phe	Ile	Thr	Glu	Thr	Val	Phe	Glu	Asn	Arg	Phe	Met	His	Val	Pro	
					325					330					335		
	Glu	Leu	Ser	Arg	Met	Gly	Ala	His	Ala	Glu	Ile	Glu	Ser	Asn	Thr	Val	
					340				345					350			
	Ile	Cys	His	Gly	Val	Glu	Lys	Leu	Ser	Gly	Ala	Gln	Val	Met	Ala	Thr	
			355					360					365				
	Asp	Leu	Arg	Ala	Ser	Ala	Ser	Leu	Val	Leu	Ala	Gly	Cys	Ile	Ala	Glu	
		370					375					380					
	Gly	Thr	Thr	Val	Val	Asp	Arg	Ile	Tyr	His	Ile	Asp	Arg	Gly	Tyr	Glu	
		385				390					395					400	
	Arg	Ile	Glu	Asp	Lys	Leu	Arg	Ala	Leu	Gly	Ala	Asn	Ile	Glu	Arg	Val	
					405					410					415		
	Lys	Gly	Glu														
	<210>	35															
	<211>	1029															
	<212>	DNA															
	<213>	大肠杆菌															
	<220>																
	<221>	CDS															
	<222>	(1)..(1029)															
	<223>	大肠杆菌必需基因 murB															
	<400>	35															
	atg	aac	cac	tcc	tta	aaa	ccc	tgg	aac	aca	ttt	ggc	att	gat	cat	aat	48
	Met	Asn	His	Ser	Leu	Lys	Pro	Trp	Asn	Thr	Phe	Gly	Ile	Asp	His	Asn	
	1			5					10					15			
	gct	cag	cac	att	gta	tgt	gcc	gaa	gac	gaa	caa	caa	tta	ctc	aat	gcc	96
	Ala	Gln	His	Ile	Val	Cys	Ala	Glu	Asp	Glu	Gln	Gln	Leu	Leu	Asn	Ala	
				20					25					30			
	tgg	cag	cat	gca	acc	gca	gaa	gga	caa	ccc	gtt	ctt	att	ctg	ggg	gaa	144
	Trp	Gln	His	Ala	Thr	Ala	Glu	Gly	Gln	Pro	Val	Leu	Ile	Leu	Gly	Glu	
			35				40					45					
	gga	agt	aat	gta	ctt	ttt	ctg	gaa	gac	tat	cgc	ggc	acg	gtg	atc	atc	192
	Gly	Ser	Asn	Val	Leu	Phe	Leu	Glu	Asp	Tyr	Arg	Gly	Thr	Val	Ile	Ile	
		50					55				60						
	aac	cgg	atc	aaa	ggg	atc	gaa	att	cat	gat	gaa	cct	gat	gcg	tgg	tat	240
	Asn	Arg	Ile	Lys	Gly	Ile	Glu	Ile	His	Asp	Glu	Pro	Asp	Ala	Trp	Tyr	
	65				70					75					80		
[0405]	tta	cat	gta	gga	gcc	gga	gaa	aac	tgg	cat	cgc	ctg	gta	aaa	tac	act	288
	Leu	His	Val	Gly	Ala	Gly	Glu	Asn	Trp	His	Arg	Leu	Val	Lys	Tyr	Thr	
				85					90					95			
	ttg	cag	gaa	ggg	atg	cct	ggg	ctg	gaa	aat	ctg	gca	tta	att	cct	ggg	336
	Leu	Gln	Glu	Gly	Met	Pro	Gly	Leu	Asn	Leu	Ala	Leu	Ile	Pro	Gly		
				100					105				110				
	tgt	gtc	ggc	tca	tca	cct	atc	cag	aat	att	ggg	gct	tat	ggc	gta	gaa	384
	Cys	Val	Gly	Ser	Ser	Pro	Ile	Gln	Asn	Ile	Gly	Ala	Tyr	Gly	Val	Glu	
			115				120						125				
	tta	cag	cga	gtt	tgc	gct	tat	gtt	gat	tgt	gtt	gaa	ctg	gcg	aca	ggc	432
	Leu	Gln	Arg	Val	Cys	Ala	Tyr	Val	Asp	Cys	Val	Glu	Leu	Ala	Thr	Gly	
		130					135					140					
	aag	caa	gtg	cgc	tta	act	gcc	aaa	gag	tgc	cgt	ttt	ggc	tat	cgc	gac	480
	Lys	Gln	Val	Arg	Leu	Thr	Ala	Lys	Glu	Cys	Arg	Phe	Gly	Tyr	Arg	Asp	
		145				150				155						160	
	agt	att	ttt	aaa	cat	gaa	tac	cag	gat	cgc	ttc	gcc	att	gta	gcc	gta	528
	Ser	Ile	Phe	Lys	His	Glu	Tyr	Gln	Asp	Arg	Phe	Ala	Ile	Val	Ala	Val	
				165					170					175			
	ggg	ctg	cgt	ctg	cca	aaa	gag	tgg	caa	cct	gta	cta	acg	tat	ggg	gac	576
	Gly	Leu	Arg	Leu	Pro	Lys	Glu	Trp	Gln	Pro	Val	Leu	Thr	Tyr	Gly	Asp	
				180					185					190			
	tta	act	cgt	ctg	gat	cct	aca	aca	gta	acg	cca	cag	caa	gta	ttt	aat	624
	Leu	Thr	Arg	Leu	Asp	Pro	Thr	Thr	Val	Thr	Pro	Gln	Gln	Val	Phe	Asn	
			195				200						205				
	gcg	gtg	tgt	cat	atg	cgc	acc	acc	aaa	ctc	cct	gat	cca	aaa	gtg	aat	672
	Ala	Val	Cys	His	Met	Arg	Thr	Thr	Lys	Leu	Pro	Asp	Pro	Lys	Val	Asn	
		210					215					220					
	ggc	aat	gcc	ggg	agt	ttc	aaa	aac	cct	gtt	gta	tct	gcc	gaa	acg		720
	Gly	Asn	Ala	Gly	Ser	Phe	Phe	Lys	Asn	Pro	Val	Val	Ser	Ala	Glu	Thr	
		225				230				235					240		
	gct	gaa	gca	tta	ctg	tca	caa	ttt	cca	aca	gca	cca	aat	tac	ccc	cag	768
	Ala	Glu	Ala	Leu	Leu	Ser	Gln	Phe	Pro	Thr	Ala	Pro	Asn	Tyr	Pro	Gln	
				245					250					255			
	gcg	gat	ggg	tca	gta	aaa	ctg	gca	gca	ggg	tgg	ctt	att	gat	cag	tgc	816
	Ala	Asp	Gly	Ser	Val	Lys	Leu	Ala	Ala	Gly	Trp	Leu	Ile	Asp	Gln	Cys	
				260				265						270			
	cag	cta	aaa	ggg	atg	caa	atg	ggg	ggg	gct	gcg	gtg	cac	cgt	caa	cag	864
	Gln	Leu	Lys	Gly	Met	Gln	Met	Gly	Gly	Ala	Ala	Val	His	Arg	Gln	Gln	
			275				280						285				
	gcg	tta	gtt	ctc	att	aat	gaa	gac	aat	gca	aaa	agc	gaa	gat	gtg	gtg	912

	Ala	Leu	Val	Leu	Ile	Asn	Glu	Asp	Asn	Ala	Lys	Ser	Glu	Asp	Val	Val	
	290						295					300					
	caa	ctg	gca	cac	cat	gta	aga	caa	aaa	gtg	ggg	gaa	aaa	ttt	aat	gtc	960
	Gln	Leu	Ala	His	His	Val	Arg	Gln	Lys	Val	Gly	Glu	Lys	Phe	Asn	Val	
	305					310					315					320	
	tgg	ctt	gag	cct	gaa	gtc	cgc	ttt	att	ggg	gca	tca	ggg	gaa	gtg	agc	1008
	Trp	Leu	Glu	Pro	Glu	Val	Arg	Phe	Ile	Gly	Ala	Ser	Gly	Glu	Val	Ser	
					325					330					335		
	gca	gtg	gag	aca	att	tca	tga										1029
	Ala	Val	Glu	Thr	Ile	Ser											
				340													
	<210>	36															
	<211>	342															
	<212>	PRT															
	<213>	大肠杆菌															
	<400>	36															
	Met	Asn	His	Ser	Leu	Lys	Pro	Trp	Asn	Thr	Phe	Gly	Ile	Asp	His	Asn	
	1				5					10					15		
	Ala	Gln	His	Ile	Val	Cys	Ala	Glu	Asp	Glu	Gln	Gln	Leu	Leu	Asn	Ala	
				20					25					30			
	Trp	Gln	His	Ala	Thr	Ala	Glu	Gly	Gln	Pro	Val	Leu	Ile	Leu	Gly	Glu	
			35					40					45				
	Gly	Ser	Asn	Val	Leu	Phe	Leu	Glu	Asp	Tyr	Arg	Gly	Thr	Val	Ile	Ile	
		50					55					60					
	Asn	Arg	Ile	Lys	Gly	Ile	Glu	Ile	His	Asp	Glu	Pro	Asp	Ala	Trp	Tyr	
	65					70				75						80	
	Leu	His	Val	Gly	Ala	Gly	Glu	Asn	Trp	His	Arg	Leu	Val	Lys	Tyr	Thr	
				85					90						95		
	Leu	Gln	Glu	Gly	Met	Pro	Gly	Leu	Glu	Asn	Leu	Ala	Leu	Ile	Pro	Gly	
			100					105						110			
	Cys	Val	Gly	Ser	Ser	Pro	Ile	Gln	Asn	Ile	Gly	Ala	Tyr	Gly	Val	Glu	
			115					120					125				
	Leu	Gln	Arg	Val	Cys	Ala	Tyr	Val	Asp	Cys	Val	Glu	Leu	Ala	Thr	Gly	
		130					135					140					
	Lys	Gln	Val	Arg	Leu	Thr	Ala	Lys	Glu	Cys	Arg	Phe	Gly	Tyr	Arg	Asp	
	145					150					155					160	
	Ser	Ile	Phe	Lys	His	Glu	Tyr	Gln	Asp	Arg	Phe	Ala	Ile	Val	Ala	Val	
				165					170						175		
[0406]	Gly	Leu	Arg	Leu	Pro	Lys	Glu	Trp	Gln	Pro	Val	Leu	Thr	Tyr	Gly	Asp	
				180				185						190			
	Leu	Thr	Arg	Leu	Asp	Pro	Thr	Thr	Val	Thr	Pro	Gln	Gln	Val	Phe	Asn	
		195					200						205				
	Ala	Val	Cys	His	Met	Arg	Thr	Thr	Lys	Leu	Pro	Asp	Pro	Lys	Val	Asn	
		210					215					220					
	Gly	Asn	Ala	Gly	Ser	Phe	Phe	Lys	Asn	Pro	Val	Val	Ser	Ala	Glu	Thr	
	225					230					235					240	
	Ala	Glu	Ala	Leu	Leu	Ser	Gln	Phe	Pro	Thr	Ala	Pro	Asn	Tyr	Pro	Gln	
				245						250					255		
	Ala	Asp	Gly	Ser	Val	Lys	Leu	Ala	Ala	Gly	Trp	Leu	Ile	Asp	Gln	Cys	
			260					265						270			
	Gln	Leu	Lys	Gly	Met	Gln	Met	Gly	Gly	Ala	Ala	Val	His	Arg	Gln	Gln	
		275						280					285				
	Ala	Leu	Val	Leu	Ile	Asn	Glu	Asp	Asn	Ala	Lys	Ser	Glu	Asp	Val	Val	
		290					295					300					
	Gln	Leu	Ala	His	His	Val	Arg	Gln	Lys	Val	Gly	Glu	Lys	Phe	Asn	Val	
	305					310					315					320	
	Trp	Leu	Glu	Pro	Glu	Val	Arg	Phe	Ile	Gly	Ala	Ser	Gly	Glu	Val	Ser	
					325					330					335		
	Ala	Val	Glu	Thr	Ile	Ser											
				340													
	<210>	37															
	<211>	1359															
	<212>	DNA															
	<213>	大肠杆菌															
	<220>																
	<221>	CDS															
	<222>	(1)..(1359)															
	<223>	大肠杆菌必需基因 murF															
	<400>	37															
	atg	att	agc	gta	acc	ctt	agc	caa	ctt	acc	gac	att	ctc	aac	ggg	gaa	48
	Met	Ile	Ser	Val	Thr	Leu	Ser	Gln	Leu	Thr	Asp	Ile	Leu	Asn	Gly	Glu	
	1			5						10					15		
	ctg	caa	ggg	gca	gat	atc	acc	ctt	gat	gct	gta	acc	act	gat	acc	cga	96
	Leu	Gln	Gly	Ala	Asp	Ile	Thr	Leu	Asp	Ala	Val	Thr	Thr	Asp	Thr	Arg	
				20					25					30			
	aaa	ctg	acg	cgg	ggc	tgc	ctg	ttt	gtt	gcc	ctg	aaa	ggc	gaa	cgt	ttt	144
	Lys	Leu	Thr	Pro	Gly	Cys	Leu	Phe	Val	Ala	Leu	Lys	Gly	Glu	Arg	Phe	
			35					40					45				

[0407]	gat	gcc	cac	gat	ttt	gcc	gac	cag	gcg	aaa	gct	ggc	ggc	gca	ggc	gca	192
	Asp	Ala	His	Asp	Phe	Ala	Asp	Gln	Ala	Lys	Ala	Gly	Gly	Ala	Gly	Ala	
	50						55					60					
	cta	ctg	ggt	agc	cgt	ccg	ctg	gac	atc	gac	ctg	ccg	cag	tta	atc	gtc	240
	Leu	Leu	Val	Ser	Arg	Pro	Leu	Asp	Ile	Asp	Leu	Pro	Gln	Leu	Ile	Val	
	65					70					75				80		
	aag	gat	acg	cgt	ctg	gcg	ttt	ggt	gaa	ctg	gct	gca	tgg	ggt	cgc	cag	288
	Lys	Asp	Thr	Arg	Leu	Ala	Phe	Gly	Glu	Leu	Ala	Ala	Trp	Val	Arg	Gln	
				85					90						95		
	caa	ggt	ccg	gcg	cgc	gtg	ggt	gct	ctg	acg	ggg	tcc	tcc	ggc	aaa	acc	336
	Gln	Val	Pro	Ala	Arg	Val	Val	Ala	Leu	Thr	Gly	Ser	Ser	Gly	Lys	Thr	
				100					105					110			
	tcc	ggt	aaa	gag	atg	acg	gcg	gcg	att	tta	agc	cag	tgc	ggc	aac	acg	384
	Ser	Val	Lys	Glu	Met	Thr	Ala	Ala	Ile	Leu	Ser	Gln	Cys	Gly	Asn	Thr	
			115						120				125				
	ctt	tat	acg	gca	ggc	aat	ctc	aac	aac	gac	atc	ggt	gta	ccg	atg	acg	432
	Leu	Tyr	Thr	Ala	Gly	Asn	Leu	Asn	Asn	Asp	Ile	Gly	Val	Pro	Met	Thr	
	130					135						140					
	ctg	ttg	cgc	tta	acg	ccg	gaa	tac	gat	tac	gca	ggt	att	gaa	ctt	ggc	480
	Leu	Leu	Arg	Leu	Thr	Pro	Glu	Tyr	Asp	Tyr	Ala	Val	Ile	Glu	Leu	Gly	
	145					150					155				160		
	gcg	aac	cat	cag	ggc	gaa	ata	gcc	tgg	act	gtg	agt	ctg	act	cgc	ccg	528
	Ala	Asn	His	Gln	Gly	Glu	Ile	Ala	Trp	Thr	Val	Ser	Leu	Thr	Arg	Pro	
				165						170				175			
	gaa	gct	gcg	ctg	gtc	aac	aac	ctg	gca	gcg	gcg	cat	ctg	gaa	ggt	ttt	576
	Glu	Ala	Ala	Leu	Val	Asn	Asn	Leu	Ala	Ala	Ala	His	Leu	Glu	Gly	Phe	
				180					185					190			
	ggc	tgc	ctt	gcg	ggt	gtc	gcg	aaa	gcg	aaa	ggt	gaa	atc	ttt	agc	ggc	624
	Gly	Ser	Leu	Ala	Gly	Val	Ala	Lys	Ala	Lys	Gly	Glu	Ile	Phe	Ser	Gly	
			195					200					205				
	ctg	ccg	gaa	aac	ggt	atc	gcc	att	atg	aac	gcc	gac	aac	aac	gac	tgg	672
	Leu	Pro	Glu	Asn	Gly	Ile	Ala	Ile	Met	Asn	Ala	Asp	Asn	Asn	Asp	Trp	
	210					215						220					
	ctg	aac	tgg	cag	agc	gta	att	ggc	tca	cgc	aaa	gtg	tgg	cgt	ttc	tca	720
	Leu	Asn	Trp	Gln	Ser	Val	Ile	Gly	Ser	Arg	Lys	Val	Trp	Arg	Phe	Ser	
	225					230					235				240		
	ccc	aat	gcc	gcc	aac	agc	gat	ttc	acc	gcc	acc	aat	atc	cat	gtg	acc	768
	Pro	Asn	Ala	Ala	Asn	Ser	Asp	Phe	Thr	Ala	Thr	Asn	Ile	His	Val	Thr	
					245				250					255			
	tgc	cac	ggt	acg	gaa	ttt	acc	cta	caa	acc	cca	acc	ggt	agc	gtc	gat	816
	Ser	His	Gly	Thr	Glu	Phe	Thr	Leu	Gln	Thr	Pro	Thr	Gly	Ser	Val	Asp	
				260					265				270				
	ggt	ctg	ctg	ccg	ttg	ccg	ggg	cgt	cac	aat	att	gcg	aat	gcg	ctg	gca	864
	Val	Leu	Leu	Pro	Leu	Pro	Gly	Arg	His	Asn	Ile	Ala	Asn	Ala	Leu	Ala	
			275					280					285				
	gcc	gct	gcg	ctc	tcc	atg	tcc	gtg	ggc	gca	acg	ctt	gat	gct	atc	aaa	912
	Ala	Ala	Ala	Leu	Ser	Met	Ser	Val	Gly	Ala	Thr	Leu	Asp	Ala	Ile	Lys	
	290						295					300					
	gcg	ggg	ctg	gca	aat	ctg	aaa	gct	ggt	cca	ggc	cgt	ctg	ttc	ccc	atc	960
	Ala	Gly	Leu	Ala	Asn	Leu	Lys	Ala	Val	Pro	Gly	Arg	Leu	Phe	Pro	Ile	
	305					310					315				320		
	caa	ctg	gca	gaa	aac	cag	ttg	ctg	ctc	gac	gac	tcc	tac	aac	gcc	aat	1008
	Gln	Leu	Ala	Glu	Gln	Leu	Leu	Leu	Leu	Asp	Ser	Tyr	Asn	Ala	Asn		
				325						330				335			
	gtc	ggt	tca	atg	act	gca	gca	gtc	cag	gta	ctg	gct	gaa	atg	ccg	ggc	1056
	Val	Gly	Ser	Met	Thr	Ala	Ala	Val	Gln	Val	Leu	Ala	Glu	Met	Pro	Gly	
				340					345					350			
	tac	cgc	gtg	ctg	gtg	gtg	ggc	gat	atg	gcg	gaa	ctg	ggc	gct	gaa	agc	1104
	Tyr	Arg	Val	Leu	Val	Val	Gly	Asp	Met	Ala	Glu	Leu	Gly	Ala	Glu	Ser	
			355					360					365				
	gaa	gcc	tgc	cat	gta	cag	gtg	ggc	gag	gcg	gca	aaa	gct	gct	ggt	att	1152
	Glu	Ala	Cys	His	Val	Gln	Val	Gly	Glu	Ala	Ala	Lys	Ala	Ala	Gly	Ile	
	370					375						380					
	gac	cgc	gtg	tta	agc	gtg	ggt	aaa	caa	agc	cat	gct	atc	agc	acc	gcc	1200
	Asp	Arg	Val	Leu	Ser	Val	Gly	Lys	Gln	Ser	His	Ala	Ile	Ser	Thr	Ala	
	385					390					395				400		
	agc	ggc	ggt	ggc	gaa	cat	ttt	gct	gat	aaa	act	gcg	tta	att	acg	cgt	1248
	Ser	Gly	Val	Gly	Glu	His	Phe	Ala	Asp	Lys	Thr	Ala	Leu	Ile	Thr	Arg	
				405						410					415		
	ctt	aaa	tca	ctg	att	gct	gag	caa	cag	gta	att	acg	att	tta	ggt	aag	1296
	Leu	Lys	Ser	Leu	Ile	Ala	Glu	Gln	Gln	Val	Ile	Thr	Ile	Leu	Val	Lys	
				420					425					430			
	ggt	tca	cgt	agt	gcc	gcc	atg	gaa	gag	gta	gta	cgc	gct	tta	cag	gag	1344
	Gly	Ser	Arg	Ser	Ala	Ala	Met	Glu	Glu	Val	Val	Arg	Ala	Leu	Gln	Glu	
			435					440					445				
	aat	ggg	aca	tgt	tag												1359
	Asn	Gly	Thr	Cys													
	450																

76

	acg cgg gtc aca atc atc cct aat aat gtt cct ccg cat cgt ccc cag	144
	Thr Arg Val Thr Ile Ile Pro Asn Asn Val Pro Pro His Arg Pro Gln	
	35 40 45	
	ccg gaa gcg aac agc gtg cag cgt aaa cac atg ctt gaa ctg gcg att	192
	Pro Glu Ala Asn Ser Val Gln Arg Lys His Met Leu Glu Leu Ala Ile	
	50 55 60	
	gcc gac aag cca tta ttt act ctt gat gaa cgc gag cta aag cgc aat	240
	Ala Asp Lys Pro Leu Phe Thr Leu Asp Glu Arg Glu Leu Lys Arg Asn	
	65 70 75 80	
	gcc ccc tct tac act gcg caa aca ctg aaa gag tgg ccg cag gaa caa	288
	Ala Pro Ser Tyr Thr Ala Gln Thr Leu Lys Glu Trp Arg Gln Glu Gln	
	85 90 95	
	gga ccg gac gtg ccg ctg gcg ttt att att ggt cag gat tca ttg ctg	336
	Gly Pro Asp Val Pro Leu Ala Phe Ile Ile Gly Gln Asp Ser Leu Leu	
	100 105 110	
	acc ttt ccg acc tgg tac gaa tac gaa acg ata ctc gac aat gca cat	384
	Thr Phe Pro Thr Trp Tyr Glu Tyr Glu Thr Ile Leu Asp Asn Ala His	
	115 120 125	
	ttg atc gtc tgt ccg cgt cca ggt tac cca ctt gaa atg gcg caa ccg	432
	Leu Ile Val Cys Arg Arg Pro Gly Tyr Pro Leu Glu Met Ala Gln Pro	
	130 135 140	
	caa tac cag caa tgg ctg gaa gat cat ttg aca cat aac ccg gaa gat	480
	Gln Tyr Gln Gln Trp Leu Glu Asp His Leu Thr His Asn Pro Glu Asp	
	145 150 155 160	
	ctt cac ctt cag cct gcc ggt aaa att tat ctg gct gaa acg ccg tgg	528
	Leu His Leu Gln Pro Ala Gly Lys Ile Tyr Leu Ala Glu Thr Pro Trp	
	165 170 175	
	ttt aac atc tcg gcg acc atc atc cgc gaa cgt ttg caa aac ggt gaa	576
	Phe Asn Ile Ser Ala Thr Ile Ile Arg Glu Arg Leu Gln Asn Gly Glu	
	180 185 190	
	tcg tgt gag gat tta ttg ccg gaa ccg gta ttg act tac att aac caa	624
	Ser Cys Glu Asp Leu Leu Pro Glu Pro Val Leu Thr Tyr Ile Asn Gln	
	195 200 205	
	caa ggc ttg tat cgc tga	642
	Gln Gly Leu Tyr Arg	
	210	
	<210> 40	
	<211> 213	
	<212> PRT	
[0409]	<213> 大肠杆菌	
	<400> 40	
	Met Lys Ser Leu Gln Ala Leu Phe Gly Gly Thr Phe Asp Pro Val His	
	1 5 10 15	
	Tyr Gly His Leu Lys Pro Val Glu Thr Leu Ala Asn Leu Ile Gly Leu	
	20 25 30	
	Thr Arg Val Thr Ile Ile Pro Asn Asn Val Pro Pro His Arg Pro Gln	
	35 40 45	
	Pro Glu Ala Asn Ser Val Gln Arg Lys His Met Leu Glu Leu Ala Ile	
	50 55 60	
	Ala Asp Lys Pro Leu Phe Thr Leu Asp Glu Arg Glu Leu Lys Arg Asn	
	65 70 75 80	
	Ala Pro Ser Tyr Thr Ala Gln Thr Leu Lys Glu Trp Arg Gln Glu Gln	
	85 90 95	
	Gly Pro Asp Val Pro Leu Ala Phe Ile Ile Gly Gln Asp Ser Leu Leu	
	100 105 110	
	Thr Phe Pro Thr Trp Tyr Glu Tyr Glu Thr Ile Leu Asp Asn Ala His	
	115 120 125	
	Leu Ile Val Cys Arg Arg Pro Gly Tyr Pro Leu Glu Met Ala Gln Pro	
	130 135 140	
	Gln Tyr Gln Gln Trp Leu Glu Asp His Leu Thr His Asn Pro Glu Asp	
	145 150 155 160	
	Leu His Leu Gln Pro Ala Gly Lys Ile Tyr Leu Ala Glu Thr Pro Trp	
	165 170 175	
	Phe Asn Ile Ser Ala Thr Ile Ile Arg Glu Arg Leu Gln Asn Gly Glu	
	180 185 190	
	Ser Cys Glu Asp Leu Leu Pro Glu Pro Val Leu Thr Tyr Ile Asn Gln	
	195 200 205	
	Gln Gly Leu Tyr Arg	
	210	
	<210> 41	
	<211> 333	
	<212> DNA	
	<213> 大肠杆菌	
	<220>	
	<221> CDS	
	<222> (1) .. (333)	
	<223> 大肠杆菌必需基因 rplV	
	<400> 41	

	atg gaa act atc gct aaa cat cgc cat gct cgt tct tct gct cag aag	48
	Met Glu Thr Ile Ala Lys His Arg His Ala Arg Ser Ser Ala Gln Lys	
	1 5 10 15	
	gtt cgc ctt gtt gct gac ctg att cgc ggt aag aaa gtg tcg cag gct	96
	Val Arg Leu Val Ala Asp Leu Ile Arg Gly Lys Lys Val Ser Gln Ala	
	20 25 30	
	ctg gat att ttg acc tac acc aac aag aaa gcg gct gta ctg gtc aag	144
	Leu Asp Ile Leu Thr Tyr Thr Asn Lys Lys Ala Ala Val Leu Val Lys	
	35 40 45	
	aaa gtt ctg gaa tct gcc att gct aac gct gaa cac aac gat ggc gct	192
	Lys Val Leu Glu Ser Ala Ile Ala Asn Ala Glu His Asn Asp Gly Ala	
	50 55 60	
	gac att gac gat ctg aaa gtt acg aaa att ttc gta gac gaa ggc ccg	240
	Asp Ile Asp Asp Leu Lys Val Thr Lys Ile Phe Val Asp Glu Gly Pro	
	65 70 75 80	
	agc atg aag cgc att atg ccg cgt gca aaa ggt cgt gca gat cgc atc	288
	Ser Met Lys Arg Ile Met Pro Arg Ala Lys Gly Arg Ala Asp Arg Ile	
	85 90 95	
	ctg aag cgc acc agc cac atc act gtg gtt gtg tcc gat cgc tga	333
	Leu Lys Arg Thr Ser His Ile Thr Val Val Ser Asp Arg	
	100 105 110	
	<210> 42	
	<211> 110	
	<212> PRT	
	<213> 大肠杆菌	
	<400> 42	
	Met Glu Thr Ile Ala Lys His Arg His Ala Arg Ser Ser Ala Gln Lys	
	1 5 10 15	
	Val Arg Leu Val Ala Asp Leu Ile Arg Gly Lys Lys Val Ser Gln Ala	
	20 25 30	
	Leu Asp Ile Leu Thr Tyr Thr Asn Lys Lys Ala Ala Val Leu Val Lys	
	35 40 45	
	Lys Val Leu Glu Ser Ala Ile Ala Asn Ala Glu His Asn Asp Gly Ala	
	50 55 60	
	Asp Ile Asp Asp Leu Lys Val Thr Lys Ile Phe Val Asp Glu Gly Pro	
	65 70 75 80	
	Ser Met Lys Arg Ile Met Pro Arg Ala Lys Gly Arg Ala Asp Arg Ile	
	85 90 95	
[0410]	Leu Lys Arg Thr Ser His Ile Thr Val Val Ser Asp Arg	
	100 105 110	
	<210> 43	
	<211> 471	
	<212> DNA	
	<213> 大肠杆菌	
	<220>	
	<221> CDS	
	<222> (1)..(471)	
	<223> 大肠杆菌必需基因 rpsG	
	<400> 43	
	atg cca cgt cgt cgc gtc att ggt cag cgt aaa att ctg ccg gat ccg	48
	Met Pro Arg Arg Arg Val Ile Gly Gln Arg Lys Ile Leu Pro Asp Pro	
	1 5 10 15	
	aag ttc gga tca gaa ctg ctg gct aaa ttt gta aat atc ctg atg gta	96
	Lys Phe Gly Ser Glu Leu Leu Ala Lys Phe Val Asn Ile Leu Met Val	
	20 25 30	
	gat ggt aaa aaa tct act gct gaa tct atc gta tac agc gcg ctg gag	144
	Asp Gly Lys Lys Ser Thr Ala Glu Ser Ile Val Tyr Ser Ala Leu Glu	
	35 40 45	
	acc ctg gct cag cgc tct ggt aaa tct gaa ctg gaa gca ttc gaa gta	192
	Thr Leu Ala Gln Arg Ser Gly Lys Ser Glu Leu Glu Ala Phe Glu Val	
	50 55 60	
	gct ctc gaa aac gtg cgc ccg act gta gaa gtt aag tct cgc cgc gtt	240
	Ala Leu Glu Asn Val Arg Pro Thr Val Glu Val Lys Ser Arg Arg Val	
	65 70 75 80	
	ggt ggt tct act tat cag gta cca gtt gaa gtc cgt ccg gtt cgt cgt	288
	Gly Gly Ser Thr Tyr Gln Val Pro Val Glu Val Arg Pro Val Arg Arg	
	85 90 95	
	aat gct ctg gca atg cgt tgg atc gtt gaa gct gct cgt aaa cgc ggt	336
	Asn Ala Leu Ala Met Arg Trp Ile Val Glu Ala Ala Arg Lys Arg Gly	
	100 105 110	
	gat aaa tcc atg gct ctg cgc ctg gcg aac gaa ctt tct gat gct gca	384
	Asp Lys Ser Met Ala Leu Arg Leu Ala Asn Glu Leu Ser Asp Ala Ala	
	115 120 125	
	gaa aac aaa ggt act gca gtt aag aaa cgt gaa gac gtt cac cgt atg	432
	Glu Asn Lys Gly Thr Ala Val Lys Lys Arg Glu Asp Val His Arg Met	
	130 135 140	
	gcc gaa gcc aac aag gcg ttc gca cac tac cgt tgg taa	471
	Ala Glu Ala Asn Lys Ala Phe Ala His Tyr Arg Trp	

	145		150		155	
	<210> 44					
	<211> 156					
	<212> PRT					
	<213> 大肠杆菌					
	<400> 44					
	Met Pro Arg Arg Arg Val Ile Gly Gln Arg Lys Ile Leu Pro Asp Pro					
	1 5 10 15					
	Lys Phe Gly Ser Glu Leu Leu Ala Lys Phe Val Asn Ile Leu Met Val					
	20 25 30					
	Asp Gly Lys Lys Ser Thr Ala Glu Ser Ile Val Tyr Ser Ala Leu Glu					
	35 40 45					
	Thr Leu Ala Gln Arg Ser Gly Lys Ser Glu Leu Glu Ala Phe Glu Val					
	50 55 60					
	Ala Leu Glu Asn Val Arg Pro Thr Val Glu Val Lys Ser Arg Arg Val					
	65 70 75 80					
	Gly Gly Ser Thr Tyr Gln Val Pro Val Glu Val Arg Pro Val Arg Arg					
	85 90 95					
	Asn Ala Leu Ala Met Arg Trp Ile Val Glu Ala Ala Arg Lys Arg Gly					
	100 105 110					
	Asp Lys Ser Met Ala Leu Arg Leu Ala Asn Glu Leu Ser Asp Ala Ala					
	115 120 125					
	Glu Asn Lys Gly Thr Ala Val Lys Lys Arg Glu Asp Val His Arg Met					
	130 135 140					
	Ala Glu Ala Asn Lys Ala Phe Ala His Tyr Arg Trp					
	145 150 155					
	<210> 45					
	<211> 2475					
	<212> DNA					
	<213> 酿酒酵母 (S. cerevisiae)					
	<220>					
	<221> CDS					
	<222> (1)..(2475)					
	<223> 酿酒酵母必需基因 FOL1 YNL256W SGDID:S000005200					
	<400> 45					
	atg tca aag cta ttt tct act gtc aat tct gca aga cat agt gta cca					48
	1 Met Ser Lys Leu Phe Ser Thr Val Asn Ser Ala Arg His Ser Val Pro					
	5 10 15					
[0411]	cta ggc ggc atg aga gat tat gtg cac att aag aaa cta gag atg aat					96
	Leu Gly Gly Met Arg Asp Tyr Val His Ile Lys Lys Leu Glu Met Asn					
	20 25 30					
	aca gtt ctt ggg cct gat tcc tgg aat caa tta atg cct cag aaa tgt					144
	Thr Val Leu Gly Pro Asp Ser Trp Asn Gln Leu Met Pro Gln Lys Cys					
	35 40 45					
	cta cta agc tta gat atg ggt aca gat ttt agt aaa tct gcg gct acg					192
	Leu Leu Ser Leu Asp Met Gly Thr Asp Phe Ser Lys Ser Ala Ala Thr					
	50 55 60					
	gat gat ttg aaa tat tct cta aat tat gca gtt att tct cgt gat ttg					240
	Asp Asp Leu Lys Tyr Ser Leu Asn Tyr Ala Val Ile Ser Arg Asp Leu					
	65 70 75 80					
	acg aat ttc gtc agc aaa aaa aag aat tgg ggt tct gtt tct aat ttg					288
	Thr Asn Phe Val Ser Lys Lys Lys Asn Trp Gly Ser Val Ser Asn Leu					
	85 90 95					
	gct aaa tct gtg tct caa ttt gtt atg gac aaa tat tct ggt gtc gag					336
	Ala Lys Ser Val Ser Gln Phe Val Met Asp Lys Tyr Ser Gly Val Glu					
	100 105 110					
	tgt ctg aat tta gaa gtg cag gcg gat aca acg cat att aga agt gac					384
	Cys Leu Asn Leu Glu Val Gln Ala Asp Thr Thr His Ile Arg Ser Asp					
	115 120 125					
	cac ata tct tgt att att caa caa gaa aga ggg aat cca gaa tca cag					432
	His Ile Ser Cys Ile Ile Gln Gln Glu Arg Gly Asn Pro Glu Ser Gln					
	130 135 140					
	gaa ttt gac gtt gtt agg ata tct gag tta aaa atg ttg act ttg att					480
	Glu Phe Asp Val Val Arg Ile Ser Glu Leu Lys Met Leu Thr Leu Ile					
	145 150 155					
	ggt gtt ttc acc ttt gag aga ctt aag aaa cag tat gta act ttg gat					528
	Gly Val Phe Thr Phe Glu Arg Leu Lys Lys Gln Tyr Val Thr Leu Asp					
	165 170 175					
	ata aag ttg cct tgg cca aag aaa gcc gaa ttg cca ccg cca gtg caa					576
	Ile Lys Leu Pro Trp Pro Lys Lys Ala Glu Leu Pro Pro Val Gln					
	180 185 190					
	agc ata att gat aac gtt gtc aag ttt gtg gag gaa tca aat ttc aag					624
	Ser Ile Ile Asp Asn Val Val Lys Phe Val Glu Glu Ser Asn Phe Lys					
	195 200 205					
	act gtg gaa gct ctt gta gaa tct gtg tca gct gtt att gcc cat aac					672
	Thr Val Glu Ala Leu Val Glu Ser Val Ser Ala Val Ile Ala His Asn					
	210 215 220					
	gag tat ttt caa aag ttt cca gat tcg cct ttg gtg gtg aag gtt ttg					720

[0412]	Glu	Tyr	Phe	Gln	Lys	Phe	Pro	Asp	Ser	Pro	Leu	Val	Val	Lys	Val	Leu	
	225					230					235					240	
	aaa	tta	aac	gca	atc	aca	gcc	aca	gaa	ggt	ggt	ggt	gta	agc	tgt	att	768
	Lys	Leu	Asn	Ala	Ile	Thr	Ala	Thr	Glu	Gly	Val	Gly	Val	Ser	Cys	Ile	
					245					250					255		
	aga	gag	ccc	agg	gag	att	gcg	atg	gta	aat	att	cca	tat	ctt	tcc	tcc	816
	Arg	Glu	Pro	Arg	Glu	Ile	Ala	Met	Val	Asn	Ile	Pro	Tyr	Leu	Ser	Ser	
				260					265					270			
	ata	cat	gaa	tgc	tct	gat	att	aag	ttc	caa	ttg	tct	tca	tca	caa	aac	864
	Ile	His	Glu	Ser	Ser	Asp	Ile	Lys	Phe	Gln	Leu	Ser	Ser	Ser	Gln	Asn	
			275					280					285				
	act	cct	att	gag	ggt	aaa	aat	aca	tggt	aaa	aga	gcg	ttt	tta	gcg	ttt	912
	Thr	Pro	Ile	Glu	Gly	Lys	Asn	Thr	Trp	Lys	Arg	Ala	Phe	Leu	Ala	Phe	
		290					295					300					
	ggt	tca	aac	att	ggg	gac	cgt	ttc	aaa	cac	att	caa	atg	gcg	ttg	caa	960
	Gly	Ser	Asn	Ile	Gly	Asp	Arg	Phe	Lys	His	Ile	Gln	Met	Ala	Leu	Gln	
	305					310					315				320		
	tta	tta	tca	agg	gaa	aaa	acg	ggt	aaa	tta	cgg	aat	att	tgc	tct	att	1008
	Leu	Leu	Ser	Arg	Glu	Lys	Thr	Val	Lys	Leu	Arg	Asn	Ile	Ser	Ser	Ile	
					325				330					335			
	ttt	gaa	agt	gaa	cca	atg	tat	ttc	aaa	gat	caa	acc	cct	ttc	atg	aat	1056
	Phe	Glu	Ser	Glu	Pro	Met	Tyr	Phe	Lys	Asp	Gln	Thr	Pro	Phe	Met	Asn	
				340					345					350			
	ggg	tgt	ggt	gag	gtg	gag	aca	tta	ctg	acc	cca	agc	gaa	tta	aaa		1104
	Gly	Cys	Val	Glu	Val	Glu	Thr	Leu	Leu	Thr	Pro	Ser	Glu	Leu	Leu	Lys	
		355						360					365				
	ttg	tgt	aaa	aaa	att	gaa	tat	gaa	gag	ttg	caa	aga	gtc	aag	cat	ttt	1152
	Leu	Cys	Lys	Lys	Ile	Glu	Tyr	Glu	Glu	Leu	Gln	Arg	Val	Lys	His	Phe	
		370					375					380					
	gat	aat	ggt	cgc	aga	aca	ata	gat	ctg	gat	att	ggt	atg	ttt	ttg	aat	1200
	Asp	Asn	Gly	Pro	Arg	Thr	Ile	Asp	Leu	Asp	Ile	Val	Met	Phe	Leu	Asn	
		385				390					395				400		
	agc	gcc	gga	gaa	gat	att	ata	gta	aat	gaa	cgc	gat	ttg	aat	ata	ccg	1248
	Ser	Ala	Gly	Glu	Asp	Ile	Ile	Val	Asn	Glu	Pro	Asp	Leu	Asn	Ile	Pro	
				405						410				415			
	cat	cct	aga	atg	ctg	gag	agg	act	ttc	ggt	ctt	gag	ccg	tta	tgt	gaa	1296
	His	Pro	Arg	Met	Glu	Glu	Arg	Thr	Phe	Val	Leu	Glu	Pro	Leu	Cys	Glu	
				420					425					430			
	tta	ata	tcc	ccc	ggt	cac	ctt	cat	cct	gtg	aca	gcg	gaa	ccc	att	gta	1344
	Leu	Ile	Ser	Pro	Val	His	Leu	His	Pro	Val	Thr	Ala	Glu	Pro	Ile	Val	
			435					440					445				
	gac	cat	tta	aaa	cag	tta	tac	gac	aaa	cag	cat	gat	gaa	gat	acc	tta	1392
	Asp	His	Leu	Lys	Gln	Leu	Tyr	Asp	Lys	Gln	His	Asp	Glu	Asp	Thr	Leu	
		450					455					460					
	tggt	aaa	tta	ggt	cca	ttg	cct	tat	cgt	agt	ggt	gtg	gag	cct	aga	ttt	1440
	Trp	Lys	Leu	Val	Pro	Leu	Pro	Tyr	Arg	Ser	Gly	Val	Glu	Pro	Arg	Phe	
		465				470					475				480		
	ttg	aaa	ttc	aag	acc	gct	aca	aaa	ctt	gac	gaa	ttt	act	gga	gaa	aca	1488
	Leu	Lys	Phe	Lys	Thr	Ala	Thr	Lys	Leu	Asp	Glu	Phe	Thr	Gly	Glu	Thr	
					485					490				495			
	aac	aga	att	act	ggt	tca	cct	aca	tat	atc	atg	gct	atc	ttc	aac	gct	1536
	Asn	Arg	Ile	Thr	Val	Ser	Pro	Thr	Tyr	Ile	Met	Ala	Ile	Phe	Asn	Ala	
				500					505					510			
	aca	cca	gat	tca	ttt	tcc	gat	gga	ggt	gag	cat	ttt	gcg	gac	att	gaa	1584
	Thr	Pro	Asp	Ser	Phe	Ser	Asp	Gly	Gly	Glu	His	Phe	Ala	Asp	Ile	Glu	
			515					520					525				
	agt	caa	ttg	aat	gat	atc	att	aaa	ttg	tgt	aaa	gac	gca	tta	tat	ttg	1632
	Ser	Gln	Leu	Asn	Asp	Ile	Ile	Lys	Leu	Cys	Lys	Asp	Ala	Leu	Tyr	Leu	
		530					535					540					
	cat	gag	agc	gtc	atc	atc	gac	ggt	gga	ggg	tgt	tct	acc	agg	cct	aac	1680
	His	Glu	Ser	Val	Ile	Ile	Asp	Val	Gly	Gly	Cys	Ser	Thr	Arg	Pro	Asn	
		545				550					555				560		
	tct	att	cag	gcg	tct	gag	gaa	gaa	gaa	ata	cgc	agg	tct	atc	cca	tta	1728
	Ser	Ile	Gln	Ala	Ser	Glu	Glu	Glu	Glu	Ile	Arg	Arg	Ser	Ile	Pro	Leu	
					565					570					575		
	att	aag	gcc	att	aga	gaa	agc	act	gag	tta	cgc	caa	gat	aaa	gtc	ata	1776
	Ile	Lys	Ala	Ile	Arg	Glu	Ser	Thr	Glu	Leu	Pro	Gln	Asp	Lys	Val	Ile	
				580					585					590			
	cta	tcc	att	gat	act	tat	cgt	tcc	aat	gtc	gct	aaa	gaa	gcg	att	aaa	1824
	Leu	Ser	Ile	Asp	Thr	Tyr	Arg	Ser	Asn	Val	Ala	Lys	Glu	Ala	Ile	Lys	
			595					600					605				
	ggt	gga	gtg	gat	att	att	aat	gat	att	tgc	gga	ggt	tta	ttt	gac	agc	1872
	Val	Gly	Val	Asp	Ile	Ile	Asn	Asp	Ile	Ser	Gly	Gly	Leu	Phe	Asp	Ser	
		610					615					620					
	aac	atg	ttt	gcc	gta	att	gca	gag	aac	cca	gaa	att	tgt	tat	att	tta	1920
	Asn	Met	Phe	Ala	Val	Ile	Ala	Glu	Asn	Pro	Glu	Ile	Cys	Tyr	Ile	Leu	
		625				630					635					640	
	tca	cac	aca	cgt	ggt	gat	att	tca	acg	atg	aat	agg	ctg	gcg	cat	tac	1968

	Ser	His	Thr	Arg	Gly	Asp	Ile	Ser	Thr	Met	Asn	Arg	Leu	Ala	His	Tyr	
					645					650					655		
	gaa	aat	ttt	gca	ttg	ggt	gat	tct	att	cag	caa	gaa	ttt	gtt	cat	aat	2016
	Glu	Asn	Phe	Ala	Leu	Gly	Asp	Ser	Ile	Gln	Gln	Glu	Phe	Val	His	Asn	
				660					665					670			
	acc	gac	att	cag	cag	cta	gac	gac	ttg	aaa	gac	aaa	aca	gtg	tta	atc	2064
	Thr	Asp	Ile	Gln	Gln	Leu	Asp	Asp	Leu	Lys	Asp	Lys	Thr	Val	Leu	Ile	
			675					680					685				
	agg	aat	gtt	ggt	caa	gaa	att	ggc	gaa	agg	tat	atc	aaa	gcg	att	gat	2112
	Arg	Asn	Val	Gly	Gln	Glu	Ile	Gly	Glu	Arg	Tyr	Ile	Lys	Ala	Ile	Asp	
		690					695					700					
	aat	gga	gta	aag	cgc	tgg	caa	att	cta	atc	gac	cct	gga	ctt	ggt	ttt	2160
	Asn	Gly	Val	Lys	Arg	Trp	Gln	Ile	Leu	Ile	Asp	Pro	Gly	Leu	Gly	Phe	
	705				710						715				720		
	gct	aag	acc	tgg	aag	caa	aac	tta	caa	att	att	aga	cat	atc	ccc	att	2208
	Ala	Lys	Thr	Trp	Lys	Gln	Asn	Leu	Gln	Ile	Ile	Arg	His	Ile	Pro	Ile	
				725					730					735			
	tta	aag	aac	tac	tca	ttc	acc	atg	aac	tca	aac	aat	tcg	caa	gtg	tat	2256
	Leu	Lys	Asn	Tyr	Ser	Phe	Thr	Met	Asn	Ser	Asn	Asn	Ser	Gln	Val	Tyr	
				740					745					750			
	gtt	aac	ctc	aga	aat	atg	ccc	gtt	tta	ttg	ggt	cca	tcg	cgc	aaa	aaa	2304
	Val	Asn	Leu	Arg	Asn	Met	Pro	Val	Leu	Leu	Gly	Pro	Ser	Arg	Lys	Lys	
			755				760					765					
	ttc	att	gga	cat	atc	aca	aaa	gat	gtg	gat	gcg	aag	caa	aga	gac	ttt	2352
	Phe	Ile	Gly	His	Ile	Thr	Lys	Asp	Val	Asp	Ala	Lys	Gln	Arg	Asp	Phe	
		770				775					780						
	gct	act	gga	gcg	gtg	gta	gcg	tcg	tgt	att	ggt	ttc	ggc	agc	gac	atg	2400
	Ala	Thr	Gly	Ala	Val	Val	Ala	Ser	Cys	Ile	Gly	Phe	Gly	Ser	Asp	Met	
	785				790					795				800			
	gtt	agg	gtc	cat	gac	gtt	aaa	aat	tgt	tcg	aag	agc	att	aaa	tta	gca	2448
	Val	Arg	Val	His	Asp	Val	Lys	Asn	Cys	Ser	Lys	Ser	Ile	Lys	Leu	Ala	
				805						810				815			
	gat	gct	att	tat	aaa	ggt	ttg	gaa	taa								2475
	Asp	Ala	Ile	Tyr	Lys	Gly	Leu	Glu									
				820													
	<210>	46															
	<211>	824															
	<212>	PRT															
[0413]	<213>	酿酒酵母															
	<400>	46															
	Met	Ser	Lys	Leu	Phe	Ser	Thr	Val	Asn	Ser	Ala	Arg	His	Ser	Val	Pro	
	1			5					10					15			
	Leu	Gly	Gly	Met	Arg	Asp	Tyr	Val	His	Ile	Lys	Lys	Leu	Glu	Met	Asn	
				20					25					30			
	Thr	Val	Leu	Gly	Pro	Asp	Ser	Trp	Asn	Gln	Leu	Met	Pro	Gln	Lys	Cys	
			35					40					45				
	Leu	Leu	Ser	Leu	Asp	Met	Gly	Thr	Asp	Phe	Ser	Lys	Ser	Ala	Ala	Thr	
		50					55					60					
	Asp	Asp	Leu	Lys	Tyr	Ser	Leu	Asn	Tyr	Ala	Val	Ile	Ser	Arg	Asp	Leu	
	65					70				75					80		
	Thr	Asn	Phe	Val	Ser	Lys	Lys	Lys	Asn	Trp	Gly	Ser	Val	Ser	Asn	Leu	
				85					90					95			
	Ala	Lys	Ser	Val	Ser	Gln	Phe	Val	Met	Asp	Lys	Tyr	Ser	Gly	Val	Glu	
				100					105					110			
	Cys	Leu	Asn	Leu	Glu	Val	Gln	Ala	Asp	Thr	Thr	His	Ile	Arg	Ser	Asp	
			115					120					125				
	His	Ile	Ser	Cys	Ile	Ile	Gln	Gln	Glu	Arg	Gly	Asn	Pro	Glu	Ser	Gln	
		130					135					140					
	Glu	Phe	Asp	Val	Val	Arg	Ile	Ser	Glu	Leu	Lys	Met	Leu	Thr	Leu	Ile	
	145					150					155				160		
	Gly	Val	Phe	Thr	Phe	Glu	Arg	Leu	Lys	Lys	Gln	Tyr	Val	Thr	Leu	Asp	
				165						170				175			
	Ile	Lys	Leu	Pro	Trp	Pro	Lys	Lys	Ala	Glu	Leu	Pro	Pro	Pro	Val	Gln	
				180					185					190			
	Ser	Ile	Ile	Asp	Asn	Val	Val	Lys	Phe	Val	Glu	Glu	Ser	Asn	Phe	Lys	
		195						200					205				
	Thr	Val	Glu	Ala	Leu	Val	Glu	Ser	Val	Ser	Ala	Val	Ile	Ala	His	Asn	
		210					215					220					
	Glu	Tyr	Phe	Gln	Lys	Phe	Pro	Asp	Ser	Pro	Leu	Val	Val	Lys	Val	Leu	
	225					230					235					240	
	Lys	Leu	Asn	Ala	Ile	Thr	Ala	Thr	Glu	Gly	Val	Gly	Val	Ser	Cys	Ile	
				245						250				255			
	Arg	Glu	Pro	Arg	Glu	Ile	Ala	Met	Val	Asn	Ile	Pro	Tyr	Leu	Ser	Ser	
				260					265					270			
	Ile	His	Glu	Ser	Ser	Asp	Ile	Lys	Phe	Gln	Leu	Ser	Ser	Ser	Gln	Asn	
		275						280					285				
	Thr	Pro	Ile	Glu	Gly	Lys	Asn	Thr	Trp	Lys	Arg	Ala	Phe	Leu	Ala	Phe	
		290					295					300					

[0414]

Gly Ser Asn Ile Gly Asp Arg Phe Lys His Ile Gln Met Ala Leu Gln
 305 310 315 320
 Leu Leu Ser Arg Glu Lys Thr Val Lys Leu Arg Asn Ile Ser Ser Ile
 325 330 335
 Phe Glu Ser Glu Pro Met Tyr Phe Lys Asp Gln Thr Pro Phe Met Asn
 340 345 350
 Gly Cys Val Glu Val Glu Thr Leu Thr Pro Ser Glu Leu Leu Lys
 355 360 365
 Leu Cys Lys Lys Ile Glu Tyr Glu Glu Leu Gln Arg Val Lys His Phe
 370 375 380
 Asp Asn Gly Pro Arg Thr Ile Asp Leu Asp Ile Val Met Phe Leu Asn
 385 390 395 400
 Ser Ala Gly Glu Asp Ile Ile Val Asn Glu Pro Asp Leu Asn Ile Pro
 405 410 415
 His Pro Arg Met Leu Glu Arg Thr Phe Val Leu Glu Pro Leu Cys Glu
 420 425 430
 Leu Ile Ser Pro Val His Leu His Pro Val Thr Ala Glu Pro Ile Val
 435 440 445
 Asp His Leu Lys Gln Leu Tyr Asp Lys Gln His Asp Glu Asp Thr Leu
 450 455 460
 Trp Lys Leu Val Pro Leu Pro Tyr Arg Ser Gly Val Glu Pro Arg Phe
 465 470 475 480
 Leu Lys Phe Lys Thr Ala Thr Lys Leu Asp Glu Phe Thr Gly Glu Thr
 485 490 495
 Asn Arg Ile Thr Val Ser Pro Thr Tyr Ile Met Ala Ile Phe Asn Ala
 500 505 510
 Thr Pro Asp Ser Phe Ser Asp Gly Glu His Phe Ala Asp Ile Glu
 515 520 525
 Ser Gln Leu Asn Asp Ile Ile Lys Leu Cys Lys Asp Ala Leu Tyr Leu
 530 535 540
 His Glu Ser Val Ile Ile Asp Val Gly Gly Cys Ser Thr Arg Pro Asn
 545 550 555 560
 Ser Ile Gln Ala Ser Glu Glu Glu Glu Ile Arg Arg Ser Ile Pro Leu
 565 570 575
 Ile Lys Ala Ile Arg Glu Ser Thr Glu Leu Pro Gln Asp Lys Val Ile
 580 585 590
 Leu Ser Ile Asp Thr Tyr Arg Ser Asn Val Ala Lys Glu Ala Ile Lys
 595 600 605
 Val Gly Val Asp Ile Ile Asn Asp Ile Ser Gly Gly Leu Phe Asp Ser
 610 615 620
 Asn Met Phe Ala Val Ile Ala Glu Asn Pro Glu Ile Cys Tyr Ile Leu
 625 630 635 640
 Ser His Thr Arg Gly Asp Ile Ser Thr Met Asn Arg Leu Ala His Tyr
 645 650 655
 Glu Asn Phe Ala Leu Gly Asp Ser Ile Gln Gln Glu Phe Val His Asn
 660 665 670
 Thr Asp Ile Gln Gln Leu Asp Asp Leu Lys Asp Lys Thr Val Leu Ile
 675 680 685
 Arg Asn Val Gly Gln Glu Ile Gly Glu Arg Tyr Ile Lys Ala Ile Asp
 690 695 700
 Asn Gly Val Lys Arg Trp Gln Ile Leu Ile Asp Pro Gly Leu Gly Phe
 705 710 715 720
 Ala Lys Thr Trp Lys Gln Asn Leu Gln Ile Ile Arg His Ile Pro Ile
 725 730 735
 Leu Lys Asn Tyr Ser Phe Thr Met Asn Ser Asn Ser Gln Val Tyr
 740 745 750
 Val Asn Leu Arg Asn Met Pro Val Leu Leu Gly Pro Ser Arg Lys Lys
 755 760 765
 Phe Ile Gly His Ile Thr Lys Asp Val Asp Ala Lys Gln Arg Asp Phe
 770 775 780
 Ala Thr Gly Ala Val Val Ala Ser Cys Ile Gly Phe Gly Ser Asp Met
 785 790 795 800
 Val Arg Val His Asp Val Lys Asn Cys Ser Lys Ser Ile Lys Leu Ala
 805 810 815
 Asp Ala Ile Tyr Lys Gly Leu Glu
 820
 <210> 47
 <211> 669
 <212> DNA
 <213> 酿酒酵母
 <220>
 <221> CDS
 <222> (1) .. (669)
 <223> 酿酒酵母必需基因 MED7 YOL135C SGDID:S000005495
 <400> 47
 atg tcc aat gat cct ggt aat gag gtt agc tcc ttg tac cca cct cca
 Met Ser Asn Asp Pro Gly Asn Glu Val Ser Ser Leu Tyr Pro Pro Pro
 1 5 10 15

48

	ccc	cca	tac	gtg	aag	ttt	ttc	aca	cag	tcc	aat	ctg	gaa	aaa	tta	cca	96
	Pro	Pro	Tyr	Val	Lys	Phe	Phe	Thr	Gln	Ser	Asn	Leu	Glu	Lys	Leu	Pro	
				20					25					30			
	aaa	tac	aaa	gaa	aag	aag	gcc	gca	agt	gcg	aag	cag	acg	gct	cca	aat	144
	Lys	Tyr	Lys	Glu	Lys	Lys	Ala	Ala	Ser	Ala	Lys	Gln	Thr	Ala	Pro	Asn	
			35				40						45				
	aac	agt	aac	ggc	ggc	agt	gaa	gag	gag	att	aca	tgt	gcg	cta	gat	tac	192
	Asn	Ser	Asn	Gly	Gly	Ser	Glu	Glu	Glu	Ile	Thr	Cys	Ala	Leu	Asp	Tyr	
		50				55						60					
	cta	atc	cca	ccc	ccc	atg	ccc	aaa	aat	cag	cag	tac	cga	gcg	ttt	ggg	240
	Leu	Ile	Pro	Pro	Pro	Met	Pro	Lys	Asn	Gln	Gln	Tyr	Arg	Ala	Phe	Gly	
		65				70					75					80	
	agt	ata	tgg	cag	gtg	aag	gac	caa	ctg	ccc	gat	tta	gag	agc	atg	gga	288
	Ser	Ile	Trp	Gln	Val	Lys	Asp	Gln	Leu	Pro	Asp	Leu	Glu	Ser	Met	Gly	
				85						90					95		
	ctc	acc	caa	ctt	tat	aag	aaa	tcc	aca	gaa	aac	gaa	agc	aca	aat	tat	336
	Leu	Thr	Gln	Leu	Tyr	Lys	Lys	Ser	Thr	Glu	Asn	Glu	Ser	Thr	Asn	Tyr	
				100					105					110			
	caa	tac	aaa	atc	caa	gag	cta	cgc	aaa	cta	ttg	aaa	tcc	ctc	ctg	cta	384
	Gln	Tyr	Lys	Ile	Gln	Glu	Leu	Arg	Lys	Leu	Leu	Lys	Ser	Leu	Leu	Leu	
			115				120						125				
	aat	tac	ttg	gaa	ctc	atc	ggc	gta	ttg	agc	ata	aat	ccg	gat	atg	tat	432
	Asn	Tyr	Leu	Glu	Leu	Ile	Gly	Val	Leu	Ser	Ile	Asn	Pro	Asp	Met	Tyr	
			130				135						140				
	gag	cgc	aaa	gtg	gag	aac	ata	agg	acc	att	ctt	gtc	aat	att	cac	cac	480
	Glu	Arg	Lys	Val	Glu	Asn	Ile	Arg	Thr	Ile	Leu	Val	Asn	Ile	His	His	
		145				150					155				160		
	cta	ctt	aat	gag	tac	aga	ccc	cat	cag	tcc	cga	gag	tcc	ctt	att	atg	528
	Leu	Leu	Asn	Glu	Tyr	Arg	Pro	His	Gln	Ser	Arg	Glu	Ser	Leu	Ile	Met	
				165						170				175			
	cta	ttg	gaa	gaa	caa	tta	gaa	tac	aag	cgc	ggg	gaa	att	cgc	gag	atc	576
	Leu	Leu	Glu	Glu	Gln	Leu	Glu	Tyr	Lys	Arg	Gly	Glu	Ile	Arg	Glu	Ile	
				180					185					190			
	gag	caa	gta	tgc	aaa	cag	gta	cac	gac	aaa	ctc	acc	agc	ata	cag	gac	624
	Glu	Gln	Val	Cys	Lys	Gln	Val	His	Asp	Lys	Leu	Thr	Ser	Ile	Gln	Asp	
			195				200						205				
	acc	tta	aga	acc	gga	tca	caa	tct	cct	cca	tcg	tcg	tcc	caa	tga		669
	Thr	Leu	Arg	Thr	Gly	Ser	Gln	Ser	Pro	Pro	Ser	Ser	Ser	Gln			
		210					215						220				
[0415]	<210>	48															
	<211>	222															
	<212>	PRT															
	<213>	酿酒酵母															
	<400>	48															
	Met	Ser	Asn	Asp	Pro	Gly	Asn	Glu	Val	Ser	Ser	Leu	Tyr	Pro	Pro	Pro	
	1				5					10					15		
	Pro	Pro	Tyr	Val	Lys	Phe	Phe	Thr	Gln	Ser	Asn	Leu	Glu	Lys	Leu	Pro	
				20					25					30			
	Lys	Tyr	Lys	Glu	Lys	Lys	Ala	Ala	Ser	Ala	Lys	Gln	Thr	Ala	Pro	Asn	
			35				40						45				
	Asn	Ser	Asn	Gly	Gly	Ser	Glu	Glu	Glu	Ile	Thr	Cys	Ala	Leu	Asp	Tyr	
		50				55						60					
	Leu	Ile	Pro	Pro	Pro	Met	Pro	Lys	Asn	Gln	Gln	Tyr	Arg	Ala	Phe	Gly	
		65				70					75					80	
	Ser	Ile	Trp	Gln	Val	Lys	Asp	Gln	Leu	Pro	Asp	Leu	Glu	Ser	Met	Gly	
				85						90					95		
	Leu	Thr	Gln	Leu	Tyr	Lys	Lys	Ser	Thr	Glu	Asn	Glu	Ser	Thr	Asn	Tyr	
				100					105					110			
	Gln	Tyr	Lys	Ile	Gln	Glu	Leu	Arg	Lys	Leu	Leu	Lys	Ser	Leu	Leu	Leu	
			115				120						125				
	Asn	Tyr	Leu	Glu	Leu	Ile	Gly	Val	Leu	Ser	Ile	Asn	Pro	Asp	Met	Tyr	
		130					135						140				
	Glu	Arg	Lys	Val	Glu	Asn	Ile	Arg	Thr	Ile	Leu	Val	Asn	Ile	His	His	
		145				150					155				160		
	Leu	Leu	Asn	Glu	Tyr	Arg	Pro	His	Gln	Ser	Arg	Glu	Ser	Leu	Ile	Met	
				165						170				175			
	Leu	Leu	Glu	Glu	Gln	Leu	Glu	Tyr	Lys	Arg	Gly	Glu	Ile	Arg	Glu	Ile	
				180					185					190			
	Glu	Gln	Val	Cys	Lys	Gln	Val	His	Asp	Lys	Leu	Thr	Ser	Ile	Gln	Asp	
			195				200						205				
	Thr	Leu	Arg	Thr	Gly	Ser	Gln	Ser	Pro	Pro	Ser	Ser	Ser	Gln			
		210					215						220				
	<210>	49															
	<211>	723															
	<212>	DNA															
	<213>	酿酒酵母															
	<220>																
	<221>	CDS															

<222> (1)..(723)																
<223> 酿酒酵母必需基因 RRP40 YOL142W SGDID:S000005502																
<400> 49																
atg	tct	acg	ttc	ata	ttc	cct	ggt	gat	agc	ttt	cct	gta	gac	cct	act	48
Met	Ser	Thr	Phe	Ile	Phe	Pro	Gly	Asp	Ser	Phe	Pro	Val	Asp	Pro	Thr	
1			5						10				15			
aca	cct	gtt	aaa	ctg	ggc	cct	ggc	ata	tat	tgt	gac	ccc	aat	act	caa	96
Thr	Pro	Val	Lys	Leu	Gly	Pro	Gly	Ile	Tyr	Cys	Asp	Pro	Asn	Thr	Gln	
			20					25				30				
gaa	ata	cga	cct	gtt	aat	aca	ggt	gtt	ttg	cat	gtt	tcc	gct	aag	ggg	144
Glu	Ile	Arg	Pro	Val	Asn	Thr	Gly	Val	Leu	His	Val	Ser	Ala	Lys	Gly	
		35				40					45					
aag	agt	ggt	gtt	cag	acc	gca	tat	ata	gac	tat	tct	agt	aag	aga	tac	192
Lys	Ser	Gly	Val	Gln	Thr	Ala	Tyr	Ile	Asp	Tyr	Ser	Ser	Lys	Arg	Tyr	
	50					55				60						
att	cca	tct	gta	aac	gat	ttt	gta	atc	ggt	gtc	att	ata	ggg	aca	ttt	240
Ile	Pro	Ser	Val	Asn	Asp	Phe	Val	Ile	Gly	Val	Ile	Ile	Gly	Thr	Phe	
65				70					75					80		
tca	gat	agc	tat	aag	gtt	tcg	ttg	caa	aat	ttc	tcc	tcc	agt	gtt	tca	288
Ser	Asp	Ser	Tyr	Lys	Val	Ser	Leu	Gln	Asn	Phe	Ser	Ser	Ser	Val	Ser	
				85					90					95		
cta	tcg	tat	atg	gct	ttt	ccc	aat	gct	tca	aag	aaa	aac	agg	cca	act	336
Leu	Ser	Tyr	Met	Ala	Phe	Pro	Asn	Ala	Ser	Lys	Lys	Asn	Arg	Pro	Thr	
			100					105					110			
ttg	caa	gtg	gga	gat	cta	gtg	tat	gcg	aga	gtt	tgt	acc	gca	gaa	aag	384
Leu	Gln	Val	Gly	Asp	Leu	Val	Tyr	Ala	Arg	Val	Cys	Thr	Ala	Glu	Lys	
		115				120						125				
gaa	cta	gaa	gcc	gaa	ata	gaa	tgt	ttt	gac	tca	act	aca	gga	cgc	gat	432
Glu	Leu	Glu	Ala	Glu	Ile	Glu	Cys	Phe	Asp	Ser	Thr	Thr	Gly	Arg	Asp	
		130				135					140					
gct	ggt	ttc	ggg	ata	ttg	gaa	gat	ggt	atg	atc	att	gac	gtg	aat	ttg	480
Ala	Gly	Phe	Gly	Ile	Leu	Glu	Asp	Gly	Met	Ile	Ile	Asp	Val	Asn	Leu	
145				150					155						160	
aat	ttc	gca	cgc	cag	ttg	ctt	ttc	aat	aat	gac	ttc	ccg	tta	tta	aaa	528
Asn	Phe	Ala	Arg	Gln	Leu	Leu	Phe	Asn	Asn	Asp	Phe	Pro	Leu	Leu	Lys	
				165				170					175			
gtg	ttg	gct	gca	cat	acc	aag	ttt	gaa	gtc	gcc	att	ggt	ctc	aat	ggg	576
Val	Leu	Ala	Ala	His	Thr	Lys	Phe	Glu	Val	Ala	Ile	Gly	Leu	Asn	Gly	
			180					185				190				
aag	atc	tgg	gtt	aag	tgc	gag	gaa	tta	tct	aac	act	tta	gct	tgt	tat	624
Lys	Ile	Trp	Val	Lys	Cys	Glu	Glu	Leu	Ser	Asn	Thr	Leu	Ala	Cys	Tyr	
		195				200					205					
aga	acc	ata	atg	gag	tgt	tgt	caa	aaa	aac	gac	acg	gca	gcg	ttc	aag	672
Arg	Thr	Ile	Met	Glu	Cys	Cys	Gln	Lys	Asn	Asp	Thr	Ala	Ala	Phe	Lys	
		210				215					220					
gat	ata	gca	aaa	aga	cag	ttt	aaa	gaa	ata	ctt	acg	gtc	aag	gag	gag	720
Asp	Ile	Ala	Lys	Arg	Gln	Phe	Lys	Glu	Ile	Leu	Thr	Val	Lys	Glu	Glu	
225					230					235				240		
tag																723
<210> 50																
<211> 240																
<212> PRT																
<213> 酿酒酵母																
<400> 50																
Met	Ser	Thr	Phe	Ile	Phe	Pro	Gly	Asp	Ser	Phe	Pro	Val	Asp	Pro	Thr	
1			5						10				15			
Thr	Pro	Val	Lys	Leu	Gly	Pro	Gly	Ile	Tyr	Cys	Asp	Pro	Asn	Thr	Gln	
			20					25				30				
Glu	Ile	Arg	Pro	Val	Asn	Thr	Gly	Val	Leu	His	Val	Ser	Ala	Lys	Gly	
		35				40					45					
Lys	Ser	Gly	Val	Gln	Thr	Ala	Tyr	Ile	Asp	Tyr	Ser	Ser	Lys	Arg	Tyr	
	50					55				60						
Ile	Pro	Ser	Val	Asn	Asp	Phe	Val	Ile	Gly	Val	Ile	Ile	Gly	Thr	Phe	
65				70					75					80		
Ser	Asp	Ser	Tyr	Lys	Val	Ser	Leu	Gln	Asn	Phe	Ser	Ser	Ser	Val	Ser	
				85					90					95		
Leu	Ser	Tyr	Met	Ala	Phe	Pro	Asn	Ala	Ser	Lys	Lys	Asn	Arg	Pro	Thr	
			100					105					110			
Leu	Gln	Val	Gly	Asp	Leu	Val	Tyr	Ala	Arg	Val	Cys	Thr	Ala	Glu	Lys	
		115				120						125				
Glu	Leu	Glu	Ala	Glu	Ile	Glu	Cys	Phe	Asp	Ser	Thr	Thr	Gly	Arg	Asp	
		130				135					140					
Ala	Gly	Phe	Gly	Ile	Leu	Glu	Asp	Gly	Met	Ile	Ile	Asp	Val	Asn	Leu	
145				150						155					160	
Asn	Phe	Ala	Arg	Gln	Leu	Leu	Phe	Asn	Asn	Asp	Phe	Pro	Leu	Leu	Lys	
				165					170					175		
Val	Leu	Ala	Ala	His	Thr	Lys	Phe	Glu	Val	Ala	Ile	Gly	Leu	Asn	Gly	
				180				185					190			

	Lys	Ile	Trp	Val	Lys	Cys	Glu	Glu	Leu	Ser	Asn	Thr	Leu	Ala	Cys	Tyr	
		195					200						205				
	Arg	Thr	Ile	Met	Glu	Cys	Cys	Gln	Lys	Asn	Asp	Thr	Ala	Ala	Phe	Lys	
		210				215						220					
	Asp	Ile	Ala	Lys	Arg	Gln	Phe	Lys	Glu	Ile	Leu	Thr	Val	Lys	Glu	Glu	
		225				230					235					240	
	<210>	51															
	<211>	1455															
	<212>	DNA															
	<213>	酿酒酵母															
	<220>																
	<221>	CDS															
	<222>	(1) .. (1455)															
	<223>	酿酒酵母必需基因 NOP8 YOL144W SGDID:S000005504															
	<400>	51															
	atg	gat	agt	gta	att	caa	aaa	aga	att	ttt	gtc	gga	aat	att	ttc	cat	48
	Met	Asp	Ser	Val	Ile	Gln	Lys	Arg	Ile	Phe	Val	Gly	Asn	Ile	Phe	His	
	1			5					10					15			
	aat	gca	gat	gat	tgt	tat	tca	gaa	cta	cta	gat	cga	ttt	gga	aaa	ttt	96
	Asn	Ala	Asp	Asp	Cys	Tyr	Ser	Glu	Leu	Leu	Asp	Arg	Phe	Gly	Lys	Phe	
				20					25					30			
	ggg	gat	tgc	caa	gat	ttt	cag	ttt	gaa	aaa	cac	aac	cac	ttt	gca	ttt	144
	Gly	Asp	Cys	Gln	Asp	Phe	Gln	Phe	Glu	Lys	His	Asn	His	Phe	Ala	Phe	
			35				40						45				
	atc	gat	ata	agg	ttt	aat	gat	gag	gca	gat	ttt	aac	aag	cta	aga	aag	192
	Ile	Asp	Ile	Arg	Phe	Asn	Asp	Glu	Ala	Asp	Phe	Asn	Lys	Leu	Arg	Lys	
		50				55						60					
	agc	ttt	aat	aat	gtt	aag	ttt	aaa	ggg	aac	att	ttg	aag	gtt	gat	gaa	240
	Ser	Phe	Asn	Asn	Val	Lys	Phe	Lys	Gly	Asn	Ile	Leu	Lys	Val	Asp	Glu	
		65			70					75				80			
	gca	aag	ccg	aat	tgg	gaa	agc	acc	tgg	gcg	gta	caa	cat	gca	aag	gac	288
	Ala	Lys	Pro	Asn	Trp	Glu	Ser	Thr	Trp	Ala	Val	Gln	His	Ala	Lys	Asp	
				85					90					95			
	ctt	aaa	gaa	gat	att	ata	cta	aat	gct	aaa	atg	aaa	aaa	aag	aat	tgg	336
	Leu	Lys	Glu	Asp	Ile	Ile	Leu	Asn	Ala	Lys	Met	Lys	Lys	Lys	Asn	Trp	
				100					105					110			
	caa	cat	tat	aag	aaa	atg	gaa	aat	gta	gcg	aaa	agt	tgg	aaa	gat	cat	384
	Gln	His	Tyr	Lys	Lys	Met	Glu	Asn	Val	Ala	Lys	Ser	Trp	Lys	Asp	His	
			115			120							125				
	aag	gaa	gtt	att	gct	ggc	aga	atg	agg	gag	gcg	ccg	aga	aaa	agg	agt	432
	Lys	Glu	Val	Ile	Ala	Gly	Arg	Met	Arg	Glu	Ala	Pro	Arg	Lys	Arg	Ser	
		130				135						140					
	caa	ttg	agg	aac	att	aca	ttc	agg	atc	aat	gtc	aat	ggg	tca	ttg	aaa	480
	Gln	Leu	Arg	Asn	Ile	Thr	Phe	Arg	Ile	Asn	Val	Asn	Gly	Ser	Leu	Lys	
		145				150					155				160		
	gtt	tac	aaa	tgc	tat	aag	act	aaa	tta	tgg	ggg	tat	gaa	agg	aac	aaa	528
	Val	Tyr	Lys	Cys	Tyr	Lys	Thr	Lys	Leu	Trp	Gly	Tyr	Glu	Arg	Lys	Lys	
				165					170					175			
	gaa	ttg	aat	gac	ctt	gtg	tac	aag	ttt	aca	aat	aat	ttc	tgg	aaa	aac	576
	Glu	Leu	Asn	Asp	Leu	Val	Tyr	Lys	Phe	Thr	Asn	Asn	Phe	Trp	Lys	Asn	
				180					185					190			
	gga	tac	aac	cac	att	gtc	gat	aga	ttg	gat	tac	agt	cgt	gct	gtt	aaa	624
	Gly	Tyr	Asn	His	Ile	Val	Asp	Arg	Leu	Asp	Tyr	Ser	Arg	Ala	Val	Lys	
			195					200					205				
	act	gtt	cga	ttc	aaa	aat	ggg	ctg	aaa	cag	tta	aca	gtg	tca	aaa	gat	672
	Thr	Val	Arg	Phe	Lys	Asn	Gly	Leu	Lys	Gln	Leu	Thr	Val	Ser	Lys	Asp	
		210				215						220					
	gaa	aac	gtt	tgc	agt	gga	gag	atg	gac	agc	gat	gag	aat	atg	tca	gaa	720
	Glu	Asn	Val	Cys	Ser	Gly	Glu	Met	Asp	Ser	Asp	Glu	Asn	Met	Ser	Glu	
		225				230					235					240	
	gag	gag	aag	gag	aaa	aat	aat	gtt	att	ctt	aat	gac	cta	tta	aag	gac	768
	Glu	Glu	Lys	Glu	Lys	Asn	Asn	Val	Ile	Leu	Asn	Asp	Leu	Leu	Lys	Asp	
				245					250					255			
	ttt	gac	ttc	gat	aag	cca	atg	acg	ttg	aat	gat	tcg	gac	gag	gaa	ctt	816
	Phe	Asp	Phe	Asp	Lys	Pro	Met	Thr	Leu	Asn	Asp	Ser	Asp	Glu	Glu	Leu	
				260					265					270			
	ctg	aca	gag	caa	cgt	aaa	gga	gag	gaa	gaa	gaa	gaa	gaa	gaa	gaa	gaa	864
	Leu	Thr	Glu	Gln	Arg	Lys	Gly	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	
			275				280						285				
	aaa	gaa	gta	aac	gct	ccc	gaa	tac	gag	aat	gtt	aat	aag	acg	aag	gac	912
	Lys	Glu	Val	Asn	Ala	Pro	Glu	Tyr	Glu	Asn	Val	Asn	Lys	Thr	Lys	Asp	
		290				295						300					
	cag	agc	acc	cta	cca	caa	gaa	aaa	cca	gag	gaa	agg	aaa	gaa	cag	gac	960
	Gln	Ser	Thr	Leu	Pro	Gln	Glu	Lys	Pro	Glu	Glu	Arg	Lys	Glu	Gln	Asp	
		305				310					315				320		
	gag	ggc	gac	ggg	cag	gag	gat	aat	gaa	ttt	ccg	act	ttc	acc	aag		1008
	Glu	Gly	Asp	Gly	Gln	Glu	Asp	Asn	Glu	Phe	Ile	Pro	Thr	Phe	Thr	Lys	
				325						330				335			

	gaa att ggt caa ggt aca ata agc aac act gag acg tta aga aac cta	1056
	Glu Ile Gly Gln Gly Thr Ile Ser Asn Thr Glu Thr Leu Arg Asn Leu	
	340 345 350	
	ttc aac ccc aac gag gca gaa cct gta tca caa ttc aaa ttg att gaa	1104
	Phe Asn Pro Asn Glu Ala Glu Pro Val Ser Gln Phe Lys Leu Ile Glu	
	355 360 365	
	gat tct gat aat gat att gat cat gcg aaa gat gtt gat gtg aac caa	1152
	Asp Ser Asp Asn Asp Ile Asp His Ala Lys Asp Val Asp Val Asn Gln	
	370 375 380	
	ctg gag gaa gaa gta agt aaa tca tct gac acc cta ggt ttg aca tcg	1200
	Leu Glu Glu Glu Val Ser Lys Ser Ser Asp Thr Leu Gly Leu Thr Ser	
	385 390 400	
	gct ccc gta cca cat gta tct aga gat aaa gat aac aaa aat ttc ttg	1248
	Ala Pro Val Pro His Val Ser Arg Asp Lys Asp Asn Lys Asn Phe Leu	
	405 410 415	
	ttt ttc cct cat ttg caa tcg cca ttt tta gta gga caa aca cag ttg	1296
	Phe Phe Pro His Leu Gln Ser Pro Phe Leu Val Gly Gln Thr Gln Leu	
	420 425 430	
	agt aaa gta agg gct cct gga aga gaa aca atg cta tcg aac tgg gat	1344
	Ser Lys Val Arg Ala Pro Gly Arg Glu Thr Met Leu Ser Asn Trp Asp	
	435 440 445	
	gag gag ttc tgg gct aac aga ggt aat tgg acc cgt gat atg agg cgc	1392
	Glu Glu Phe Trp Ala Asn Arg Gly Asn Trp Thr Arg Asp Met Arg Arg	
	450 455 460	
	aaa atg aaa gat gca ttg aag cac cgt aag agg aaa caa tca aag agc	1440
	Lys Met Lys Asp Ala Leu Lys His Arg Lys Arg Lys Gln Ser Lys Ser	
	465 470 480	
	ggg ctt ctt cta tag	1455
	Gly Leu Leu Leu	
	<210> 52	
	<211> 484	
	<212> PRT	
	<213> 酿酒酵母	
	<400> 52	
	Met Asp Ser Val Ile Gln Lys Arg Ile Phe Val Gly Asn Ile Phe His	
	1 5 10 15	
	Asn Ala Asp Asp Cys Tyr Ser Glu Leu Leu Asp Arg Phe Gly Lys Phe	
	20 25 30	
[0418]	Gly Asp Cys Gln Asp Phe Gln Phe Glu Lys His Asn His Phe Ala Phe	
	35 40 45	
	Ile Asp Ile Arg Phe Asn Asp Glu Ala Asp Phe Asn Lys Leu Arg Lys	
	50 55 60	
	Ser Phe Asn Asn Val Lys Phe Lys Gly Asn Ile Leu Lys Val Asp Glu	
	65 70 75 80	
	Ala Lys Pro Asn Trp Glu Ser Thr Trp Ala Val Gln His Ala Lys Asp	
	85 90 95	
	Leu Lys Glu Asp Ile Ile Leu Asn Ala Lys Met Lys Lys Lys Asn Trp	
	100 105 110	
	Gln His Tyr Lys Lys Met Glu Asn Val Ala Lys Ser Trp Lys Asp His	
	115 120 125	
	Lys Glu Val Ile Ala Gly Arg Met Arg Glu Ala Pro Arg Lys Arg Ser	
	130 135 140	
	Gln Leu Arg Asn Ile Thr Phe Arg Ile Asn Val Asn Gly Ser Leu Lys	
	145 150 155 160	
	Val Tyr Lys Cys Tyr Lys Thr Lys Leu Trp Gly Tyr Glu Arg Asn Lys	
	165 170 175	
	Glu Leu Asn Asp Leu Val Tyr Lys Phe Thr Asn Asn Phe Trp Lys Asn	
	180 185 190	
	Gly Tyr Asn His Ile Val Asp Arg Leu Asp Tyr Ser Arg Ala Val Lys	
	195 200 205	
	Thr Val Arg Phe Lys Asn Gly Leu Lys Gln Leu Thr Val Ser Lys Asp	
	210 215 220	
	Glu Asn Val Cys Ser Gly Glu Met Asp Ser Asp Glu Asn Met Ser Glu	
	225 230 235 240	
	Glu Glu Lys Glu Lys Asn Asn Val Ile Leu Asn Asp Leu Leu Lys Asp	
	245 250 255	
	Phe Asp Phe Asp Lys Pro Met Thr Leu Asn Asp Ser Asp Glu Glu Leu	
	260 265 270	
	Leu Thr Glu Gln Arg Lys Gly Glu Glu Glu Glu Glu Glu Glu Glu	
	275 280 285	
	Lys Glu Val Asn Ala Pro Glu Tyr Glu Asn Val Asn Lys Thr Lys Asp	
	290 295 300	
	Gln Ser Thr Leu Pro Gln Glu Lys Pro Glu Glu Arg Lys Glu Gln Asp	
	305 310 315 320	
	Glu Gly Asp Gly Gln Glu Asp Asn Glu Phe Ile Pro Thr Phe Thr Lys	
	325 330 335	
	Glu Ile Gly Gln Gly Thr Ile Ser Asn Thr Glu Thr Leu Arg Asn Leu	

```
<210> 53
<211> 1665
<212> DNA
<213> 酿酒酵母
<220>
<221> CDS
<222> (1)..(1665)
<223> 酿酒酵母必需基因 PGI1 YBR196C SGDID:S000000400
<400> 53
```

atg	tcc	aat	aac	tca	ttc	act	aac	ttc	aaa	ctg	gcc	act	gaa	ttg	cca	48
Met	Ser	Asn	Asn	Ser	Phe	Thr	Asn	Phe	Lys	Leu	Ala	Thr	Glu	Leu	Pro	
1			5						10					15		
gcc	tggtct	aag	ttgttg	caa	aaa	att	tattat	gaa	tct	caa	ggt	aag	act	ttgttg	96	
Ala	Trp	Ser	Lys	Leu	Gln	Lys	Ile	Tyr	Glu	Ser	Gln	Gly	Lys	Thr		Leu
			20					25					30			
tct	gtc	aag	caa	gaa	ttc	caa	aaa	gat	gcc	aag	cgt	ttt	gaa	aaa	ttgttg	144
Ser	Val	Lys	Gln	Glu	Phe	Gln	Lys	Asp	Ala	Lys	Arg	Phe	Glu	Lys	Leu	
			35				40					45				
aac	aag	act	ttc	acc	aac	tat	gat	ggt	tcc	aaa	atc	ttg	ttc	gac	tac	192
Asn	Lys	Thr	Phe	Thr	Asn	Tyr	Asp	Gly	Ser	Lys	Ile	Leu	Phe	Asp	Tyr	
			50			55					60					
tca	aag	aac	ttg	gtc	aac	gat	gaa	atc	att	gct	gca	ttg	att	gaa	ctgttg	240
Ser	Lys	Asn	Leu	Val	Asn	Asp	Glu	Ile	Ile	Ala	Ala	Leu	Ile	Glu	Leu	
			65		70				75					80		
gcc	aag	gag	gct	aac	gtc	acc	ggt	ttg	aga	gat	gct	atg	ttc	aaa	ggtggt	288
Ala	Lys	Glu	Ala	Asn	Val	Thr	Gly	Leu	Arg	Asp	Ala	Met	Phe	Lys	Gly	
			85						90					95		
gaa	cac	atc	aac	tcc	act	gaa	gat	cgt	gct	gtc	tac	cac	gtc	gca	ttgttg	336
Glu	His	Ile	Asn	Ser	Thr	Glu	Asp	Arg	Ala	Val	Tyr	His	Val	Ala	Leu	
			100					105					110			
aga	aac	aga	gct	aac	aag	cca	atg	tac	ggt	gat	ggt	gtc	aac	ggt	gctgct	384
Arg	Asn	Arg	Ala	Asn	Lys	Pro	Met	Tyr	Val	Asp	Gly	Val	Asn	Val	Ala	
			115				120					125				
cca	gaa	gtc	gac	tct	gtc	ttg	aag	cac	atg	aag	gag	ttc	tct	gaa	caagaa	432
Pro	Glu	Val	Asp	Ser	Val	Leu	Lys	His	Met	Lys	Glu	Phe	Ser	Glu	Gln	
			130			135					140					
gtt	cgt	tct	ggt	gaa	tggtgg	aag	ggt	tattat	acc	ggt	aag	aag	atc	acc	gatgat	480
Val	Arg	Ser	Gly	Glu	Trp	Lys	Gly	Tyr	Thr	Gly	Lys	Lys	Ile	Thr	Asp	
			145		150					155				160		
gtt	gtt	aac	atc	ggt	att	ggt	ggt	tcc	gat	ttg	ggt	cca	gtc	atg	gtcgtc	528
Val	Val	Asn	Ile	Gly	Ile	Gly	Gly	Ser	Asp	Leu	Gly	Pro	Val	Met	Val	
			165						170					175		
act	gag	gct	ttg	aag	cac	tac	gct	ggt	gtc	ttg	gat	gtc	cac	ttc	gttgtt	576
Thr	Glu	Ala	Leu	Lys	His	Tyr	Ala	Gly	Val	Leu	Asp	Val	His	Phe	Val	
			180					185					190			
tcc	aac	att	gac	ggt	act	cac	att	gct	gaa	acc	ttg	aag	ggt	ggt	gacgac	624
Ser	Asn	Ile	Asp	Gly	Thr	His	Ile	Ala	Glu	Thr	Leu	Lys	Val	Val	Asp	
			195				200					205				
cca	gaa	act	act	ttg	ttt	ttg	att	gct	tcc	aag	act	ttc	act	acc		

	260	265	270	864
ttt ggt ttc gaa agt tgg gtc ggt ggt cgt tac tct gtc tgg tcg gct				
Phe Gly Phe Glu Ser Trp Val Gly Gly Arg Tyr Ser Val Trp Ser Ala				
275	280	285		
att ggt ttg tct gtt gcc ttg tac att ggc tat gac aac ttt gag gct				912
Ile Gly Leu Ser Val Ala Leu Tyr Ile Gly Tyr Asp Asn Phe Glu Ala				
290	295	300		
ttc ttg aag ggt gct gaa gcc gtc gac aac cac ttc acc caa acc cca				960
Phe Leu Lys Gly Ala Glu Ala Val Asp Asn His Phe Thr Gln Thr Pro				
305	310	315		
ttg gaa gac aac att cca ttg ttg ggt ggt ttg tct gtc tgg tac				1008
Leu Glu Asp Asn Ile Pro Leu Leu Gly Gly Leu Leu Ser Val Trp Tyr				
325	330	335		
aac aac ttc ttt ggt gct caa acc cat ttg gct cca ttc gac caa				1056
Asn Asn Phe Phe Gly Ala Gln Thr His Leu Val Ala Pro Phe Asp Gln				
340	345	350		
tac ttg cac aga ttc cca gcc tac ttg caa caa ttg tca atg gaa tct				1104
Tyr Leu His Arg Phe Pro Ala Tyr Leu Gln Gln Leu Ser Met Glu Ser				
355	360	365		
aac ggt aag tct gtt acc aga ggt aac gtg ttt act gac tac tct act				1152
Asn Gly Lys Ser Val Thr Arg Gly Asn Val Phe Thr Asp Tyr Ser Thr				
370	375	380		
ggt tct atc ttg ttt ggt gaa cca gct acc aac gct caa cac tct ttc				1200
Gly Ser Ile Leu Phe Gly Glu Pro Ala Thr Asn Ala Gln His Ser Phe				
385	390	395		
ttc caa ttg gtt cac caa ggt acc aag ttg att cca tct gat ttc atc				1248
Phe Gln Leu Val His Gln Gly Thr Lys Leu Ile Pro Ser Asp Phe Ile				
405	410	415		
tta gct gct caa tct cat aac cca att gag aac aaa tta cat caa aag				1296
Leu Ala Ala Gln Ser His Asn Pro Ile Glu Asn Lys Leu His Gln Lys				
420	425	430		
atg ttg gct tca aac ttc ttt gct caa gct gaa gct tta atg gtt ggt				1344
Met Leu Ala Ser Asn Phe Phe Ala Gln Ala Glu Ala Leu Met Val Gly				
435	440	445		
aag gat gaa gaa caa gtt aag gct gaa ggt gcc act ggt ggt ttg gtc				1392
Lys Asp Glu Glu Gln Val Lys Ala Glu Gly Ala Thr Gly Gly Leu Val				
450	455	460		
cca cac aag gtc ttc tca ggt aac aga cca act acc tct atc ttg gct				1440
Pro His Lys Val Phe Ser Gly Asn Arg Pro Thr Thr Ser Ile Leu Ala				
465	470	475		
caa aag att act cca gct act ttg ggt gct ttg att gcc tac tac gaa				1488
Gln Lys Ile Thr Pro Ala Thr Leu Gly Ala Leu Ile Ala Tyr Tyr Glu				
485	490	495		
cat gtt act ttc act gaa ggt gcc att tgg aat atc aac tct ttc gac				1536
His Val Thr Phe Thr Glu Gly Ala Ile Trp Asn Ile Asn Ser Phe Asp				
500	505	510		
caa tgg ggt gtt gaa ttg ggt aaa gtc ttg gct aaa gtc ggc aag				1584
Gln Trp Gly Val Glu Leu Gly Lys Val Leu Ala Lys Val Ile Gly Lys				
515	520	525		
gaa ttg gac aac tcc tcc acc att tct acc cac gat gct tct acc aac				1632
Glu Leu Asp Asn Ser Ser Thr Ile Ser Thr His Asp Ala Ser Thr Asn				
530	535	540		
ggt tta atc aat caa ttc aag gaa tgg atg tga				1665
Gly Leu Ile Asn Gln Phe Lys Glu Trp Met				
545	550			
<210> 54				
<211> 554				
<212> PRT				
<213> 酿酒酵母				
<400> 54				
Met Ser Asn Asn Ser Phe Thr Asn Phe Lys Leu Ala Thr Glu Leu Pro				
1 5				
Ala Trp Ser Lys Leu Gln Lys Ile Tyr Glu Ser Gln Gly Lys Thr Leu				
20				
Ser Val Lys Gln Glu Phe Gln Lys Asp Ala Lys Arg Phe Glu Lys Leu				
35				
Asn Lys Thr Phe Thr Asn Tyr Asp Gly Ser Lys Ile Leu Phe Asp Tyr				
50				
Ser Lys Asn Leu Val Asn Asp Glu Ile Ile Ala Leu Ile Glu Leu				
65 70				
Ala Lys Glu Ala Asn Val Thr Gly Leu Arg Asp Ala Met Phe Lys Gly				
85				
Glu His Ile Asn Ser Thr Glu Asp Arg Ala Val Tyr His Val Ala Leu				
100				
Arg Asn Arg Ala Asn Lys Pro Met Tyr Val Asp Gly Val Asn Val Ala				
115				
Pro Glu Val Asp Ser Val Leu Lys His Met Lys Glu Phe Ser Glu Gln				
130				
135				
140				

[0421]

Val	Arg	Ser	Gly	Glu	Trp	Lys	Gly	Tyr	Thr	Gly	Lys	Lys	Ile	Thr	Asp		
145					150					155					160		
Val	Val	Asn	Ile	Gly	Ile	Gly	Gly	Ser	Asp	Leu	Gly	Pro	Val	Met	Val		
				165					170					175			
Thr	Glu	Ala	Leu	Lys	His	Tyr	Ala	Gly	Val	Leu	Asp	Val	His	Phe	Val		
				180				185					190				
Ser	Asn	Ile	Asp	Gly	Thr	His	Ile	Ala	Glu	Thr	Leu	Lys	Val	Val	Asp		
		195					200					205					
Pro	Glu	Thr	Thr	Leu	Phe	Leu	Ile	Ala	Ser	Lys	Thr	Phe	Thr	Thr	Ala		
	210					215					220						
Glu	Thr	Ile	Thr	Asn	Ala	Asn	Thr	Ala	Lys	Asn	Trp	Phe	Leu	Ser	Lys		
	225				230					235					240		
Thr	Gly	Asn	Asp	Pro	Ser	His	Ile	Ala	Lys	His	Phe	Ala	Ala	Leu	Ser		
				245					250					255			
Thr	Asn	Glu	Thr	Glu	Val	Ala	Lys	Phe	Gly	Ile	Asp	Thr	Lys	Asn	Met		
				260				265					270				
Phe	Gly	Phe	Glu	Ser	Trp	Val	Gly	Arg	Tyr	Ser	Val	Trp	Ser	Ala			
		275					280				285						
Ile	Gly	Leu	Ser	Val	Ala	Leu	Tyr	Ile	Gly	Tyr	Asp	Asn	Phe	Glu	Ala		
	290					295					300						
Phe	Leu	Lys	Gly	Ala	Glu	Ala	Val	Asp	Asn	His	Phe	Thr	Gln	Thr	Pro		
	305				310					315					320		
Leu	Glu	Asp	Asn	Ile	Pro	Leu	Leu	Gly	Gly	Leu	Leu	Ser	Val	Trp	Tyr		
				325					330					335			
Asn	Asn	Phe	Phe	Gly	Ala	Gln	Thr	His	Leu	Val	Ala	Pro	Phe	Asp	Gln		
			340					345					350				
Tyr	Leu	His	Arg	Phe	Pro	Ala	Tyr	Leu	Gln	Gln	Leu	Ser	Met	Glu	Ser		
		355					360					365					
Asn	Gly	Lys	Ser	Val	Thr	Arg	Gly	Asn	Val	Phe	Thr	Asp	Tyr	Ser	Thr		
	370					375					380						
Gly	Ser	Ile	Leu	Phe	Gly	Glu	Pro	Ala	Thr	Asn	Ala	Gln	His	Ser	Phe		
	385				390					395					400		
Phe	Gln	Leu	Val	His	Gln	Gly	Thr	Lys	Leu	Ile	Pro	Ser	Asp	Phe	Ile		
				405					410					415			
Leu	Ala	Ala	Gln	Ser	His	Asn	Pro	Ile	Glu	Asn	Lys	Leu	His	Gln	Lys		
			420					425					430				
Met	Leu	Ala	Ser	Asn	Phe	Phe	Ala	Gln	Ala	Glu	Ala	Leu	Met	Val	Gly		
		435					440					445					
Lys	Asp	Glu	Glu	Gln	Val	Lys	Ala	Glu	Gly	Ala	Thr	Gly	Gly	Leu	Val		
	450					455					460						
Pro	His	Lys	Val	Phe	Ser	Gly	Asn	Arg	Pro	Thr	Thr	Ser	Ile	Leu	Ala		
	465				470					475					480		
Gln	Lys	Ile	Thr	Pro	Ala	Thr	Leu	Gly	Ala	Leu	Ile	Ala	Tyr	Tyr	Glu		
				485					490					495			
His	Val	Thr	Phe	Thr	Glu	Gly	Ala	Ile	Trp	Asn	Ile	Asn	Ser	Phe	Asp		
				500				505					510				
Gln	Trp	Gly	Val	Glu	Leu	Gly	Lys	Val	Leu	Ala	Lys	Val	Ile	Gly	Lys		
		515					520					525					
Glu	Leu	Asp	Asn	Ser	Ser	Thr	Ile	Ser	Thr	His	Asp	Ala	Ser	Thr	Asn		
	530					535					540						
Gly	Leu	Ile	Asn	Gln	Phe	Lys	Glu	Trp	Met								
	545				550												

<210> 55
 <211> 759
 <212> DNA
 <213> 酿酒酵母
 <220>
 <221> CDS
 <222> (1)..(759)
 <223> 酿酒酵母必需基因 NEP1 EMG1 YLR186W SGDID:S000004176
 <400> 55

atg	gtc	gaa	gat	tcc	aga	gtt	aga	gac	gcc	ctc	aaa	ggt	ggt	gat	cag		48
Met	Val	Glu	Asp	Ser	Arg	Val	Arg	Asp	Ala	Leu	Lys	Gly	Gly	Asp	Gln		
				5					10					15			
aag	gca	tta	cgg	gcc	tct	ttg	gtt	cct	caa	gca	cct	cct	gtc	ttg	aca		96
Lys	Ala	Leu	Pro	Ala	Ser	Leu	Val	Pro	Gln	Ala	Pro	Pro	Val	Leu	Thr		
				20				25					30				
tca	aag	gat	aag	att	act	aag	cgg	atg	att	gtg	gta	tta	gcg	atg	gca		144
Ser	Lys	Asp	Lys	Ile	Thr	Lys	Arg	Met	Ile	Val	Val	Leu	Ala	Met	Ala		
		35				40					45						
tcc	ctc	gag	aca	cac	aag	ata	tcg	tcc	aac	ggg	cct	ggt	ggt	gac	aaa		192
Ser	Leu	Glu	Thr	His	Lys	Ile	Ser	Ser	Asn	Gly	Pro	Gly	Gly	Asp	Lys		
		50				55					60						
tat	gtc	ctt	ttg	aac	tgt	gac	gac	cat	caa	ggg	tta	tta	aaa	aaa	atg		240
Tyr	Val	Leu	Leu	Asn	Cys	Asp	Asp	His	Gln	Gly	Leu	Leu	Lys	Lys	Met		
		65			70					75				80			
ggg	aga	gac	att	agt	gaa	gca	aga	cct	gat	att	acc	cac	caa	tgt	ctt		288
Gly	Arg	Asp	Ile	Ser	Glu	Ala	Arg	Pro	Asp	Ile	Thr	His	Gln	Cys	Leu		

	ttg	acg	ttg	cta	gat	tct	cca	atc	aac	aaa	gcc	gga	aag	ctg	cag	gtc	336
	Leu	Thr	Leu	Leu	Asp	Ser	Pro	Ile	Asn	Lys	Ala	Gly	Lys	Leu	Gln	Val	
				100					105					110			
	tat	att	caa	aca	agt	cga	gga	att	ctg	atc	gag	gtt	aac	ccc	act	gtt	384
	Tyr	Ile	Gln	Thr	Ser	Arg	Gly	Ile	Leu	Ile	Glu	Val	Asn	Pro	Thr	Val	
				115					120					125			
	cgt	ata	cca	aga	act	ttc	aaa	aga	ttt	tca	ggt	tta	atg	gtt	cag	tta	432
	Arg	Ile	Pro	Arg	Thr	Phe	Lys	Arg	Phe	Ser	Gly	Leu	Met	Val	Gln	Leu	
				130					135					140			
	cta	cat	aag	ctt	tct	atc	aga	tcg	gta	aat	tct	gaa	gaa	aag	tta	ctt	480
	Leu	His	Lys	Leu	Ser	Ile	Arg	Ser	Val	Asn	Ser	Glu	Glu	Lys	Leu	Leu	
						150						155				160	
	aaa	gtc	att	aag	aac	cca	att	acc	gat	cac	cta	cct	act	aag	tgc	cgt	528
	Lys	Val	Ile	Lys	Asn	Pro	Ile	Thr	Asp	His	Leu	Pro	Thr	Lys	Cys	Arg	
					165									175			
	aag	gtg	aca	tta	tcc	ttt	gac	gca	cca	gtt	atc	cgc	gtt	caa	gat	tac	576
	Lys	Val	Thr	Leu	Ser	Phe	Asp	Ala	Pro	Val	Ile	Arg	Val	Gln	Asp	Tyr	
						180								190			
	atc	gaa	aaa	cta	gac	gat	gat	gaa	agt	ata	tgt	gtc	ttt	gtt	ggg	gcc	624
	Ile	Glu	Lys	Leu	Asp	Asp	Asp	Glu	Ser	Ile	Cys	Val	Phe	Val	Gly	Ala	
						195							205				
	atg	gca	aga	ggg	aaa	gat	aac	ttt	gcg	gat	gaa	tac	gtc	gac	gaa	aaa	672
	Met	Ala	Arg	Gly	Lys	Asp	Asn	Phe	Ala	Asp	Glu	Tyr	Val	Asp	Glu	Lys	
						210							220				
	gtc	ggc	ttg	tcc	aat	tac	cca	ttg	tct	gcc	tca	gtt	gca	tgt	tct	aaa	720
	Val	Gly	Leu	Ser	Asn	Tyr	Pro	Leu	Ser	Ala	Ser	Val	Ala	Cys	Ser	Lys	
						230							235			240	
	ttt	tgc	cat	ggc	gct	gaa	gat	gct	tgg	aat	att	tta	tag				759
	Phe	Cys	His	Gly	Ala	Glu	Asp	Ala	Trp	Asn	Ile	Leu					
						245											
	<210>	56															
	<211>	252															
	<212>	PRT															
	<213>	酿酒酵母															
	<400>	56															
	Met	Val	Glu	Asp	Ser	Arg	Val	Arg	Asp	Ala	Leu	Lys	Gly	Gly	Asp	Gln	
	1				5					10					15		
[0422]	Lys	Ala	Leu	Pro	Ala	Ser	Leu	Val	Pro	Gln	Ala	Pro	Pro	Val	Leu	Thr	
				20					25					30			
	Ser	Lys	Asp	Lys	Ile	Thr	Lys	Arg	Met	Ile	Val	Val	Leu	Ala	Met	Ala	
				35					40					45			
	Ser	Leu	Glu	Thr	His	Lys	Ile	Ser	Ser	Asn	Gly	Pro	Gly	Gly	Asp	Lys	
				50					55				60				
	Tyr	Val	Leu	Leu	Asn	Cys	Asp	Asp	His	Gln	Gly	Leu	Leu	Lys	Lys	Met	
						70						75				80	
	Gly	Arg	Asp	Ile	Ser	Glu	Ala	Arg	Pro	Asp	Ile	Thr	His	Gln	Cys	Leu	
						85								95			
	Leu	Thr	Leu	Leu	Asp	Ser	Pro	Ile	Asn	Lys	Ala	Gly	Lys	Leu	Gln	Val	
						100								110			
	Tyr	Ile	Gln	Thr	Ser	Arg	Gly	Ile	Leu	Ile	Glu	Val	Asn	Pro	Thr	Val	
						115								125			
	Arg	Ile	Pro	Arg	Thr	Phe	Lys	Arg	Phe	Ser	Gly	Leu	Met	Val	Gln	Leu	
														140			
	Leu	His	Lys	Leu	Ser	Ile	Arg	Ser	Val	Asn	Ser	Glu	Glu	Lys	Leu	Leu	
						150								160			
	Lys	Val	Ile	Lys	Asn	Pro	Ile	Thr	Asp	His	Leu	Pro	Thr	Lys	Cys	Arg	
						165								175			
	Lys	Val	Thr	Leu	Ser	Phe	Asp	Ala	Pro	Val	Ile	Arg	Val	Gln	Asp	Tyr	
						180								190			
	Ile	Glu	Lys	Leu	Asp	Asp	Asp	Glu	Ser	Ile	Cys	Val	Phe	Val	Gly	Ala	
						195								205			
	Met	Ala	Arg	Gly	Lys	Asp	Asn	Phe	Ala	Asp	Glu	Tyr	Val	Asp	Glu	Lys	
						210								220			
	Val	Gly	Leu	Ser	Asn	Tyr	Pro	Leu	Ser	Ala	Ser	Val	Ala	Cys	Ser	Lys	
						230										240	
	Phe	Cys	His	Gly	Ala	Glu	Asp	Ala	Trp	Asn	Ile	Leu					
						245											
	<210>	57															
	<211>	804															
	<212>	DNA															
	<213>	酿酒酵母															
	<220>																
	<221>	CDS															
	<222>	(1)..(804)															
	<223>	酿酒酵母必需基因 URA3															
	<400>	57															
	atg	tcg	aaa	gct	aca	tat	aag	gaa	cgt	gct	gct	act	cat	cct	agt	cct	48

	Met	Ser	Lys	Ala	Thr	Tyr	Lys	Glu	Arg	Ala	Ala	Thr	His	Pro	Ser	Pro	
	1				5					10					15		
	ggt	gct	gcc	aag	cta	ttt	aat	atc	atg	cac	gaa	aag	caa	aca	aac	ttg	96
	Val	Ala	Ala	Lys	Leu	Phe	Asn	Ile	Met	His	Glu	Lys	Gln	Thr	Asn	Leu	
				20					25					30			
	tgt	gct	tca	ttg	gat	ggt	cgt	acc	acc	aag	gaa	tta	ctg	gag	tta	ggt	144
	Cys	Ala	Ser	Leu	Asp	Val	Arg	Thr	Lys	Glu	Leu	Leu	Glu	Leu	Val		
			35				40						45				
	gaa	gca	tta	ggt	ccc	aaa	att	tgt	tta	cta	aaa	aca	cat	gtg	gat	atc	192
	Glu	Ala	Leu	Gly	Pro	Lys	Ile	Cys	Leu	Leu	Lys	Thr	His	Val	Asp	Ile	
			50				55					60					
	ttg	act	gat	ttt	tcc	atg	gag	ggc	aca	ggt	aag	ccg	cta	aag	gca	tta	240
	Leu	Thr	Asp	Phe	Ser	Met	Glu	Gly	Thr	Val	Lys	Pro	Leu	Lys	Ala	Leu	
			65			70					75				80		
	tcc	gcc	aag	tac	aat	ttt	tta	ctc	ttc	gaa	gac	aga	aaa	ttt	gct	gac	288
	Ser	Ala	Lys	Tyr	Asn	Phe	Leu	Leu	Phe	Glu	Asp	Arg	Lys	Phe	Ala	Asp	
					85					90				95			
	att	ggt	aat	aca	gtc	aaa	ttg	cag	tac	tct	gcg	ggt	gta	tac	aga	ata	336
	Ile	Gly	Asn	Thr	Val	Lys	Leu	Gln	Tyr	Ser	Ala	Gly	Val	Tyr	Arg	Ile	
				100					105					110			
	gca	gaa	tgg	gca	gac	att	acg	aat	gca	cac	ggt	gtg	gtg	ggc	cca	ggt	384
	Ala	Glu	Trp	Ala	Asp	Ile	Thr	Asn	Ala	His	Gly	Val	Val	Gly	Pro	Gly	
			115				120						125				
	att	ggt	agc	ggt	ttg	aag	cag	gcg	gca	gaa	gaa	gta	aca	aag	gaa	cct	432
	Ile	Val	Ser	Gly	Leu	Lys	Gln	Ala	Ala	Glu	Glu	Val	Thr	Lys	Glu	Pro	
			130				135					140					
	aga	ggc	ctt	ttg	atg	tta	gca	gaa	ttg	tca	tgc	aag	ggc	tcc	cta	tct	480
	Arg	Gly	Leu	Leu	Met	Leu	Ala	Glu	Leu	Ser	Cys	Lys	Gly	Ser	Leu	Ser	
			145			150					155				160		
	act	gga	gaa	tat	act	aag	ggt	act	ggt	gac	att	gcg	aag	agc	gac	aaa	528
	Thr	Gly	Glu	Tyr	Thr	Lys	Gly	Thr	Val	Asp	Ile	Ala	Lys	Ser	Asp	Lys	
					165					170				175			
	gat	ttt	ggt	atc	ggc	ttt	att	gct	caa	aga	gac	atg	ggt	gga	aga	gat	576
	Asp	Phe	Val	Ile	Gly	Phe	Ile	Ala	Gln	Arg	Asp	Met	Gly	Gly	Arg	Asp	
				180					185					190			
	gaa	ggt	tac	gat	tgg	ttg	att	atg	aca	ccc	ggt	gtg	ggt	tta	gat	gac	624
	Glu	Gly	Tyr	Asp	Trp	Leu	Ile	Met	Thr	Pro	Gly	Val	Gly	Leu	Asp	Asp	
			195					200					205				
[0423]	aag	gga	gac	gca	ttg	ggt	caa	cag	tat	aga	acc	gtg	gat	gat	gtg	gtc	672
	Lys	Gly	Asp	Ala	Leu	Gly	Gln	Gln	Tyr	Arg	Thr	Val	Val	Asp	Val	Val	
			210				215					220					
	tct	aca	gga	tct	gac	att	att	att	ggt	gga	aga	gga	cta	ttt	gca	aag	720
	Ser	Thr	Gly	Ser	Asp	Ile	Ile	Ile	Val	Gly	Arg	Gly	Leu	Phe	Ala	Lys	
			225			230					235				240		
	gga	agg	gat	gct	aag	gta	gag	ggt	gaa	cgt	tac	aga	aaa	gca	ggc	tggt	768
	Gly	Arg	Asp	Ala	Lys	Val	Glu	Gly	Glu	Arg	Tyr	Arg	Lys	Ala	Gly	Trp	
					245				250					255			
	gaa	gca	tat	ttg	aga	aga	tgc	ggc	cag	caa	aac	taa					804
	Glu	Ala	Tyr	Leu	Arg	Arg	Cys	Gly	Gln	Gln	Asn						
				260					265								
	<210>	58															
	<211>	267															
	<212>	PRT															
	<213>	酿酒酵母															
	<400>	58															
	Met	Ser	Lys	Ala	Thr	Tyr	Lys	Glu	Arg	Ala	Ala	Thr	His	Pro	Ser	Pro	
	1				5					10					15		
	Val	Ala	Ala	Lys	Leu	Phe	Asn	Ile	Met	His	Glu	Lys	Gln	Thr	Asn	Leu	
				20					25					30			
	Cys	Ala	Ser	Leu	Asp	Val	Arg	Thr	Lys	Lys	Glu	Leu	Leu	Glu	Leu	Val	
			35				40						45				
	Glu	Ala	Leu	Gly	Pro	Lys	Ile	Cys	Leu	Leu	Lys	Thr	His	Val	Asp	Ile	
			50				55					60					
	Leu	Thr	Asp	Phe	Ser	Met	Glu	Gly	Thr	Val	Lys	Pro	Leu	Lys	Ala	Leu	
			65			70					75				80		
	Ser	Ala	Lys	Tyr	Asn	Phe	Leu	Leu	Phe	Glu	Asp	Arg	Lys	Phe	Ala	Asp	
					85					90				95			
	Ile	Gly	Asn	Thr	Val	Lys	Leu	Gln	Tyr	Ser	Ala	Gly	Val	Tyr	Arg	Ile	
				100					105					110			
	Ala	Glu	Trp	Ala	Asp	Ile	Thr	Asn	Ala	His	Gly	Val	Val	Gly	Pro	Gly	
			115				120						125				
	Ile	Val	Ser	Gly	Leu	Lys	Gln	Ala	Ala	Glu	Glu	Val	Thr	Lys	Glu	Pro	
			130				135					140					
	Arg	Gly	Leu	Leu	Met	Leu	Ala	Glu	Leu	Ser	Cys	Lys	Gly	Ser	Leu	Ser	
			145			150					155				160		
	Thr	Gly	Glu	Tyr	Thr	Lys	Gly	Thr	Val	Asp	Ile	Ala	Lys	Ser	Asp	Lys	
					165					170				175			
	Asp	Phe	Val	Ile	Gly	Phe	Ile	Ala	Gln	Arg	Asp	Met	Gly	Gly	Arg	Asp	

	Glu	Gly	Tyr 195	Asp	Trp	Leu	Ile	Met 200	Thr	Pro	Gly	Val	Gly 205	Leu	Asp	Asp	
	Lys	Gly 210	Asp	Ala	Leu	Gly	Gln 215	Gln	Tyr	Arg	Thr	Val 220	Asp	Asp	Val	Val	
	Ser 225	Thr	Gly	Ser	Asp	Ile 230	Ile	Ile	Val	Gly 235	Arg	Gly	Leu	Phe	Ala	Lys 240	
	Gly	Arg	Asp	Ala	Lys 245	Val	Glu	Gly	Glu	Arg 250	Tyr	Arg	Lys	Ala	Gly 255	Trp	
	Glu	Ala	Tyr	Leu 260	Arg	Arg	Cys	Gly	Gln 265	Gln	Asn						
<210>	59																
<211>	1095																
<212>	DNA																
<213>	酿酒酵母																
<220>																	
<221>	CDS																
<222>	(1)..(1095)																
<223>	酿酒酵母必需基因 LEU2																
<400>	59																
atg	tct	gcc	cct	aag	aag	atc	gtc	gtt	ttg	cca	ggg	gac	cac	ggt	ggt		48
Met	Ser	Ala	Pro	Lys	Lys	Ile	Val	Val	Leu	Pro	Gly	Asp	His	Val	Gly		
1			5						10					15			
caa	gaa	atc	aca	gcc	gaa	gcc	att	aag	ggt	ctt	aaa	gct	att	tct	gat		96
Gln	Glu	Ile	Thr	Ala	Glu	Ala	Ile	Lys	Val	Leu	Lys	Ala	Ile	Ser	Asp		
			20					25					30				
ggt	cgt	tcc	aat	gtc	aag	ttc	gat	ttc	gaa	aat	cat	tta	att	ggt	ggt		144
Val	Arg	Ser	Asn	Val	Lys	Phe	Asp	Phe	Glu	Asn	His	Leu	Ile	Gly	Gly		
		35				40						45					
gct	gct	atc	gat	gct	aca	ggt	gtc	cca	ctt	cca	gat	gag	gcg	ctg	gaa		192
Ala	Ala	Ile	Asp	Ala	Thr	Gly	Val	Pro	Leu	Pro	Asp	Glu	Ala	Leu	Glu		
		50				55					60						
gcc	tcc	aag	aag	ggt	gat	gcc	ggt	ttg	tta	ggt	gct	gtg	gct	ggt	cct		240
Ala	Ser	Lys	Lys	Val	Asp	Ala	Val	Leu	Leu	Gly	Ala	Val	Ala	Gly	Pro		
		65			70					75				80			
aaa	tgg	ggt	acc	ggt	agt	ggt	aga	cct	gaa	caa	ggt	tta	cta	aaa	atc		288
Lys	Trp	Gly	Thr	Gly	Ser	Val	Arg	Pro	Glu	Gln	Gly	Leu	Leu	Lys	Ile		
			85					90						95			
cgT	aaa	gaa	ctt	caa	ttg	tac	gcc	aac	tta	aga	cca	tgt	aac	ttt	gca		336
Arg	Lys	Glu	Leu	Gln	Leu	Tyr	Ala	Asn	Leu	Arg	Pro	Cys	Asn	Phe	Ala		
			100					105					110				
tcc	gac	tct	ctt	tta	gac	tta	tct	cca	atc	aag	cca	caa	ttt	gct	aaa		384
Ser	Asp	Ser	Leu	Leu	Asp	Leu	Ser	Pro	Ile	Lys	Pro	Gln	Phe	Ala	Lys		
		115					120					125					
ggt	act	gac	ttc	ggt	ggt	gtc	aga	gaa	tta	gtg	gga	ggt	att	tac	ttt		432
Gly	Thr	Asp	Phe	Val	Val	Val	Arg	Glu	Leu	Val	Gly	Gly	Ile	Tyr	Phe		
		130				135					140						
ggt	aag	aga	aag	gaa	gac	gat	ggt	gat	ggt	gtc	gct	tgg	gat	agt	gaa		480
Gly	Lys	Arg	Lys	Glu	Asp	Asp	Gly	Asp	Gly	Val	Ala	Trp	Asp	Ser	Glu		
		145			150					155					160		
caa	tac	acc	ggt	cca	gaa	gtg	caa	aga	atc	aca	aga	atg	gcc	gct	ttc		528
Gln	Tyr	Thr	Val	Pro	Glu	Val	Gln	Arg	Ile	Thr	Arg	Met	Ala	Ala	Phe		
			165					170					175				
atg	gcc	cta	caa	cat	gag	cca	cca	ttg	cct	att	tgg	tcc	ttg	gat	aaa		576
Met	Ala	Leu	Gln	His	Glu	Pro	Pro	Leu	Pro	Ile	Trp	Ser	Leu	Asp	Lys		
		180						185					190				
gct	aat	ctt	ttg	gcc	tct	tca	aga	tta	tgg	aga	aaa	act	gtg	gag	gaa		624
Ala	Asn	Leu	Leu	Ala	Ser	Ser	Arg	Leu	Trp	Arg	Lys	Thr	Val	Glu	Glu		
		195					200					205					
acc	atc	aag	aac	gaa	ttc	cct	aca	ttg	aag	ggt	caa	cat	caa	ttg	att		672
Thr	Ile	Lys	Asn	Glu	Phe	Pro	Thr	Leu	Lys	Val	Gln	His	Gln	Leu	Ile		
		210				215					220						
gat	tct	gcc	gcc	atg	atc	cta	ggt	aag	aac	cca	acc	cac	cta	aat	ggt		720
Asp	Ser	Ala	Ala	Met	Ile	Leu	Val	Lys	Asn	Pro	Thr	His	Leu	Asn	Gly		
		225			230					235					240		
att	ata	atc	acc	agc	aac	atg	ttt	ggt	gat	atc	atc	tcc	gat	gaa	gcc		768
Ile	Ile	Ile	Thr	Ser	Asn	Met	Phe	Gly	Asp	Ile	Ile	Ser	Asp	Glu	Ala		
			245						250					255			
tcc	ggt	atc	cca	ggt	tcc	ttg	ggt	ttg	ttg	cca	tct	gcg	tcc	ttg	gcc		816
Ser	Val	Ile	Pro	Gly	Ser	Leu	Gly	Leu	Leu	Pro	Ser	Ala	Ser	Leu	Ala		
			260					265					270				
tct	ttg	cca	gac	aag	aac	acc	gca	ttt	ggt	ttg	tac	gaa	cca	tgc	cac		864
Ser	Leu	Pro	Asp	Lys	Asn	Thr	Ala	Phe	Gly	Leu	Tyr	Glu	Pro	Cys	His		
		275				280						285					
ggt	tct	gct	cca	gat	ttg	cca	aag	aat	aag	ggt	gac	cct	atc	gcc	act		912
Gly	Ser	Ala	Pro	Asp	Leu	Pro	Lys	Asn	Lys	Val	Asp	Pro	Ile	Ala	Thr		
		290				295					300						
atc	ttg	tct	gct	gca	atg	atg	ttg	aaa	ttg	tca	ttg	aac	ttg	cct	gaa		960

	Ile	Leu	Ser	Ala	Ala	Met	Met	Leu	Lys	Leu	Ser	Leu	Asn	Leu	Pro	Glu	
	305					310					315					320	
	gaa	ggt	aag	gcc	att	gaa	gat	gca	gtt	aaa	aag	gtt	ttg	gat	gca	ggt	1008
	Glu	Gly	Lys	Ala	Ile	Glu	Asp	Ala	Val	Lys	Lys	Val	Leu	Asp	Ala	Gly	
					325					330					335		
	atc	aga	act	ggt	gat	tta	ggt	ggt	tcc	aac	agt	acc	acc	gaa	gtc	ggt	1056
	Ile	Arg	Thr	Gly	Asp	Leu	Gly	Gly	Ser	Asn	Ser	Thr	Thr	Glu	Val	Gly	
				340					345					350			
	gat	gct	gtc	gcc	gaa	gaa	gtt	aag	aaa	atc	ctt	gct	taa				1095
	Asp	Ala	Val	Ala	Glu	Glu	Val	Lys	Lys	Ile	Leu	Ala					
				355				360									
	<210>	60															
	<211>	364															
	<212>	PRT															
	<213>	酿酒酵母															
	<400>	60															
	Met	Ser	Ala	Pro	Lys	Lys	Ile	Val	Val	Leu	Pro	Gly	Asp	His	Val	Gly	
	1				5					10					15		
	Gln	Glu	Ile	Thr	Ala	Glu	Ala	Ile	Lys	Val	Leu	Lys	Ala	Ile	Ser	Asp	
				20					25					30			
	Val	Arg	Ser	Asn	Val	Lys	Phe	Asp	Phe	Glu	Asn	His	Leu	Ile	Gly	Gly	
				35				40					45				
	Ala	Ala	Ile	Asp	Ala	Thr	Gly	Val	Pro	Leu	Pro	Asp	Glu	Ala	Leu	Glu	
				50			55					60					
	Ala	Ser	Lys	Lys	Val	Asp	Ala	Val	Leu	Leu	Gly	Ala	Val	Ala	Gly	Pro	
				65		70				75					80		
	Lys	Trp	Gly	Thr	Gly	Ser	Val	Arg	Pro	Glu	Gln	Gly	Leu	Leu	Lys	Ile	
				85						90					95		
	Arg	Lys	Glu	Leu	Gln	Leu	Tyr	Ala	Asn	Leu	Arg	Pro	Cys	Asn	Phe	Ala	
				100					105					110			
	Ser	Asp	Ser	Leu	Leu	Asp	Leu	Ser	Pro	Ile	Lys	Pro	Gln	Phe	Ala	Lys	
				115				120					125				
	Gly	Thr	Asp	Phe	Val	Val	Val	Arg	Glu	Leu	Val	Gly	Gly	Ile	Tyr	Phe	
				130			135					140					
	Gly	Lys	Arg	Lys	Glu	Asp	Asp	Gly	Asp	Gly	Val	Ala	Trp	Asp	Ser	Glu	
				145		150					155					160	
	Gln	Tyr	Thr	Val	Pro	Glu	Val	Gln	Arg	Ile	Thr	Arg	Met	Ala	Ala	Phe	
				165					170					175			
[0425]	Met	Ala	Leu	Gln	His	Glu	Pro	Pro	Leu	Pro	Ile	Trp	Ser	Leu	Asp	Lys	
				180					185					190			
	Ala	Asn	Leu	Leu	Ala	Ser	Ser	Arg	Leu	Trp	Arg	Lys	Thr	Val	Glu	Glu	
				195				200					205				
	Thr	Ile	Lys	Asn	Glu	Phe	Pro	Thr	Leu	Lys	Val	Gln	His	Gln	Leu	Ile	
				210			215					220					
	Asp	Ser	Ala	Ala	Met	Ile	Leu	Val	Lys	Asn	Pro	Thr	His	Leu	Asn	Gly	
				225		230					235					240	
	Ile	Ile	Ile	Thr	Ser	Asn	Met	Phe	Gly	Asp	Ile	Ile	Ser	Asp	Glu	Ala	
				245					250					255			
	Ser	Val	Ile	Pro	Gly	Ser	Leu	Gly	Leu	Leu	Pro	Ser	Ala	Ser	Leu	Ala	
				260					265				270				
	Ser	Leu	Pro	Asp	Lys	Asn	Thr	Ala	Phe	Gly	Leu	Tyr	Glu	Pro	Cys	His	
				275				280					285				
	Gly	Ser	Ala	Pro	Asp	Leu	Pro	Lys	Asn	Lys	Val	Asp	Pro	Ile	Ala	Thr	
				290			295					300					
	Ile	Leu	Ser	Ala	Ala	Met	Met	Leu	Lys	Leu	Ser	Leu	Asn	Leu	Pro	Glu	
				305		310				315						320	
	Glu	Gly	Lys	Ala	Ile	Glu	Asp	Ala	Val	Lys	Lys	Val	Leu	Asp	Ala	Gly	
				325						330				335			
	Ile	Arg	Thr	Gly	Asp	Leu	Gly	Gly	Ser	Asn	Ser	Thr	Thr	Glu	Val	Gly	
				340					345					350			
	Asp	Ala	Val	Ala	Glu	Glu	Val	Lys	Lys	Ile	Leu	Ala					
				355				360									
	<210>	61															
	<211>	675															
	<212>	DNA															
	<213>	酿酒酵母															
	<220>																
	<221>	CDS															
	<222>	(1)..(675)															
	<223>	酿酒酵母必需基因 TRP1															
	<400>	61															
	atg	tct	ggt	att	aat	ttc	aca	ggt	agt	tct	ggt	cca	ttg	gtg	aaa	gtt	48
	Met	Ser	Val	Ile	Asn	Phe	Thr	Gly	Ser	Ser	Gly	Pro	Leu	Val	Lys	Val	
				5					10						15		
	tgc	ggc	ttg	cag	agc	aca	gag	gcc	gca	gaa	tgt	gct	cta	gat	tcc	gat	96
	Cys	Gly	Leu	Gln	Ser	Thr	Glu	Ala	Ala	Glu	Cys	Ala	Leu	Asp	Ser	Asp	
				20					25				30				
	gct	gac	ttg	ctg	ggt	att	ata	tgt	gtg	ccc	aat	aga	aag	aga	aca	att	144

[0426]

```

Ala Asp Leu Leu Gly Ile Ile Cys Val Pro Asn Arg Lys Arg Thr Ile
   35   40   45
gac ccg gtt att gca agg aaa att tca agt ctt gta aaa gca tat aaa 192
Asp Pro Val Ile Ala Arg Lys Ile Ser Ser Leu Val Lys Ala Tyr Lys
   50   55   60
aat agt tca ggc act ccg aaa tac ttg gtt ggc gtg ttt cgt aat caa 240
Asn Ser Ser Gly Thr Pro Lys Tyr Leu Val Gly Val Phe Arg Asn Gln
   65   70   75   80
cct aag gag gat gtt ttg gct ctg gtc aat gat tac ggc att gat atc 288
Pro Lys Glu Asp Val Leu Ala Leu Val Asn Asp Tyr Gly Ile Asp Ile
   85   90   95
gtc caa ctg cat gga gat gag tcg tgg caa gaa tac caa gag ttc ctc 336
Val Gln Leu His Gly Asp Glu Ser Trp Gln Glu Tyr Gln Glu Phe Leu
  100  105  110
ggg ttg cca gtt att aaa aga ctc gta ttt cca aaa gac tgc aac ata 384
Gly Leu Pro Val Ile Lys Arg Leu Val Phe Pro Lys Asp Cys Asn Ile
  115  120  125
cta ctc agt gca gct tca cag aaa cct cat tcg ttt att ccc ttg ttt 432
Leu Leu Ser Ala Ala Ser Gln Lys Pro His Ser Phe Ile Pro Leu Phe
  130  135  140
gat tca gaa gca ggt ggg aca ggt gaa ctt ttg gat tgg aac tcg att 480
Asp Ser Glu Ala Gly Gly Thr Gly Glu Leu Leu Asp Trp Asn Ser Ile
  145  150  155  160
tct gac tgg gtt gga agg caa gag agc ccc gaa agc tta cat ttt atg 528
Ser Asp Trp Val Gly Arg Gln Glu Ser Pro Glu Ser Leu His Phe Met
  165  170  175
tta gct ggt gga ctg acg cca gaa aat gtt ggt gat gcg ctt aga tta 576
Leu Ala Gly Gly Leu Thr Pro Glu Asn Val Gly Asp Ala Leu Arg Leu
  180  185  190
aat ggc gtt att ggt gtt gat gta agc gga ggt gtg gag aca aat ggt 624
Asn Gly Val Ile Gly Val Asp Val Ser Gly Gly Val Glu Thr Asn Gly
  195  200  205
gta aaa gac tct aac aaa ata gca aat ttc gtc aaa aat gct aag aaa 672
Val Lys Asp Ser Asn Lys Ile Ala Asn Phe Val Lys Asn Ala Lys Lys
  210  215  220
tag
<210> 62
<211> 224
<212> PRT
<213> 酿酒酵母
<400> 62
Met Ser Val Ile Asn Phe Thr Gly Ser Ser Gly Pro Leu Val Lys Val
   1   5  10  15
Cys Gly Leu Gln Ser Thr Glu Ala Ala Glu Cys Ala Leu Asp Ser Asp
  20  25  30
Ala Asp Leu Leu Gly Ile Ile Cys Val Pro Asn Arg Lys Arg Thr Ile
  35  40  45
Asp Pro Val Ile Ala Arg Lys Ile Ser Ser Leu Val Lys Ala Tyr Lys
  50  55  60
Asn Ser Ser Gly Thr Pro Lys Tyr Leu Val Gly Val Phe Arg Asn Gln
  65  70  75  80
Pro Lys Glu Asp Val Leu Ala Leu Val Asn Asp Tyr Gly Ile Asp Ile
  85  90  95
Val Gln Leu His Gly Asp Glu Ser Trp Gln Glu Tyr Gln Glu Phe Leu
 100 105 110
Gly Leu Pro Val Ile Lys Arg Leu Val Phe Pro Lys Asp Cys Asn Ile
 115 120 125
Leu Leu Ser Ala Ala Ser Gln Lys Pro His Ser Phe Ile Pro Leu Phe
 130 135 140
Asp Ser Glu Ala Gly Gly Thr Gly Glu Leu Leu Asp Trp Asn Ser Ile
 145 150 155 160
Ser Asp Trp Val Gly Arg Gln Glu Ser Pro Glu Ser Leu His Phe Met
 165 170 175
Leu Ala Gly Gly Leu Thr Pro Glu Asn Val Gly Asp Ala Leu Arg Leu
 180 185 190
Asn Gly Val Ile Gly Val Asp Val Ser Gly Gly Val Glu Thr Asn Gly
 195 200 205
Val Lys Asp Ser Asn Lys Ile Ala Asn Phe Val Lys Asn Ala Lys Lys
 210 215 220
<210> 63
<211> 657
<212> DNA
<213> 酿酒酵母
<220>
<221> CDS
<222> (1) .. (657)
<223> 酿酒酵母必需基因 HIS3
<400> 63

```

	atg	aca	gag	cag	aaa	gcc	cta	gta	aag	cgt	att	aca	aat	gaa	acc	aag	48
	Met	Thr	Glu	Gln	Lys	Ala	Leu	Val	Lys	Arg	Ile	Thr	Asn	Glu	Thr	Lys	
	1				5					10					15		
	att	cag	att	gcg	atc	tct	tta	aag	ggg	ggg	ccc	cta	gcg	ata	gag	cac	96
	Ile	Gln	Ile	Ala	Ile	Ser	Leu	Lys	Gly	Gly	Pro	Leu	Ala	Ile	Glu	His	
				20					25					30			
	tcg	atc	ttc	cca	gaa	aaa	gag	gca	gaa	gca	gta	gca	gaa	cag	gcc	aca	144
	Ser	Ile	Phe	Pro	Glu	Lys	Glu	Ala	Glu	Ala	Val	Ala	Glu	Gln	Ala	Thr	
			35				40			45							
	caa	tcg	caa	gtg	att	aac	gtc	cac	aca	ggg	ata	ggg	ttt	ctg	gac	cat	192
	Gln	Ser	Gln	Val	Ile	Asn	Val	His	Thr	Gly	Ile	Gly	Phe	Leu	Asp	His	
		50				55				60							
	atg	ata	cat	gct	ctg	gcc	aag	cat	tcc	ggc	tcg	tcg	cta	atc	gtt	gag	240
	Met	Ile	His	Ala	Leu	Ala	Lys	His	Ser	Gly	Trp	Ser	Leu	Ile	Val	Glu	
	65				70					75					80		
	tgc	att	ggg	gac	tta	cac	ata	gac	gac	cat	cac	acc	act	gaa	gac	tgc	288
	Cys	Ile	Gly	Asp	Leu	His	Ile	Asp	Asp	His	His	Thr	Thr	Glu	Asp	Cys	
				85					90					95			
	ggg	att	gct	ctc	ggg	caa	gct	ttt	aaa	gag	gcc	cta	ctg	gcg	cgt	gga	336
	Gly	Ile	Ala	Leu	Gly	Gln	Ala	Phe	Lys	Glu	Ala	Leu	Leu	Ala	Arg	Gly	
				100					105					110			
	gta	aaa	agg	ttt	gga	tca	gga	ttt	gcg	cct	ttg	gat	gag	gca	ctt	tcc	384
	Val	Lys	Arg	Phe	Gly	Ser	Gly	Phe	Ala	Pro	Leu	Asp	Glu	Ala	Leu	Ser	
			115				120						125				
	aga	gcg	gtg	gta	gat	ctt	tcg	aac	agg	ccg	tac	gca	gtt	gtc	gaa	ctt	432
	Arg	Ala	Val	Val	Asp	Leu	Ser	Asn	Arg	Pro	Tyr	Ala	Val	Val	Glu	Leu	
		130					135					140					
	ggg	ttg	caa	agg	gag	aaa	gta	gga	gat	ctc	tct	tgc	gag	atg	atc	ccg	480
	Gly	Leu	Gln	Arg	Glu	Lys	Val	Gly	Asp	Leu	Ser	Cys	Glu	Met	Ile	Pro	
	145				150					155				160			
	cat	ttt	ctt	gaa	agc	ttt	gca	gag	gct	agc	aga	att	acc	ctc	cac	gtt	528
	His	Phe	Leu	Glu	Ser	Phe	Ala	Glu	Ala	Ser	Arg	Ile	Thr	Leu	His	Val	
				165					170					175			
	gat	tgt	ctg	cga	ggc	aag	aat	gat	cat	cac	cgt	agt	gag	agt	gcg	ttc	576
	Asp	Cys	Leu	Arg	Gly	Lys	Asn	Asp	His	His	Arg	Ser	Glu	Ser	Ala	Phe	
				180					185					190			
	aag	gct	ctt	gcg	gtt	gcc	ata	aga	gaa	gcc	acc	tcg	ccc	aat	ggg	acc	624
	Lys	Ala	Leu	Ala	Val	Ala	Ile	Arg	Glu	Ala	Thr	Ser	Pro	Asn	Gly	Thr	
			195					200					205				
[0427]	aac	gat	gtt	ccc	tcc	acc	aaa	ggg	gtt	ctt	atg						657
	Asn	Asp	Val	Pro	Ser	Thr	Lys	Gly	Val	Leu	Met						
		210					215										
	<210>	64															
	<211>	219															
	<212>	PRT															
	<213>	酿酒酵母															
	<400>	64															
	Met	Thr	Glu	Gln	Lys	Ala	Leu	Val	Lys	Arg	Ile	Thr	Asn	Glu	Thr	Lys	
	1				5					10					15		
	Ile	Gln	Ile	Ala	Ile	Ser	Leu	Lys	Gly	Gly	Pro	Leu	Ala	Ile	Glu	His	
				20					25					30			
	Ser	Ile	Phe	Pro	Glu	Lys	Glu	Ala	Glu	Ala	Val	Ala	Glu	Gln	Ala	Thr	
			35				40						45				
	Gln	Ser	Gln	Val	Ile	Asn	Val	His	Thr	Gly	Ile	Gly	Phe	Leu	Asp	His	
		50				55				60							
	Met	Ile	His	Ala	Leu	Ala	Lys	His	Ser	Gly	Trp	Ser	Leu	Ile	Val	Glu	
	65				70					75					80		
	Cys	Ile	Gly	Asp	Leu	His	Ile	Asp	Asp	His	His	Thr	Thr	Glu	Asp	Cys	
				85					90					95			
	Gly	Ile	Ala	Leu	Gly	Gln	Ala	Phe	Lys	Glu	Ala	Leu	Leu	Ala	Arg	Gly	
			100						105					110			
	Val	Lys	Arg	Phe	Gly	Ser	Gly	Phe	Ala	Pro	Leu	Asp	Glu	Ala	Leu	Ser	
			115					120					125				
	Arg	Ala	Val	Val	Asp	Leu	Ser	Asn	Arg	Pro	Tyr	Ala	Val	Val	Glu	Leu	
		130					135					140					
	Gly	Leu	Gln	Arg	Glu	Lys	Val	Gly	Asp	Leu	Ser	Cys	Glu	Met	Ile	Pro	
	145				150					155					160		
	His	Phe	Leu	Glu	Ser	Phe	Ala	Glu	Ala	Ser	Arg	Ile	Thr	Leu	His	Val	
				165					170					175			
	Asp	Cys	Leu	Arg	Gly	Lys	Asn	Asp	His	His	Arg	Ser	Glu	Ser	Ala	Phe	
			180						185					190			
	Lys	Ala	Leu	Ala	Val	Ala	Ile	Arg	Glu	Ala	Thr	Ser	Pro	Asn	Gly	Thr	
			195					200					205				
	Asn	Asp	Val	Pro	Ser	Thr	Lys	Gly	Val	Leu	Met						
		210					215										
	<210>	65															
	<211>	290															
	<212>	DNA															

<213> 大肠杆菌
 <220>
 <221> 启动子
 <222> (1) .. (290)
 <223> 负荷感应启动子 p (htpG)
 <400> 65
 gccaccaacc ccacggttga aggtgaagct accgctaact acattgccga gctttgcgcg 60
 caatatgacg tggaagccag cogaatcgct catggcgtag cggttggcgg cgaactggaa 120
 atggtcgatg gcaccacgct gtcacaactcc cttgccgggc gtcataagat tcgtttttta 180
 gcaaacgaga gcaggatcac ctgctctcgc ttgaaattat tctcccttgt ccccatctct 240
 ccctcatcct gtttttaacc ttaaaatggc attattgagg tagacctaca 290
 <210> 66
 <211> 257
 <212> DNA
 <213> 大肠杆菌
 <220>
 <221> 启动子
 <222> (1) .. (257)
 <223> 负荷感应启动子 p (ibpA)
 <400> 66
 atcccttgct gaaaataaca tcatcattac gtgcgactgt ggcggtatc gcactttaac 60
 gtttcgtgct gcccctcag tctatgcaat agaccataaa ctgcaaaaaa aagtccgctg 120
 ataaggcttg aaaagttcat ttccagacc atttttacat cgtagccgat gaggacgcgc 180
 ctgatgggtg ttctggctac ctgacctgtc cattgtggaa ggtcttacat tctcgctgat 240
 ttcaggagct attgatt 257
 <210> 67
 <211> 298
 <212> DNA
 <213> 大肠杆菌
 <220>
 <221> 启动子
 <222> (1) .. (298)
 <223> 负荷感应启动子 p (clpB)
 <400> 67
 gataagtatc tggcggtat tttctgctt gcccggcagc gtctggcgag cgttgggtgtt 60
 gaacaaattt tcggcggcga ccgttgtaca tatacggaaa atgagacttt cttctcttat 120
 gctcgcgaca agaccaccg gcgtatggca agtttcatat ggctgatata acctaaagaa 180
 tcaagacgat ccggtacgcg tgattttctt ttcacattaa tctgggtcaat aaccttgaat 240
 aattgaggga tgacctcatt taatctccag tagcaacttt gatccgttat gggaggag 298
 <210> 68
 <211> 300
 <212> DNA
 <213> 大肠杆菌
 <220>
 <221> 启动子
 <222> (1) .. (300)
 <223> 负荷感应启动子 p (yccV)
 <400> 68
 ggttataaag acatcaacat gctggcgatt cagctgctga atgaaggcgg tattctctctg 60
 actttctcct gttccggtct gatgaccagc gatttatttc agaaaatcat ccggtatgcc 120
 gcaattgatg ccggtcgtga tgtacaattt atagagcagt tccgtcaggc agccgatcat 180
 ccggtgatcg ctacctatcc ggaagggtta tatctgaaag ggtttgcttg tcgctcatg 240
 taacttgaaa tgtggaattt tacccttata tcaaaggtaa agatttcccg ggaggtgact 300
 <210> 69
 <211> 300
 <212> DNA
 <213> 大肠杆菌
 <220>
 <221> 启动子
 <222> (1) .. (300)
 <223> 负荷感应启动子 p (grpE)
 <400> 69
 tctccgcgag cgtgccagtt ttcacattct tcagttgcag ttcgtgagcg atttgttgct 60
 caacgatgac ctcgtaacct tttgtgcaca gccagcggtg gagcatttca tgtgttgtca 120
 gtgcagtggg gtgcggtggg tgtcccacaa tgccaatata cttgaaatga ttattcattt 180
 ttccgaggtc cttgttgcca agattgatga caatgtgagt gcttcccttg aaacctgaa 240
 actgatccc ataataagcg aagtttagcga gatgaatgcg aaaaaaacgc ggagaaattc 300
 <210> 70
 <211> 300
 <212> DNA
 <213> 大肠杆菌
 <220>
 <221> 启动子
 <222> (1) .. (300)
 <223> 负荷感应启动子 p (ycjX)

[0429]

```

<400> 70
ttggctgata ctgtaattct tcaatggccc gtaacacttt ttcgcgcgct tccggggaga 60
tattagtttg cttattcagc acgcgtgata cgggtggattt tgatacgctt gcaaccctgg 120
caatatcata aatagtaggg gacataggtc aggaactccg tccgaacgtg aatgctgcac 180
atcttatgga gttcgcgcgt cgatgacaac acgggcagtt gataatcaat ggcttgcccc 240
ccacattcat atccttacga atgatttttt ttctgggtctt cagagcagca caggacagca 300
<210> 71
<211> 300
<212> DNA
<213> 大肠杆菌
<220>
<221> 启动子
<222> (1) .. (300)
<223> 负荷感应启动子 p (ldhA)
<400> 71
gggaaccac agccccgagc tcatcagcag cgtcaacggc acaagaataa tcagtaataa 60
cagcgcgaga acggctttat atttaccag catgggtagt taatatcctg atttagcgaa 120
aaattaagca ttcaatacgg gtattgtggc atgttttaacc gttcagttga aggttgccgc 180
tacactaagc atagttgttg atgaattttt caatatcgcc atagctttca attatatattg 240
aaattttgtg aaatattttt agtagcttaa atgtgattca acatcactgg agaaagtctt 300
<210> 72
<211> 300
<212> DNA
<213> 大肠杆菌
<220>
<221> 启动子
<222> (1) .. (300)
<223> 负荷感应启动子 p (mutM)
<400> 72
ggcgaaaagt actaagtact tagaggaaat aaatcatggc taaagggtatt cgtgagaaaa 60
tcaagctggg ttcttctgct ggtactggtc acttctatac cactacgaag aacaaacgta 120
ctaagccgga aaaactggaa ctgaaaaaat tccgatccag tgttcgccag cacgtgatct 180
acaaagaagc gaaaatcaaa taattctcgc ttgatgtaa caaaaaacct cgctccggcg 240
gggttttttg ttatctgctt gcccccatat tgactgcac tgttcattcc tggagatgct 300
<210> 73
<211> 300
<212> DNA
<213> 大肠杆菌
<220>
<221> 启动子
<222> (1) .. (300)
<223> 负荷感应启动子 p (ybbN)
<400> 73
cccattcgta ctgcgttcac cgacaacgtc aaccagacgc aaagtataa accagtcgaa 60
aatccccgca tcgccgccgc ctttacgttg ggaccggaag cgggtggtgga caaagtacgc 120
catgctttta ttagecgaga gccgaagatg cgtatccgg taacgctggg gacctgggca 180
gtaatggtgc ttaagcgctt gctgccgggg cgctgatgg acaaaatatt gcaggggtga 240
gttgaagcgc gcgcttaagc ccccatgtca aagaaaatat cgacaacaga gagcgactcc 300
<210> 74
<211> 298
<212> DNA
<213> 大肠杆菌
<220>
<221> 启动子
<222> (1) .. (298)
<223> 负荷感应启动子 p (prlC)
<400> 74
agttcttgtt tgagtccact catgggttga tgtccgattg cgcctaaatt ttgggcaact 60
gcgtagattt tcgatggtag cacaatcaga ttccgttatg acggtgatga agaaattgcg 120
atgaaatgtg aggtgaatca gggttttcac ccgattttgt gctgatcaga attttttttc 180
tttttcccc ttgaaggggc gaagcctcat cccattttct ctggtcacca gccgggaaac 240
cacgtaagct ccggcgctac ccataacaga tacggacttt ctcaaaggag agttatca 298
<210> 75
<211> 298
<212> DNA
<213> 大肠杆菌
<220>
<221> 启动子
<222> (1) .. (298)
<223> 负荷感应启动子 p (groES)
<400> 75
agttcttgtt tgagtccact catgggttga tgtccgattg cgcctaaatt ttgggcaact 60
gcgtagattt tcgatggtag cacaatcaga ttccgttatg acggtgatga agaaattgcg 120
atgaaatgtg aggtgaatca gggttttcac ccgattttgt gctgatcaga attttttttc 180
tttttcccc ttgaaggggc gaagcctcat cccattttct ctggtcacca gccgggaaac 240
cacgtaagct ccggcgctac ccataacaga tacggacttt ctcaaaggag agttatca 298

```

[0430]

```

<210> 76
<211> 392
<212> DNA
<213> 大肠杆菌
<220>
<221> 启动子
<222> (1) .. (392)
<223> 负荷感应启动子 p (fxsA)
<220>
<221> 尚未归类的特征
<222> (328) .. (328)
<223> N 是 A、T、C 或 G
<400> 76
gatccccgga atattcatat gttaaacaca cagaatatat gtgggttcga atgtttttcg 60
accgacgatt atccccctgca tcgaccgaat acccgagatc atatgctgct tgaggatttc 120
taccgtaatc tggatcactt taagtgtcgg tttttacccc ttaattatta atttgtgaaa 180
tagatcaccg ctttgggatt actacaaaaa atagttgcgc aaacatcttg aaattttgct 240
aatgaccaca atataagctg aacgcgattc gcaaccatt caggtagccg gggtttaccg 300
gctgctatta caggagaaac ctacttgnaa ggtatgaaac tctttgccc aagggtactca 360
ctggacctta gccatcctca cgtaatgggg at 392
<210> 77
<211> 300
<212> DNA
<213> 大肠杆菌
<220>
<221> 启动子
<222> (1) .. (300)
<223> 负荷感应启动子 p (htpX)
<400> 77
agattaccag gagccggatc cttatctgga tgagacggtg aatatcgcac tcgatctggc 60
gaagcttgaa aaagccagac ccgcggaaca acccgctccc gtcaagtaat atcaatcagg 120
cacaagaaat tgtgcctgat tttttaacag ccgcaagatg ccgtaaatca gatgctacaa 180
aatgtaaaat tgtgtctttc ttgtgactta cgcactatcc agacttgaaa atagtcgcgt 240
aaccatacgt atgtgggtat cgcataattgc gttttgttaa actgaggtaa aaagaaaatt 300
<210> 78
<211> 275
<212> DNA
<213> 大肠杆菌
<220>
<221> 启动子
<222> (1) .. (275)
<223> 负荷感应启动子 p (rrnB)
<400> 78
gcaaacacgc cgcgcgggtca gcgggggttct cctgagaatt ccggcgagaga aagcaaaaat 60
aaatgcttga ctctgtagcg ggaaaagcgta ttatgcacac ccgcgcgcgc tgagaaaaag 120
cgaagcggca ctgctcttta acaattttatc agacaattctg tgtgggcact cgaagatac 180
gattcttaac gtgcgaagac gaaaaatgaa taccaagtct caagagtga cgcgtaattc 240
attacgaagt ttaattcttt gagcgtcaaaa ctttt 275
<210> 79
<211> 296
<212> DNA
<213> 大肠杆菌
<220>
<221> 启动子
<222> (1) .. (296)
<223> 负荷感应启动子 p (rrnE)
<400> 79
aatgcgcctc cgttgagacg acaacgtgaa acacttcaca ggatgggtcgg aacaacgaag 60
agaaaaaaat cctgaaattc aggggttgact ctgaaaagagg aaagcgtaat atacgccacc 120
tcgcgacagt gagctgaaag ccgcgtcgca actgctcttt aacaattttat cagacaattct 180
gtgtgggcac tcgaagatac ggattcttaa cgtcgcaaga cgaaaaatga ataccaagtc 240
tcaagagtga acacgtaatt cattacgaag ttttaattctt tgagcatcaa actttt 296
<210> 80
<211> 300
<212> DNA
<213> 大肠杆菌
<220>
<221> 启动子
<222> (1) .. (300)
<223> 负荷感应启动子 p (cspD)
<400> 80
cggttatcgg cagaacgccc tgccgctgac aagagcttat ttgcctgacc aatctaactcc 60
acatgataga taactatcat ctatcgttgg catcagcgac atctgtcaca ttctgtcaa 120
tagcgtaaac tgcttcaaat ttttgattca tacctgcccc ctgtccctct tcaaattgctt 180
gacggctcgc cctaattctc taaattgtat ttctagagtt ggcgaggttt tgaacagccc 240

```


	cctctctgac	cccggtttat	tccatcttac	ttgtataaga	tttgcaagg	atgtcgaagc	300
<210>	81						
<211>	299						
<212>	DNA						
<213>	大肠杆菌						
<220>							
<221>	启动子						
<222>	(1) .. (299)						
<223>	负荷感应启动子 p (katE)						
<400>	81						
	taaatccatc	tttatcgcat	gttctggcgg	cgcggttcc	gtgcgtgga	catagctaata	60
	aatctggcgg	ttttgctggc	ggagcggttt	cttcattact	ggcttcacta	aacgcataatt	120
	aaaaatcaga	aaaactgtag	tttagccgat	ttagccctg	tacgtccgc	tttgctgta	180
	tttcataaca	ccgtttccag	aatagtctcc	gaagcgggat	ctggctggtg	gtctatagtt	240
	agagagtttt	ttgacaaaa	cagcggccct	ttcagtaata	aattaaggag	acgagttca	299
<210>	82						
<211>	300						
<212>	DNA						
<213>	大肠杆菌						
<220>							
<221>	启动子						
<222>	(1) .. (300)						
<223>	负荷感应启动子 p (xthA)						
<400>	82						
	caaaccatata	gcattttaaag	tgcataataa	gtgaatacgt	ttgcgatgtg	ggtgaataaaa	60
	aagaataaaa	aacgcaatgt	tatgcagaag	taaaataata	ttctggaatt	gtgatcattg	120
	acgaaattta	ctggaaatta	ctgcgccatt	ctgacgcagc	gcgcacaaa	agcgggcatt	180
	ttttgcgcca	togttgacat	cattaacaat	catcgatcaa	atcacttaac	aacaggcggt	240
	aagcaacgcg	aaattctgct	accatccacg	cactctttat	ctgaataaat	ggcagcgact	300
<210>	83						
<211>	301						
<212>	DNA						
<213>	大肠杆菌						
<220>							
<221>	启动子						
<222>	(1) .. (301)						
<223>	负荷感应启动子 p (uspE)						
<400>	83						
	taatatctgg	gcctgacggt	agaaaccatc	agccgtctgc	tgggtcgctt	ccagaaaagc	60
	ggtatgctgg	cagtcaaaag	taaatacatc	accatcgaaa	ataacgatgc	gctggcccag	120
	cttgctggtc	atacgcgtaa	cgttgcctga	tttttccgca	taactcacta	tccttctgtc	180
	atatcattaa	atttttctga	tttattgatc	tggcagaagg	ttcatcactg	tttcattcac	240
	cagatatggg	ttaatctttt	aattacaaac	tgcgttgaca	gttggtgtaa	ggagaccctg	300
	t						301
<210>	84						
<211>	300						
<212>	DNA						
<213>	大肠杆菌						
<220>							
<221>	启动子						
<222>	(1) .. (300)						
<223>	负荷感应启动子 p (gadB)						
<400>	84						
	taattttata	aatgcgttca	aaataataat	caagtactaa	tagtgatatt	ttaaggctctg	60
	attttttacgt	gataattcag	gaggcacaga	atgcgcataa	aaataacagc	ataaaacacc	120
	ttaccaccac	ccaagaattt	catattgtat	tgtttttcaa	tgaaaaaata	ttattcgogt	180
	aatatctcac	gataaataac	attaggattt	tgttatttaa	acaagagtcc	tttgcaacttg	240
	cttactttat	cgataaatcc	tactttttta	atgcgatcca	atcattttta	ggagttttaa	300
<210>	85						
<211>	300						
<212>	DNA						
<213>	大肠杆菌						
<220>							
<221>	启动子						
<222>	(1) .. (300)						
<223>	负荷感应启动子 p (ahpC)						
<400>	85						
	tagatcaggt	gattgcccctt	tgtttatgag	ggtgttgtaa	tccatgtcgt	tgttgcatttt	60
	gtaagggcaa	cacctcagcc	tgagggcagg	cactgaagat	accaaagggt	agttcagatt	120
	acacggtcac	ctggaaaagg	ggccatttta	ctttttatcg	ccgctggcgg	tgcaaagtcc	180
	acaaagtgtg	cttacgaagg	ttgtaaggta	aaacttatcg	atttgataat	ggaaacgcac	240
	tagccgaatc	ggcaaaaaatt	ggttacctta	catctcatcg	aaaacacgga	ggaagtatag	300
<210>	86						
<211>	300						
<212>	DNA						

<213> 大肠杆菌
 <220>
 <221> 启动子
 <222> (1)..(300)
 <223> 负荷感应启动子 p(katG)
 <400> 86
 gtgatcacaa atttttaaaca gagcacaaaa tgctgcctcg aaatgagggc gggaaaataa 60
 gggttatcagc cttgtttttct ccttcattac ttgaaggata tgaagctaaa accctttttt 120
 ataaagcatt tgtccgaatt cggacataat caaaaaagct taattaagat caatttgatc 180
 tacatctctt taaccaacaa tatgtaagat ctcaactatc gcatccgtgg ataatattcaa 240
 ttataacttc tctctaacgc tgtgtatcgt aacggttaaca ctgtagaggg gagcacattg 300
 <210> 87
 <211> 300
 <212> DNA
 <213> 大肠杆菌
 <220>
 <221> 启动子
 <222> (1)..(300)
 <223> 负荷感应启动子 p(grxA)
 <400> 87
 tagaatctgc acttctggcc gagaaaatgg ctcaggcagc gacagggagt cgcttaccga 60
 cagcagcgcc accgccagta gcatcattcc gataaattcc agtatcaaca cgccttttagg 120
 caattttaccg atcgcgcgca tacgcttccc tctgcaaagt gagccttcag tctaaaactt 180
 ttcaactgtat tgtgttttaac agttatagct ttttagcaatt aatgcaacag gttaaaccta 240
 ctttcagcga atacattttta gcgtgatcat tacaggcata aatctatgag gagagaaata 300
 <210> 88
 <211> 299
 <212> DNA
 <213> 大肠杆菌
 <220>
 <221> 启动子
 <222> (1)..(299)
 <223> 负荷感应启动子 p(oxyS)
 <400> 88
 taactgatcc accagcagca ttcccgcctg ggtgaacaac actttacggc tgggtccgctc 60
 cagcaacatc acgcccagct catcttccag cttacgaatt tgcccgtctaa gcgtcggctg 120
 gctaacgtgg caggaatctg ccgcacgcgcg aaaaatggcgg tgttcagcca atgccaccag 180
 gtactcaaga tcacgaatat tcattatcca tctcccatcg ccacgatagt tcatggcgat 240
 aggtagaata gcaatgaacg attatcccta tcaagcattc tgactgataa ttgctcaca 299
 <210> 89
 <211> 300
 <212> DNA
 <213> 大肠杆菌
 <220>
 <221> 启动子
 <222> (1)..(300)
 <223> 负荷感应启动子 p(poxB)
 <400> 89
 taggttgctg ctgcctgccg tgcgggtggt tgcggctgct gtaagacaaa agtgggtttcc 60
 ggtgaatata cggtgagcag cacaatgacg ctgaccgacg ccgaaatcgc tgaagggttac 120
 gtactggcct gctcctgcc a tccgcagggg gatttggttc tgcataatc gccttatgcc 180
 cgatgatatt cctttcatcg ggctatttta ccggttagtgc ctccctttctc tcccatccct 240
 tccccctccg tcagatgaac taaacttggt accgttatca cattcaggag atggagaacc 300
 <210> 90
 <211> 299
 <212> DNA
 <213> 大肠杆菌
 <220>
 <221> 启动子
 <222> (1)..(299)
 <223> 负荷感应启动子 p(trxC)
 <400> 90
 tagcgggttg tgggtacggt tatcagaatc atcatcactg cggacataca tcactttgac 60
 cttgcccgtt ttacctttca ttctatcggt catgtttttc tccaccagcg ctgcgcggaag 120
 cgcgcagatt acccgaagtc cgcgcgggttc gccatgattt cgtaccaaag cctgcgacta 180
 tcatacctat tgaataaaac agattgttgt ctggaacaat gtccccgata atatgtaaca 240
 tattagaaac ataccggcgt cggttgccgat aagtctcctt actcatcccg aggttagtt 299
 <210> 91
 <211> 1000
 <212> DNA
 <213> 酿酒酵母
 <220>
 <221> 启动子
 <222> (1)..(1000)
 <223> 负荷感应启动子 pKAR2

```

<400> 91
gtgggagtc aatcaaatccc aatcgacgct attactggac aagtcaactt tattcgcaaa 60
acaagtcttg taatcatcac caaaagtctc aatgaatgct tcttcattaa gatactttga 120
ccattccata ggcggaacg tgttttttate caacactctg tcaactcttg accgtggctt 180
tgggcgccctt tcatTTTTgtc ttacataaaa gtatctcagc agcattattg cttcttttct 240
caagatcccc ggtatactct cgtaccctgc cgcactattg ttcttgggca ctaatgtaca 300
cgtatcatga tttactgaca agacagtacc gttgcctcca aatctctcgt tgccacaacc 360
gaacaccacc tttccaatgt ctaatttgctt gagagcagat gcacacatta tacacggttc 420
tacagtaaca tatagagtaa tgtcttttgaa cacgtcaaca actcctcggg agcccaacat 480
cgccttgatc tgatcgatcc ccatgaactc agcatgtgct actccagtta atgacttggt 540
cgtatcgctt atgccataag ccatcacctg gccagttggc gtatgtacaa agatgcaagc 600
taccggtgtc tcatcgtggt caagagcgta tctagccaaa cggacagctg tctcatatg 660
tttaatatgc tgcatagtgt gagtctctta gtttttaccg cagccaccag ccgcttctcg 720
agcaaatgt agatccatt aggactcat attcatctaa ttttgctatg ttactgtcaa 780
ctttctattt taatagaacc ttctggaaat ttcacccggc gcggcacccg aggaactgga 840
cagcgtgtcg aaaaagttgc ttttttatat aaaggacacg aaaagggttc tctggaagat 900
ataaatatgg ctatgtaatt ctaaagatta acgtgttact gttttacttt tttaaagtc 960
ccaagagtag tctcaaggga aaaagcgtat caaacatacc 1000
<210> 92
<211> 1000
<212> DNA
<213> 酿酒酵母
<220>
<221> 启动子
<222> (1)..(1000)
<223> 负荷感应启动子 pPDI1
<400> 92
caggctcgta gatgtgacgt tctcatcgct tagcaaatca tcgggaatct tggacttcat 60
ttgctccatg gagaaagtat gatctccatg caagttaaca gaacatatac ggaaattggg 120
accaccagg tcggcggtga gtaaaacacc gcgtccgctc ccgttgggtg acccggtgac 180
gaacgccgga atcatagga gacctttgtc cgaggccaat gtgtggcctt cctttgggtg 240
agctagacct ttttccattt gttcgatgaa gtaagcagtt aattcgtcca gcttctcggg 300
ggtaacctcg aaatcgctgc agatctggct cacggcctgg atgaccgctc tctcagtggt 360
tttgtgaag tcgtcgaatg acatcttgtg tatgatagag ttgtattagt ggtgggtgtg 420
tggttttcgg aagctgagtt gggcactgat ggggggggtg ggggttctgt tgatagcgga 480
tatttatatg ggcgcggttt cattagtggc cacgttttga gcgtggtgtg acaccacgcc 540
caagataggg ggccccctga aaggggttga ccgtccgtcg gcacaccact tataatgcgg 600
ggtgcaagcg ccgctgctaa aatttttttt ttttccattt ttgtcgttat ttttatttcc 660
cgttttttgt tttttttgat tttttcggag cgacaaacct ttcgaaacac gtgtcctgaa 720
aattatcctg ggctgcacgt gataatatgt taccctgtcg ggcggcgctt ctttttccct 780
tttctctcac tagtctctt ttccaatttg ccacgtgta ccattttgtt gtgctgttac 840
aaccacaaca aaacgaaaaa ccggtatgga catacatata tatatatata tatatatata 900
ttttgttacg cgtgcatttt cttgttgcaa gcagcatgtc taattggtaa ttttaaagct 960
gccaagctct acataaagaa aaacatacat ctatcccggt 1000
<210> 93
<211> 1000
<212> DNA
<213> 酿酒酵母
<220>
<221> 启动子
<222> (1)..(1000)
<223> 负荷感应启动子 pSSA1
<400> 93
cactcatatc ggggggtctta cttcccatca taaacacgga ccaaagaatt aaattcatca 60
ttcacataag cgaaaaaaaag taatcattat tagttaactt ccctttacat gtaaccatag 120
gaaaaaaa aaagaaagaa aaacaacgat ctataaagta aactaaatcg gaacataccc 180
ttcaatgtat gacttgtcag ccaaagaagc gtttaattgt ttcaaagttt caatcttggg 240
gaaatcggtg gatgccatta tttcggttga tttcgggttg gactctcag tgctgaaagg 300
tagaaagcat gaggtaaaaa gataaagagt agaaggaaaa aaaaaagaa caactaaata 360
tagtggttag ataaccttaa actattttatt tagaggcgac aatgtgaaaa attttttctt 420
caaaggctcg gttgtcgaca aattgttacg ttgtgctttg atttctaaag cgcttcttca 480
cctgcagggt ctgagcccta agaaaaaaa tttccttggg tgaaaatggc ggaaaaaaa 540
aattcagaaa aagaaataaa gcacgtgtgc gcggtgtgtg gatgatggtt tcatcattgt 600
caacggcatt ttcgttcttg tggattgttg taaacttttc agaacattct agaaagaaag 660
cacacggaac gttttagaagc tgtcattttg gttttttctc cagatttttag ttgagaaagt 720
aattaaatta tttctctttt tccagaacgt tccatcggtg gcaaaaggga gagaaagaa 780
caaaaaagaa gggggggccat ttagattagc tgatcgtttc gaggacttca aggttatata 840
aggggtggat tgatgtatct tcgagaaggg attgagttgt agtattcggt cccaattctt 900
acttaagttg ttttattttt tctatttgta agataagcac atcaaaagaa aagtaatcaa 960
gtattacaag aaacaaaaat tcaagtaaat aacagataat 1000
<210> 94
<211> 1000
<212> DNA
<213> 酿酒酵母
<220>
<221> 启动子
<222> (1)..(1000)

```

<223> 负荷感应启动子 pFPR2		
<400> 94		
ctgatctttt	cacgcagtta	ggtaaggaaa
actctacaac	gttcccacac	cacgttagat
ttgctatcca	taatatgatc	aataatctga
agtacgcgaa	attagaaaaa	ccacaaccac
aagattaccg	tgaaggtact	gccaaagccaa
aaaagtccaa	tggtttataaa	atagttggga
acaaggatgt	ccttgtcaaa	tattacgcga
ctatttaaga	agaaattgca	aatgtcttag
tgatcgccga	agtagattca	ggggcaaatg
caaccattgc	tttgtatcct	gccggaaata
gaaatttgga	agatgttttc	gaatttatca
aggcaattta	tgataaattg	caccaggcca
tacatgatga	attataatca	ataaataaag
tactcgcatc	gtttctggaa	gaaaatagct
tctcaccagt	taccggttta	tgctcttgaa
aaatacagg	ttatatgaga	cccattacta
cgttttcaat	cacattctag	tttacagtaa
<210> 95		
<211> 1000		
<212> DNA		
<213> 酿酒酵母		
<220>		
<221> 启动子		
<222> (1) .. (1000)		
<223> 负荷感应启动子 pRPL3		
<400> 95		
cctttctctgc	gggatagctt	ttgcttcagg
tttccgacaa	tttctcacta	tgccctctatc
tttctaaact	ctgcacacta	gaccgagaca
gacggaaaaa	ctagaagtaa	tcttatgcgg
gaatgacact	acatactttg	atacggatgt
gatgcctgtg	ctgctgccat	tacgtacagg
tcacttgctg	cttttttttg	tgatgaagt
tcagccaatt	acgttatcat	tctaaagtaa
ttcaagaatt	gctcgataat	tgccaacaaa
aacacagaaa	tgatgcaaac	aatttattaa
ggtggaggga	ttaacgcatt	tcagacagct
tgctgcctgt	tcgcagtcac	tggttgcaaa
aatggcaata	ttctttttta	ggttaaccgg
gtttttcaag	aagcgacgca	caactgtttt
tcgtccatga	aaaaaatttt	tcattttgtc
attttattag	ttgtctttgt	caatctctac
tgctcatctc	agaacaacac	agttactaca
<210> 96		
<211> 1000		
<212> DNA		
<213> 酿酒酵母		
<220>		
<221> 启动子		
<222> (1) .. (1000)		
<223> 负荷感应启动子 pRPL6A		
<400> 96		
attctcgtgc	aagcagtttt	gcgcgtactt
ttaccgtggt	tgctgcggaa	aacggttgcca
aaggtggtag	actaaaacac	atatacaaaa
ttgccgtatt	tgatgaagag	tgctcgttca
gtgcaatttc	ccatcaaaga	ctatcgaggg
tggttgaccc	tcaaattcaa	caagatggcc
tgattgtaaa	gattaattac	ggtaaagaca
ttcctcaaga	atacacgcag	taataaacag
ggcacatttg	gtgtgttaac	aggtgtcttt
tatataacgt	tatatattact	ataatatatt
ctcctgttac	tgcttaggga	gaagaaagaa
attaaaatgt	atggattttg	tatatatttc
gcggcccgcc	tagaccattg	cggttgccgc
ggcgacaac	agaagcgaga	gacacgtttc
gtttcccaact	ggagtgtgct	ttgttccctt
aatggaaggg	ttacatgaat	agattaaatg
tacattagcg	gaaacaacaa	agtaatagcc
<210> 97		
<211> 1000		
<212> DNA		
<213> 酿酒酵母		
<220>		
<221> 启动子		
attgcattaa	60	
gagagaacag	120	
ccagaagaag	180	
cagttggtaa	240	
ccaaaagaac	300	
catgatgatg	360	
aggtttgcgc	420	
gataaaatct	480	
acaggatatc	540	
aataaaatta	600	
attgacggcc	660	
gaagataccg	720	
aacatctgat	780	
tcacgatatt	840	
gcgcattctt	900	
ggagtgcctt	960	
	1000	
tttctgaaaa	tcaatatatt	60
agaaacactt	tctaccctct	120
gcccagagaga	aaatccggag	180
ccagaacgcc	cagaaaaacgt	240
ccgctttttt	ctattttatc	300
caaacaagaa	tagcactttt	360
ttctgttagt	gccgctcgta	420
tgtgcaaacc	ttagtagacg	480
agccgggaag	ttaagtcctc	540
aatgcaacag	caaggagaaa	600
aaagaaagca	ctcgcaaac	660
tactcacgca	cactggaatg	720
atataagta	cgttttgtac	780
tttttttttt	tcagtgatca	840
tcctggatat	ataaaatac	900
tcattatttc	atttcggttt	960
		1000

<222> (1) .. (1000)
 <223> pRPL28
 <220>
 <221> 启动子
 <222> (1) .. (1000)
 <223> 负荷感应启动子 pRPL28
 <400> 97
 gggatagccc tccatgggaa tgtcaaattc tgagcacgaa agcacctgct gaggtgcatt 60
 taattccgac aagtggtaac caaattgtat tgaaccgatt gaggctacta taattgcata 120
 acataacgcc ttgtgaattt tatgcttggt ttgccttggc atgctaacat tggatagcaa 180
 cgatgacta gatagaatcc tattcatact atgctaogta cctgtttaac tcttctcatt 240
 ttatccggtt tctttcttca ccgttccctc taagtctctt atttttttca taacagaatg 300
 atcaccocct tccactttgcc gccaatataa tattaacaca caagaaataa gtatgaggta 360
 gtttgctctc gaaaaaacca agtaatatgt caaaaaatga aaaaaagaaa acaaatacca 420
 aattatgcca ccgttacctt acgtttcoat gttaatcatc gtttactgcc gcctatgagc 480
 gtaagcta atgtataaaga aacaagctat aatattgtta aatatagttg atcaacagca 540
 ttgtaaatga tacaagagac gaggtgggaat gaaccttatg aaatgcgtat tatatataaa 600
 ctgtaataag agctaagttg aattgaaatc tacgatactt gatgttgaca ttatagcact 660
 agttcccagg aaaccctttc gaaaaacaca gcaaaaacaa gagtactgta accaatgtaa 720
 catctgtaca ccagggaccc acacattacc aaaatcaaaa ttatttttct aatgcctgtt 780
 atttttccta tttttctctc ggccgctgaa tagcccgag agacgcaa acatatttctc 840
 gcagtttttc gcttggttaa tgcgtatttt ccagatagg ttcaaacctt tcatctgtat 900
 cccgtatat taaagtggcg tttgctttct ccgttgattt ttttctctct tagtgatttt 960
 tttgcattaa atcccagaac aatcatccaa ctaatcaaga 1000
 <210> 98
 <211> 580
 <212> DNA
 <213> 酿酒酵母
 <220>
 <221> 启动子
 <222> (1) .. (580)
 <223> 负荷感应启动子 pOGG1
 <400> 98
 tgcaatcagg ttcttcccaa tcatccgaat cagagagcaa atctaggtct tottttgaaa 60
 gtatggaagg atgcttcaat tgtgcttgaa taaagtattt agacctgaaa ccacgactac 120
 tcatagaaaa cgatagtttg gcgtgcgata cgtttttgag taaaagaaaa ctgttgaaac 180
 tacagataaa tggccttctt ggtataatat gattcaatgt tgatcttctc cactttggca 240
 tgtaaaattg attgatacaa gtaagcta atcgctcaatgg ataataattac aaataaaaaat 300
 atcgaggcaa taccagtatt tcaatgtctg gccaaataaa tagtttagta cagcaatttg 360
 attaacgtta aagttactct cctctttgaa taatgtcgtg gttaaaaagt aaagtatttg 420
 tttatacatt aagtgaagaa aattaagacg cgctttctct taatgtaacg ccaaaaaattt 480
 tcaagagatg agctgaacag catcacaata ctactatttc cagcggaaga aggcatttga 540
 agcgtcttga ttcataattg cgatttttatt tatcaaccag 580
 <210> 99
 <211> 1000
 <212> DNA
 <213> 酿酒酵母
 <220>
 <221> 启动子
 <222> (1) .. (1000)
 <223> 负荷感应启动子 pRAD51
 <400> 99
 ccgcaataaa gggcttcccg gacttggaat taatacctgc taccaccggc cccacaaagt 60
 atggaccgaa tcttcttttca taaagtgaac ttgacactaa ttgggtaaac gtttcaggct 120
 caatggccct tcttctcttc aacttatata gggttcgtctt gtagcggaac atctcgttca 180
 gagtagtcac gtccgtagcc agaccggtaa tccctaagaa cacatgcca tagtggaaga 240
 ttttttcaaa cttgttcgac acacctagtg actggcttcc gagacgcaa tcacaggcaa 300
 tggccacaca gtctttacct gtcacgcaa cgacgatacc accgttaata gaacttggat 360
 ccgacatttt ttttttatgc tttattcact gttcaatatt ttcaccacaa ttcgcaagaa 420
 acgcaactta cttcgaaact acggttcaaa cttacttagc agcttccga tttaatggc 480
 ctttctacta tgccataaac tctttcttcc tctcttttca tcgcccctgc atttgcactt 540
 ttttgccacc ggcagtgcc tccggtcaca tgactacacc acgttaatag cgatctggct 600
 tatcatgtgc acagagtaaa ttaaaatgga cggtaaatgt tggaaatgca ccactaccgt 660
 tcttcaacca atctagttaa gctatcctgc aacagggtggc cttcttgagc attccctgag 720
 cattccaacc ggttgatca gtgttttctc accgtctcac catatcccac gactaggcca 780
 cacttcgtta cctatgcta cgcgtcattt ccgctatttc tgtcctgggt tgtttacagt 840
 acgcgtgggt ggaccataaa ggggaatagt ggggactgga gaaaaaattt tctcagttac 900
 ttcttctatc ttccgtagtt tccatatact agtagttgag tgtagcgaca aagagcagac 960
 gtagttattt gttaaaggcc tactaatttg ttatcgtcat 1000
 <210> 100
 <211> 1000
 <212> DNA
 <213> 酿酒酵母
 <220>
 <221> 启动子
 <222> (1) .. (1000)

```

<223> 负荷感应启动子 pRAD54
<400> 100
tgcgctacgg ttctgtgccgc tcgtgggaac cccacgcaaa acatattatt cgcttctctc 60
tgctgacaac tccgggtttac gttataccgt attagatca ctataagggt tccttcggga 120
ggagggggga ggggaagaat gtacatcgtc ataaggcctt tatgggtgtga agtgggtttt 180
gcgtggaaaa ttctgttttca atgatataga gcccacgcac atacgtacat actagtggcc 240
aaaagcgtgg ggtggggcga caaagctaca ctggttaaaat acaggattct atgaacaata 300
acaacaacca gctcacgttg ctgaacagcc gaggtcagcc gatgcaaccg aggtttccaa 360
agtagcattt ctgtgctagc tatgtctgtg ggtttacatt taatgggtgc tggttccagc 420
ttcatgtgct tgcatgtgat gtccgtgcaga tggtaagaag attctgaaag ccgcgcctagg 480
agaaaaatat tctgtctgaa gatctgtcct cttaagtaga aagcgtgaaa ttgttgctgt 540
cttgccattac tactcaacgc gtacgcaaat gcgtctactg cacctgcacg ataaagctta 600
tgtatcaaaa atttaacatc ttgaaaatac acaagtggtg caaagatgtg tcacgttctg 660
gacctgagtg gtgccatgta tgctatttaa catgcaaagg ggaagaccct tccgccttac 720
tgcaataata aaaagtattt tacgcgttac ccaatatagc aaagtttcgc gcaaaaaaaa 780
aaataaaaaa caattacaaa caaaaagaaa aaaaaggaaa taatagaaga tctaactgaa 840
gcgaaggcca aaactcttct cacttgacgt aatagccgat acaaaatcta gagcagcaac 900
ttttctcttt cttcactaaa gctgctacga aagtatagaa aaatcaaacy ctcagaactt 960
agctctattt caaggtacca tatatatatt cttataactg 1000

<210> 101
<211> 5776
<212> DNA
<213> 合成
<220>
<221> 尚未归类的特征
<222> (1)..(5776)
<223> 质粒 pEG34
<220>
<221> CDS
<222> (463)..(1758)
<223> hGH-甘氨酸接头-GFP
<400> 101
atcgatgcat gcgccaata cgcaaacgcg ctctcccgcg gcgttgcccg attcattaat 60
gcagctggca cgacaggttt ccgactggga aagcgggcag tgagcgcgca gcaattaatg 120
tgagtttagt cactcattag gcaccccagg ctttacactt tatgcttcg gctcgtatgt 180
tgtgtggaat tgtgagcgga taacaatttc acacaggaaa cagctatgac catgattacg 240
ccaagcgcgc aattaaccct cactaaaggg aacaaaagct gggtagccgg cccccctcg 300
aggctgacgg tatcgataag cttgatatac gctgaattca ttaaagagga gaaaggtacc 360
atgggtctgg gcaagaaaaa gaaactgctg ctgctgttac tgagcagcgc agttgcagca 420
agctttatga gcctgaccat tagcctgcct ggtgttcagg ca ttt ccg acc att 474
                                     Phe Pro Thr Ile
                                     1

ccg ctg agc cgt ctg ttt gat aat gca atg ctg cgt gca cat cgt ctg 522
Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu Arg Ala His Arg Leu
5 10 15 20
cat cag ctg gca ttt gat acc tat caa gaa ttt gaa gaa gcg tat atc 570
His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe Glu Glu Ala Tyr Ile
25 30 35
ccg aaa gag cag aaa tat agc ttc ctg cag aat ccg cag acc agc ctg 618
Pro Lys Glu Gln Lys Tyr Ser Phe Leu Gln Asn Pro Gln Thr Ser Leu
40 45 50
tgt ttt agc gaa agc att ccg aca ccg agc aat cgt gaa gaa acc cag 666
Cys Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn Arg Glu Glu Thr Gln
55 60 65
cag aaa agc aat ctg gaa ctg ctg cgt att agt ctg ctg ctg att cag 714
Gln Lys Ser Asn Leu Glu Leu Leu Arg Ile Ser Leu Leu Leu Ile Gln
70 75 80
agc tgg ctg gaa ccg gtg cag ttt ctg cgt agc gtt ttt gca aat agc 762
Ser Trp Leu Glu Pro Val Gln Phe Leu Arg Ser Val Phe Ala Asn Ser
85 90 95 100
ctg gtt tat ggt gca agc gat agc aat gtt tat gat ctg ctg aaa gat 810
Leu Val Tyr Gly Ala Ser Asp Ser Asn Val Tyr Asp Leu Leu Lys Asp
105 110 115
ctg gaa gaa ggt att cag acc ctg atg ggt cgt ctg gaa gat ggt tca 858
Leu Glu Glu Gly Ile Gln Thr Leu Met Gly Arg Leu Glu Asp Gly Ser
120 125 130
ccg cgt acc ggt cag atc ttt aaa cag acc tat agc aaa ttc gat acc 906
Pro Arg Thr Gly Gln Ile Phe Lys Gln Thr Tyr Ser Lys Phe Asp Thr
135 140 145
aac agc cat aat gat gat gcc ctg ctg aaa aac tat ggt ctg ctg tat 954
Asn Ser His Asn Asp Asp Ala Leu Leu Lys Asn Tyr Gly Leu Leu Tyr
150 155 160
tgt ttc cgc aaa gat atg gat aaa gtg gaa acc ttt ctg cgc att gtt 1002
Cys Phe Arg Lys Asp Met Asp Lys Val Glu Thr Phe Leu Arg Ile Val
165 170 175 180
cag tgt cgt agc gtt gaa ggt agc tgt ggt ttt ggt ggc atg cgt aaa 1050
Gln Cys Arg Ser Val Glu Gly Ser Cys Gly Phe Gly Gly Met Arg Lys

```


[0437]

ggc gaa gag ctg ttc act ggt ttc gtc act att ctg gtg gaa ctg gat	1098
Gly Glu Glu Leu Phe Thr Gly Phe Val Thr Ile Leu Val Glu Leu Asp	
200	
ggg gat gtc aac ggt cat aag ttt tcc gtg cgt ggc gag ggt gaa ggt	1146
Gly Asp Val Asn Gly His Lys Phe Ser Val Arg Gly Glu Gly Glu Gly	
215	
gac gca act aat ggt aaa ctg acg ctg aag ttc atc tgt act act ggt	1194
Asp Ala Thr Asn Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly	
230	
aaa ctg ccg gta cct tgg ccg act ctg gta acg ctg act tat ggt	1242
Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly	
245	
gtt cag tgc ttt gct cgt tat ccg gac cac atg aag cag cat gac ttc	1290
Val Gln Cys Phe Ala Arg Tyr Pro Asp His Met Lys Gln His Asp Phe	
265	
ttc aag tcc gcc atg ccg gaa ggc tat gtg cag gaa cgc acg att tcc	1338
Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Ser	
280	
ttt aag gat gac ggc acg tac aaa acg cgt gcg gaa gtg aaa ttt gaa	1386
Phe Lys Asp Asp Gly Thr Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu	
295	
ggc gat acc ctg gta aac cgc att gag ctg aaa ggc att gac ttt aaa	1434
Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys	
310	
gaa gac ggc aat atc ctg ggc cat aag ctg gaa tac aat ttt aac agc	1482
Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Phe Asn Ser	
325	
cac aat gtt tac atc acc gcc gat aaa caa aaa aat ggc att aaa gcg	1530
His Asn Val Tyr Ile Thr Ala Asp Lys Gln Lys Asn Gly Ile Lys Ala	
345	
aat ttt aaa att cgc cac aac gtg gag gat ggc agc gtg cag ctg gct	1578
Asn Phe Lys Ile Arg His Asn Val Glu Asp Gly Ser Val Gln Leu Ala	
360	
gat cac tac cag caa aac act cca atc ggt gat ggt cct gtt ctg ctg	1626
Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu	
375	
cca gac aat cac tat ctg agc acg caa agc gtt ctg tct aaa gat ccg	1674
Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Val Leu Ser Lys Asp Pro	
390	
aac gag aaa cgc gat cac atg gtt ctg ctg gag ttc gta acc gca gcg	1722
Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val Thr Ala Ala	
405	
ggc atc acg cat ggt atg gat gaa ctg tac aaa taa gtcgacctgc	1768
Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys	
425	
aggcatgcaa gcttggtgt tttggcggat gagagaagat tttcagcctg atacagatta	1828
aatcagaacg cagaagcggg ctgataaaac agaatttgcc tggcgccagt agcgcgggtg	1888
ccccacctga ccccatgccc aactcagaag tgaacgcgag tagcgccgat ggtagtgtg	1948
ggctctccca tgcgagagta gggaactgcc aggcatacaa taaaacgaaa ggctcagtcg	2008
aaagactggg cctttcggtt tatctgttgt ttgtcgtgta acgctctcct gagtaggaca	2068
aatccgcccgg gageggattt gaacgttgcg aagcaacggc ccggagggtg gcggggcagg	2128
cgcccgccat aaactgccag gcatcaaatt aagcagaagg ccactcctgac ggatggcctt	2188
tttgcggttc taaaaactct tttgtttatt ttcttaaata cattcaaata tgtatccgct	2248
catgagacaa taaccctgat aaatgcttca ataatttga aaaagggaaga gtatgagtat	2308
tcaacatttc cgtgtcgccc ttattccctt ttttgcggga ttttgccttc ctgtttttgc	2368
tcacccagaa acgctggtga aagtaaaaga tgctgaagat cagttgggtg cagcaaaacta	2428
ttaactggcg aactacttac tctagcttcc cggcaacaat taatagactg gatggaggcg	2488
gataaagtgt caggaccact tctgcgctcg gcccttccgg ctggctggtt tattgctgat	2548
aaatctggag ccggtgagcg tgggtctcgc ggtatcattg cagcactggg gccagatgg	2608
aagccctccc gtatcgtagt tatctacacg acggggagtc aggcaactat ggatgaacga	2668
aatagacaga tcgctgagat aggtgcctca ctgattaagc atttgtaact gtcagacca	2728
gtttactcat atatacttta gattgattta cgcgcctcgt agcggcgcat taagcgcgg	2788
gggtgtggtg gttacgcgca gcgtgaccgc tacacttgcc agcgccttag cgcccgctcc	2848
tttcgctttc ttcccttccct ttctcgccac gttcgcggcg tttccccgtc aagctctaaa	2908
tcgggggctc cctttagggt tcogatttag tgctttacgg cacctcgacc ccaaaaaact	2968
tgatttgggt gatggttcac gtagtgggccc atcgcctga tagacggtt ttgcgctttt	3028
gacgttggag tccacgttct ttaatatgtg actcttgctt caaacttgaa caacactcaa	3088
ccctattctg ggctattctt ttgatttata agggattttg ccgatttcgg cctattgggt	3148
aaaaaatgag ctgattttaac aaaaaatttaa cgcgaatttt aacaaaaat taacgtttac	3208
aattttaaaag gatctagggtg aagatccctt ttgataatct catgaccaa atcccttaac	3268
gtgagttttc gttccactga gcgtcagacc ccgtagaaaa gatcaaagg tcttcttgag	3328
atcctttttt tctgcgcgta atctgctgct tgcaaacaaa aaaaccaccg ctaccagcg	3388
tggtttgttt gcgggatcaa gagctacca ctctttttcc gaaggtaact ggcttcagca	3448
gagcgcagat accaaatact gtcccttctag tgtagccgta gttaggccac cacttcaaga	3508
actctgtagc accgcctaca tacctcgctc tgctaatct gttaccagtg gggcatttga	3568
gaagcacacg gtcacactgc ttccggtagt caataaaccg gtaaaccagc aatagacata	3628
agcggctatt taacgacct gcctgaacc gacgaccggg tcgaatttgc tttcgaattt	3688

```

ctgccattca tccgcttatt atcacttatt caggcgtagc accaggcgtt taagggcacc 3748
aataactgcc ttaaaaaaat tacgccccgc cctgccactc atcgccagta tgttgtaatt 3808
cattaagcat tctgcccaga tggaagccat cacagacggc atgatgaacc tgaatcgcca 3868
gcgccatcag caccttgctc ccttgcgat aatatttgcc catggtgaaa acgggggcga 3928
agaagtgtgc catattggcc acgtttaaat caaaactggg gaaactcacc cagggattgg 3988
ctgagacgaa aaacatattc tcaataaacc ctttagggaa ataggccagg ttttcaccgt 4048
aacacgccac atcttgcgaa tatatgtgta gaaactgcgc gaaatcgctg tggatttcac 4108
tccagagcga tgaaaaacgt tcagtttgct catggaaaac ggtgtaacaa ggggtgaacac 4168
tatcccatat caccagctca ccgtctttca ttgccatacg gaattccgga tgagcattca 4228
tcaggcgggc aagaatgtga ataaaggccg gataaaactt gtgcttattt ttctttacgg 4288
tctttaaaaa ggccgtaata tccagctgaa cggctcgggt ataggtacat tgagcaactg 4348
actgaaatgc ctcaaaatgt tctttacgat gccattggga tatatcaacg gtggtatata 4408
cagtgtattt tttctccatt ttagcttccct tagctcctga aaatctcgat aactcaaaaa 4468
atacgccggg tagtgatctt atttcattat ggtgaaagtt ggaacctctt acgtgccgat 4528
caacgtctca ttttcgcca aagttggccc agggcttccc ggtatcaaca gggacaccag 4588
gatttattta ttctgcgaag tgatcttccg tcacaggtat ttattcggcg caaagtgcgt 4648
cgggtgatgc tgcgaactta ctgatttagt gtatgatggt gtttttgagg tgctccagt 4708
gcttctgttt ctatcagctg tccctcctgt tcagctactg acgggggtggg gcgtaacggc 4768
aaaagcaccg ccggacatca gcgctagcgg agtgataact ggcttactat gttggcactg 4828
atgagggtgt cagtgaagtg cttcatgtgg caggagaaaa aaggctgcac cgggtgcgtc 4888
gcagaatcgt tgatacagga tatattccgc ttcctcgctc actgactcgc tgcgtccggt 4948
cgttcgactg ccgagcagcg aaatggctta cgaacggggc ggagatttcc tggagatgc 5008
caggaaagata cttaacaggg aagtgaaggg gccgcggcaa agccgttttt aaataggctc 5068
cgccccctg acaagcatca cgaaatctga cgtcctaaat agtggtggcg acatccgaca 5128
ggactataaa gataccaggc gttccccctt ggcggtccc tctgctgctc tctgttccct 5188
gcctttcggg ttaccggtgt cattccgctg ttatggccgc gtttgtctca ttccacgct 5248
gacactcagt tccgggtagg cagttcgctc caagctggac tgtatgcacg aacccccgt 5308
tcagtcggac cgctgcgcct tatccggtaa ctatcgtctt gactccaacc cggaaagaca 5368
tgcaaaagca ccactggcag cagccactgg taattgattt agaggagtta gtcttgaagt 5428
catgcgcggg ttaaggctaa actgaaagga caagtttgg tgactgcgtc cctccaagcc 5488
agttacctcg gttcaaagag ttggtagctc agagaacctt cgaaaaaccg ccctgcaagg 5548
cggttttttc gttttcagag caagagatta cgcgagacc aaaacgatct caagaagatc 5608
atcttattaa tcagataaaa tatttgctca tgagcccgaa gtggcgagcc cgtcttccc 5668
catcgctgat gtcggcgata taggcgccag caaccgcacc tgtggcgccg gtgatgccg 5728
ccacgatcgc tccggcgtag aggatctgct catgtttgac agcttatc 5776

```

[0438]

```

<210> 102
<211> 431
<212> PRT
<213> 合成
<400> 102
Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu Arg
1 5 10 15
Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe Glu
20 25 30
Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser Phe Leu Gln Asn Pro
35 40 45
Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn Arg
50 55 60
Glu Glu Thr Gln Gln Lys Ser Asn Leu Glu Leu Arg Ile Ser Leu
65 70 75 80
Leu Leu Ile Gln Ser Trp Leu Glu Pro Val Gln Phe Leu Arg Ser Val
85 90 95
Phe Ala Asn Ser Leu Val Tyr Gly Ala Ser Asp Ser Asn Val Tyr Asp
100 105 110
Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln Thr Leu Met Gly Arg Leu
115 120 125
Glu Asp Gly Ser Pro Arg Thr Gly Gln Ile Phe Lys Gln Thr Tyr Ser
130 135 140
Lys Phe Asp Thr Asn Ser His Asn Asp Asp Ala Leu Leu Lys Asn Tyr
145 150 155 160
Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp Lys Val Glu Thr Phe
165 170 175
Leu Arg Ile Val Gln Cys Arg Ser Val Glu Gly Ser Cys Gly Phe Gly
180 185 190
Gly Met Arg Lys Gly Glu Glu Leu Phe Thr Gly Phe Val Thr Ile Leu
195 200 205
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Arg Gly
210 215 220
Glu Gly Glu Gly Asp Ala Thr Asn Gly Lys Leu Thr Leu Lys Phe Ile
225 230 235 240
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
245 250 255
Leu Thr Tyr Gly Val Gln Cys Phe Ala Arg Tyr Pro Asp His Met Lys
260 265 270
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
275 280 285
Arg Thr Ile Ser Phe Lys Asp Asp Gly Thr Tyr Lys Thr Arg Ala Glu
290 295 300

```


Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys	Gly		
305					310					315					320		
Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Leu	Glu	Tyr		
				325					330					335			
Asn	Phe	Asn	Ser	His	Asn	Val	Tyr	Ile	Thr	Ala	Asp	Lys	Gln	Lys	Asn		
			340					345					350				
Gly	Ile	Lys	Ala	Asn	Phe	Lys	Ile	Arg	His	Asn	Val	Glu	Asp	Gly	Ser		
		355					360					365					
Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Gly	Asp	Gly		
	370					375					380						
Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Val	Leu		
385					390					395					400		
Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	Leu	Glu	Phe		
			405						410					415			
Val	Thr	Ala	Ala	Gly	Ile	Thr	His	Gly	Met	Asp	Glu	Leu	Tyr	Lys			
			420					425					430				
<210>	103																
<211>	660																
<212>	DNA																
<213>	合成																
<220>																	
<221>	CDS																
<222>	(1)..(660)																
<223>	氯霉素_乙酰转移酶																
<400>	103																
atg	gag	aaa	aaa	atc	act	gga	tat	acc	acc	gtt	gat	ata	tcc	caa	tgg		48
Met	Glu	Lys	Lys	Ile	Thr	Gly	Tyr	Thr	Thr	Val	Asp	Ile	Ser	Gln	Trp		
1				5				10					15				
cat	cgt	aaa	gaa	cat	ttt	gag	gca	ttt	cag	tca	gtt	gct	caa	tgt	acc		96
His	Arg	Lys	Glu	His	Phe	Glu	Ala	Phe	Gln	Ser	Val	Ala	Gln	Cys	Thr		
			20				25					30					
tat	aac	cag	acc	gtt	cag	ctg	gat	att	acg	gcc	ttt	tta	aag	acc	gta		144
Tyr	Asn	Gln	Thr	Val	Gln	Leu	Asp	Ile	Thr	Ala	Phe	Leu	Lys	Thr	Val		
		35				40					45						
aag	aaa	aat	aag	cac	aag	ttt	tat	ccg	gcc	ttt	att	cac	att	ctt	gcc		192
Lys	Lys	Asn	Lys	His	Lys	Phe	Tyr	Pro	Ala	Phe	Ile	His	Ile	Leu	Ala		
	50					55			60								
cgc	ctg	atg	aat	gct	cat	ccg	gaa	ttc	cgt	atg	gca	atg	aaa	gac	ggc		240
Arg	Leu	Met	Asn	Ala	His	Pro	Glu	Phe	Arg	Met	Ala	Met	Lys	Asp	Gly		
65				70				75							80		
gag	ctg	gtg	ata	tggt	gat	agt	gtt	cac	cct	tgt	tac	acc	gtt	ttc	cat		288
Glu	Leu	Val	Ile	Trp	Asp	Ser	Val	His	Pro	Cys	Tyr	Thr	Val	Phe	His		
				85				90					95				
gag	caa	act	gaa	acg	ttt	tca	tcg	ctc	ttg	agt	gaa	tac	cac	gac	gat		336
Glu	Gln	Thr	Glu	Thr	Phe	Ser	Ser	Leu	Trp	Ser	Glu	Tyr	His	Asp	Asp		
			100					105					110				
ttc	cgg	cag	ttt	cta	cac	ata	tat	tcg	caa	gat	gtg	gcg	tgt	tac	ggc		384
Phe	Arg	Gln	Phe	Leu	His	Ile	Tyr	Ser	Gln	Asp	Val	Ala	Cys	Tyr	Gly		
		115					120					125					
gaa	aac	ctg	gcc	tat	ttc	cct	aaa	ggg	ttt	att	gag	aat	atg	ttt	ttc		432
Glu	Asn	Leu	Ala	Tyr	Phe	Pro	Lys	Gly	Phe	Ile	Glu	Asn	Met	Phe	Phe		
	130					135			140								
gtc	tca	gcc	aat	ccc	ttg	gtg	agt	ttc	acc	agt	ttt	gat	tta	aac	gtg		480
Val	Ser	Ala	Asn	Pro	Trp	Val	Ser	Phe	Thr	Ser	Phe	Asp	Leu	Asn	Val		
145					150				155						160		
gcc	aat	atg	gac	aac	ttc	gcc	ccc	gtt	ttc	acc	atg	ggc	aaa	tat			528
Ala	Asn	Met	Asp	Asn	Phe	Phe	Ala	Pro	Val	Phe	Thr	Met	Gly	Lys	Tyr		
			165					170					175				
tat	acg	caa	ggc	gac	aag	gtg	ctg	atg	ccg	ctg	gcg	att	cag	gtt	cat		576
Tyr	Thr	Gln	Gly	Asp	Lys	Val	Leu	Met	Pro	Leu	Ala	Ile	Gln	Val	His		
			180					185					190				
cat	gcc	gtc	tgt	gat	ggc	ttc	cat	gtc	ggc	aga	atg	ctt	aat	gaa	tta		624
His	Ala	Val	Cys	Asp	Gly	Phe	His	Val	Gly	Arg	Met	Leu	Asn	Glu	Leu		
		195				200						205					
caa	cag	tac	tgc	gat	gag	tggt	cag	ggc	ggg	gcg	taa						660
Gln	Gln	Tyr	Cys	Asp	Glu	Trp	Gln	Gly	Gly	Ala							
		210				215											
<210>	104																
<211>	219																
<212>	PRT																
<213>	合成																
<400>	104																
Met	Glu	Lys	Lys	Ile	Thr	Gly	Tyr	Thr	Thr	Val	Asp	Ile	Ser	Gln	Trp		
1				5					10					15			
His	Arg	Lys	Glu	His	Phe	Glu	Ala	Phe	Gln	Ser	Val	Ala	Gln	Cys	Thr		
			20					25					30				
Tyr	Asn	Gln	Thr	Val	Gln	Leu	Asp	Ile	Thr	Ala	Phe	Leu	Lys	Thr	Val		
		35				40						45					

Lys	Lys	Asn	Lys	His	Lys	Phe	Tyr	Pro	Ala	Phe	Ile	His	Ile	Leu	Ala
50						55					60				
Arg	Leu	Met	Asn	Ala	His	Pro	Glu	Phe	Arg	Met	Ala	Met	Lys	Asp	Gly
65					70					75				80	
Glu	Leu	Val	Ile	Trp	Asp	Ser	Val	His	Pro	Cys	Tyr	Thr	Val	Phe	His
			85						90					95	
Glu	Gln	Thr	Glu	Thr	Phe	Ser	Ser	Leu	Trp	Ser	Glu	Tyr	His	Asp	Asp
			100					105					110		
Phe	Arg	Gln	Phe	Leu	His	Ile	Tyr	Ser	Gln	Asp	Val	Ala	Cys	Tyr	Gly
		115					120					125			
Glu	Asn	Leu	Ala	Tyr	Phe	Pro	Lys	Gly	Phe	Ile	Glu	Asn	Met	Phe	Phe
	130					135					140				
Val	Ser	Ala	Asn	Pro	Trp	Val	Ser	Phe	Thr	Ser	Phe	Asp	Leu	Asn	Val
145					150					155				160	
Ala	Asn	Met	Asp	Asn	Phe	Phe	Ala	Pro	Val	Phe	Thr	Met	Gly	Lys	Tyr
			165						170					175	
Tyr	Thr	Gln	Gly	Asp	Lys	Val	Leu	Met	Pro	Leu	Ala	Ile	Gln	Val	His
		180						185					190		
His	Ala	Val	Cys	Asp	Gly	Phe	His	Val	Gly	Arg	Met	Leu	Asn	Glu	Leu
		195					200					205			
Gln	Gln	Tyr	Cys	Asp	Glu	Trp	Gln	Gly	Gly	Ala					
	210					215									
<210>	105														
<211>	4347														
<212>	DNA														
<213>	合成														
<220>															
<221>	尚未归类的特征														
<222>	(1) .. (4347)														
<223>	质粒 pEG0														
<400>	105														
atcgatgcat	gcgccaata	cgcaaaccgc	ctctccccgc	gcgttgcccg	attcattaat										60
gcagctggca	cgacaggttt	cccgaactga	aagcgggcag	tgagcgcaac	gcaattaatg										120
tgagtttagct	cactcattag	gcaccccagg	ctttacactt	tatgcttcog	gctcgtatgt										180
tgtgtggaat	tgtgagcgga	taacaatttc	acacaggaaa	cagctatgac	catgattacg										240
ccaagcgcgc	aattaaccct	cactaaaggg	aacaaaagct	gggtaccggg	ccccccctcg										300
aggctcgacgg	tatcgataag	cttgatatcg	aattcctgca	ggcatgcaag	cttggtctgtt										360
ttggcggatg	agagaagatt	ttcagcctga	tacagattaa	atcagaacgc	agaagcgggtc										420
tgataaaaaca	gaattttgcct	ggcggcagta	gcgcgggtgg	cccacctgac	cccatgccga										480
actcagaagt	gaaacgcctg	agcgccgatg	gtagtgtggg	gtctcccat	gcgagagtag										540
ggaactgccca	ggcatcaaat	aaaacgaaaag	gctcagtcga	aagactgggc	ctttcgtttt										600
atctgttgtt	tgctcggtgaa	cgctctcctg	agtaggacaa	atccgcgggg	agcggatttg										660
aacgttgcga	agcaacggcc	cgagggttgg	cgggcaggac	gcccggccata	aactgccagg										720
catcaaaatta	agcagaaggc	catcctgacg	gatggccctt	ttgcgtttct	acaaactott										780
ttgtttatatt	ttctaaatac	attcaaatat	gtatccgctc	atgagacaat	aaccctgata										840
aatgcttcaa	taatatggaa	aaaggaagag	tatgagtatt	caacatttcc	gtgtcgcctt										900
tattcccttt	tttgcgccat	tttgcccttc	tgtttttgct	caccagaaa	cgctggtgaa										960
agtaaaagat	gctgaagatc	agttgggtgc	agcaaaactat	taactggcga	actacttact										1020
ctagcttccc	ggcaacaatt	aatagactgg	atggaggcgg	ataaagttgc	aggaccactt										1080
ctgcgttcgg	cccttcgggc	tggtctggtt	attgctgata	aactctggagc	cggtgacgct										1140
gggtctcgcg	gtatcattgc	agcactgggg	ccagatggta	agccctcccg	tatcgtagtt										1200
atctacacga	cggggagtc	ggcaactatg	gatgaacgaa	atagacagat	cgctgagata										1260
ggtgcctcac	tgattaaagc	ttggtaactg	tcagaccaag	tttactcata	attactttag										1320
attgattttac	gcgccttgta	gcggcgcttg	aagcgcggcg	ggtgtggtgg	ttacgcgcag										1380
cgtgacccgt	acacttgcca	gcgccttagc	gcccgcctct	ttcgctttct	tcccttccct										1440
tctcgccacg	ttcgccggct	ttccccgtca	agctctaaat	cgggggctcc	ctttagggtt										1500
ccgattttagt	gcttttacggc	acctcgaccc	caaaaaactt	gatttgggtg	atggttcacg										1560
tagtggggcca	tcgcccctgat	agacgggtttt	tcgccccttg	acgttggagt	ccacgttctt										1620
taatagtgg	ctcttgtttcc	aaacttgaac	aacactcaac	cctatctcgg	gctattcttt										1680
tgattttataa	gggatttttgc	cgatttccggc	ctatttggtta	aaaaatgagc	tgattttaaca										1740
aaaattttaac	gcgaattttta	acaaaatatt	aacgtttaca	attttaaagg	atctaggtga										1800
agatccctttt	tgataatctc	atgacaaaaa	tcctttaacg	tgagttttcg	ttccactgag										1860
cgtcagaccc	cgtagaaaag	atcaaaggat	cttcttgaga	tccttttttt	ctgcgcgtaa										1920
tctgctgctt	gcaaacaaaa	aaaccaccgc	taccagcggt	ggtttgtttg	ccggatcaag										1980
agctaccaac	tctttttccg	aaggtaactg	gcttcagcag	agcgcagata	ccaaatactg										2040
tcctttctagt	gtagccgtag	ttaggccacc	acttcaagaa	ctctgtagca	ccgcctacat										2100
acctcgtctt	gctaattcctg	ttaccagtgg	ggcattttgag	aagcacacgg	tcacactgct										2160
tcgggtagtc	aataaacggg	taaaccagca	atagacataa	gcggctattt	aacgaccttg										2220
ccctgaaccg	acgaccgggt	cgaatttgcct	ttcgaatttc	tgccattcat	ccgcttatta										2280
tcactttatc	aggcgtagca	ccaggcgttt	aagggcacca	ataactgcct	taaaaaaatt										2340
acgccccggc	ctgccactca	tcgcagtaact	gttgtaattc	attaagcatt	ctgccgacat										2400
ggaagccatc	acagacggca	tgatgaacct	gaatcgccag	cggcatcagc	acctgtgcgc										2460
cttgcgtata	atatttgcgc	atggtgaaaa	cggggggcgaa	gaagttgtoc	atattggcca										2520
cgttttaaatc	aaaactgggtg	aaactcaccc	agggtattggc	tgagacgaaa	aacataattct										2580
caataaaccc	tttagggaaa	taggccaggt	tttcaccgta	acacgccaca	tcttgcgaa										2640
atatgtgttag	aaactgccgg	aaatcgctgt	ggtattcact	ccagagcgat	gaaaacgttt										2700
cagttttgctc	atggaaaaacg	gtgtaacaag	ggtgaacact	atcccatatc	accagctcac										2760
cgtcttttcat	tgccatacgg	aattccggat	gagcattcat	caggcgggca	agaatgtgaa										2820

[0440]

taaaggccgg	ataaaacttg	tgtttatatt	tctttacggg	ctttaaaaag	gccgtaatat	2880
ccagctgaac	gggtctggta	taggtacatt	gagcaactga	ctgaaatgcc	tcaaaatggt	2940
ctttacgatg	ccattgggat	atatcaacgg	tggatatcc	agtgaatttt	ttctccattt	3000
tagcttcctt	agctcctgaa	aatctcgata	actcaaaaaa	tacgcccggg	agtgatctta	3060
tttcattatg	gtgaaagttg	gaacctctta	cgtgccgatc	aacgtctcat	tttcgcaaaa	3120
agttggccca	gggcttcccg	gtatcaacag	ggacaccagg	atttattttat	tctgcgaagt	3180
gatctttccg	cacaggtatt	tattcggcgc	aaagtgcgtc	gggtgatgct	gccaaacttac	3240
tgatttagtg	tatgatgggt	tttttgaggt	gctccagtgg	cttctgtttc	tatcagctgt	3300
ccctcctgtt	cagctactga	cgggggtggg	cgtaacggca	aaagcaccgc	cggacatcag	3360
cgctagcgga	gtgtatactg	gcttactatg	ttggcactga	tgagggtgtc	agtgaagtgc	3420
ttcatgtggc	aggagaaaaa	aggctgcacc	ggtgcgtcag	cagaatatgt	gatacaggat	3480
atattccgct	tctcgcgtca	ctgactcgct	acgctcggtc	gttcgactgc	ggcgagcgga	3540
aatggccttac	gaacggggcg	gagattttcct	ggaagatgcc	aggaagatac	ttaacaggga	3600
agtgaagagg	ccgcggaaca	gccgtttttc	cataggctcc	gccccctga	caagcatcac	3660
gaaatctgac	gctcaaatca	gtggtggcga	aaccgcacag	gactataaag	ataccaggcg	3720
tttccccctg	gcggtccctt	cgtgcgtctt	cctgttcctg	cctttcgggt	taccgggtgc	3780
attccgctgt	tatggcgcgt	tttgtctcat	tccacgcctg	acactcagtt	ccgggtaggc	3840
agttcgcctc	aagctggact	gtatgcacga	accccccggt	cagtcgcgacc	gctgcgcctt	3900
atccggtaac	tatcgtcttg	agtccaaccc	ggaagacat	gcaaaagcac	cactggcagc	3960
agccactggg	aattgattta	gaggagttag	tcttgaagtc	atgcgcgggt	taaggctaaa	4020
ctgaaaggac	aagtttttgg	gactgcgctc	ctccaagcca	gttacctcgg	ttcaaagagt	4080
tggtagctca	gagaaccttc	gaaaaacgcg	cctgcaaggc	ggttttttcg	ttttcagagc	4140
aagagattac	gcgcagacca	aaacgatctc	aagaagatca	tcttattaat	cagataaaat	4200
atgtgtctcat	gagcccgaa	tgccgagccc	gatcttcccc	atcgggtgat	tccggcgatat	4260
agggcgccagc	aaccgcacct	gtggcgccgg	tgatgccggc	cacgatgcgt	ccggcgtaga	4320
ggatctgctc	atgtttgaca	gcttattc				4347
<210>	106					
<211>	6076					
<212>	DNA					
<213>	合成					
<220>						
<221>	尚未归类的特征					
<222>	(1) .. (6076)					
<223>	质粒 pENDU5CAM					
<220>						
<221>	CDS					
<222>	(108) .. (911)					
<223>	ompA-溶葡萄球菌酶					
<400>	106					
gatctcgatc	ccgcgaaatt	aatacgactc	actatagggg	aattgtgagc	ggataacaat	60
tcccctctag	aaataatttt	gtttaacttt	aagaaggaga	tatacat	atg aaa aag	116
					Met Lys Lys	
					1	
aca gct atc gcg att gca gtg gca ctg gct ggt ttc gct acc gta gcg						164
Thr Ala Ile Ala Ile Ala Val Ala Leu Ala Gly					Phe Ala Thr Val Ala	
5					15	
cag gct gca gca acc cat gaa cat agc gca cag tgg ctg aat aat tac						212
Gln Ala Ala Ala Thr His Glu His Ser Ala Gln					Trp Leu Asn Asn Tyr	
20					30	
aaa aaa ggt tat ggc tat ggt ccg tat ccg ctg ggt att aat ggt ggt						260
Lys Lys Gly Tyr Gly Tyr Gly Pro Tyr					Leu Gly Ile Asn Gly Gly	
40					45	
atg cat tat ggt gtg gat ttc ttt atg aat att ggc aca ccg gtt aaa						308
Met His Tyr Gly Val Asp Phe Phe Met Asn Ile					Gly Thr Pro Val Lys	
55					60	
gca att agc agc ggt aaa att gtt gaa gca ggt tgg agc aat tat ggt						356
Ala Ile Ser Ser Gly Lys Ile Val Glu Ala Gly					Trp Ser Asn Tyr Gly	
70					80	
ggg ggt aat cag att ggc ctg att gaa aat gat ggt gtt cat cgt cag						404
Gly Gly Asn Gln Ile Gly Leu Ile Glu Asn Asp					Gly Val His Arg Gln	
85					95	
tgg tat atg cat ctg agc aaa tac aat gtg aaa gtg ggc gat tat gtt						452
Trp Tyr Met His Leu Ser Lys Tyr Asn Val Lys					Val Gly Asp Tyr Val	
100					110	
aaa gcc ggt cag att att ggt tgg agc ggt agc acc ggt tat agt acc						500
Lys Ala Gly Gln Ile Ile Gly Trp Ser Gly Ser					Thr Gly Tyr Ser Thr	
120					125	
gca ccg cat ctg cat ttt cag cgt atg gtt aat agc ttt agc aat agc						548
Ala Pro His Leu His Phe Gln Arg Met Val Asn Ser					Phe Ser Asn Ser	
135					140	
acc gca cag gat ccg atg ccg ttt ctg aaa agc gca ggt tat ggt aaa						596
Thr Ala Gln Asp Pro Met Pro Phe Leu Lys Ser					Ala Gly Tyr Gly Lys	
150					160	
gca ggc ggt aca gtt acc ccg aca ccg aat acc ggt tgg aaa acc aac						644
Ala Gly Gly Thr Val Thr Pro Thr Pro Asn Thr					Gly Trp Lys Thr Asn	
165					175	
aaa tat ggc acc ctg tat aaa agc gaa agc gca agc ttt acc ccg aat						692

Lys	Tyr	Gly	Thr	Leu	Tyr	Lys	Ser	Glu	Ser	Ala	Ser	Phe	Thr	Pro	Asn	
180					185				190						195	
aca	gat	att	atc	acc	cgt	acc	acc	ggg	ccg	ttt	cgt	agc	atg	ccg	cag	740
Thr	Asp	Ile	Ile		Thr	Arg	Thr	Gly	Pro	Phe	Arg	Ser	Met	Pro	Gln	
				200					205					210		
agc	ggg	gtt	ctg	aaa	gca	ggg	cag	acc	att	cat	tat	gat	gaa	gtg	atg	788
Ser	Gly	Val	Leu	Lys	Ala	Gly	Gln	Thr	Ile	His	Tyr	Asp	Glu	Val	Met	
				215				220						225		
aaa	cag	gat	ggg	cat	gtt	tgg	gtt	ggg	tat	acc	ggg	aat	agc	ggg	cag	836
Lys	Gln	Asp	Gly	His	Val	Trp	Val	Gly	Tyr	Thr	Gly	Asn	Ser	Gly	Gln	
		230						235					240			
cgt	att	tat	ctg	ccg	gtt	cgt	acc	tgg	aat	aaa	agc	acc	aat	aca	ctg	884
Arg	Ile	Tyr	Leu	Pro	Val	Arg	Thr	Trp	Asn	Lys	Ser	Thr	Asn	Thr	Leu	
		245				250						255				
ggg	gtt	ctg	tgg	ggc	acc	atc	aaa	taa	tgagatccgg	ctgctaacaa						931
Gly	Val	Leu	Trp	Gly	Thr	Ile	Lys									
260				265												
agccccgaaag	gaagctgagt	tggtctgctgc	caccgctgag	caataactag	cataaccct											991
tggggctct	aaacgggtct	tgaggggttt	tttgctgaaa	ggaggaacta	tatccggatt											1051
ggcgaaatggg	acgcgccctg	tagcggcgca	ttaagcgcgg	cgggtgtggt	gggtacgcgc											1111
cgcgtgaccg	ctacacttgc	cagcgcccta	gcgccgctc	ctttcgcttt	cttcccttcc											1171
tttctcgcca	cggttcgccg	ctttccccgt	caagctctaa	atcgggggct	cccttttaggg											1231
ttccgattta	gtgctttacg	gcacctcgac	cccaaaaaac	ttgattaggg	tgatgggtta											1291
ctgagtgggc	catcgccctg	atagacgggt	tttcgcccct	tgacgttgga	gtccacgttc											1351
tttaatagtg	gactcttggt	ccaaactgga	acaacactca	accctatctc	gggtctattct											1411
tttgatttat	aagggtatttt	gccgatttgc	gcctatttgg	taaaaaatga	gctgatttaa											1471
caaaaatttta	acgcgaatttt	taacaaaata	ttacggttta	caatttcaag	tggcactttt											1531
cggggaaatg	tgccgcgaac	ccctatttgt	ttatttttct	aaatacatte	aaatatgtat											1591
ccgctcaata	actgccttaa	aaaaattacg	ccccgcctg	ccactcatcg	cagtactgtt											1651
gtaatttcatt	aagcattctg	ccgacatgga	agccatca	gacggcatga	tgaacctgaa											1711
tcgccagcgg	catcagcacc	ttgtgcctt	gcgtataata	tttgcccatg	gtgaaaacgg											1771
gggcgaagaa	gttggtccata	ttggccacgt	ttaaatcaaa	actggtgaaa	ctcaccacgg											1831
gattggctga	gacgaaaaac	atatctctca	taaacctttt	agggaaatag	gccagggttt											1891
caccgttaaca	gcgcacatct	tgcgaaatata	tggttagaaa	ctgccggaaa	tcgtcgtggg											1951
attcactcca	gagcgatgaa	aacgttttcag	tttgctcatg	gaaaaacggt	taacaagggt											2011
gaacactatc	ccatatcacc	agctcaccgt	ctttcattgc	catacgggat	tcgggatgag											2071
caactcatcg	cgccgcaaga	atgtgaataa	agcccgatga	aaacttctgc	ttatttttct											2131
ttacgggtctt	taaaaaaggcc	gtaatatcca	gctgaacggg	ctgggttatag	gtacattgag											2191
caactgactg	aaatgcctca	aaatgttctt	tacgatgcca	ttgggatata	tcaacgggtg											2251
tatatccagg	gatttttttct	tccatttttag	cttccttagc	tcctgaaaat	ctcgataact											2311
caaaaaatac	gcccggtagt	gatcttattt	cattatgggt	aaagtgggaa	cctcttacgt											2371
gccgatcaac	gtctcatttt	cgccaaaagt	tgcccgagg	cttcccggta	tcaacaggga											2431
caccaggatt	tattttattct	gcgaagtgat	cttcgctcac	aggattttat	toggcgcaaa											2491
gtgcgtcatt	gttcatgacc	aaaatccctt	aacgtgagtt	ttcgttccac	tgagcgtcag											2551
accccgtaga	aaagattggt	catgacccaa	atcccttaac	gtgatcaaa	gatcttcttg											2611
agatcctttt	tttctgcgcg	taactctgct	cttgcaaaac	aaaaaacacc	cgtaccacgc											2671
gggtggttgt	ttgcgggac	aagagctacc	aactcttttt	ccgaaggtaa	ctggttctcag											2731
cagagcgcag	ataccaaata	ctgtccttct	agtgtagccg	tagttaggcc	accacttcaa											2791
gaactctgta	gcaccgccta	catacctcgc	tctgctaate	ctgttaccag	tggctgctgc											2851
cagtgccgct	aagtcgtgtc	ttaccgggtt	ggactcaaga	cgatagttag	tcgaataagg											2911
gcagcggctg	ggctgaacgg	gggttctgtg	cacacagccc	agcttgagac	gaacgacct											2971
caccgaactg	agatacctac	agcgtgagct	atgagaaaag	gccacgcttc	ccgaaggag											3031
aaaggcggac	aggtatccgg	taagcggcag	ggtcggaaca	ggagagcgca	cgagggagc											3091
tccaggggga	aacgcctggt	atcttttatag	tcctgtcggg	tttcgccacc	tctgacttga											3151
gcgtcgattt	ttgtgatgct	cgtcaggggg	gcggagccta	tggaaaaacg	ccagcaacgc											3211
ggccttttta	cggttctctg	ccttttctgc	gccttttctg	cacatgttct	ttcctgcgtt											3271
atccctgat	tctgtggata	accgtattac	gccttttgag	tgagctgata	ccgctcgccg											3331
cagccgaacg	accgagcgca	gcgagtcagt	gagcgaggaa	gcggaagagc	gcctgatgcg											3391
gtattttctc	cttacgcate	tgtgcgggat	ttcacaccgc	atataatggt	cactctcagt											3451
acaatctgct	ctgatccgc	atagttaaag	cagtatacac	tcgcctatcg	ctacgtgact											3511
gggtcatggc	tgccgccccg	caccgcacca	caccgcctga	cgcgccctga	cgggcttctc											3571
tgctcccggc	atccgcttac	agacaagctg	tgaccgtctc	cgggagctgc	atgtgtcaga											3631
ggttttcacc	gtcatcaccc	aaacgcgcga	ggcagctgcg	gtaaaagctc	tcagcgtggg											3691
cgtgaagcga	ttcacagatg	tctgcctggt	caccgcgctc	cagctcgttg	agtttctcca											3751
gaagcgttaa	tgtctggctt	ctgataaaag	gggcatgtt	aaggcggtt	ttttcctggt											3811
tggtcactga	tgctccgtg	taagggggat	ttctgttcat	gggggtaatg	ataccgatga											3871
aacgagagag	gatgctcacg	atacgggtta	ctgatgatga	acatgcccgg	ttactggaac											3931
gttctgaggg	taaacaaactg	gcgggtatga	tgccgaggga	ccagagaaaa	atcacatcag											3991
gtcaatgcca	gcgcttcgtt	aatacagatg	taggtgttcc	acagggtagc	cagcagcatg											4051
ctgcgatgca	gatccgggaa	ataatgggtg	agggcgctga	cttccgcgtt	tcagactttt											4111
acgaaacacg	gaaaccgaag	accatttcag	ttgttgcctc	ggctgcagac	gtttttgcagc											4171
agcagtcgct	tcacgttcgc	tcgcgtatcg	gtgattcatt	ctgctaacca	gtaaggcaac											4231
cccgcagacc	tagccgggtc	ctcaacgaca	ggagcacgat	catgcgcacc	cgtggggccg											4291
ccatgcgggc	gataatggcc	tgcttctcgc	cgaaacgttt	gggtggcggga	ccagtgcagc											4351
aggtcttgagc	gagggcgctg	aagattccga	ataccgcaag	cgacaggccg	atcatcgtcg											4411
cgtccagcgc	aaagcgggtc	tcgccgaaaa	tgaccagag	cgtgcggcgc	acctgtccta											4471
cgagttgcat	gataaagaag	acagtcataa	gtgcggcgac	gatagtcagc	ccccgcgcc											4531
accggaaggga	gctgactggg	ttgaaggctc	tcaaggcgat	cggctcagat	ccccgtgcct											4591

[0442]

```

aatgagttag ctaacttaca ttaattgcgt tgcgctcact gcccgctttc cagtcgggaa 4651
acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc ggggagagggc ggtttgcgta 4711
ttggggcgcca ggggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 4771
accgcctggc cctgagagag ttgcagcaag cgggtccacgc tggtttgccc cagcaggcga 4831
aaatcctgtt tgatgggtgtg taacggcggg atataacatg agctgtcttc ggtatcgtcg 4891
tatcccaacta ccgagatata cgcaccaacg cgcagcccg actcggtaat ggcgcgcatt 4951
gcgcccacgc ccactctgat gttggcaacc agcatcgag tgggaacgat gccctcattc 5011
agcatttgta tggtttggtg aaaaccggac atggcactcc agtcgccttc ccgttcctgt 5071
atcggtgtaa tttgattgct agtgagatat ttatgccagc cagccagacg cagacgcgcc 5131
gagacagaac ttaatggggc cgctaacacg gcgatttgct ggtgacccaa tgcgaccaga 5191
tgctccacgc ccagtcgcgt accgtcttca tgggagaaaa taatactgtt gatgggtgtc 5251
tggtcagaga catcaagaaa taacgcggga acattagtgc aggcagcttc cacagcaatg 5311
gcattcctgt catccagcgg atagttaatg atcagcccac tgacgcgttg cgcgagaaga 5371
ttgtgcaccg ccgctttaca ggcttcgacg ccgcttcgtt ctaccatcga caccaccacg 5431
ctggcaccca gttgatcggc gcgagattta atcgccgcga caatttgcca cggcgcgtgc 5491
agggccagac tggaggtggc aacgccaatc agcaacgact gtttgcccg cagttgtgtg 5551
gccacgggt tgggaatgta attcagctcc gccatcgccg cttccacttt ttcccgcgtt 5611
ttcgcagaaa cgtgggtggc ctggttcacc acgcgggaaa cgggtctgata atagacaccg 5671
gcatactctg cgacatcgta taacgttact ggtttcacat tcaccaccct gaattgactc 5731
tcttcggggc gctatcatgc cataccgcga aaggttttgc gccattcgat ggtgtccggg 5791
atctcgacgc tctcccttat gcgactcctg cattaggaag cagcccagta gtaggttgag 5851
gccgttgagc accgcgcgcg caaggaatgg tgcattgcaag gagatggcgc ccaacagctc 5911
cccgccacg gggcctgcc ccatcccac gccgaaacaa gcgctcatga gcccgagtg 5971
gcgagcccg tcttcccat cgtgatgtc ggcatatag gcgcagcaa cgcacctgt 6031
ggcgccggtg atgccggcca cgatgcgtcc ggcgtagagg atcga 6076

```

<210> 107
 <211> 267
 <212> PRT
 <213> 合成
 <400> 107

[0443]

```

Met Lys Lys Thr Ala Ile Ala Ile Val Ala Leu Ala Gly Phe Ala
1      5      10
Thr Val Ala Gln Ala Ala Ala Thr His Glu His Ser Ala Gln Trp Leu
20     25     30
Asn Asn Tyr Lys Lys Gly Tyr Gly Tyr Gly Pro Tyr Pro Leu Gly Ile
35     40     45
Asn Gly Gly Met His Tyr Gly Val Asp Phe Phe Met Asn Ile Gly Thr
50     55     60
Pro Val Lys Ala Ile Ser Ser Gly Lys Ile Val Glu Ala Gly Trp Ser
65     70     75
Asn Tyr Gly Gly Gly Asn Gln Ile Gly Leu Ile Glu Asn Asp Gly Val
85     90     95
His Arg Gln Trp Tyr Met His Leu Ser Lys Tyr Asn Val Lys Val Gly
100    105    110
Asp Tyr Val Lys Ala Gly Gln Ile Ile Gly Trp Ser Gly Ser Thr Gly
115    120    125
Tyr Ser Thr Ala Pro His Leu His Phe Gln Arg Met Val Asn Ser Phe
130    135    140
Ser Asn Ser Thr Ala Gln Asp Pro Met Pro Phe Leu Lys Ser Ala Gly
145    150    155
Tyr Gly Lys Ala Gly Gly Thr Val Thr Pro Thr Pro Asn Thr Gly Trp
165    170    175
Lys Thr Asn Lys Tyr Gly Thr Leu Tyr Lys Ser Glu Ser Ala Ser Phe
180    185    190
Thr Pro Asn Thr Asp Ile Ile Thr Arg Thr Thr Gly Pro Phe Arg Ser
195    200    205
Met Pro Gln Ser Gly Val Leu Lys Ala Gly Gln Thr Ile His Tyr Asp
210    215    220
Glu Val Met Lys Gln Asp Gly His Val Trp Val Gly Tyr Thr Gly Asn
225    230    235
Ser Gly Gln Arg Ile Tyr Leu Pro Val Arg Thr Trp Asn Lys Ser Thr
245    250    255
Asn Thr Leu Gly Val Leu Trp Gly Thr Ile Lys
260    265

```

<210> 108
 <211> 8593
 <212> DNA
 <213> 合成
 <220>
 <221> 尚未归类的特征
 <222> (1) .. (8593)
 <223> 质粒 pMevT
 <220>
 <221> CDS
 <222> (380) .. (1564)
 <223> 乙酰辅酶 A_乙酰转移酶
 <220>


```

<221> CDS
<222> (1578)..(3053)
<223> 3-羟基-3-甲基戊二酰-辅酶 A_合酶
<220>
<221> CDS
<222> (3067)..(4575)
<223> N-末端_截短_3-羟基-3-甲基戊二酰_辅酶 A_Reductase
<400> 108
atcgatgcat gcgcccgaata cgcaaaccgc ctctccccgc gcgttgcccg attcattaat 60
gcagctggca cgacaggttt ccgactgga aagcgggcag tgagcgcaac gcaattaatg 120
tgagtttagct cactcattag gcaccccagg ctttacactt tatgcttcg gctcgtatgt 180
tgtgtggaat tgtgagcgga taacaatttc acacaggaaa cagctatgac catgattacg 240
ccaagcgcgc aattaaccct cactaaaggg aacaaaagct gggtagccgg cccccctcg 300
aggtcgacgg tatcgataag cttgatatcg aattcctgca gcccggggat cctctagagt 360
cgactaggag gaataataaa atg aaa aat tgt gtc atc gtc agt gcg gta cgt 412
                               Met Lys Asn Cys Val Ile Val Ser Ala Val Arg
                               1      5      10
act gct atc ggt agt ttt aac ggt tca ctc gct tcc acc agc gcc atc 460
Thr Ala Ile Gly Ser Phe Asn Gly Ser Leu Ala Ser Thr Ser Ala Ile
                               15      20      25
gac ctg ggg gcg aca gta att aaa gcc gcc att gaa cgt gca aaa atc 508
Asp Leu Gly Ala Thr Val Ile Lys Ala Ala Ile Glu Arg Ala Lys Ile
                               30      35      40
gat tca caa cac gtt gat gaa gtg att atg ggt aac gtg tta caa gcc 556
Asp Ser Gln His Val Asp Glu Val Ile Met Gly Asn Val Leu Gln Ala
                               45      50      55
ggg ctg ggg caa aat ccg gcg cgt cag gca ctg tta aaa agc ggg ctg 604
Gly Leu Gly Gln Asn Pro Ala Arg Gln Ala Leu Leu Lys Ser Gly Leu
                               60      65      70      75
gca gaa acg gtg tgc gga ttc acg gtc aat aaa gta tgt ggt tgc ggt 652
Ala Glu Thr Val Cys Gly Phe Thr Val Asn Lys Val Cys Gly Ser Gly
                               80      85      90
ctt aaa agt gtg gcg ctt gcc gcc cag gcc att cag gca ggt cag gcg 700
Leu Lys Ser Val Ala Leu Ala Ala Gln Ala Ile Gln Ala Gly Gln Ala
                               95      100      105
cag agc att gtg gcg ggg ggt atg gaa aat atg agt tta gcc ccc tac 748
Gln Ser Ile Val Ala Gly Gly Met Glu Asn Met Ser Leu Ala Pro Tyr
                               110      115      120
tta ctc gat gca aaa gca cgc tct ggt tat cgt ctt gga gac gga cag 796
Leu Leu Asp Ala Lys Ala Arg Ser Gly Tyr Arg Leu Gly Asp Gly Gln
                               125      130      135
gtt tat gac gta atc ctg cgc gat ggc ctg atg tgc gcc acc cat ggt 844
Val Tyr Asp Val Ile Leu Arg Asp Gly Leu Met Cys Ala Thr His Gly
                               140      145      150      155
tat cat atg ggg att acc gcc gaa aac gtg gct aaa gag tac gga att 892
Tyr His Met Gly Ile Thr Ala Glu Asn Val Ala Lys Glu Tyr Gly Ile
                               160      165      170
acc cgt gaa atg cag gat gaa ctg gcg cta cat tca cag cgt aaa gcg 940
Thr Arg Glu Met Gln Asp Glu Leu Ala Leu His Ser Gln Arg Lys Ala
                               175      180      185
gca gcc gca att gag tcc ggt gct ttt aca gcc gaa atc gtc ccg gta 988
Ala Ala Ala Ile Glu Ser Gly Ala Phe Thr Ala Glu Ile Val Pro Val
                               190      195      200
aat gtt gtc act cga aag aaa acc ttc gtc ttc agt caa gac gaa ttc 1036
Asn Val Val Thr Arg Lys Lys Thr Phe Val Phe Ser Gln Asp Glu Phe
                               205      210      215
ccg aaa gcg aat tca acg gct gaa gcg tta ggt gca ttg cgc ccg gcc 1084
Pro Lys Ala Asn Ser Thr Ala Glu Ala Leu Gly Ala Leu Arg Pro Ala
                               220      225      230      235
ttc gat aaa gca gga aca gtc acc gct ggg aac gcg tct ggt att aac 1132
Phe Asp Lys Ala Gly Thr Val Thr Ala Gly Asn Ala Ser Gly Ile Asn
                               240      245      250
gac ggt gct gcc gct ctg gtg att atg gaa gaa tct gcg gcg ctg gca 1180
Asp Gly Ala Ala Ala Leu Val Ile Met Glu Glu Ser Ala Ala Leu Ala
                               255      260      265
gca ggc ctt acc ccc ctg gct cgc att att aaa agt tat gcc agc ggt ggc 1228
Ala Gly Leu Thr Pro Leu Ala Arg Ile Lys Ser Tyr Ala Ser Gly Gly
                               270      275      280
gtg ccc ccc gca ttg atg ggt atg ggg cca gta cct gcc acg caa aaa 1276
Val Pro Pro Ala Leu Met Gly Met Gly Pro Val Pro Ala Thr Gln Lys
                               285      290      295
gcg tta caa ctg gcg ggg ctg caa ctg gcg gat att gat ctc att gag 1324
Ala Leu Gln Leu Ala Gly Leu Gln Leu Ala Asp Ile Asp Leu Ile Glu
                               300      305      310      315
gct aat gaa gca ttt gct gca cag ttc ctt gcc gtt ggg aaa aac ctg 1372
Ala Asn Glu Ala Phe Ala Ala Gln Phe Leu Ala Val Gly Lys Asn Leu
                               320      325      330

```

[0445]	ggc ttt gat tct gag aaa gtg aat gtc aac ggc ggg gcc atc gcg ctc	1420
	Gly Phe Asp Ser Glu Lys Val Asn Val Asn Gly Gly Ala Ile Ala Leu	
	335 340 345	
	ggg cat cct atc ggt gcc agt ggt gct cgt att ctg gtc aca cta tta	1468
	Gly His Pro Ile Gly Ala Ser Gly Ala Arg Ile Leu Val Thr Leu Leu	
	350 355 360	
	cat gcc atg cag gca cgc gat aaa acg ctg ggg ctg gca aca ctg tgc	1516
	His Ala Met Gln Ala Arg Asp Lys Thr Leu Gly Leu Ala Thr Leu Cys	
	365 370 375	
	att ggc ggc ggt cag gga att gcg atg gtg att gaa cgg ttg aat taa	1564
	Ile Gly Gly Gly Gln Gly Ile Ala Met Val Ile Glu Arg Leu Asn	
	380 385 390	
	ggaggacagc taa atg aaa ctc tca act aaa ctt tgt tgg tgt ggt att	1613
	Met Lys Leu Ser Thr Lys Leu Cys Trp Cys Gly Ile	
	395 400 405	
	aaa gga aga ctt agg ccg caa aag caa caa caa tta cac aat aca aac	1661
	Lys Gly Arg Leu Arg Pro Gln Lys Gln Gln Gln Leu His Asn Thr Asn	
	410 415 420	
	ttg caa atg act gaa cta aaa aaa caa aag acc gct gaa caa aaa acc	1709
	Leu Gln Met Thr Glu Leu Lys Lys Gln Lys Thr Ala Glu Gln Lys Thr	
	425 430 435	
	aga cct caa aat gtc ggt att aaa ggt atc caa att tac atc cca act	1757
	Arg Pro Gln Asn Val Gly Ile Lys Gly Ile Gln Tyr Ile Pro Thr	
	440 445 450	
	caa tgt gtc aac caa tct gag cta gag aaa ttt gat ggc gtt tct caa	1805
	Gln Cys Val Asn Gln Ser Glu Leu Glu Lys Phe Asp Gly Val Ser Gln	
	455 460 470	
	ggt aaa tac aca att ggt ctg ggc caa acc aac atg tct ttt gtc aat	1853
	Gly Lys Tyr Thr Ile Gly Leu Gly Gln Thr Asn Met Ser Phe Val Asn	
	475 480 485	
	gac aga gaa gat atc tac tcg atg tcc cta act gtt ttg tct aag ttg	1901
	Asp Arg Glu Asp Ile Tyr Ser Met Ser Leu Thr Val Leu Ser Lys Leu	
	490 495 500	
	atc aag agt tac aac atc gac acc aac aaa att ggt aga tta gaa gtc	1949
	Ile Lys Ser Tyr Asn Ile Asp Thr Asn Lys Ile Gly Arg Leu Glu Val	
	505 510 515	
	ggt act gaa act ctg att gac aag tcc aag tct gtc aag tct gtc ttg	1997
	Gly Thr Glu Thr Leu Ile Asp Lys Ser Lys Ser Val Lys Ser Val Leu	
	520 525 530	
	atg caa ttg ttt ggt gaa aac act gac gtc gaa ggt att gac acg ctt	2045
	Met Gln Leu Phe Gly Glu Asn Thr Asp Val Glu Gly Ile Asp Thr Leu	
	535 540 550	
	aat gcc tgt tac ggt ggt acc aac gcg ttg ttc aac tct ttg aac tgg	2093
	Asn Ala Cys Tyr Gly Gly Thr Asn Ala Leu Phe Asn Ser Leu Asn Trp	
	555 560 565	
	att gaa tct aac gca tgg gat ggt aga gac gcc att gta gtt tgc ggt	2141
	Ile Glu Ser Asn Ala Trp Asp Gly Arg Asp Ala Ile Val Val Cys Gly	
	570 575 580	
	gat att gcc atc tac gat aag ggt gcc gca aga cca acc ggt ggt gcc	2189
	Asp Ile Ala Ile Tyr Asp Lys Gly Ala Ala Arg Pro Thr Gly Gly Ala	
	585 590 595	
	ggt act gtt gct atg tgg atc ggt cct gat gct cca att gta ttt gac	2237
	Gly Thr Val Ala Met Trp Ile Gly Pro Asp Ala Pro Ile Val Phe Asp	
	600 605 610	
	tct gta aga gct tct tac atg gaa cac gcc tac gat ttt tac aag cca	2285
	Ser Val Arg Ala Ser Tyr Met Glu His Ala Tyr Asp Phe Tyr Lys Pro	
	615 620 630	
	gat ttc acc agc gaa tat cct tac gtc gat ggt cat ttt tca tta act	2333
	Asp Phe Thr Ser Glu Tyr Pro Tyr Val Asp Gly His Phe Ser Leu Thr	
	635 640 645	
	tgt tac gtc aag gct ctt gat caa gtt tac aag agt tat tcc aag aag	2381
	Cys Tyr Val Lys Ala Leu Asp Gln Val Tyr Lys Ser Tyr Ser Lys Lys	
	650 655 660	
	gct att tct aaa ggg ttg gtt agc gat ccc gct ggt tcg gat gct ttg	2429
	Ala Ile Ser Lys Gly Leu Val Ser Asp Pro Ala Gly Ser Asp Ala Leu	
	665 670 675	
	aac gtt ttg aaa tat ttc gac tac aac gtt ttc cat gtt cca acc tgt	2477
	Asn Val Leu Lys Tyr Phe Asp Tyr Asn Val Phe His Val Pro Thr Cys	
	680 685 690	
	aaa ttg gtc aca aaa tca tac ggt aga tta cta tat aac gat ttc aga	2525
	Lys Leu Val Thr Lys Ser Tyr Gly Arg Leu Leu Tyr Asn Asp Phe Arg	
	695 700 710	
	gcc aat cct caa ttg ttc cca gaa gtt gac gcc gaa tta gct act cgc	2573
	Ala Asn Pro Gln Leu Phe Pro Glu Val Asp Ala Glu Leu Ala Thr Arg	
	715 720 725	
	gat tat gac gaa tct tta acc gat aag aac att gaa aaa act ttt gtt	2621
	Asp Tyr Asp Glu Ser Leu Thr Asp Lys Asn Ile Glu Lys Thr Phe Val	
	730 735 740	

	aat gtt gct aag cca ttc cac aaa gag aga gtt gcc caa tct ttg att	2669
	Asn Val Ala Lys Pro Phe His Lys Glu Arg Val Ala Gln Ser Leu Ile	
	745 750 755	
	gtt cca aca aac aca ggt aac atg tac acc gca tct gtt tat gcc gcc	2717
	Val Pro Thr Asn Thr Gly Asn Met Tyr Thr Ala Ser Val Tyr Ala Ala	
	760 765 770	
	ttt gca tct cta tta aac tat gtt gga tct gac gac tta caa ggc aag	2765
	Phe Ala Ser Leu Leu Asn Tyr Val Gly Ser Asp Asp Leu Gln Gly Lys	
	775 780 785 790	
	cgt gtt ggt tta ttt tct tac ggt tcc ggt tta gct gca tct cta tat	2813
	Arg Val Gly Leu Phe Ser Tyr Gly Ser Gly Leu Ala Ala Ser Leu Tyr	
	795 800 805	
	tct tgc aaa att gtt ggt gac gtc caa cat att atc aag gaa tta gat	2861
	Ser Cys Lys Ile Val Gly Asp Val Gln His Ile Ile Lys Glu Leu Asp	
	810 815 820	
	att act aac aaa tta gcc aag aga atc acc gaa act cca aag gat tac	2909
	Ile Thr Asn Lys Leu Ala Lys Arg Ile Thr Glu Thr Pro Lys Asp Tyr	
	825 830 835	
	gaa gct gcc atc gaa ttg aga gaa aat gcc cat ttg aag aag aac ttc	2957
	Glu Ala Ala Ile Glu Leu Arg Glu Asn Ala His Leu Lys Lys Asn Phe	
	840 845 850	
	aaa cct caa ggt tcc att gag cat ttg caa agt ggt gtt tac tac ttg	3005
	Lys Pro Gln Gly Ser Ile Glu His Leu Gln Ser Gly Val Tyr Tyr Leu	
	855 860 865 870	
	acc aac atc gat gac aaa ttt aga aga tct tac gat gtt aaa aaa taa	3053
	Thr Asn Ile Asp Asp Lys Phe Arg Arg Ser Tyr Asp Val Lys Lys	
	875 880 885	
	ggaggattac act atg gtt tta acc aat aaa aca gtc att tct gga tcg	3102
	Met Val Leu Thr Asn Lys Thr Val Ile Ser Gly Ser	
	890 895	
	aaa gtc aaa agt tta tca tct gcg caa tcg agc tca tca gga cct tca	3150
	Lys Val Lys Ser Leu Ser Ser Ala Gln Ser Ser Ser Ser Gly Pro Ser	
	900 905 910	
	tca tct agt gag gaa gat gat tcc cgc gat att gaa agc ttg gat aag	3198
	Ser Ser Ser Glu Glu Asp Asp Ser Arg Asp Ile Glu Ser Leu Asp Lys	
	915 920 925	
	aaa ata cgt cct tta gaa gaa tta gaa gca tta agt agt gga aat	3246
	Lys Ile Arg Pro Leu Glu Glu Leu Glu Ala Leu Leu Ser Ser Gly Asn	
	930 935 940 945	
	aca aaa caa ttg aag aac aaa gag gtc gct gcc ttg gtt att cac ggt	3294
	Thr Lys Gln Leu Lys Asn Lys Glu Val Ala Ala Leu Val Ile His Gly	
	950 955 960	
	aag tta cct ttg tac gct ttg gag aaa tta ggt gat act acg aga	3342
	Lys Leu Pro Leu Tyr Ala Leu Glu Lys Lys Leu Gly Asp Thr Thr Arg	
	965 970 975	
	gcg gtt gcg gta cgt agg aag gct ctt tca att ttg gca gaa gct cct	3390
	Ala Val Ala Val Arg Arg Lys Ala Leu Ser Ile Leu Ala Glu Ala Pro	
	980 985 990	
	gta tta gca tct gat cgt tta cca tat aaa aat tat gac tac gac cgc	3438
	Val Leu Ala Ser Asp Arg Leu Pro Tyr Lys Asn Tyr Asp Tyr Asp Arg	
	995 1000 1005	
	gta ttt ggc gct tgt tgt gaa aat gtt ata ggt tac atg cct ttg	3483
	Val Phe Gly Ala Cys Cys Glu Asn Val Ile Gly Tyr Met Pro Leu	
	1010 1015 1020	
	ccc gtt ggt gtt ata ggc ccc ttg gtt atc gat ggt aca tct tat	3528
	Pro Val Gly Val Ile Gly Pro Leu Val Ile Asp Gly Thr Ser Tyr	
	1025 1030 1035	
	cat ata cca atg gca act aca gag ggt tgt ttg gta gct tct gcc	3573
	His Ile Pro Met Ala Thr Thr Glu Gly Cys Leu Val Ala Ser Ala	
	1040 1045 1050	
	atg cgt ggc tgt aag gca atc aat gct ggc ggt ggt gca aca act	3618
	Met Arg Gly Cys Lys Ala Ile Asn Ala Gly Gly Gly Ala Thr Thr	
	1055 1060 1065	
	gtt tta act aag gat ggt atg aca aga ggc cca gta gtc cgt ttc	3663
	Val Leu Thr Lys Asp Gly Met Thr Arg Gly Pro Val Val Arg Phe	
	1070 1075 1080	
	cca act ttg aaa aga tct ggt gcc tgt aag ata tgg tta gac tca	3708
	Pro Thr Leu Lys Arg Ser Gly Ala Cys Lys Ile Trp Leu Asp Ser	
	1085 1090 1095	
	gaa gag gga caa aac gca att aaa aaa gct ttt aac tct aca tca	3753
	Glu Glu Gly Gln Asn Ala Ile Lys Lys Ala Phe Asn Ser Thr Ser	
	1100 1105 1110	
	aga ttt gca cgt ctg caa cat att caa act tgt cta gca gga gat	3798
	Arg Phe Ala Arg Leu Gln His Ile Gln Thr Cys Leu Ala Gly Asp	
	1115 1120 1125	
	tta ctc ttc atg aga ttt aga aca act act ggt gac gca atg ggt	3843
	Leu Leu Phe Met Arg Phe Arg Thr Thr Thr Gly Asp Ala Met Gly	
	1130 1135 1140	

[0447]

atg aat atg att tct aaa ggt gtc gaa tac tca tta aag caa atg	3888
Met Asn Met Ile Ser Lys Gly Val Glu Tyr Ser Leu Lys Gln Met	
1145 1150 1155	
gta gaa gag tat ggc tgg gaa gat atg gag gtt gtc tcc gtt tct	3933
Val Glu Glu Tyr Gly Trp Glu Asp Met Glu Val Val Ser Val Ser	
1160 1165 1170	
ggt aac tac tgt acc gac aaa aaa cca gct gcc atc aac tgg atc	3978
Gly Asn Tyr Cys Thr Asp Lys Lys Pro Ala Ala Ile Asn Trp Ile	
1175 1180 1185	
gaa ggt cgt ggt aag agt gtc gtc gca gaa gct act att cct ggt	4023
Glu Gly Arg Gly Lys Ser Val Val Ala Glu Ala Thr Ile Pro Gly	
1190 1195 1200	
gat gtt gtc aga aaa gtg tta aaa agt gat gtt tcc gca ttg gtt	4068
Asp Val Val Arg Lys Val Leu Lys Ser Asp Val Ser Ala Leu Val	
1205 1210 1215	
gag ttg aac att gct aag aat ttg gtt gga tct gca atg gct ggg	4113
Glu Leu Asn Ile Ala Lys Asn Leu Val Gly Ser Ala Met Ala Gly	
1220 1225 1230	
tct gtt ggt gga ttt aac gca cat gca gct aat tta gtg aca gct	4158
Ser Val Gly Gly Phe Asn Ala His Ala Ala Asn Leu Val Thr Ala	
1235 1240 1245	
gtt ttc ttg gca tta gga caa gat cct gca caa aat gtt gaa agt	4203
Val Phe Leu Ala Leu Gly Gln Asp Pro Ala Gln Asn Val Glu Ser	
1250 1255 1260	
tcc aac tgt ata aca ttg atg aaa gaa gtg gac ggt gat ttg aga	4248
Ser Asn Cys Ile Thr Leu Met Lys Glu Val Asp Gly Asp Leu Arg	
1265 1270 1275	
att tcc gta tcc atg cca tcc atc gaa gta ggt acc atc ggt ggt	4293
Ile Ser Val Ser Met Pro Ser Ile Glu Val Gly Thr Ile Gly Gly	
1280 1285 1290	
ggt act gtt cta gaa cca caa ggt gcc atg ttg gac tta tta ggt	4338
Gly Thr Val Leu Glu Pro Gln Gly Ala Met Leu Asp Leu Leu Gly	
1295 1300 1305	
gta aga ggc ccg cat gct acc gct cct ggt acc aac gca cgt caa	4383
Val Arg Gly Pro His Ala Thr Ala Pro Gly Thr Asn Ala Arg Gln	
1310 1315 1320	
tta gca aga ata gtt gcc tgt gcc gtc ttg gca ggt gaa tta tcc	4428
Leu Ala Arg Ile Val Ala Cys Ala Val Leu Ala Gly Glu Leu Ser	
1325 1330 1335	
tta tgt gct gcc cta gca gcc ggc cat ttg gtt caa agt cat atg	4473
Leu Cys Ala Ala Leu Ala Ala Gly His Leu Val Gln Ser His Met	
1340 1345 1350	
acc cac aac agg aaa cct gct gaa cca aca aaa cct aac aat ttg	4518
Thr His Asn Arg Lys Pro Ala Glu Pro Thr Lys Pro Asn Asn Leu	
1355 1360 1365	
gac gcc act gat ata aat cgt ttg aaa gat ggg tcc gtc acc tgc	4563
Asp Ala Thr Asp Ile Asn Arg Leu Lys Asp Gly Ser Val Thr Cys	
1370 1375 1380	
att aaa tcc taa gtcgacctgc aggcattgcaa gcttggtgt tttggcggat	4615
Ile Lys Ser	
1385	
gagagaagat tttcagcctg atacagatta aatcagaacg cagaagcggg ctgataaaaac	4675
agaattttgcc tggcggcagt agcgcggtgg tcccacctga ccccatgccc aactcagaag	4735
tgaaacgcgc tagcgccgat ggtagtgtgg ggtctcccca tgcgagagta gggaactgcc	4795
aggcatcaaa taaaacgaaa ggctcagtcg aaagactggg cctttcgttt tatctgttgt	4855
ttgtcggtga acgctctcct gagtaggaca aatccgcggg gacgggattt gaacgttgcg	4915
aagcaacggc ccggagggtg gcgggcagga cgcccgccat aaactgccag gcatcaaatt	4975
aagcagaagg ccatacctgac ggtatggcctt tttgcgtttc tacaaaactct tttgtttatt	5035
tttctaaata cattcaaata tgtatccgct catgagacaa taacctgat aaatgcttca	5095
ataatattga aaaaggaaga gtatgagtat tcaacatttc cgtgtcgccc ttattccctt	5155
ttttgcggca ttttgccttc ctgtttttgc tcaccagaaa acgctggtga aagtaaaaga	5215
tgctgaagat cagttgggtg cagcaaaact ttaactggcg aactacttac tctagcttcc	5275
cggcaacaat taatagactg gatggaggcg gataaagtgg caggaccact tctgcgctcg	5335
gcccttccgg ctggctggtt tattgtctgat aaatctggag ccggtgagcg tgggtctcgc	5395
ggtatcattg cagcactggg gccagatggt aagccctccc gtatcgtagt tatctacacg	5455
acggggagtc aggcactat ggatgaacga aatagacaga tcgctgagat aggtgcctca	5515
ctgattaagc attggtaact gtcagaccaa gtttactcat atatacttta gattgattta	5575
cgcgccctgt agcggcgcat taagcgcggc ggggtgtgtg gttacgcgca gcgtgaccgc	5635
tacacttgc agcgccctag cgccgcctcc tttcgtttc ttcccttcc tctcgccac	5695
gttcgcggc tttcccgctc aagctctaaa tcggggggctc ccttttaggt tccgatttag	5755
tgctttacgg cacctcgacc ccaaaaaact tgatttgggt gatggttcac gtatggggcc	5815
atcgccctga tagacggttt ttgcgccctt gacgttggag tccacgttct ttaatatgtg	5875
actcttgctt caaacttgaa ccaacttcaa ccttatctcg ggctattctt ttgatttata	5935
agggattttg ccgatttctg cctattgggt aaaaaatgag ctgatttaac aaaaatttaa	5995
cgcgaaatttt aacaaaaatat taacgtttac aatttaaaag gatctagggtg aagatccttt	6055
ttgataatct catgaccaaa atcccttaac gtgagttttc gttccactga cgctcagacc	6115
ccgtagaaaa gatcaaagga tcttcttgag atcctttttt tctgcgcgta atctgctgct	6175
tgcaaacaaaa aaaaccaccg ctaccagcgg tggtttggtt gccggatcaa gagctaccaa	6235

[0448]

```

ctcttttttcc gaaggtaact ggcttcagca gagcgcagat accaaatact gtcctttctag 6295
tgtagccgta gttaggccac cacttcaaga actctgtagc accgcctaca tacctcgctc 6355
tgctaatacct gttaccagtg gggcatttga gaagcacacg gtcacactgc ttccggtagt 6415
caataaaaccg gtaaaccagc aatagacata agcggctatt taacgacct gccctgaacc 6475
gacgaccggg tcgaatttgc tttcgaattt ctgccattca tccgcttatt atcacttatt 6535
caggcgtagc accaggcggt taagggcacc aataactgcc ttaaaaaaat tacgccccgc 6595
cctgccactc atcgcagtac tgttgtaatt cattaagcat tctgccgaca tgggaagccat 6655
cacagacggc atgatgaacc tgaatcgcca gcggcatcag caccttgctg ccttgcgat 6715
aatattttgcc catggtgaaa acggggggcg agaagtgtgc catattggcc acgttttaat 6775
caaaactgggt gaaactcacc cagggatttg ctgagacgaa aaacataatc tcaataaacc 6835
cttttagggaa ataggccagg ttttcaccgt tcaggcgggc aagaaatgtg ataaaggccg 6895
gaaactgcgc gaaatcgctg tgggtattcac tccagagcga tgaaaacggt tcagtttgct 6955
catggaaaaac ggtgtaacaa ggggtgaacac tatcccatat caccagctca ccgctctttca 7015
ttgccatagc gaattccgga tgagcattca tcaggcgggc aagaatgtga ataaaggccg 7075
gataaaactt gtgcttattt tcttttacgg tctttaaaaa ggccgtaata tccagctgaa 7135
cggctcgtgt ataggtacat tgagcaactg actgaaatgc ctcaaaatgt tctttacgat 7195
gccattggga tatatcaacg gtggtatatc cagtgtattt tttctccatt tttagcttct 7255
tagctcctga aaatctcgat aactcaaaaa atacgcccgg tagtgatctt atttcattat 7315
ggtgaaagtt ggaacctctt acgtgccgat caacgtctca ttttcgccaa aagttggccc 7375
agggcttccc ggtatcaaca gggacaccag gatttattta ttctgcgaag tgatcttccg 7435
tcacaggtat ttattcggcg caaagtgcgt cgggtgatgc tgccaactta ctatcagctg 7495
gtatgatggt gtttttgagg tgctccagtg gcttctgttt ctatcagctg tccctcctgt 7555
tcagctactg accggggtggt gcgtaacggc aaaagcaccg ccggacatca cgcgtagcgg 7615
agtgtactat gcttactatc gttggcactg atgaggggtg cagtgaagtg cctcatgtgg 7675
caggagaaaa aaggctgcac cggtgcgtea gcagaatatg tgatacagga tatattccgc 7735
ttcctcgctc actgactcgc tacgctcggg cgttcgactg cggcgagcgg aaatggctta 7795
cgaacggggc ggagatttcc tggaaagatg caggaagata cttaacaggg aagtgaagg 7855
gccgcggcaa agccgttttt ccataggctc cgcccccttg acaagcatca cgaatctga 7915
cgctcaaatc agtgggtggc aaacccgaca ggactataaa gataccaggc gtttccccct 7975
ggcgctcccc tctgctcctc tctgttccct gcctttcggt ttaccggtgt cattccgctg 8035
ttatggccgc gtttgctcga ttccacgcct gacactcagt tccgggtagg cagttcgctc 8095
caagctggac tgtatgcacg aacccccctg tcagtcgcac cgctgcgcct tatccggtaa 8155
ctatcgtctt gagtccaacc cggaagaca tgcaaaagca ccaactggcg cagcactgg 8215
taattgattt agaggagtta gtottgaagt catcgccggg ttaaggctaa actgaaagga 8275
caagtttttg tgactgcgct cctccaagcc agttacctcg gttcaaagag ttggtagctc 8335
agagaacctt cgaaaaaccg ccctgcaagg cggttttttc gttttcagag caagagatta 8395
cgcgacagac aaaacgatct caagaagatc atcttattaa tcagataaaa tatttgctca 8455
tgagcccgaa gtggcgagcc cgaatctccc catcggtgat gtcggcgata taggcgccag 8515
caaccgcacc tgtggcgccg gtgatgccgg ccacgatgcg tccggcgtag aggatctgct 8575
catgtttgac agcttatac 8593

```

<210> 109
 <211> 394
 <212> PRT
 <213> 合成
 <400> 109

```

Met Lys Asn Cys Val Ile Val Ser Ala Val Arg Thr Ala Ile Gly Ser
1      5      10      15
Phe Asn Gly Ser Leu Ala Ser Thr Ser Ala Ile Asp Leu Gly Ala Thr
20     25     30
Val Ile Lys Ala Ala Ile Glu Arg Ala Lys Ile Asp Ser Gln His Val
35     40     45
Asp Glu Val Ile Met Gly Asn Val Leu Gln Ala Gly Leu Gly Gln Asn
50     55     60
Pro Ala Arg Gln Ala Leu Leu Lys Ser Gly Leu Ala Glu Thr Val Cys
65     70     75     80
Gly Phe Thr Val Asn Lys Val Cys Gly Ser Gly Leu Lys Ser Val Ala
85     90     95
Leu Ala Ala Gln Ala Ile Gln Ala Gly Gln Ala Gln Ser Ile Val Ala
100    105    110
Gly Gly Met Glu Asn Met Ser Leu Ala Pro Tyr Leu Leu Asp Ala Lys
115    120    125
Ala Arg Ser Gly Tyr Arg Leu Gly Asp Gly Gln Val Tyr Asp Val Ile
130    135    140
Leu Arg Asp Gly Leu Met Cys Ala Thr His Gly Tyr His Met Gly Ile
145    150    155    160
Thr Ala Glu Asn Val Ala Lys Glu Tyr Gly Ile Thr Arg Glu Met Gln
165    170    175
Asp Glu Leu Ala Leu His Ser Gln Arg Lys Ala Ala Ala Ile Glu
180    185    190
Ser Gly Ala Phe Thr Ala Glu Ile Val Pro Val Asn Val Val Thr Arg
195    200    205
Lys Lys Thr Phe Val Phe Ser Gln Asp Glu Phe Pro Lys Ala Asn Ser
210    215    220
Thr Ala Glu Ala Leu Gly Ala Leu Arg Pro Ala Phe Asp Lys Ala Gly
225    230    235    240
Thr Val Thr Ala Gly Asn Ala Ser Gly Ile Asn Asp Gly Ala Ala Ala
245    250    255
Leu Val Ile Met Glu Glu Ser Ala Ala Leu Ala Ala Gly Leu Thr Pro

```

260 265 270
 Leu Ala Arg Ile Lys Ser Tyr Ala Ser Gly Gly Val Pro Pro Ala Leu
 275 280 285
 Met Gly Met Gly Pro Val Pro Ala Thr Gln Lys Ala Leu Gln Leu Ala
 290 295 300
 Gly Leu Gln Leu Ala Asp Ile Asp Leu Ile Glu Ala Asn Glu Ala Phe
 305 310 315 320
 Ala Ala Gln Phe Leu Ala Val Gly Lys Asn Leu Gly Phe Asp Ser Glu
 325 330 335
 Lys Val Asn Val Asn Gly Gly Ala Ile Ala Leu Gly His Pro Ile Gly
 340 345 350
 Ala Ser Gly Ala Arg Ile Leu Val Thr Leu Leu His Ala Met Gln Ala
 355 360 365
 Arg Asp Lys Thr Leu Gly Leu Ala Thr Leu Cys Ile Gly Gly Gln
 370 375 380
 Gly Ile Ala Met Val Ile Glu Arg Leu Asn
 385 390
 <210> 110
 <211> 491
 <212> PRT
 <213> 合成
 <400> 110
 Met Lys Leu Ser Thr Lys Leu Cys Trp Cys Gly Ile Lys Gly Arg Leu
 1 5 10 15
 Arg Pro Gln Lys Gln Gln Gln Leu His Asn Thr Asn Leu Gln Met Thr
 20 25 30
 Glu Leu Lys Lys Gln Lys Thr Ala Glu Gln Lys Thr Arg Pro Gln Asn
 35 40 45
 Val Gly Ile Lys Gly Ile Gln Ile Tyr Ile Pro Thr Gln Cys Val Asn
 50 55 60
 Gln Ser Glu Leu Glu Lys Phe Asp Gly Val Ser Gln Gly Lys Tyr Thr
 65 70 75 80
 Ile Gly Leu Gly Gln Thr Asn Met Ser Phe Val Asn Asp Arg Glu Asp
 85 90 95
 Ile Tyr Ser Met Ser Leu Thr Val Leu Ser Lys Leu Ile Lys Ser Tyr
 100 105 110
 Asn Ile Asp Thr Asn Lys Ile Gly Arg Leu Glu Val Gly Thr Glu Thr
 115 120 125
 [0449] Leu Ile Asp Lys Ser Lys Ser Val Lys Ser Val Leu Met Gln Leu Phe
 130 135 140
 Gly Glu Asn Thr Asp Val Glu Gly Ile Asp Thr Leu Asn Ala Cys Tyr
 145 150 155 160
 Gly Gly Thr Asn Ala Leu Phe Asn Ser Leu Asn Trp Ile Glu Ser Asn
 165 170 175
 Ala Trp Asp Gly Arg Asp Ala Ile Val Val Cys Gly Asp Ile Ala Ile
 180 185 190
 Tyr Asp Lys Gly Ala Ala Arg Pro Thr Gly Gly Ala Gly Thr Val Ala
 195 200 205
 Met Trp Ile Gly Pro Asp Ala Pro Ile Val Phe Asp Ser Val Arg Ala
 210 215 220
 Ser Tyr Met Glu His Ala Tyr Asp Phe Tyr Lys Pro Asp Phe Thr Ser
 225 230 235 240
 Glu Tyr Pro Tyr Val Asp Gly His Phe Ser Leu Thr Cys Tyr Val Lys
 245 250 255
 Ala Leu Asp Gln Val Tyr Lys Ser Tyr Ser Lys Lys Ala Ile Ser Lys
 260 265 270
 Gly Leu Val Ser Asp Pro Ala Gly Ser Asp Ala Leu Asn Val Leu Lys
 275 280 285
 Tyr Phe Asp Tyr Asn Val Phe His Val Pro Thr Cys Lys Leu Val Thr
 290 295 300
 Lys Ser Tyr Gly Arg Leu Leu Tyr Asn Asp Phe Arg Ala Asn Pro Gln
 305 310 315 320
 Leu Phe Pro Glu Val Asp Ala Glu Leu Ala Thr Arg Asp Tyr Asp Glu
 325 330 335
 Ser Leu Thr Asp Lys Asn Ile Glu Lys Thr Phe Val Asn Val Ala Lys
 340 345 350
 Pro Phe His Lys Glu Arg Val Ala Gln Ser Leu Ile Val Pro Thr Asn
 355 360 365
 Thr Gly Asn Met Tyr Thr Ala Ser Val Tyr Ala Ala Phe Ala Ser Leu
 370 375 380
 Leu Asn Tyr Val Gly Ser Asp Asp Leu Gln Gly Lys Arg Val Gly Leu
 385 390 395 400
 Phe Ser Tyr Gly Ser Gly Leu Ala Ala Ser Leu Tyr Ser Cys Lys Ile
 405 410 415
 Val Gly Asp Val Gln His Ile Ile Lys Glu Leu Asp Ile Thr Asn Lys
 420 425 430
 Leu Ala Lys Arg Ile Thr Glu Thr Pro Lys Asp Tyr Glu Ala Ala Ile
 435 440 445

[0450]

```

Glu Leu Arg Glu Asn Ala His Leu Lys Lys Asn Phe Lys Pro Gln Gly
 450      455      460
Ser Ile Glu His Leu Gln Ser Gly Val Tyr Tyr Leu Thr Asn Ile Asp
 465      470      475      480
Asp Lys Phe Arg Arg Ser Tyr Asp Val Lys Lys
      485      490
<210> 111
<211> 502
<212> PRT
<213> 合成
<400> 111
Met Val Leu Thr Asn Lys Thr Val Ile Ser Gly Ser Lys Val Lys Ser
 1      5      10      15
Leu Ser Ser Ala Gln Ser Ser Ser Ser Gly Pro Ser Ser Ser Glu
      20      25      30
Glu Asp Asp Ser Arg Asp Ile Glu Ser Leu Asp Lys Lys Ile Arg Pro
      35      40      45
Leu Glu Glu Leu Glu Ala Leu Leu Ser Ser Gly Asn Thr Lys Gln Leu
      50      55      60
Lys Asn Lys Glu Val Ala Ala Leu Val Ile His Gly Lys Leu Pro Leu
      65      70      75      80
Tyr Ala Leu Glu Lys Lys Leu Gly Asp Thr Thr Arg Ala Val Ala Val
      85      90      95
Arg Arg Lys Ala Leu Ser Ile Leu Ala Glu Ala Pro Val Leu Ala Ser
      100      105      110
Asp Arg Leu Pro Tyr Lys Asn Tyr Asp Tyr Asp Arg Val Phe Gly Ala
      115      120      125
Cys Cys Glu Asn Val Ile Gly Tyr Met Pro Leu Pro Val Gly Val Ile
      130      135      140
Gly Pro Leu Val Ile Asp Gly Thr Ser Tyr His Ile Pro Met Ala Thr
      145      150      155      160
Thr Glu Gly Cys Leu Val Ala Ser Ala Met Arg Gly Cys Lys Ala Ile
      165      170      175
Asn Ala Gly Gly Gly Ala Thr Thr Val Leu Thr Lys Asp Gly Met Thr
      180      185      190
Arg Gly Pro Val Val Arg Phe Pro Thr Leu Lys Arg Ser Gly Ala Cys
      195      200      205
Lys Ile Trp Leu Asp Ser Glu Glu Gly Gln Asn Ala Ile Lys Lys Ala
      210      215      220
Phe Asn Ser Thr Ser Arg Phe Ala Arg Leu Gln His Ile Gln Thr Cys
      225      230      235      240
Leu Ala Gly Asp Leu Leu Phe Met Arg Phe Arg Thr Thr Thr Gly Asp
      245      250      255
Ala Met Gly Met Asn Met Ile Ser Lys Gly Val Glu Tyr Ser Leu Lys
      260      265      270
Gln Met Val Glu Glu Tyr Gly Trp Glu Asp Met Glu Val Ser Val
      275      280      285
Ser Gly Asn Tyr Cys Thr Asp Lys Lys Pro Ala Ala Ile Asn Trp Ile
      290      295      300
Glu Gly Arg Gly Lys Ser Val Val Ala Glu Ala Thr Ile Pro Gly Asp
      305      310      315      320
Val Val Arg Lys Val Leu Lys Ser Asp Val Ser Ala Leu Val Glu Leu
      325      330      335
Asn Ile Ala Lys Asn Leu Val Gly Ser Ala Met Ala Gly Ser Val Gly
      340      345      350
Gly Phe Asn Ala His Ala Ala Asn Leu Val Thr Ala Val Phe Leu Ala
      355      360      365
Leu Gly Gln Asp Pro Ala Gln Asn Val Glu Ser Ser Asn Cys Ile Thr
      370      375      380
Leu Met Lys Glu Val Asp Gly Asp Leu Arg Ile Ser Val Ser Met Pro
      385      390      395      400
Ser Ile Glu Val Gly Thr Ile Gly Gly Gly Thr Val Leu Glu Pro Gln
      405      410      415
Gly Ala Met Leu Asp Leu Leu Gly Val Arg Gly Pro His Ala Thr Ala
      420      425      430
Pro Gly Thr Asn Ala Arg Gln Leu Ala Arg Ile Val Ala Cys Ala Val
      435      440      445
Leu Ala Gly Glu Leu Ser Leu Cys Ala Ala Leu Ala Ala Gly His Leu
      450      455      460
Val Gln Ser His Met Thr His Asn Arg Lys Pro Ala Glu Pro Thr Lys
      465      470      475      480
Pro Asn Asn Leu Asp Ala Thr Asp Ile Asn Arg Leu Lys Asp Gly Ser
      485      490      495
Val Thr Cys Ile Lys Ser
      500
<210> 112
<211> 32
<212> DNA

```

[0451]

```

<213> 合成
<220>
<221> 引物 结合
<222> (1) .. (32)
<223> yccV 正向引物
<400> 112
attcatatgu gggtataaaag acatcaacat gc 32
<210> 113
<211> 90
<212> DNA
<213> 合成
<220>
<221> 引物 结合
<222> (1) .. (90)
<223> yccV 反向引物
<220>
<221> 尚未归类的特征
<222> (65) .. (65)
<223> N 是 A、T、C 或 G
<400> 113
atccccatta cgtgaggatg gctaagggtcc agtgaagtac cctggggcaaa gagtttcata 60
ccttncaagt agtcacctcc cgggaaatct 90
<210> 114
<211> 32
<212> DNA
<213> 合成
<220>
<221> 引物 结合
<222> (1) .. (32)
<223> ycjX 正向引物
<400> 114
attcatatgu ttggctgata ctgtaattct tc 32
<210> 115
<211> 90
<212> DNA
<213> 合成
<220>
<221> 引物 结合
<222> (1) .. (90)
<223> ycjX 反向引物
<220>
<221> 尚未归类的特征
<222> (65) .. (65)
<223> N 是 A、T、C 或 G
<400> 115
atccccatta cgtgaggatg gctaagggtcc agtgaagtac cctggggcaaa gagtttcata 60
ccttncaagt tgctgtcctg tgctgtcctg 90
<210> 116
<211> 32
<212> DNA
<213> 合成
<220>
<221> 引物 结合
<222> (1) .. (32)
<223> mutM 正向引物
<400> 116
attcatatgu ggcgaaaagt actaagtact ta 32
<210> 117
<211> 90
<212> DNA
<213> 合成
<220>
<221> 引物 结合
<222> (1) .. (90)
<223> mutM 反向引物
<220>
<221> 尚未归类的特征
<222> (65) .. (65)
<223> N 是 A、T、C 或 G
<400> 117
atccccatta cgtgaggatg gctaagggtcc agtgaagtac cctggggcaaa gagtttcata 60
ccttncaagt agcatctcca ggaatgaaca 90
<210> 118
<211> 19

```

	<212> DNA	
	<213> 合成	
	<220>	
	<221> 引物 结合	
	<222> (1) .. (19)	
	<223> 引物 P492	
	<400> 118	
	tgcaactctt cttcaacgc	19
	<210> 119	
	<211> 19	
	<212> DNA	
	<213> 合成	
	<220>	
	<221> 引物 结合	
	<222> (1) .. (19)	
	<223> 引物 P493	
	<400> 119	
	aagaaccagg cgcttctaa	19
	<210> 120	
	<211> 19	
	<212> DNA	
	<213> 合成	
	<220>	
	<221> 引物 结合	
	<222> (1) .. (19)	
	<223> 引物 P525	
	<400> 120	
	attgatgacc tgatggcac	19
	<210> 121	
	<211> 18	
	<212> DNA	
	<213> 合成	
	<220>	
	<221> 引物 结合	
	<222> (1) .. (18)	
	<223> 引物 P526	
[0452]	<400> 121	
	gatccagact tcgaagcg	18
	<210> 122	
	<211> 90	
	<212> DNA	
	<213> 合成	
	<220>	
	<221> 引物 结合	
	<222> (1) .. (90)	
	<223> ibpA 正向引物	
	<220>	
	<221> 尚未归类的特征	
	<222> (65) .. (65)	
	<223> N 是 A、T、C 或 G	
	<400> 122	
	atccccatta cgtgaggatg gctaagggtcc agtgaagtac cctgggcaaa gagtttcata	60
	ccttncaagt aatcaatagc tcctgaaatc	90
	<210> 123	
	<211> 31	
	<212> DNA	
	<213> 合成	
	<220>	
	<221> 引物 结合	
	<222> (1) .. (31)	
	<223> ibpA 反向引物	
	<400> 123	
	attcatatgt ttggcctgat gagttatagc g	31
	<210> 124	
	<211> 32	
	<212> DNA	
	<213> 合成	
	<220>	
	<221> 引物 结合	
	<222> (1) .. (32)	
	<223> grpE 正向引物	
	<400> 124	
	attcatatgu tctccgcgag cgtgccagtt tt	32
	<210> 125	

<211> 90
 <212> DNA
 <213> 合成
 <220>
 <221> 引物 结合
 <222> (1) .. (90)
 <223> grpE 反向引物
 <220>
 <221> 尚未归类的特征
 <222> (65) .. (65)
 <223> N 是 A、T、C 或 G
 <400> 125
 atccccatta cgtgaggatg gctaagggtcc agtgaagtac cctgggcaaa gagtttcata 60
 ccttncaagt gaatttctcc gcgttttttt 90
 <210> 126
 <211> 32
 <212> DNA
 <213> 合成
 <220>
 <221> 引物 结合
 <222> (1) .. (32)
 <223> ldhA 正向引
 <400> 126
 attcatatgu gggaaccac agcccgagcg tc 32
 <210> 127
 <211> 90
 <212> DNA
 <213> 合成
 <220>
 <221> 引物 结合
 <222> (1) .. (90)
 <223> ldhA 反向引物
 <220>
 <221> 尚未归类的特征
 <222> (65) .. (65)
 <223> N 是 A、T、C 或 G
 <400> 127
 atccccatta cgtgaggatg gctaagggtcc agtgaagtac cctgggcaaa gagtttcata 60
 ccttncaagt aagactttct ccagtgatgt 90
 <210> 128
 <211> 122
 <212> DNA
 <213> 合成
 <220>
 <221> 引物 结合
 <222> (1) .. (122)
 <223> ybbN 正向引物
 <220>
 <221> 尚未归类的特征
 <222> (97) .. (97)
 <223> N 是 A、T、C 或 G
 <400> 128
 attcatatgu cccattcgta ctgcgttcac cgatcccat tacgtgagga tggctaaggt 60
 ccagtgaagt accctgggca aagagtttca taccttncaa gtggagtcgc tctctgttgt 120
 cg 122
 <210> 129
 <211> 90
 <212> DNA
 <213> 合成
 <220>
 <221> 引物 结合
 <222> (1) .. (90)
 <223> ybbN 反向引物
 <220>
 <221> 尚未归类的特征
 <222> (65) .. (65)
 <223> N 是 A、T、C 或 G
 <400> 129
 atccccatta cgtgaggatg gctaagggtcc agtgaagtac cctgggcaaa gagtttcata 60
 ccttncaagt ggagtcgctc tctgttgtcg 90
 <210> 130
 <211> 32
 <212> DNA
 <213> 合成

	<220>		
	<221>	引物_结合	
	<222>	(1) .. (32)	
	<223>	fxsA 正向引物	
	<400>	130	
		attcatatgu taaacacaca gaatatatgt gg	32
	<210>	131	
	<211>	90	
	<212>	DNA	
	<213>	合成	
	<220>		
	<221>	引物_结合	
	<222>	(1) .. (90)	
	<223>	fxsA 反向引物	
	<220>		
	<221>	尚未归类的特征	
	<222>	(65) .. (65)	
	<223>	N 是 A、T、C 或 G	
	<400>	131	
		atccccatta cgtgaggatg gctaagggtcc agtgaagtac cctggggcaaa gagtttcata	60
		ccttncaagt aggttttctcc tgtaatagca	90
	<210>	132	
	<211>	31	
	<212>	DNA	
	<213>	合成	
	<220>		
	<221>	引物_结合	
	<222>	(1) .. (31)	
	<223>	rrnB 正向引物	
	<400>	132	
		attcatatgu gcaaacacgc cgccgggtca g	31
	<210>	133	
	<211>	90	
	<212>	DNA	
	<213>	合成	
[0454]	<220>		
	<221>	引物_结合	
	<222>	(1) .. (90)	
	<223>	rrnB 反向引物	
	<220>		
	<221>	尚未归类的特征	
	<222>	(65) .. (65)	
	<223>	N 是 A、T、C 或 G	
	<400>	133	
		atccccatta cgtgaggatg gctaagggtcc agtgaagtac cctggggcaaa gagtttcata	60
		ccttncaagt aaaagtttga cgctcaaaga	90
	<210>	134	
	<211>	35	
	<212>	DNA	
	<213>	合成	
	<220>		
	<221>	引物_结合	
	<222>	(1) .. (35)	
	<223>	rrnE 正向引物	
	<400>	134	
		attcatatgu aatgcgcctc cggtgagacg acaac	35
	<210>	135	
	<211>	90	
	<212>	DNA	
	<213>	合成	
	<220>		
	<221>	引物_结合	
	<222>	(1) .. (90)	
	<223>	rrnE 反向引物	
	<220>		
	<221>	尚未归类的特征	
	<222>	(65) .. (65)	
	<223>	N 是 A、T、C 或 G	
	<400>	135	
		atccccatta cgtgaggatg gctaagggtcc agtgaagtac cctggggcaaa gagtttcata	60
		ccttncaagt aaaagtttga tgctcaaaga	90
	<210>	136	
	<211>	32	
	<212>	DNA	

[0455]

```

<213> 合成
<220>
<221> 引物 结合
<222> (1) .. (32)
<223> poxB 正向引物
<400> 136
attcatatgu taggttgctg ctgcctgccg tg
<210> 137
<211> 90
<212> DNA
<213> 合成
<220>
<221> 引物 结合
<222> (1) .. (90)
<223> poxB 反向引物
<220>
<221> 尚未归类的特征
<222> (65) .. (65)
<223> N 是 A、T、C 或 G
<400> 137
atccccatta cgtgaggatg gctaagggtcc agtgaagtac cctgggcaaa gagtttcata
ccttncaagt gggtctccat ctctgaatg
<210> 138
<211> 32
<212> DNA
<213> 合成
<220>
<221> 引物 结合
<222> (1) .. (32)
<223> cspD 正向引物
<400> 138
attcatatgu cggttatcgg cagaacgccc tg
<210> 139
<211> 90
<212> DNA
<213> 合成
<220>
<221> 引物 结合
<222> (1) .. (90)
<223> cspD 反向引物
<220>
<221> 尚未归类的特征
<222> (65) .. (65)
<223> N 是 A、T、C 或 G
<400> 139
atccccatta cgtgaggatg gctaagggtcc agtgaagtac cctgggcaaa gagtttcata
ccttncaagt gcttcgacat ccttcgcaaa
<210> 140
<211> 1902
<212> DNA
<213> 合成
<220>
<221> 尚未归类的特征
<222> (1) .. (1902)
<223> 整合序列 s9_pmutM_folP
<220>
<221> CDS
<222> (299) .. (1093)
<223> kanR
<220>
<221> 启动子
<222> (1533) .. (1832)
<223> 负荷感应启动子 p(mutM)
<220>
<221> RBS
<222> (1833) .. (1838)
<223> rbs 目录
<220>
<221> 尚未归类的特征
<222> (1838) .. (1838)
<223> N 是、T、C 或 G
<220>
<221> 尚未归类的特征
<222> (1844) .. (1902)

```

32

60

90

32

60

90

<223> folP 基因的开始
 <400> 140
 aagaaccagg cgcttctaac aattctggcg acaatggtag tccaaaggct cctcgtccgg 60
 ttgatgaacc gcgtacgccg aacccgggta acaccatgtc agagcagtta ggcgacaagt 120
 aagttcccgcc atcagatgac tgtatttcta ccgaaaaccc cggggcgctgc tccgggggttt 180
 tttcttatca attcatacca gggataacat cagcgattgt gtaggctgga gctgcttcga 240
 agttcctata ctttctagag aataggaact tcggaatagg aacttcaaga tccccctta 298
 tta gaa gaa ctc gtc aag aag gcg ata gaa ggc gat gcg ctg cga atc 346
 Leu Glu Glu Leu Val Lys Lys Ala Ile Glu Gly Asp Ala Leu Arg Ile
 1 5 10 15
 ggg agc ggc gat acc gta aag cac gag gaa gcg gtc agc cca ttc gcc 394
 Gly Ser Gly Asp Thr Val Lys His Glu Glu Ala Val Ser Pro Phe Ala
 20 25 30
 gcc aag ctc ttc agc aat atc acg ggt agc caa cgc tat gtc ctg ata 442
 Ala Lys Leu Phe Ser Asn Ile Thr Gly Ser Gln Arg Tyr Val Leu Ile
 35 40 45
 gcg gtc cgc cac acc cag ccg gcc aca gtc gat gaa tcc aga aaa gcg 490
 Ala Val Arg His Thr Gln Pro Ala Thr Val Asp Glu Ser Arg Lys Ala
 50 55 60
 gcc att ttc cac cat gat att cgg caa gca ggc atc gcc atg ggt cac 538
 Ala Ile Phe His His Asp Ile Arg Gln Ala Gly Ile Ala Met Gly His
 65 70 75 80
 gac gag atc ctc gcc gtc ggg cat gcg cgc ctt gag cct ggc gaa cag 586
 Asp Glu Ile Leu Ala Val Gly His Ala Arg Leu Glu Pro Gly Glu Gln
 85 90 95
 ttc ggc tgg cgc gag ccc ctg atg ctc ttc gtc cag atc atc ctg atc 634
 Phe Gly Trp Arg Glu Pro Leu Met Leu Phe Val Gln Ile Ile Leu Ile
 100 105 110
 gac aag acc ggc ttc cat ccg agt acg tgc tgc ctc gat gcg atg ttt 682
 Asp Lys Thr Gly Phe His Pro Ser Thr Cys Ser Leu Asp Ala Met Phe
 115 120 125
 cgc ttg gtg gtc gaa tgg gca ggt agc cgg atc aag cgt atg cag ccg 730
 Arg Leu Val Val Glu Trp Ala Gly Ser Arg Ile Lys Arg Met Gln Pro
 130 135 140
 ccg cat tgc atc agc cat gat gga tac ttt ctc ggc agg agc aag gtg 778
 Pro His Cys Ile Ser His Asp Gly Tyr Phe Leu Gly Arg Ser Lys Val
 145 150 155 160
 aga tga cag gag atc ctg ccc cgg cac ttc gcc caa tag cag cca gtc 826
 Arg Gln Glu Ile Leu Pro Arg His Phe Ala Gln Gln Pro Val
 165 170 175
 cct tcc cgc ttc agt gac aac gtc gag cac agc tgc gca agg aac gcc 874
 Pro Ser Arg Phe Ser Asp Asn Val Glu His Ser Cys Ala Arg Asn Ala
 180 185 190
 cgt cgt ggc cag cca cga tag ccg cgc tgc ctc gtc ctg cag ttc att 922
 Arg Arg Gly Gln Pro Arg Pro Arg Cys Leu Val Leu Gln Phe Ile
 195 200 205
 cag ggc acc gga cag gtc ggt ctt gac aaa aag aac cgg gcg ccc ctg 970
 Gln Gly Thr Gly Gln Val Gly Leu Asp Lys Lys Asn Arg Ala Pro Leu
 210 215 220
 cgc tga cag ccg gaa cac ggc ggc atc aga gca gcc gat tgt ctg ttg 1018
 Arg Gln Pro Glu His Gly Gly Ile Arg Ala Ala Asp Cys Leu Leu
 225 230 235
 tgc cca gtc ata gcc gaa tag cct ctc cac cca agc ggc cgg aga acc 1066
 Cys Pro Val Ile Ala Glu Pro Leu His Pro Ser Gly Arg Arg Thr
 240 245 250
 tgc gtg caa tcc atc ttg ttc aat cat gcgaaacgat cctcatcctg 1113
 Cys Val Gln Ser Ile Leu Phe Asn His
 255 260
 tctcttgatc agatcttgat cccctgcgcc atcagatcct tggcggaag aaagccatcc 1173
 agttttacttt gcagggttc ccaaccttac cagagggcgc cccagctggc aattccggtt 1233
 cgcttgctgt ccataaaaacc gccaggtcta gctatcgcca tgtaagcca ctgcaagcta 1293
 cctgctttct ctttgcgctt gcgttttccc ttgtccagat agcccagtag ctgacattca 1353
 tccgggggtca gcaccgtttc tgcggactgg ctttotaagt gttccgcttc ctttagcagc 1413
 ccttgccccc tgagtgtctg cggcagcgtg agcttcaaaa gcgctctgaa gttcctatac 1473
 tttctagaga ataggaactt cgaactgcag gtcgacggat ccccggaata ttcataatgtg 1533
 gcgaaaagta ctaagtactt agaggaaata aatcatggct aaaggtatc gtgagaaaat 1593
 caagctggtt tcttctgctg gtactggtca cttctatacc actacgaaga acaaacgtac 1653
 taagccggaa aaactggaac tgaaaaaatt cgtaccagtt gttcgccagc acgtgatcta 1713
 caaaggaagc aaaatcaaat aattctcgct ttgatgtaac aaaaaacctc gctccggcgg 1773
 gggtttttgt tatctgcttg ccccatatct gactgcatct gttcattcct ggagatgcta 1833
 cttgnaaggt atgaaactct ttgcccaggg tacttcaactg gaccttagcc atcctcacgt 1893
 aatggggat 1902
 <210> 141
 <211> 161
 <212> PRT
 <213> 合成
 <400> 141
 Leu Glu Glu Leu Val Lys Lys Ala Ile Glu Gly Asp Ala Leu Arg Ile

[0457]

```

1      5      10      15
Gly Ser Gly Asp Thr Val Lys His Glu Glu Ala Val Ser Pro Phe Ala
20      25      30
Ala Lys Leu Phe Ser Asn Ile Thr Gly Ser Gln Arg Tyr Val Leu Ile
35      40      45
Ala Val Arg His Thr Gln Pro Ala Thr Val Asp Glu Ser Arg Lys Ala
50      55      60
Ala Ile Phe His His Asp Ile Arg Gln Ala Gly Ile Ala Met Gly His
65      70      75      80
Asp Glu Ile Leu Ala Val Gly His Ala Arg Leu Glu Pro Gly Glu Gln
85      90      95
Phe Gly Trp Arg Glu Pro Leu Met Leu Phe Val Gln Ile Ile Leu Ile
100      105
Asp Lys Thr Gly Phe His Pro Ser Thr Cys Ser Leu Asp Ala Met Phe
115      120      125
Arg Leu Val Val Glu Trp Ala Gly Ser Arg Ile Lys Arg Met Gln Pro
130      135      140
Pro His Cys Ile Ser His Asp Gly Tyr Phe Leu Gly Arg Ser Lys Val
145      150      155      160
Arg
<210> 142
<211> 10
<212> PRT
<213> 合成
<400> 142
Gln Glu Ile Leu Pro Arg His Phe Ala Gln
1      5      10
<210> 143
<211> 25
<212> PRT
<213> 合成
<400> 143
Gln Pro Val Pro Ser Arg Phe Ser Asp Asn Val Glu His Ser Cys Ala
1      5      10      15
Arg Asn Ala Arg Arg Gly Gln Pro Arg
20      25
<210> 144
<211> 26
<212> PRT
<213> 合成
<400> 144
Pro Arg Cys Leu Val Leu Gln Phe Ile Gln Gly Thr Gly Gln Val Gly
1      5      10      15
Leu Asp Lys Lys Asn Arg Ala Pro Leu Arg
20      25
<210> 145
<211> 20
<212> PRT
<213> 合成
<400> 145
Gln Pro Glu His Gly Gly Ile Arg Ala Ala Asp Cys Leu Leu Cys Pro
1      5      10      15
Val Ile Ala Glu
20
<210> 146
<211> 18
<212> PRT
<213> 合成
<400> 146
Pro Leu His Pro Ser Gly Arg Arg Thr Cys Val Gln Ser Ile Leu Phe
1      5      10      15
Asn His
<210> 147
<211> 1347
<212> DNA
<213> 地衣芽孢杆菌 (Bacillus licheniformis)
<220>
<221> CDS
<222> (1) .. (1347)
<223> 地衣芽孢杆菌必需基因 glmM
<400> 147
atg ggc aag tat ttt ggt aca gac ggt gta aga ggc gtg gca aac agt
Met Gly Lys Tyr Phe Gly Thr Asp Gly Val Arg Gly Val Ala Asn Ser
1      5      10      15
gaa ctt aca cct gag ctg gcc ttt aaa gtc gga cgc ttt ggc gga tat
Glu Leu Thr Pro Glu Leu Ala Phe Lys Val Gly Arg Phe Gly Gly Tyr
20      25      30

```

[0458]

gtc cta aca aaa gat aag gag cgt ccc aag gtt ctg atc ggc cgt gac	144
Val Leu Thr Lys Asp Lys Glu Arg Pro Lys Val Leu Ile Gly Arg Asp	
35 40 45	
acg cgc att tca ggg cat atg ctg gaa ggg gcc ctt gta gca ggc ctt	192
Thr Arg Ile Ser Gly His Met Leu Glu Gly Ala Leu Val Ala Gly Leu	
50 55 60	
ctt tcg ata gga gca gaa gtc atg cgt ctc ggc gtg att tcg acg ccc	240
Leu Ser Ile Gly Ala Glu Val Met Arg Leu Gly Val Ile Ser Thr Pro	
65 70 75 80	
gga gtc gca tat ttg acg aag gct atg gac gcg gag gcg ggt gtg atg	288
Gly Val Ala Tyr Leu Thr Lys Ala Met Asp Ala Glu Ala Gly Val Met	
85 90 95	
att tcc gct tcc cac aac cct gtc cag gac aac gga att aag ttt ttc	336
Ile Ser Ala Ser His Asn Pro Val Gln Asp Asn Gly Ile Lys Phe Phe	
100 105 110	
ggc ggc gac ggc ttt aag ctt tcc gat gaa cag gag ctt gaa atc gag	384
Gly Gly Asp Gly Phe Lys Leu Ser Asp Glu Gln Glu Leu Glu Ile Glu	
115 120 125	
cgt ctg atg gat cag ccg gaa gat cac ctg cca agg cct gta ggc gct	432
Arg Leu Met Asp Gln Pro Glu Asp His Leu Pro Arg Pro Val Gly Ala	
130 135 140	
gat ctg ggc atg gtg aac gac tac ttt gaa ggc gga cag aaa tat ttg	480
Asp Leu Gly Met Val Asn Asp Tyr Phe Glu Gly Gly Gln Lys Tyr Leu	
145 150 155 160	
cag ttt ctt aag cag tct gct gat gaa gac ttc aca ggc att cat gtc	528
Gln Phe Leu Lys Gln Ser Ala Asp Glu Asp Phe Thr Gly Ile His Val	
165 170 175	
gcg ctt gac tgt gcg cac ggg gcg aca tct tcg ctg gcg acc cac ctg	576
Ala Leu Asp Cys Ala His Gly Ala Thr Ser Ser Leu Ala Thr His Leu	
180 185 190	
ttt gcc gat ctt gac gca gat gta tcg acg atg ggg act tct ccg aac	624
Phe Ala Asp Leu Asp Ala Asp Val Ser Thr Met Gly Thr Ser Pro Asn	
195 200 205	
gga tta aac atc aat gac gga gtc gga tca acc cac cct gaa gcc ctg	672
Gly Leu Asn Ile Asn Asp Gly Val Gly Ser Thr His Pro Glu Ala Leu	
210 215 220	
gca gaa ttt gtg aag gaa aag gga gct gac gtc ggc atg gcg ttt gac	720
Ala Glu Phe Val Lys Glu Lys Gly Ala Asp Val Gly Met Ala Phe Asp	
225 230 235 240	
ggg gat ggc gac cgc tta atc gct gtc gac gaa aaa ggg aat atc gta	768
Gly Asp Gly Asp Arg Leu Ile Ala Val Asp Glu Lys Gly Asn Ile Val	
245 250 255	
gac ggc gat caa atc atg tac ata tgc gcg aag tac ttg aag agc gag	816
Asp Gly Asp Gln Ile Met Tyr Ile Cys Ala Lys Tyr Leu Lys Ser Glu	
260 265 270	
ggg cgc ctg acg gac aat acc gtt gtc tcc acg gtg atg agc aac ctc	864
Gly Arg Leu Thr Asp Asn Thr Val Val Ser Thr Val Met Ser Asn Leu	
275 280 285	
ggt ttc tac aaa gcg ctt gaa gcg gaa ggc att aaa agc gtt cag aca	912
Gly Phe Tyr Lys Ala Leu Glu Ala Glu Gly Ile Lys Ser Val Gln Thr	
290 295 300	
gcg gtc ggc gac cgc tat gta gtg gag gcg atg aaa aag ggc gga ttt	960
Ala Val Gly Asp Arg Tyr Val Val Glu Ala Met Lys Lys Gly Gly Phe	
305 310 315 320	
acc ctc ggc gga gaa cag tcg gga cac ctg att ttc ctt gat tac aat	1008
Thr Leu Gly Gly Glu Gln Ser Gly His Leu Ile Phe Leu Asp Tyr Asn	
325 330 335	
acg aca ggc gac gga ctg ctt tct gcg atc atg ctg atg aac acg att	1056
Thr Thr Gly Asp Gly Leu Leu Ser Ala Ile Met Leu Met Asn Thr Ile	
340 345 350	
aaa atg aca ggc aag ccg ctg tct gag ctt gcg gca gag atg cag aag	1104
Lys Met Thr Gly Lys Pro Leu Ser Glu Leu Ala Ala Glu Met Gln Lys	
355 360 365	
ttt ccg cag ctc ctc ttg aat gtg aaa gta acc gat aaa cat aaa gta	1152
Phe Pro Gln Leu Leu Leu Asn Val Lys Val Thr Asp Lys His Lys Val	
370 375 380	
acg gaa aac gaa aag gta aaa gcg gtc atc gaa gaa gtt gag aag gaa	1200
Thr Glu Asn Glu Lys Val Lys Ala Val Ile Glu Glu Val Glu Lys Glu	
385 390 395 400	
atg aac ggt gac ggc cgc atc ctt gtg cgc cca tcc ggc acc gag ccg	1248
Met Asn Gly Asp Gly Arg Ile Leu Val Arg Pro Ser Gly Thr Glu Pro	
405 410 415	
ctt gtc cgt gtc atg gct gaa gcg aaa acg aag gag ctt tgc gag aag	1296
Leu Val Arg Val Met Ala Glu Ala Lys Thr Lys Glu Leu Cys Glu Lys	
420 425 430	
tac gta ggc cgc att gca gat gtc gta aaa gct gaa atg gca gag	1344
Tyr Val Gly Arg Ile Ala Asp Val Val Lys Ala Glu Met Gly Ala Glu	
435 440 445	

taa
 <210> 148
 <211> 448
 <212> PRT
 <213> 地衣芽孢杆菌
 <400> 148

Met Gly Lys Tyr Phe Gly Thr Asp Gly Val Arg Gly Val Ala Asn Ser
 1 5 10 15
 Glu Leu Thr Pro Glu Leu Ala Phe Lys Val Gly Arg Phe Gly Tyr
 20 25 30
 Val Leu Thr Lys Asp Lys Glu Arg Pro Lys Val Leu Ile Gly Arg Asp
 35 40 45
 Thr Arg Ile Ser Gly His Met Leu Glu Gly Ala Leu Val Ala Gly Leu
 50 55 60
 Leu Ser Ile Gly Ala Glu Val Met Arg Leu Gly Val Ile Ser Thr Pro
 65 70 75 80
 Gly Val Ala Tyr Leu Thr Lys Ala Met Asp Ala Glu Ala Gly Val Met
 85 90 95
 Ile Ser Ala Ser His Asn Pro Val Gln Asp Asn Gly Ile Lys Phe Phe
 100 105 110
 Gly Gly Asp Gly Phe Lys Leu Ser Asp Glu Gln Glu Leu Glu Ile Glu
 115 120 125
 Arg Leu Met Asp Gln Pro Glu Asp His Leu Pro Arg Pro Val Gly Ala
 130 135 140
 Asp Leu Gly Met Val Asn Asp Tyr Phe Glu Gly Gln Lys Tyr Leu
 145 150 155 160
 Gln Phe Leu Lys Gln Ser Ala Asp Glu Asp Phe Thr Gly Ile His Val
 165 170 175
 Ala Leu Asp Cys Ala His Gly Ala Thr Ser Ser Leu Ala Thr His Leu
 180 185 190
 Phe Ala Asp Leu Asp Ala Asp Val Ser Thr Met Gly Thr Ser Pro Asn
 195 200 205
 Gly Leu Asn Ile Asn Asp Gly Val Gly Ser Thr His Pro Glu Ala Leu
 210 215 220
 Ala Glu Phe Val Lys Glu Lys Gly Ala Asp Val Gly Met Ala Phe Asp
 225 230 235 240
 Gly Asp Gly Asp Arg Leu Ile Ala Val Asp Glu Lys Gly Asn Ile Val
 245 250 255
 Asp Gly Asp Gln Ile Met Tyr Ile Cys Ala Lys Tyr Leu Lys Ser Glu
 260 265 270
 Gly Arg Leu Thr Asp Asn Thr Val Val Ser Thr Val Met Ser Asn Leu
 275 280 285
 Gly Phe Tyr Lys Ala Leu Glu Ala Glu Gly Ile Lys Ser Val Gln Thr
 290 295 300
 Ala Val Gly Asp Arg Tyr Val Val Glu Ala Met Lys Lys Gly Gly Phe
 305 310 315 320
 Thr Leu Gly Gly Glu Gln Ser Gly His Leu Ile Phe Leu Asp Tyr Asn
 325 330 335
 Thr Thr Gly Asp Gly Leu Leu Ser Ala Ile Met Leu Met Asn Thr Ile
 340 345 350
 Lys Met Thr Gly Lys Pro Leu Ser Glu Leu Ala Ala Glu Met Gln Lys
 355 360 365
 Phe Pro Gln Leu Leu Leu Asn Val Lys Val Thr Asp Lys His Lys Val
 370 375 380
 Thr Glu Asn Glu Lys Val Lys Ala Val Ile Glu Glu Val Glu Lys Glu
 385 390 395 400
 Met Asn Gly Asp Gly Arg Ile Leu Val Arg Pro Ser Gly Thr Glu Pro
 405 410 415
 Leu Val Arg Val Met Ala Glu Ala Lys Thr Lys Glu Leu Cys Glu Lys
 420 425 430 435
 Tyr Val Gly Arg Ile Ala Asp Val Val Lys Ala Glu Met Ala Glu
 435 440 445

<210> 149
 <211> 282
 <212> DNA
 <213> 地衣芽孢杆菌
 <220>
 <221> CDS
 <222> (1)..(282)
 <223> 地衣芽孢杆菌必需基因 ylaN
 <400> 149

ttg gcg tct gag att gta gtc gac cat cgt gaa aaa gca ctt gct ctg
 1 5 10 15
 Leu Ala Ser Glu Ile Val Val Asp His Arg Glu Lys Ala Leu Ala Leu
 tta aaa cgt gat gct gat aaa atc tta aag ctg att caa gtg caa atg
 20 25 30 35
 Leu Lys Arg Asp Ala Asp Lys Ile Leu Lys Leu Ile Gln Val Gln Met
 gac aac tta acg atg cct caa tgt cct ctt tat gaa gag gtt tta gat
 40 45 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175 180 185 190 195 200 205 210 215 220 225 230 235 240 245 250 255 260 265 270 275 280 282

	Asp	Asn	Leu	Thr	Met	Pro	Gln	Cys	Pro	Leu	Tyr	Glu	Glu	Val	Leu	Asp	
			35					40					45				
	aca	caa	atg	ttc	gga	ctc	tct	aga	gaa	att	gat	ttt	gct	gtc	cgc	ctt	192
	Thr	Gln	Met	Phe	Gly	Leu	Ser	Arg	Glu	Ile	Asp	Phe	Ala	Val	Arg	Leu	
			50					55					60				
	gga	tta	gtg	gat	gaa	gaa	gag	ggg	aaa	gaa	ctg	ctt	tac	agg	ctg	gag	240
	Gly	Leu	Val	Asp	Glu	Glu	Glu	Gly	Lys	Glu	Leu	Leu	Tyr	Arg	Leu	Glu	
			65			70					75				80		
	cgt	gaa	ttg	tct	gca	tta	cat	gat	gca	ttt	aca	aaa	aaa	taa			282
	Arg	Glu	Leu	Ser	Ala	Leu	His	Asp	Ala	Phe	Thr	Lys	Lys				
					85					90							
	<210>	150															
	<211>	93															
	<212>	PRT															
	<213>	地衣芽孢杆菌															
	<400>	150															
	Leu	Ala	Ser	Glu	Ile	Val	Val	Asp	His	Arg	Glu	Lys	Ala	Leu	Ala	Leu	
	1				5					10					15		
	Leu	Lys	Arg	Asp	Ala	Asp	Lys	Ile	Leu	Lys	Leu	Ile	Gln	Val	Gln	Met	
					20				25					30			
	Asp	Asn	Leu	Thr	Met	Pro	Gln	Cys	Pro	Leu	Tyr	Glu	Glu	Val	Leu	Asp	
			35					40					45				
	Thr	Gln	Met	Phe	Gly	Leu	Ser	Arg	Glu	Ile	Asp	Phe	Ala	Val	Arg	Leu	
			50					55				60					
	Gly	Leu	Val	Asp	Glu	Glu	Glu	Gly	Lys	Glu	Leu	Leu	Tyr	Arg	Leu	Glu	
			65			70					75				80		
	Arg	Glu	Leu	Ser	Ala	Leu	His	Asp	Ala	Phe	Thr	Lys	Lys				
					85					90							
	<210>	151															
	<211>	219															
	<212>	DNA															
	<213>	地衣芽孢杆菌															
	<220>																
	<221>	CDS															
	<222>	(1)..(219)															
	<223>	地衣芽孢杆菌必需基因 infA															
	<400>	151															
[0460]	atg	gcg	aaa	gac	gat	gta	att	gaa	gtg	gaa	ggg	act	gta	gtc	gaa	acg	48
	Met	Ala	Lys	Asp	Asp	Val	Ile	Glu	Val	Glu	Gly	Thr	Val	Val	Glu	Thr	
	1				5					10					15		
	ctg	cca	aac	gca	atg	ttc	aaa	gtt	gaa	ctt	gag	aat	ggc	cac	acg	ggt	96
	Leu	Pro	Asn	Ala	Met	Phe	Lys	Val	Glu	Leu	Glu	Asn	Gly	His	Thr	Val	
					20				25					30			
	ttg	gct	cac	gta	tct	ggg	aaa	atc	cg	atg	cac	ttc	att	cgc	att	tta	144
	Leu	Ala	His	Val	Ser	Gly	Lys	Ile	Arg	Met	His	Phe	Ile	Arg	Ile	Leu	
			35					40					45				
	cct	gga	gac	aaa	gtt	acg	gta	gaa	tta	tct	cca	tat	gac	tta	act	cgt	192
	Pro	Gly	Asp	Lys	Val	Thr	Val	Glu	Leu	Ser	Pro	Tyr	Asp	Leu	Thr	Arg	
			50					55				60					
	ggg	aga	att	acg	tac	cgt	tac	aaa	taa								219
	Gly	Arg	Ile	Thr	Tyr	Arg	Tyr	Lys									
					70												
	<210>	152															
	<211>	72															
	<212>	PRT															
	<213>	地衣芽孢杆菌															
	<400>	152															
	Met	Ala	Lys	Asp	Asp	Val	Ile	Glu	Val	Glu	Gly	Thr	Val	Val	Glu	Thr	
	1				5					10					15		
	Leu	Pro	Asn	Ala	Met	Phe	Lys	Val	Glu	Leu	Glu	Asn	Gly	His	Thr	Val	
					20				25					30			
	Leu	Ala	His	Val	Ser	Gly	Lys	Ile	Arg	Met	His	Phe	Ile	Arg	Ile	Leu	
			35					40					45				
	Pro	Gly	Asp	Lys	Val	Thr	Val	Glu	Leu	Ser	Pro	Tyr	Asp	Leu	Thr	Arg	
			50					55				60					
	Gly	Arg	Ile	Thr	Tyr	Arg	Tyr	Lys									
					70												
	<210>	153															
	<211>	870															
	<212>	DNA															
	<213>	地衣芽孢杆菌															
	<220>																
	<221>	CDS															
	<222>	(1)..(870)															
	<223>	地衣芽孢杆菌必需基因 dapA															
	<400>	153															
	atg	aac	ttc	gga	aat	atc	gca	acc	gcg	atg	gtt	acg	ccc	ttt	gac	aaa	48

	Met	Asn	Phe	Gly	Asn	Ile	Ala	Thr	Ala	Met	Val	Thr	Pro	Phe	Asp	Lys	
	1				5					10					15		
	aat	gaa	aat	atc	gat	ttt	caa	aaa	ctg	tca	aag	ctg	atc	gat	tat	tta	96
	Asn	Glu	Asn	Ile	Asp	Phe	Gln	Lys	Leu	Ser	Lys	Leu	Ile	Asp	Tyr	Leu	
				20					25					30			
	cta	aac	aat	ggg	acg	gat	tcc	ctg	gtc	ggt	gcc	gga	aca	acc	gga	gaa	144
	Leu	Asn	Asn	Gly	Thr	Asp	Ser	Leu	Val	Val	Ala	Gly	Thr	Thr	Gly	Glu	
				35				40					45				
	tcg	ccg	acg	ctt	tcc	gaa	gaa	gaa	aag	gta	gcg	ctc	att	caa	tac	tct	192
	Ser	Pro	Thr	Leu	Ser	Glu	Glu	Glu	Lys	Val	Ala	Leu	Ile	Gln	Tyr	Ser	
				50				55				60					
	gtc	aag	gaa	gcg	gca	ggc	cgt	gcc	cca	atc	atc	gcc	ggc	acg	gga	agc	240
	Val	Lys	Glu	Ala	Ala	Gly	Arg	Ala	Pro	Ile	Ile	Ala	Gly	Thr	Gly	Ser	
				65			70					75				80	
	aac	aac	acg	aag	gcg	tcg	atc	aag	ctg	acg	aaa	aaa	gct	gaa	gaa	gcc	288
	Asn	Asn	Thr	Lys	Ala	Ser	Ile	Lys	Leu	Thr	Lys	Lys	Ala	Glu	Glu	Ala	
					85					90					95		
	ggc	gct	gac	gct	gtc	atg	ctc	gtc	act	ccg	tac	tac	aac	aag	cct	tct	336
	Gly	Ala	Asp	Ala	Val	Met	Leu	Val	Thr	Pro	Tyr	Tyr	Asn	Lys	Pro	Ser	
				100					105					110			
	cag	gaa	ggc	atg	tac	cgc	cat	ttc	aga	gcg	att	gcg	gag	gaa	aca	tcg	384
	Gln	Glu	Gly	Met	Tyr	Arg	His	Phe	Arg	Ala	Ile	Ala	Glu	Glu	Thr	Ser	
				115				120					125				
	ctg	cct	gtc	atg	ctt	tat	aat	gtg	ccg	gga	aga	acg	gct	tcg	ctc		432
	Leu	Pro	Val	Met	Leu	Tyr	Asn	Val	Pro	Gly	Arg	Thr	Ala	Ala	Ser	Leu	
				130				135				140					
	gca	cca	gag	acg	aca	atc	cgc	ctt	gcg	gag	atc	ccg	aat	att	atc	gcc	480
	Ala	Pro	Glu	Thr	Thr	Ile	Arg	Leu	Ala	Glu	Ile	Pro	Asn	Ile	Ile	Ala	
						150				155					160		
	att	aaa	gaa	gcg	agc	ggc	gac	cta	gat	gcg	ata	aca	aaa	atc	gta	gcc	528
	Ile	Lys	Glu	Ala	Ser	Gly	Asp	Leu	Asp	Ala	Ile	Thr	Lys	Ile	Val	Ala	
					165				170					175			
	gaa	acc	cct	gag	gac	ttt	gca	gtt	tat	tca	gga	gat	gac	agc	ctg	aca	576
	Glu	Thr	Pro	Glu	Asp	Phe	Ala	Val	Tyr	Ser	Gly	Asp	Asp	Ser	Leu	Thr	
					180				185				190				
	ctg	ccg	gcg	ctg	tcc	gtg	gga	gcg	aga	gga	atc	gtt	tcc	gtt	gca	tca	624
	Leu	Pro	Ala	Leu	Ser	Val	Gly	Ala	Arg	Gly	Ile	Val	Ser	Val	Ala	Ser	
				195				200					205				
[0461]	cac	atc	atc	gga	ccg	gaa	atg	cag	gaa	atg	atc	aaa	cat	tat	acc	gag	672
	His	Ile	Ile	Gly	Pro	Glu	Met	Gln	Glu	Met	Ile	Lys	His	Tyr	Thr	Glu	
				210				215				220					
	ggg	aac	acg	gct	cag	gct	gtt	atc	cat	caa	aag	ctg	ctt	ccg	ctg		720
	Gly	Asn	Thr	Ala	Gln	Ala	Ala	Leu	Ile	His	Gln	Lys	Leu	Leu	Pro	Leu	
				225				230				235			240		
	atg	aaa	ggg	ctt	ttt	gcc	gca	ccg	aac	ccg	tct	cca	tta	aag	acg	gct	768
	Met	Lys	Gly	Leu	Phe	Ala	Ala	Pro	Asn	Pro	Ser	Pro	Leu	Lys	Thr	Ala	
					245					250				255			
	ctt	cag	ctg	aag	ggt	ctt	gac	gtc	ggg	tcc	gtc	cgg	ctc	cct	ttg	att	816
	Leu	Gln	Leu	Lys	Gly	Leu	Asp	Val	Gly	Ser	Val	Arg	Leu	Pro	Leu	Ile	
				260					265					270			
	ccg	cta	aat	gaa	gat	gaa	cga	ctc	cgt	tta	agc	agc	ctg	atg	aac	ggc	864
	Pro	Leu	Asn	Glu	Asp	Glu	Arg	Leu	Arg	Leu	Ser	Ser	Leu	Met	Asn	Gly	
				275				280					285				
	ctg	taa															870
	Leu																
	<210>	154															
	<211>	289															
	<212>	PRT															
	<213>	地衣芽孢杆菌															
	<400>	154															
	Met	Asn	Phe	Gly	Asn	Ile	Ala	Thr	Ala	Met	Val	Thr	Pro	Phe	Asp	Lys	
	1				5					10					15		
	Asn	Glu	Asn	Ile	Asp	Phe	Gln	Lys	Leu	Ser	Lys	Leu	Ile	Asp	Tyr	Leu	
				20					25					30			
	Leu	Asn	Asn	Gly	Thr	Asp	Ser	Leu	Val	Val	Ala	Gly	Thr	Thr	Gly	Glu	
				35				40					45				
	Ser	Pro	Thr	Leu	Ser	Glu	Glu	Glu	Lys	Val	Ala	Leu	Ile	Gln	Tyr	Ser	
				50				55				60					
	Val	Lys	Glu	Ala	Ala	Gly	Arg	Ala	Pro	Ile	Ile	Ala	Gly	Thr	Gly	Ser	
				65			70					75				80	
	Asn	Asn	Thr	Lys	Ala	Ser	Ile	Lys	Leu	Thr	Lys	Lys	Ala	Glu	Glu	Ala	
					85					90					95		
	Gly	Ala	Asp	Ala	Val	Met	Leu	Val	Thr	Pro	Tyr	Tyr	Asn	Lys	Pro	Ser	
				100					105					110			
	Gln	Glu	Gly	Met	Tyr	Arg	His	Phe	Arg	Ala	Ile	Ala	Glu	Glu	Thr	Ser	
				115				120					125				
	Leu	Pro	Val	Met	Leu	Tyr	Asn	Val	Pro	Gly	Arg	Thr	Ala	Ala	Ser	Leu	

130

[0463]	caa gaa cct gag aga aat gaa tgc ttc ttg caa cac aaa gat gac aac	1399
	Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn	
	120 125 130	
	cca aac ctc ccc cga ttg gtg aga cca gag gtt gat gtg atg tgc act	1447
	Pro Asn Leu Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr	
	135 140 145	
	gct ttt cat gac aat gaa gag aca ttt ttg aaa aaa tac tta tat gaa	1495
	Ala Phe His Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu	
	150 155 160	
	att gcc aga aga cat cct tac ttt tat gcc ccg gaa ctc ctt ttc ttt	1543
	Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe	
	170 175 180	
	gct aaa agg tat aaa gct gct ttt aca gaa tgt tgc caa gct gct gat	1591
	Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp	
	185 190 195	
	aaa gct gcc tgc ctg ttg cca aag ctc gat gaa ctt cgg gat gaa ggg	1639
	Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly	
	200 205 210	
	aag gct tcg tct gcc aaa cag aga ctc aag tgt gcc agt ctc caa aaa	1687
	Lys Ala Ser Ser Ala Lys Lys Arg Leu Lys Cys Ala Ser Leu Gln Lys	
	215 220 225	
	ttt gga gaa aga gct ttc aaa gca tgg gca gta gct cgc ctg agc cag	1735
	Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln	
	230 235 240 245	
	aga ttt ccc aaa gct gag ttt gca gaa gtt tcc aag tta gtg aca gat	1783
	Arg Phe Pro Lys Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp	
	250 255 260	
	ctt acc aaa gtc cac acg gaa tgc tgc cat gga gat ctg ctt gaa tgt	1831
	Leu Thr Lys Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys	
	265 270 275	
	gct gat gac agg gcg gac ctt gcc aag tat atc tgt gaa aat caa gat	1879
	Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp	
	280 285 290	
	tcg atc tcc agt aaa ctg aag gaa tgc tgt gaa aaa cct ctg ttg gaa	1927
	Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu	
	295 300 305	
	aaa tcc cac tgc att gcc gaa gtg gaa aat gat gag atg cct gct gac	1975
	Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp	
	310 315 320 325	
	ttg cct tca tta gct gct gat ttt gtt gaa agt aag gat gtt tgc aaa	2023
	Leu Pro Ser Leu Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys	
	330 335 340	
	aac tat gct gag gca aag gat gtc ttc ctg ggc atg ttt ttg tat gaa	2071
	Asn Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu	
	345 350 355	
	tat gca aga agg cat cct gat tac tct gtc gtg ctg ctg aga ctt	2119
	Tyr Ala Arg Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu	
	360 365 370	
	gcc aag aca tat gaa acc act cta gag aag tgc tgt gcc gct gca gat	2167
	Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp	
	375 380 385	
	cct cat gaa tgc tat gcc aaa gtg ttc gat gaa ttt aaa cct ctt gtg	2215
	Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val	
	390 395 400 405	
	gaa gag cct cag aat tta atc aaa caa aat tgt gag ctt ttt gag cag	2263
	Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln	
	410 415 420	
	ctt gga gag tac aaa ttc cag aat gcg cta tta gtt cgt tac acc aag	2311
	Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys	
	425 430 435	
	aaa gta ccc caa gtg tca act cca act ctt gta gag gtc tca aga aac	2359
	Lys Val Pro Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn	
	440 445 450	
	cta gga aaa gtg ggc agc aaa tgt tgt aaa cat cct gaa gca aaa aga	2407
	Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg	
	455 460 465	
	atg ccc tgt gca gaa gac tat cta tcc gtg gtc aac cag tta tgt	2455
	Met Pro Cys Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys	
	470 475 480 485	
	gtg ttg cat gag aaa acg cca gta agt gac aga gtc acc aaa tgc tgc	2503
	Val Leu His Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys	
	490 495 500	
	aca gaa tcc ttg gtg aac agg cga cca tgc ttt tca gct ctg gaa gtc	2551
	Thr Glu Ser Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val	
	505 510 515	
	gat gaa aca tac gtt ccc aaa gag ttt aat gct gaa aca acc ttc	2599
	Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe	
	520 525 530	

	cat	gca	gat	ata	tgc	aca	ctt	tct	gag	aag	gag	aga	caa	atc	aag	aaa	2647
	His	Ala	Asp	Ile	Cys	Thr	Leu	Ser	Glu	Lys	Glu	Arg	Gln	Ile	Lys	Lys	
		535					540					545					
	caa	act	gca	ctt	gtt	gag	ctc	gtg	aaa	cac	aag	ccc	aag	gca	aca	aaa	2695
	Gln	Thr	Ala	Leu	Val	Glu	Leu	Val	Lys	His	Lys	Pro	Lys	Ala	Thr	Lys	
		550				555					560					565	
	gag	caa	ctg	aaa	gct	gtt	atg	gat	gat	ttc	gca	gct	ttt	gta	gag	aag	2743
	Glu	Gln	Leu	Lys	Ala	Val	Met	Asp	Asp	Phe	Ala	Ala	Phe	Val	Glu	Lys	
					570					575					580		
	tgc	tgc	aag	gct	gac	gat	aag	gag	acc	ttt	gcc	gag	gag	ggg	aaa		2791
	Cys	Cys	Lys	Ala	Asp	Asp	Lys	Glu	Thr	Cys	Phe	Ala	Glu	Glu	Gly	Lys	
				585					590					595			
	aaa	ctt	gtt	gct	gca	agt	caa	gct	gcc	tta	ggc	tta	taa				2830
	Lys	Leu	Val	Ala	Ala	Ser	Gln	Ala	Ala	Leu	Gly	Leu					
			600					605									
	<210>	156															
	<211>	609															
	<212>	PRT															
	<213>	合成															
	<400>	156															
	Met	Lys	Trp	Val	Thr	Phe	Ile	Ser	Leu	Leu	Phe	Leu	Phe	Ser	Ser	Ala	
	1			5						10					15		
	Tyr	Ser	Arg	Gly	Val	Phe	Arg	Arg	Asp	Ala	His	Lys	Ser	Glu	Val	Ala	
				20					25					30			
	His	Arg	Phe	Lys	Asp	Leu	Gly	Glu	Glu	Asn	Phe	Lys	Ala	Leu	Val	Leu	
			35					40					45				
	Ile	Ala	Phe	Ala	Gln	Tyr	Leu	Gln	Gln	Cys	Pro	Phe	Glu	Asp	His	Val	
			50				55					60					
	Lys	Leu	Val	Asn	Glu	Val	Thr	Glu	Phe	Ala	Lys	Thr	Cys	Val	Ala	Asp	
			65			70					75				80		
	Glu	Ser	Ala	Glu	Asn	Cys	Asp	Lys	Ser	Leu	His	Thr	Leu	Phe	Gly	Asp	
				85						90					95		
	Lys	Leu	Cys	Thr	Val	Ala	Thr	Leu	Arg	Glu	Thr	Tyr	Gly	Glu	Met	Ala	
				100					105					110			
	Asp	Cys	Cys	Ala	Lys	Gln	Glu	Pro	Glu	Arg	Asn	Glu	Cys	Phe	Leu	Gln	
			115					120					125				
	His	Lys	Asp	Asp	Asn	Pro	Asn	Leu	Pro	Arg	Leu	Val	Arg	Pro	Glu	Val	
		130					135					140					
[0464]	Asp	Val	Met	Cys	Thr	Ala	Phe	His	Asp	Asn	Glu	Glu	Thr	Phe	Leu	Lys	
		145				150					155				160		
	Lys	Tyr	Leu	Tyr	Glu	Ile	Ala	Arg	Arg	His	Pro	Tyr	Phe	Tyr	Ala	Pro	
				165						170					175		
	Glu	Leu	Leu	Phe	Ala	Lys	Arg	Tyr	Lys	Ala	Ala	Phe	Thr	Glu	Cys		
				180				185					190				
	Cys	Gln	Ala	Ala	Asp	Lys	Ala	Ala	Cys	Leu	Leu	Pro	Lys	Leu	Asp	Glu	
			195					200					205				
	Leu	Arg	Asp	Glu	Gly	Lys	Ala	Ser	Ser	Ala	Lys	Gln	Arg	Leu	Lys	Cys	
		210					215					220					
	Ala	Ser	Leu	Gln	Lys	Phe	Gly	Glu	Arg	Ala	Phe	Lys	Ala	Trp	Ala	Val	
		225				230					235				240		
	Ala	Arg	Leu	Ser	Gln	Arg	Phe	Pro	Lys	Ala	Glu	Phe	Ala	Glu	Val	Ser	
				245						250					255		
	Lys	Leu	Val	Thr	Asp	Leu	Thr	Lys	Val	His	Thr	Glu	Cys	Cys	His	Gly	
				260					265					270			
	Asp	Leu	Leu	Glu	Cys	Ala	Asp	Asp	Arg	Ala	Asp	Leu	Ala	Lys	Tyr	Ile	
		275						280					285				
	Cys	Glu	Asn	Gln	Asp	Ser	Ile	Ser	Ser	Lys	Leu	Lys	Glu	Cys	Cys	Glu	
		290					295					300					
	Lys	Pro	Leu	Leu	Glu	Lys	Ser	His	Cys	Ile	Ala	Glu	Val	Glu	Asn	Asp	
		305				310					315				320		
	Glu	Met	Pro	Ala	Asp	Leu	Pro	Ser	Leu	Ala	Ala	Asp	Phe	Val	Glu	Ser	
				325						330				335			
	Lys	Asp	Val	Cys	Lys	Asn	Tyr	Ala	Glu	Ala	Lys	Asp	Val	Phe	Leu	Gly	
				340					345					350			
	Met	Phe	Leu	Tyr	Glu	Tyr	Ala	Arg	Arg	His	Pro	Asp	Tyr	Ser	Val	Val	
		355						360					365				
	Leu	Leu	Leu	Arg	Leu	Ala	Lys	Thr	Tyr	Glu	Thr	Thr	Leu	Glu	Lys	Cys	
		370					375					380					
	Cys	Ala	Ala	Ala	Asp	Pro	His	Glu	Cys	Tyr	Ala	Lys	Val	Phe	Asp	Glu	
		385				390					395				400		
	Phe	Lys	Pro	Leu	Val	Glu	Glu	Pro	Gln	Asn	Leu	Ile	Lys	Gln	Asn	Cys	
				405						410				415			
	Glu	Leu	Phe	Glu	Gln	Leu	Gly	Glu	Tyr	Lys	Phe	Gln	Asn	Ala	Leu	Leu	
				420				425						430			
	Val	Arg	Tyr	Thr	Lys	Lys	Val	Pro	Gln	Val	Ser	Thr	Pro	Thr	Leu	Val	
			435					440					445				
	Glu	Val	Ser	Arg	Asn	Leu	Gly	Lys	Val	Gly	Ser	Lys	Cys	Cys	Lys	His	
		450					455					460					

Pro Glu Ala Lys Arg Met Pro Cys Ala Glu Asp Tyr Leu Ser Val Val
 465 470 475 480
 Leu Asn Gln Leu Cys Val Leu His Glu Lys Thr Pro Val Ser Asp Arg
 485 490 495
 Val Thr Lys Cys Cys Thr Glu Ser Leu Val Asn Arg Arg Pro Cys Phe
 500 505 510
 Ser Ala Leu Glu Val Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala
 515 520 525
 Glu Thr Phe Thr Phe His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu
 530 535 540
 Arg Gln Ile Lys Lys Gln Thr Ala Leu Val Glu Leu Val Lys His Lys
 545 550 555 560
 Pro Lys Ala Thr Lys Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala
 565 570 575
 Ala Phe Val Glu Lys Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe
 580 585 590
 Ala Glu Glu Gly Lys Lys Leu Val Ala Ala Ser Gln Ala Leu Gly
 595 600 605

Leu

<210> 157

<211> 3505

<212> DNA

<213> 合成

<220>

<221> 尚未归类的特征

<222> (1) .. (3505)

<223> 构建体 KU70 (上) -pPDI1-kanMX-终止子-KU70

<220>

<221> 尚未归类的特征

<222> (1) .. (1002)

<223> 上游区域

<220>

<221> 启动子

<222> (1002) .. (1542)

<223> 启动子 pPDI1

<220>

<221> CDS

<222> (1543) .. (2352)

<223> kanMX

<220>

<221> 终止子

<222> (2353) .. (2555)

<223> 终止子

<220>

<221> 尚未归类的特征

<222> (2556) .. (3505)

<223> 下游区域

<400> 157

gctcttcaga	agcctaactt	taatggggta	aaatatcagc	agtaggctga	atatagagca	60
tggcaacggt	tacgtcaaca	gagtttaata	gtgaaaacta	cgatagggtc	aggcctgtat	120
accagacga	gttgatcag	caattagtgg	attatcatgt	tgggtgctaaa	ggattgtgtg	180
tcgatgttgg	atgtggctct	ggccaggcaa	cgttcaccct	gaaaaagtac	tttgacaagg	240
tcataggatc	tgatatctca	gaaaatcagt	tggcagtggc	taggaaaaga	cagcctgctg	300
gaattgaatt	taggttagga	acgggcgagg	atttttcttg	gctaaccggag	acgccagatg	360
taatcacgcg	cgctgaatgc	ctgcactggg	tagaccaca	gaaattcggt	gcaaatgtcg	420
ccaattctct	aagggatcat	ggcaccttga	gttactggct	ttatacagaa	ccaatttttc	480
agaacgaaag	ggccaatcag	gtttacaaca	aattcaccta	cgggtctgac	tacctgggcc	540
catattggga	tcctggccgt	acacatttca	gaaatcattt	aaaggagtgg	aatcacatct	600
tactggatag	cgagcttttt	gacgaagtga	aaattttctaa	ttttaaacia	gaggaagggg	660
tcaaaaacgg	agatatctta	tacttggaag	aagagatgac	aatcagtgat	ttcatcaatt	720
ttgtatctag	ttggccttct	gtgttttcgt	ggaagcagca	acgaggaaaag	gagggtatcc	780
tagatgattt	ttacaacgaa	ctgaacgact	gctttgaggg	gggtaacatg	aaagtaatat	840
ggaactccgt	cctagtattt	gccaggagga	agcaaagggt	tgtataggct	ttagtactta	900
tagaggaaac	ggggttacgt	gcaagcgcgc	atgcctgagc	tttgaggggg	gggactttca	960
catctcttct	tctcacactt	agccctaaca	cagagatcaa	ctactcaagg	cagcaccaaa	1020
agacgccaga	aatttatatg	tccaggaaaa	aacaactctc	tttactacaa	aatttgaaca	1080
ggtgtcctgc	ctgcggaaac	tacaaaacgat	cacacaccct	ctgcatgcat	tgcgtaggac	1140
aaatcaggag	acatttgaac	gactotgttc	ctcaacagga	ggcatttctg	gaagagtgtg	1200
ttaatccttt	ggatgagaag	attcttttatc	caggaaagaa	agaactgccc	gatgaacgaa	1260
ctttacgtaa	gaaggagtgg	ctgaagagaa	gaccccgaa	actccctgtt	gaatagaaca	1320
cgaacactgt	aaatagaata	aaagaaaact	tggatagtag	aacttcaatg	tagtggtttct	1380
attgtctttac	gcggtctttt	agatttgcatt	ccccagaatg	gaatcgtcca	tctttctcaa	1440
cccactcaaa	gataatctac	cagacatacc	tacgccctcc	atoccagcac	cacgtcgcca	1500
tcacccttaa	aacttcaata	attgaacacg	tactgatttc	ca atg ggt	aag gaa	1554
				Met Gly	Lys Glu	
				1		

[0465]

[0466]

```

aag act cac gtt tgc agg ccg cga tta aat tcc aac atg gat gct gat 1602
Lys Thr His Val Ser Arg Pro Arg Leu Asn Ser Asn Met Asp Ala Asp
5 10 15 20
tta tat ggg tat aaa tgg gct cgc gat aat gtc ggg caa tca ggt gcg 1650
Leu Tyr Gly Tyr Lys Trp Ala Arg Asp Asn Val Gly Gln Ser Gly Ala
25 30 35
aca atc tat cga ttg tat ggg aag ccc gat gcg cca gag ttg ttt ctg 1698
Thr Ile Tyr Arg Leu Tyr Gly Lys Pro Asp Ala Pro Glu Leu Phe Leu
40 45 50
aaa cat ggc aaa ggt agc gtt gcc aat gat gtt aca gat gag atg gtc 1746
Lys His Gly Lys Gly Ser Val Ala Asn Asp Val Thr Asp Glu Met Val
55 60 65
aga cta aac tgg ctg acg gaa ttt atg cct ctt ccg acc atc aag cat 1794
Arg Leu Asn Trp Leu Thr Glu Phe Met Pro Leu Pro Thr Ile Lys His
70 75 80
ttt atc cgt act cct gat gat gca tgg tta ctc acc act gcg atc ccc 1842
Phe Ile Arg Thr Pro Asp Ala Trp Leu Leu Thr Thr Ala Ile Pro
85 90 95 100
ggc aaa aca gca ttc cag gta tta gaa gaa tat cct gat tca ggt gaa 1890
Gly Lys Thr Ala Phe Gln Val Leu Glu Glu Tyr Pro Asp Ser Gly Glu
105 110 115
aat att gtt gat gcg ctg gca gtg ttc ctg cgc ccg ttg cat tgc att 1938
Asn Ile Val Ala Leu Ala Val Phe Leu Arg Arg Leu His Ser Ile
120 125 130
cct gtt tgt aat tgt cct ttt aac agc gat cgc gta ttt cgt ctc gct 1986
Pro Val Cys Asn Cys Pro Phe Asn Ser Asp Arg Val Phe Arg Leu Ala
135 140 145
cag gcg caa tca cga atg aat aac ggt ttg gtt gat gcg agt gat ttt 2034
Gln Ala Gln Ser Arg Met Asn Asn Gly Leu Val Asp Ala Ser Asp Phe
150 155 160
gat gac gag cgt aat ggc tgg cct gtt gaa caa gtc tgg aaa gaa atg 2082
Asp Asp Glu Arg Asn Gly Trp Pro Val Glu Gln Val Trp Lys Glu Met
165 170 175 180
cat aag ctt ttg cca ttc tca ccg gat tca gtc gtc act cat ggt gat 2130
His Lys Leu Leu Pro Phe Ser Pro Asp Ser Val Val Thr His Gly Asp
185 190 195
ttc tca ctt gat aac ctt att ttt gac gag ggg aaa tta ata tgt 2178
Phe Ser Leu Asp Asn Leu Ile Phe Asp Glu Gly Lys Leu Ile Gly Cys
200 205 210
att gat gtt gga cga gtc gga atc gca gac cga tac cag gat ctt gcc 2226
Ile Asp Val Gly Arg Val Gly Ile Ala Asp Arg Tyr Gln Asp Leu Ala
215 220 225
atc cta tgg aac tgc ctc ggt gag ttt tct cct tca tta cag aaa ccg 2274
Ile Leu Trp Asn Cys Leu Gly Glu Phe Ser Pro Ser Leu Gln Lys Arg
230 235 240
ctt ttt caa aaa tat ggt att gat aat cct gat atg aat aaa ttg cag 2322
Leu Phe Gln Lys Tyr Gly Ile Asp Asn Pro Asp Met Asn Lys Leu Gln
245 250 255 260
ttt cat ttg atg ctc gat gag ttt ttc taa tcagtactga caataaaaaag 2372
Phe His Leu Met Leu Asp Glu Phe Phe
265
attcttgttt tcaagaactt gtcatttgta tagttttttt atattgtagt tgttctattt 2432
taatcaaatg ttagcgtgat ttatatattt tttgcctcg acatcatctg cccagatgcg 2492
aagttaagtg cgcagaaagt aatatcatgc gtcaatcgta tgtgaatgct ggtcgcgtata 2552
ctgttcatct tagagaatgt cagtggaggt atttctgagc taatcattac cttgcccggt 2612
acaggaatag ggtgttacct tattaattac gacggtggtc aaaacgacga aatttaccoc 2672
atttttgagt tacaagacct gaatttggaa atgatgaaac aattgtacca agtcttggag 2732
gaccatgtaa gtgggcttaa tctctcgcag aagcaattcc caattgaaca cagtaaacccg 2792
ttatcagcca ctcgtttctt tcaacttaagg tctctttttt acatggcgaa gactcataag 2852
cgtactggaa gacattacaa cttgaaaaag attttcttgt tcactaataa cgataaacct 2912
tacaatggaa actctcagct gagagttccc ttgaagaaaa ccctggctga ttacaatgac 2972
gtagacatta ctttgattcc gtttcttctg aacaagcctt caggtgtcaa gtttgacaag 3032
acggaatact cagaaatttt gttctatgat aaagatgctt gttcgatgtc aattgaggag 3092
atccgccaac gaatttctag acataaggag atcaagcggg tttacttcac ctgtcctttg 3152
aaaatcgcaa ataacttgtg catttctgtg aaaggttatt ctatgtttta tcatgaaact 3212
ccaaggaaga tcaaatttgt cgtcaatgag ggttcaactt tcaaagatgt ggagacaaaa 3272
tctcagtttg tcgatccaac atccggaaaa gagttttcca gtgaacagct gatcaaagca 3332
tatectctag gtgcgatgc ttacattcct ttaaaactcag agcaagtcaa aacaataaat 3392
cgatttaatg atatcatcaa tatccctctt ttggaaattc taggtttcag ggatatact 3452
aattggttgc cacagtatca gtttggcaaa gcacgcgtttt tatcccttaa taa 3505
<210> 158
<211> 269
<212> PRT
<213> 合成
<400> 158
Met Gly Lys Glu Lys Thr His Val Ser Arg Pro Arg Leu Asn Ser Asn
1 5 10 15
Met Asp Ala Asp Leu Tyr Gly Tyr Lys Trp Ala Arg Asp Asn Val Gly

```

20 25 30
 Gln Ser Gly Ala Thr Ile Tyr Arg Leu Tyr Gly Lys Pro Asp Ala Pro
 35 40 45
 Glu Leu Phe Leu Lys His Gly Lys Gly Ser Val Ala Asn Asp Val Thr
 50 55 60
 Asp Glu Met Val Arg Leu Asn Trp Leu Thr Glu Phe Met Pro Leu Pro
 65 70 75 80
 Thr Ile Lys His Phe Ile Arg Thr Pro Asp Asp Ala Trp Leu Leu Thr
 85 90 95
 Thr Ala Ile Pro Gly Lys Thr Ala Phe Gln Val Leu Glu Glu Tyr Pro
 100 105 110
 Asp Ser Gly Glu Asn Ile Val Asp Ala Leu Ala Val Phe Leu Arg Arg
 115 120 125
 Leu His Ser Ile Pro Val Cys Asn Cys Pro Phe Asn Ser Asp Arg Val
 130 135 140
 Phe Arg Leu Ala Gln Ala Gln Ser Arg Met Asn Asn Gly Leu Val Asp
 145 150 155 160
 Ala Ser Asp Phe Asp Asp Glu Arg Asn Gly Trp Pro Val Glu Gln Val
 165 170 175
 Trp Lys Glu Met His Lys Leu Leu Pro Phe Ser Pro Asp Ser Val Val
 180 185 190
 Thr His Gly Asp Phe Ser Leu Asp Asn Leu Ile Phe Asp Glu Gly Lys
 195 200 205
 Leu Ile Gly Cys Ile Asp Val Gly Arg Val Gly Ile Ala Asp Arg Tyr
 210 215 220
 Gln Asp Leu Ala Ile Leu Trp Asn Cys Leu Gly Glu Phe Ser Pro Ser
 225 230 235 240
 Leu Gln Lys Arg Leu Phe Gln Lys Tyr Gly Ile Asp Asn Pro Asp Met
 245 250 255
 Asn Lys Leu Gln Phe His Leu Met Leu Asp Glu Phe Phe
 260 265

<210> 159

<211> 3533

<212> DNA

<213> 合成

<220>

<221> 尚未归类的特征

<222> (1) .. (3533)

[0467]

<223> KU70 (上) -pFPR2-kanMX-终止子-KU70 构建体

<220>

<221> 尚未归类的特征

<222> (1) .. (1002)

<223> KU70 上游区域

<220>

<221> 启动子

<222> (1003) .. (1570)

<223> 启动子 pFPR2

<220>

<221> CDS

<222> (1571) .. (2380)

<223> kanMX

<220>

<221> 终止子

<222> (2381) .. (2583)

<223> 终止子

<220>

<221> 尚未归类的特征

<222> (2584) .. (3533)

<223> KU70 下游区域

<400> 159

gctcttcaga	agcctaactt	taatggggta	aaatatcagc	agtaggctga	atatagagca	60
tggcaacgtt	tacgtcaaca	gagtttaata	gtgaaaacta	cgatagggtc	aggcctgtat	120
accagacga	gttgatcag	caattagtgg	attatcatgt	tgggtgctaaa	ggattgtgtg	180
tcgatgttgg	atgtggctct	ggccaggcaa	cgttcaccct	gaaaaagtac	tttgacaagg	240
tcataggatc	tgatatctca	gaaaatcagt	tggcagtggc	taggaaaaga	cagcctgctg	300
gaattgaatt	taggttagga	acgggcgagg	atttttcttg	gctaaccggag	acgccagatg	360
taatcaccgc	cgctgaatgc	ctgcactggg	tagaccacaa	gaaattcgtt	gcaaattgtcg	420
ccaattctct	aagggatcat	ggcaccttga	gttactggct	ttatacagaa	ccaatttttc	480
agaacgaaag	ggccaatcag	gtttacaaca	aattcaccta	cgggtctgac	tacctggggc	540
catattggga	tcctggccgt	acacatttca	gaaatcatat	aaaggagttg	aatcacatct	600
tactggatag	cgagcttttt	gacgaagtga	aaattttctaa	ttttaaacaa	gaggaagggg	660
tcaaaaacgg	agatatctta	tacttggaag	aagagatgac	aatcagtgat	ttcatcaatt	720
ttgtatctag	ttggccttct	gtgttttctg	ggaagcagca	acgaggaaaag	gaggggtatcc	780
tagatgattt	ttacaacgaa	ctgaacgact	gctttgaggg	gggtaacatg	aaagtaatat	840
ggaactccgt	cctagtattt	gccaggagga	agcaaagggt	tgtataggct	ttagtactta	900
tagaggaaac	ggggttacgt	gcaagcgcgc	atgcctgagc	tttgaggggg	gggactttca	960

[0468]

```

catctcttct tctcacactt agccctaaca cagagatcaa ctacaagcgg caattcctac 1020
agaaagtgtt tgcaaagttc ctaatctcgg tttgtcacca agtcactttt ctacgaaccc 1080
tgagtttccct cactgaggaat caagttttga aatgcccgat ttgtcgcagt ttgatagttt 1140
tgatgaattt tttaatggaa tctttccaga attttaaaga gtcaagaata tttactttgt 1200
gagcgaacga acatgaaaac cgaaagaata gttgatgatg atgttgaagc aatgcaagta 1260
ccatataaat taggagaagt cctttgtgta ttggcgattc cttttgatgc aagtgagcgt 1320
aaatggttaa caggctgata catttactca agctagttag gaaagccatt gcgaaatcag 1380
tttatcatag tagtggacgt cacaatggaa ctcatctctg cgaatcgttg tagaataact 1440
gtaggttaaat atgcaattct aacttctggt ctaaaatgtt cagtgcacaa gcttatttct 1500
gtagtcgcga agaattgctt ctgaggagac ggagcgagcc atgtattctc gccaaactga 1560
caagaacgcc atg ggt aag gaa aag act cac gtt tcg agg ccg cga tta 1609
Met Gly Lys Glu Lys Thr His Val Ser Arg Pro Arg Leu
1 5 10
aat tcc aac atg gat gct gat tta tat ggg tat aaa tgg gct cgc gat 1657
Asn Ser Asn Met Asp Ala Asp Leu Tyr Gly Tyr Lys Trp Ala Arg Asp
15 20 25
aat gtc ggg caa tca ggt gcg aca atc tat cga ttg tat ggg aag ccc 1705
Asn Val Gly Gln Ser Gly Ala Thr Ile Tyr Arg Leu Tyr Gly Lys Pro
30 35 40 45
gat gcg cca gag ttg ttt ctg aaa cat ggc aaa ggt agc gtt gcc aat 1753
Asp Ala Pro Glu Leu Phe Leu Lys His Gly Lys Gly Ser Val Ala Asn
50 55 60
gat gtt aca gat gag atg gtc aga cta aac tgg ctg acg gaa ttt atg 1801
Asp Val Thr Asp Glu Met Val Arg Leu Asn Trp Leu Thr Glu Phe Met
65 70 75
cct ctt ccg acc atc aag cat ttt atc cgt act cct gat gat gca tgg 1849
Pro Leu Pro Thr Ile Lys His Phe Ile Arg Thr Pro Asp Asp Ala Trp
80 85 90
tta ctc acc act gcg atc ccc ggc aaa aca gca ttc cag gta tta gaa 1897
Leu Leu Thr Thr Ala Ile Pro Gly Lys Thr Ala Phe Gln Val Leu Glu
95 100 105
gaa tat cct gat tca ggt gaa aat att gtt gat gcg ctg gca gtg ttc 1945
Glu Tyr Pro Asp Ser Gly Glu Asn Ile Val Asp Ala Leu Ala Val Phe
110 115 120 125
ctg cgc cgg ttg cat tcg att cct gtt tgt aat tgt cct ttt aac agc 1993
Leu Arg Arg Leu His Ser Ile Pro Val Cys Asn Cys Pro Phe Asn Ser
130 135 140
gat cgc gta ttt cgt ctc gct cag gcg caa tca cga atg aat aac ggt 2041
Asp Arg Val Phe Arg Leu Ala Gln Ala Gln Ser Arg Met Asn Asn Gly
145 150 155
ttg gtt gat gcg agt gat ttt gat gac gag cgt aat ggc tgg cct gtt 2089
Leu Val Asp Ala Ser Asp Phe Asp Asp Glu Arg Asn Gly Trp Pro Val
160 165 170
gaa caa gtc tgg aaa gaa atg cat aag ctt ttg cca ttc tca ccg gat 2137
Glu Gln Val Trp Lys Glu Met His Lys Leu Leu Pro Phe Ser Pro Asp
175 180 185
tca gtc gtc act cat ggt gat ttc tca ctt gat aac ctt att ttt gac 2185
Ser Val Val Thr His Gly Asp Phe Ser Leu Asp Asn Leu Ile Phe Asp
190 195 200 205
gag ggg aaa tta ata ggt tgt att gat gtt gga cga gtc gga atc gca 2233
Glu Gly Lys Leu Ile Gly Cys Ile Asp Val Gly Arg Val Gly Ile Ala
210 215 220
gac cga tac cag gat ctt gcc atc cta tgg aac tgc ctc ggt gag ttt 2281
Asp Arg Tyr Gln Asp Leu Ala Ile Leu Trp Asn Cys Leu Gly Glu Phe
225 230 235
tct cct tca tta cag aaa cgg ctt ttt caa aaa tat ggt att gat aat 2329
Ser Pro Ser Leu Gln Lys Arg Leu Phe Gln Lys Tyr Gly Ile Asp Asn
240 245 250
cct gat atg aat aaa ttg cag ttt cat ttg atg ctc gat gag ttt ttc 2377
Pro Asp Met Asn Lys Leu Gln Phe His Leu Met Leu Asp Glu Phe Phe
255 260 265
taa tcagtactga caataaaaag attcttgttt tcaagaactt gtcatttga 2430
tagttttttt atattgtagt tgttctatatt taatcaaatg ttagcgtgat ttatatattt 2490
tttcgcctcg acatcatctg cccagatgcg aagttaagtg cgcagaaagt aatatcatgc 2550
gtcaatcgta tgtgaatgct ggtcgcgtata ctgttcatct tagagaatgt cagttaggtt 2610
atttctgagc taatcattac cttgcccggt acaggaatag ggtgttacct tattaattac 2670
gacggtggtc aaaacgacga aattttaccc atttttgagt tacaagacct gaatttggaa 2730
atgatgaaac aattgtacca agtcttgag gaccatgtaa gtgggcttaa tctctcagag 2790
aagcaattcc caattgaaca cagtaaaccg ttatcagcca ctctgttctt tctcttaagg 2850
tctctttttt acatggcgaa gactcataag cgtactggaa gacattacaa cttgaaaaag 2910
attttcttgt tcaactataa cgataaacct tacaatggaa actctcagct gagagttccc 2970
ttgaagaaaa ccctggctga ttacaatgac gttagacatta ctttgattcc gtttcttctg 3030
aacaagcctt caggtgtcaa gtttgacaag acggaatact cagaaatttt gttctatgat 3090
aaagatgctt gttcgatgtc aattgaggag atccgccaac gaatttctag acataaggag 3150
atcaagcggg tttacttcac ctgtcctttg aaaatcgcaa ataacttggt catctctgtg 3210
aaagggttat ctatgtttta tcatgaaact ccaagggaaga tcaaatgtgt cgtcaatgag 3270
ggttcaactt tcaaagatgt ggagacaaaa tctcagtttg tcgatccaac atccggaaaa 3330
gagttttcca gtgaacagct gatcaaagca tatctcttag gtgccgatgc ttacattcct 3390

```

```

ttaaactcag agcaagtcaa aacaataaat cgatttaatg atatcatcaa tatcccctct 3450
ttggaaattc taggtttcag ggatataatc aattgggtgc cacagtatca gtttggcaaa 3510
gcacgttttt tatcccctaa taa 3533
<210> 160
<211> 269
<212> PRT
<213> 合成
<400> 160
Met Gly Lys Glu Lys Thr His Val Ser Arg Pro Arg Leu Asn Ser Asn
1 5 10 15
Met Asp Ala Asp Leu Tyr Gly Tyr Lys Trp Ala Arg Asp Asn Val Gly
20 25 30
Gln Ser Gly Ala Thr Ile Tyr Arg Leu Tyr Gly Lys Pro Asp Ala Pro
35 40 45
Glu Leu Phe Leu Lys His Gly Lys Gly Ser Val Ala Asn Asp Val Thr
50 55 60
Asp Glu Met Val Arg Leu Asn Trp Leu Thr Glu Phe Met Pro Leu Pro
65 70 75 80
Thr Ile Lys His Phe Ile Arg Thr Pro Asp Asp Ala Trp Leu Leu Thr
85 90 95
Thr Ala Ile Pro Gly Lys Thr Ala Phe Gln Val Leu Glu Glu Tyr Pro
100 105 110
Asp Ser Gly Glu Asn Ile Val Asp Ala Leu Ala Val Phe Leu Arg Arg
115 120 125
Leu His Ser Ile Pro Val Cys Asn Cys Pro Phe Asn Ser Asp Arg Val
130 135 140
Phe Arg Leu Ala Gln Ala Gln Ser Arg Met Asn Asn Gly Leu Val Asp
145 150 155 160
Ala Ser Asp Phe Asp Asp Glu Arg Asn Gly Trp Pro Val Glu Gln Val
165 170 175
Trp Lys Glu Met His Lys Leu Leu Pro Phe Ser Pro Asp Ser Val Val
180 185 190
Thr His Gly Asp Phe Ser Leu Asp Asn Leu Ile Phe Asp Glu Gly Lys
195 200 205
Leu Ile Gly Cys Ile Asp Val Gly Arg Val Gly Ile Ala Asp Arg Tyr
210 215 220
Gln Asp Leu Ala Ile Leu Trp Asn Cys Leu Gly Glu Phe Ser Pro Ser
225 230 235 240
Leu Gln Lys Arg Leu Phe Gln Lys Tyr Gly Ile Asp Asn Pro Asp Met
245 250 255
Asn Lys Leu Gln Phe His Leu Met Leu Asp Glu Phe Phe
260 265
<210> 161
<211> 3632
<212> DNA
<213> 合成
<220>
<221> 尚未归类的特征
<222> (1) .. (3632)
<223> 构建体 KU70 (上) -pRPL3-kanMX-终止子-KU70 (下)
<220>
<221> 尚未归类的特征
<222> (1) .. (1002)
<223> KU70 上游区域
<220>
<221> 启动子
<222> (1003) .. (1669)
<223> pRPL3 启动子
<220>
<221> CDS
<222> (1670) .. (2479)
<223> kanMX
<220>
<221> 终止子
<222> (2480) .. (2682)
<223> 终止子
<220>
<221> 尚未归类的特征
<222> (2480) .. (2682)
<223> KU70 下游区域
<400> 161
gctcttcaga agcctaactt taatggggta aaatatcagc agtaggctga atatagagca 60
tggaacggtt tacgtcaaca gagtttaata gtgaaaacta cgatagggtc aggccgtgat 120
accagacga gttgtatcag caattagtgg attatcatgt tgggtgctaaa ggattgtgtg 180
tcgatgttgg atgtggctct ggccaggcaa cgttcaccct gaaaaagtac tttgacaagg 240
tcataggatc tgatatctca gaaaatcagt tggcagtggc taggaaaaga cagcctgctg 300

```


gaattgaatt	taggttagga	acgggcgagg	atttttcttg	gctaacggag	acgccagatg	360
taatcaccgc	cgctgaatgc	ctgcactggg	tagaccacac	gaaattcggt	gcaaattgtcg	420
ccaattctct	aagggatcat	ggcaccttga	gttactggct	ttatacagaa	ccaatttttc	480
agaacgaaa	ggccaatcag	gtttacaaca	aattcaccta	cggtgtctgac	tacctgggcc	540
catattggga	tcttgccgt	acacatttca	gaaatcattt	aaaggagttg	aatcacatct	600
tactggatag	cgagcttttt	gacgaagtga	aaatttctaa	ttttaaacaa	gaggaagggg	660
tcaaaaacgg	agatatctta	tacttggaaa	aagagatgac	aatcagtgat	ttcatcaatt	720
ttgtatctag	ttggccttct	gtgttttctg	ggaagcagca	acgaggaaa	gaggggtatcc	780
tagatgattt	ttacaacgaa	ctgaacgact	gctttgaggg	gggtaacatg	aaagtaatat	840
ggaactccgt	cctagttatt	gccaggagga	agcaaagggt	tgtataggct	ttagtactta	900
ttggcgaagt	gggtttacgt	gcaagcgcgc	atgcctgagc	tttgaggggg	gggactttca	960
catctcttct	tctcacactt	agccctaaca	cagagatcaa	ctccaacgac	ttgaggaaaat	1020
gacgattatt	tcttaagata	ttcatgggtga	ggttttccaa	atgtgatctg	aaatgtaatc	1080
ttggcgaagt	tggttcataa	aacgtcaaaa	gaccacctga	agataaaaaa	taaagtggaa	1140
ggagaagaaa	aagaggtttt	tttccagatg	agtctatgac	taagatttct	cttcaaagaa	1200
ttttttcaag	gtttgagagg	gatgaaagcg	aggctgctgc	acgattatcg	atgtcattca	1260
atagggaggt	agaagagaac	tgaggagaaa	aaggctacaa	ctgaacctaa	aactaactga	1320
atttggaatg	aagcctatta	cacgagaggt	gtggcgccct	tccaaagtca	tcagggttcaa	1380
gtttattgctt	tctggagcaa	taaagtatgt	caatcccata	cggggtctcc	accctttatg	1440
aattacttat	ctaccggcta	ggtgttgcac	cgcggaccag	cccaatcggg	cccaatttagt	1500
tgaaaaggtt	tttagcgcga	gtctagttat	taccgcgata	gtctaattgct	gaaaattttct	1560
cggacatttt	tgatccgtac	gtatatattca	gtacggagaa	tcatacaaaa	gaatctttca	1620
atgccaccaa	aaaaaaagac	atttttcacc	atcgcgcatt	cacagtagc	atg ggt aag	1678
					Met Gly Lys	
					1	
gaa aag act cac gtt tgc agg ccg cga tta aat tcc aac atg gat gct	1726					
Glu Lys Thr His Val Ser Arg Pro Arg Leu Asn Ser Asn Met Asp Ala						
5 10 15						
gat tta tat ggg tat aaa tgg gct cgc gat aat gtc ggg caa tca ggt	1774					
Asp Leu Tyr Gly Tyr Lys Trp Ala Arg Asp Asn Val Gly Gln Ser Gly						
20 25 30 35						
gcg aca atc tat cga ttg tat ggg aag ccc gat gcg cca gag ttg ttt	1822					
Ala Thr Ile Tyr Arg Leu Tyr Gly Lys Pro Asp Ala Pro Glu Leu Phe						
40 45 50						
ctg aaa cat ggc aaa ggt agc gtt gcc aat gat gtt aca gat gag atg	1870					
Leu Lys His Gly Lys Gly Ser Val Ala Asn Asp Val Thr Asp Glu Met						
55 60 65						
gtc aga cta aac tgg ctg acg gaa ttt atg cct ctt ccg acc atc aag	1918					
Val Arg Leu Asn Trp Leu Thr Glu Phe Met Pro Leu Pro Thr Ile Lys						
70 75 80						
cat ttt atc cgt act cct gat gat gca tgg tta ctc acc act gcg atc	1966					
His Phe Ile Arg Thr Pro Asp Asp Ala Trp Leu Thr Thr Ala Ile						
85 90 95						
ccc ggc aaa aca gca ttc cag gta tta gaa gaa tat cct gat tca ggt	2014					
Pro Gly Lys Thr Ala Phe Gln Val Leu Glu Glu Tyr Pro Asp Ser Gly						
100 105 110 115						
gaa aat att gtt gat gcg ctg gca gtg ttc ctg cgc cgg ttg cat tcg	2062					
Glu Asn Ile Val Asp Ala Leu Ala Val Phe Leu Arg Arg Leu His Ser						
120 125 130						
att cct gtt tgt aat tgt cct ttt aac agc gat cgc gta ttt cgt ctc	2110					
Ile Pro Val Cys Asn Cys Pro Phe Asn Ser Asp Arg Val Phe Arg Leu						
135 140 145						
gct cag gcg caa tca cga atg aat aac ggt ttg gtt gat gcg agt gat	2158					
Ala Gln Ala Gln Ser Arg Met Asn Asn Gly Leu Val Asp Ala Ser Asp						
150 155 160						
ttt gat gac gag cgt aat ggc tgg cct gtt gaa caa gtc tgg aaa gaa	2206					
Phe Asp Asp Glu Arg Asn Gly Trp Pro Val Glu Gln Val Trp Lys Glu						
165 170 175						
atg cat aag ctt ttg cca ttc tca ccg gat tca gtc gtc act cat ggt	2254					
Met His Lys Leu Leu Pro Phe Ser Pro Asp Ser Val Val Thr His Gly						
180 185 190 195						
gat ttc tca ctt gat aac ctt att ttt gac gag ggg aaa tta ata ggt	2302					
Asp Phe Ser Leu Asp Asn Leu Ile Phe Asp Glu Gly Lys Leu Ile Gly						
200 205 210 215						
tgt att gat gtt gga cga gtc gga atc gca gac cga tac cag gat ctt	2350					
Cys Ile Asp Val Gly Arg Val Gly Ile Ala Asp Arg Tyr Gln Asp Leu						
215 220 225						
gcc atc cta tgg aac tgc ctc ggt gag ttt tct cct tca tta cag aaa	2398					
Ala Ile Leu Trp Asn Cys Leu Gly Glu Phe Ser Pro Ser Leu Gln Lys						
230 235 240						
cgg ctt ttt caa aaa tat ggt att gat aat cct gat atg aat aaa ttg	2446					
Arg Leu Phe Gln Lys Tyr Gly Ile Asp Asn Pro Asp Met Asn Lys Leu						
245 250 255						
cag ttt cat ttg atg ctc gat gag ttt ttc taa tcagtactga caataaaaag	2499					
Gln Phe His Leu Met Leu Asp Glu Phe Phe						
260 265						
attcttgttt tcaagaactt gtcatttgta tagttttttt atattgtagt tgttctatctt	2559					
taatcaaatg ttagcgtgat ttatatatttt ttctgcctcg acatcatctg cccagatgcg	2619					

```

aagttaagtg cgcagaaagt aatatcatgc gtcaatcgta tgtgaatgct ggtcgctata 2679
ctgttcatct tagagaatgt cagtgagggt atttctgagc taatcattac cttgcccggg 2739
acaggaatag ggtgttacct tattaattac gacggtgggtc aaaacgacga aatttaccct 2799
atttttgagt tacaagacct gaatttggaa atgatgaaac aattgtacca agtcttgag 2859
gaccatgtaa gtgggcttaa tcctctcgag aagcaattcc caattgaaca cagtaaaccg 2919
ttatcagcca ctctgttctt tcacttaagg tctctttttt acatggcgaa gactcataag 2979
cgtactggaa gacattacaa cttgaaaaag atttcttctg tcaactaataa cgataaacct 3039
tacaatggaa actctcagct gagagtcccc ttgaagaaaa ccttggtctga ttacaatgac 3099
gtagacatta ctttgattcc gtttcttctg aacaagcctt cagggtgtcaa gtttgacaag 3159
acggaatact cagaaatttt gttctatgat aaagatgctt gttcgatgtc aattgaggag 3219
tctcagtttg tcgatccaac atccggaaaa gagttttcca gtgaacagct gatcaaagca 3279
aaaatcgcaa ataacttggt catttctgtg aaaggttatt ctatgtttta tcatgaaact 3339
ccaaggaaga tcaaatttgt cgtcaatgag ggttcaactt tcaaagatgt ggagacaaaa 3399
tctcagtttg tcgatccaac atccggaaaa gagttttcca gtgaacagct gatcaaagca 3459
tctcctctag gtgccgatgc ttacattcct ttaaactcag agcaagtcaa aacaataaat 3519
cgatttaatg atatcatcaa tatcccctct ttggaaattc taggtttcag ggatataatc 3579
aattggttgc cacagtatca gtttggcaaa gcacgtgttt tatcccctaa taa 3632
<210> 162
<211> 269
<212> PRT
<213> 合成
<400> 162
Met Gly Lys Glu Lys Thr His Val Ser Arg Pro Arg Leu Asn Ser Asn
1 5 10 15
Met Asp Ala Asp Leu Tyr Gly Tyr Lys Trp Ala Arg Asp Asn Val Gly
20 25 30
Gln Ser Gly Ala Thr Ile Tyr Arg Leu Tyr Gly Lys Pro Asp Ala Pro
35 40 45
Glu Leu Phe Leu Lys His Gly Lys Gly Ser Val Ala Asn Asp Val Thr
50 55 60
Asp Glu Met Val Arg Leu Asn Trp Leu Thr Glu Phe Met Pro Leu Pro
65 70 75 80
Thr Ile Lys His Phe Ile Arg Thr Pro Asp Ala Trp Leu Leu Thr
85 90 95
Thr Ala Ile Pro Gly Lys Thr Ala Phe Gln Val Leu Glu Glu Tyr Pro
100 105 110
Asp Ser Gly Glu Asn Ile Val Asp Ala Leu Ala Val Phe Leu Arg Arg
115 120 125
[0471] Leu His Ser Ile Pro Val Cys Asn Cys Pro Phe Asn Ser Asp Arg Val
130 135 140
Phe Arg Leu Ala Gln Ala Gln Ser Arg Met Asn Asn Gly Leu Val Asp
145 150 155 160
Ala Ser Asp Phe Asp Asp Glu Arg Asn Gly Trp Pro Val Glu Gln Val
165 170 175
Trp Lys Glu Met His Lys Leu Leu Pro Phe Ser Pro Asp Ser Val Val
180 185 190
Thr His Gly Asp Phe Ser Leu Asp Asn Leu Ile Phe Asp Glu Gly Lys
195 200 205
Leu Ile Gly Cys Ile Asp Val Gly Arg Val Gly Ile Ala Asp Arg Tyr
210 215 220
Gln Asp Leu Ala Ile Leu Trp Asn Cys Leu Gly Glu Phe Ser Pro Ser
225 230 235 240
Leu Gln Lys Arg Leu Phe Gln Lys Tyr Gly Ile Asp Asn Pro Asp Met
245 250 255
Asn Lys Leu Gln Phe His Leu Met Leu Asp Glu Phe Phe
260 265
<210> 163
<211> 3573
<212> DNA
<213> 合成
<220>
<221> 尚未归类的特征
<222> (1) .. (3573)
<223> 构建体 KU70 (上) -pRPL6A-kanMX-终止子-KU70 (下)
<220>
<221> 尚未归类的特征
<222> (1) .. (1002)
<223> KU70 上游区域
<220>
<221> 启动子
<222> (1003) .. (1610)
<223> 启动子 pRPL6A
<220>
<221> CDS
<222> (1611) .. (2420)
<223> kanMX
<220>

```


	Gly	Cys	Ile	Asp	Val	Gly	Arg	Val	Gly	Ile	Ala	Asp	Arg	Tyr	Gln	Asp	
	ctt	gcc	atc	cta	215	tg	aac	tgc	ctc	ggg	gag	ttt	tct	cct	tca	220	225
	Leu	Ala	Ile	Leu	Trp	Asn	Cys	Leu	Gly	Glu	Phe	Ser	Pro	Ser	Leu	Gln	2336
	aaa	cgg	ctt	ttt	caa	aaa	tat	ggg	att	gat	aat	cct	gat	atg	aat	aaa	2384
	Lys	Arg	Leu	Phe	Gln	Lys	Tyr	Gly	Ile	Asp	Asn	Pro	Asp	Met	Asn	Lys	
	ttg	cag	ttt	cat	ttg	atg	ctc	gat	gag	ttt	ttc	taa	tcagtactga				2430
	Leu	Gln	Phe	His	Leu	Met	Leu	Asp	Glu	Phe	Phe						
	caataaaaaag	attcttggttt	tcaagaactt	gtcattttgta	tagttttttt	atattgtagt											2490
	tgttctattt	taatacaaatg	ttagcgtgat	ttatatatttt	tttcgcctcg	acatcatctg											2550
	cccagatg	aagtttaagt	cgcagaaagt	aatatcatgc	gtcaatcgta	tgtgaatgct											2610
	ggtcgctata	ctgttcatct	tagagaatgt	cagtgagggt	atttctgagc	taatacattac											2670
	cttgcccggg	acaggaatag	gggtgttacct	tatttaattac	gacgggtggc	aaaacgacga											2730
	aattttacccc	atttttgagt	tacaagacct	gaattttggaa	atgatgaaac	aattgtacca											2790
	agtcttgagg	gaccatgtaa	gtgggcttaa	tcctctcgag	aagcaattcc	caattgaaca											2850
	cagtaaacgg	ttatcagcca	ctctgttctt	tcaacttaagg	tctctttttt	acatggcgaa											2910
	gactcataag	cgtactggaa	gacattacaa	cttgaaaaag	attttcttgt	tcactaataa											2970
	cgataaacct	tacaatggaa	actctcagct	gagagtcc	ttgaagaaaa	ccctggctga											3030
	ttacaatgac	gtagacatta	ctttgattcc	gtttcttctg	aacaagcctt	cagggtgtcaa											3090
	gtttgacaag	acggaatact	cagaaatttt	gtttctatgat	aaagatgctt	gttcgatgtc											3150
	aattgaggag	atccgccaac	gaattttctag	acataaggag	atcaagcggg	tttacttcac											3210
	ctgtcctttg	aaaatcgcaa	ataacttggt	catttctgtg	aaaggttatt	ctatgtttta											3270
	tcatgaaact	ccaaggaaga	tcaaatttgt	cgtaaatgag	ggttcaactt	tcaaagatgt											3330
	ggagacaaaa	tctcagtttg	tcgatccaac	atccggaaaa	gagttttcca	gtgaacagct											3390
	gatcaaaaga	tatctcttag	gtgcgatgc	ttacattcct	ttaaactcag	agcaagtcaa											3450
	aacaataaat	cgattttaatg	atatcatcaa	tatccctctt	ttggaaattc	taggtttcag											3510
	ggatatactt	aattggttgc	cacagtatca	gtttggcaaa	gcacgtttt	tatcccttaa											3570
	taa																3573
	<210>	164															
	<211>	269															
	<212>	PRT															
	<213>	合成															
	<400>	164															
	Met	Gly	Lys	Glu	Lys	Thr	His	Val	Ser	Arg	Pro	Arg	Leu	Asn	Ser	Asn	
	1				5					10					15		
[0473]	Met	Asp	Ala	Asp	Leu	Tyr	Gly	Tyr	Lys	Trp	Ala	Arg	Asp	Asn	Val	Gly	
					20					25					30		
	Gln	Ser	Gly	Ala	Thr	Ile	Tyr	Arg	Leu	Tyr	Gly	Lys	Pro	Asp	Ala	Pro	
					35					40					45		
	Glu	Leu	Phe	Leu	Lys	His	Gly	Lys	Gly	Ser	Val	Ala	Asn	Asp	Val	Thr	
					50					55					60		
	Asp	Glu	Met	Val	Arg	Leu	Asn	Trp	Leu	Thr	Glu	Phe	Met	Pro	Leu	Pro	
					65					70					75		
	Thr	Ile	Lys	His	Phe	Ile	Arg	Thr	Pro	Asp	Asp	Ala	Trp	Leu	Leu	Thr	
					85					90					95		
	Thr	Ala	Ile	Pro	Gly	Lys	Thr	Ala	Phe	Gln	Val	Leu	Glu	Glu	Tyr	Pro	
					100					105					110		
	Asp	Ser	Gly	Glu	Asn	Ile	Val	Asp	Ala	Leu	Ala	Val	Phe	Leu	Arg	Arg	
					115					120					125		
	Leu	His	Ser	Ile	Pro	Val	Cys	Asn	Cys	Pro	Phe	Asn	Ser	Asp	Arg	Val	
					130					135					140		
	Phe	Arg	Leu	Ala	Gln	Ala	Gln	Ser	Arg	Met	Asn	Asn	Gly	Leu	Val	Asp	
					145					150					155		
	Ala	Ser	Asp	Phe	Asp	Asp	Glu	Arg	Asn	Gly	Trp	Pro	Val	Glu	Gln	Val	
					165					170					175		
	Trp	Lys	Glu	Met	His	Lys	Leu	Leu	Pro	Phe	Ser	Pro	Asp	Ser	Val	Val	
					180					185					190		
	Thr	His	Gly	Asp	Phe	Ser	Leu	Asp	Asn	Leu	Ile	Phe	Asp	Glu	Gly	Lys	
					195					200					205		
	Leu	Ile	Gly	Cys	Ile	Asp	Val	Gly	Arg	Val	Gly	Ile	Ala	Asp	Arg	Tyr	
					210					215					220		
	Gln	Asp	Leu	Ala	Ile	Leu	Trp	Asn	Cys	Leu	Gly	Glu	Phe	Ser	Pro	Ser	
					225					230					235		
	Leu	Gln	Lys	Arg	Leu	Phe	Gln	Lys	Tyr	Gly	Ile	Asp	Asn	Pro	Asp	Met	
					245					250					255		
	Asn	Lys	Leu	Gln	Phe	His	Leu	Met	Leu	Asp	Glu	Phe	Phe				
					260					265							
	<210>	165															
	<211>	786															
	<212>	DNA															
	<213>	芽孢杆菌属 (Bacillus)															
	<220>																
	<221>	CDS															
	<222>	(1) .. (786)															
	<223>	iscU 操纵子中的必需基因 sufC															

[0474]

```

<400> 165
atg gct gct tca aca tta acg atc aaa gat ctt cac gtt gaa atc gaa      48
Met Ala Ala Ser Thr Leu Thr Ile Lys Asp Leu His Val Glu Ile Glu
1 5 10 15
ggg aaa gag atc tta aag ggt gta aac ctt gaa ata aaa ggt gga gaa      96
Gly Lys Glu Ile Leu Lys Gly Val Asn Leu Glu Ile Lys Gly Gly Glu
20 25 30
ttc cac gca gta atg ggc ccg aac gga act ggt aaa tcc act tta tca      144
Phe His Ala Val Met Gly Pro Asn Gly Thr Gly Lys Ser Thr Leu Ser
35 40 45
gct gct att atg ggg cat cct aaa tat gaa gta aca aaa ggc agc atc      192
Ala Ala Ile Met Gly His Pro Lys Tyr Glu Val Thr Lys Gly Ser Ile
50 55 60
acg ctt gac ggc aaa gat gta ctg gag atg gaa gtg gac gag cgt gcg      240
Thr Leu Asp Gly Lys Asp Val Leu Glu Met Glu Val Asp Glu Arg Ala
65 70 75 80
cag gca ggc tta ttc ctt gca atg cag tac cca agt gaa atc agc ggt      288
Gln Ala Gly Leu Phe Leu Ala Met Gln Tyr Pro Ser Glu Ile Ser Gly
85 90 95
gtg aca aat gcc gac ttc ctt cgc tca gca atc aac gcg cgc aga gaa      336
Val Thr Asn Ala Asp Phe Leu Arg Ser Ala Ile Asn Ala Arg Arg Glu
100 105 110
gaa ggc gat gaa att tct ctc atg aaa ttc atc cgc aaa atg gac gaa      384
Glu Gly Asp Glu Ile Ser Leu Met Lys Phe Ile Arg Lys Met Asp Glu
115 120 125
aac atg gag ttc ctt gaa atg gac cct gaa atg gct cag cgc tat ctt      432
Asn Met Glu Phe Leu Glu Met Asp Pro Glu Met Ala Gln Arg Tyr Leu
130 135 140
aac gaa ggc ttc tca ggc ggg gag aaa aaa cgc aac gaa atc ctt caa      480
Asn Glu Gly Phe Ser Gly Gly Glu Lys Lys Arg Asn Glu Ile Leu Gln
145 150 155 160
tta atg atg att gaa ccg aaa atc gcc atc ctt gat gaa att gac tca      528
Leu Met Met Ile Glu Pro Lys Ile Ala Ile Leu Asp Glu Ile Asp Ser
165 170 175
ggc ctt gat att gac gct ttg aaa gtt gta tca aaa ggg atc aac aaa      576
Gly Leu Asp Ile Asp Ala Leu Lys Val Val Ser Lys Gly Ile Asn Lys
180 185 190
atg cgc agc gaa aac ttc ggc tgc ctg atg atc act cac tat cag cgc      624
Met Arg Ser Glu Asn Phe Gly Cys Leu Met Ile Thr His Tyr Gln Arg
195 200 205
ctg tta aac tac atc act ccg gat gtt gtt cac gtt atg atg caa ggc      672
Leu Leu Asn Tyr Ile Thr Pro Asp Val Val His Val Met Met Gln Gly
210 215 220
cgc gtt gta aaa tcc ggc ggt gca gag ctt gca cag cgt ctt gaa gca      720
Arg Val Val Lys Ser Gly Gly Ala Glu Leu Ala Gln Arg Leu Glu Ala
225 230 235 240
gaa ggt tat gac tgg atc aaa caa gaa ctt ggc att gaa gac gaa act      768
Glu Gly Tyr Asp Trp Ile Lys Gln Glu Leu Gly Ile Glu Asp Glu Thr
245 250 255
gtt ggc caa gaa gcg taa
Val Gly Gln Glu Ala      786
260

<210> 166
<211> 261
<212> PRT
<213> 芽孢杆菌属
<400> 166
Met Ala Ala Ser Thr Leu Thr Ile Lys Asp Leu His Val Glu Ile Glu
1 5 10 15
Gly Lys Glu Ile Leu Lys Gly Val Asn Leu Glu Ile Lys Gly Gly Glu
20 25 30
Phe His Ala Val Met Gly Pro Asn Gly Thr Gly Lys Ser Thr Leu Ser
35 40 45
Ala Ala Ile Met Gly His Pro Lys Tyr Glu Val Thr Lys Gly Ser Ile
50 55 60
Thr Leu Asp Gly Lys Asp Val Leu Glu Met Glu Val Asp Glu Arg Ala
65 70 75 80
Gln Ala Gly Leu Phe Leu Ala Met Gln Tyr Pro Ser Glu Ile Ser Gly
85 90 95
Val Thr Asn Ala Asp Phe Leu Arg Ser Ala Ile Asn Ala Arg Arg Glu
100 105 110
Glu Gly Asp Glu Ile Ser Leu Met Lys Phe Ile Arg Lys Met Asp Glu
115 120 125
Asn Met Glu Phe Leu Glu Met Asp Pro Glu Met Ala Gln Arg Tyr Leu
130 135 140
Asn Glu Gly Phe Ser Gly Gly Glu Lys Lys Arg Asn Glu Ile Leu Gln
145 150 155 160
Leu Met Met Ile Glu Pro Lys Ile Ala Ile Leu Asp Glu Ile Asp Ser
165 170 175

```

				165					170					175					
	Gly	Leu	Asp	Ile	Asp	Ala	Leu	Lys	Val	Val	Ser	Lys	Gly	Ile	Asn	Lys			
				180					185					190					
	Met	Arg	Ser	Glu	Asn	Phe	Gly	Cys	Leu	Met	Ile	Thr	His	Tyr	Gln	Arg			
			195					200					205						
	Leu	Leu	Asn	Tyr	Ile	Thr	Pro	Asp	Val	Val	His	Val	Met	Met	Gln	Gly			
		210					215					220							
	Arg	Val	Val	Lys	Ser	Gly	Gly	Ala	Glu	Leu	Ala	Gln	Arg	Leu	Glu	Ala			
		225				230					235				240				
	Glu	Gly	Tyr	Asp	Trp	Ile	Lys	Gln	Glu	Leu	Gly	Ile	Glu	Asp	Glu	Thr			
				245						250					255				
	Val	Gly	Gln	Glu	Ala														
				260															
	<210>	167																	
	<211>	1314																	
	<212>	DNA																	
	<213>	芽孢杆菌属																	
	<220>																		
	<221>	CDS																	
	<222>	(1) .. (1314)																	
	<223>	iscU 操纵子中的必需基因 sufD																	
	<400>	167																	
	atg	aca	cta	ggt	aca	aaa	cta	tcc	gta	gat	cag	gag	tat	ctc	aaa	agc			48
	Met	Thr	Leu	Gly	Thr	Lys	Leu	Ser	Val	Asp	Gln	Glu	Tyr	Leu	Lys	Ser			
				5						10					15				
	ttt	tcc	gaa	aag	cac	caa	gaa	cct	gcc	tgg	ctg	aag	aac	ctg	cgc	tta			96
	Phe	Ser	Glu	Lys	His	Gln	Glu	Pro	Ala	Trp	Leu	Lys	Asn	Leu	Arg	Leu			
				20					25					30					
	cag	gct	ctt	gaa	caa	gct	gag	gat	ctg	ccg	atg	ccg	aaa	cct	gac	aaa			144
	Gln	Ala	Leu	Glu	Gln	Ala	Glu	Asp	Leu	Pro	Met	Pro	Lys	Pro	Asp	Lys			
				35				40					45						
	aca	aaa	atc	aca	aac	tgg	aac	ttc	acg	aat	ttc	gcg	aag	cat	aca	gtg			192
	Thr	Lys	Ile	Thr	Asn	Trp	Asn	Phe	Thr	Asn	Phe	Ala	Lys	His	Thr	Val			
				50			55					60							
	gat	aac	gaa	ccg	ctt	tct	tca	tta	gaa	gat	ttg	act	gat	gaa	gta	aaa			240
	Asp	Asn	Glu	Pro	Leu	Ser	Ser	Leu	Glu	Asp	Leu	Thr	Asp	Glu	Val	Lys			
				65			70				75				80				
	gcg	ctg	atc	gac	att	gaa	aac	gaa	gat	aaa	aca	tta	tac	gta	cag	cgt			288
[0475]	Ala	Leu	Ile	Asp	Ile	Glu	Asn	Glu	Asp	Lys	Thr	Leu	Tyr	Val	Gln	Arg			
					85					90				95					
	gac	cag	acg	ccg	gca	cat	ctt	tct	ctt	tct	caa	gaa	ttg	aaa	gac	aaa			336
	Asp	Gln	Thr	Pro	Ala	His	Leu	Ser	Leu	Ser	Gln	Glu	Leu	Lys	Asp	Lys			
				100					105					110					
	ggc	gtc	atc	ttc	act	gat	att	ctg	act	gct	gca	cgc	gag	cac	agc	gac			384
	Gly	Val	Ile	Phe	Thr	Asp	Ile	Leu	Thr	Ala	Ala	Arg	Glu	His	Ser	Asp			
				115			120						125						
	ttg	gtg	gag	aaa	tac	ttt	atg	aag	gac	ggc	gtg	aag	gtt	gac	gaa	cat			432
	Leu	Val	Glu	Lys	Tyr	Phe	Met	Lys	Asp	Gly	Val	Lys	Val	Asp	Glu	His			
				130			135					140							
	aag	tta	act	gca	ctt	cac	gcg	gct	ctt	gtt	aac	gga	gga	gca	ttc	ctt			480
	Lys	Leu	Thr	Ala	Leu	His	Ala	Ala	Leu	Val	Asn	Gly	Gly	Ala	Phe	Leu			
		145				150					155				160				
	tat	gtt	ccg	aaa	aat	gtt	cag	gtg	gag	acg	ccg	gtt	cag	gct	gtt	tat			528
	Tyr	Val	Pro	Lys	Asn	Val	Gln	Val	Glu	Thr	Pro	Val	Gln	Ala	Val	Tyr			
				165						170				175					
	gtc	cac	gaa	agc	aat	gat	aca	gct	cta	ttc	aac	cac	gtg	ctg	att	gtg			576
	Val	His	Glu	Ser	Asn	Asp	Thr	Ala	Leu	Phe	Asn	His	Val	Leu	Ile	Val			
				180					185					190					
	gct	gaa	gac	cac	agc	tct	gta	aca	tat	gta	gaa	aac	tac	atc	agc	act			624
	Ala	Glu	Asp	His	Ser	Ser	Val	Thr	Tyr	Val	Glu	Asn	Tyr	Ile	Ser	Thr			
				195			200						205						
	gta	aac	ccg	aaa	gac	gct	gtc	ttc	aat	att	atc	agt	gag	gtc	atc	aca			672
	Val	Asn	Pro	Lys	Asp	Ala	Val	Phe	Asn	Ile	Ile	Ser	Glu	Val	Ile	Thr			
				210			215					220							
	ggt	gat	aat	gcc	agc	gtg	aca	tac	ggt	gcg	gtg	gat	aat	ctg	tca	agc			720
	Gly	Asp	Asn	Ala	Ser	Val	Thr	Tyr	Gly	Ala	Val	Asp	Asn	Leu	Ser	Ser			
		225				230					235				240				
	ggt	gtg	aca	act	tat	gtg	aac	cgc	cgc	ggt	gct	gcg	cgc	gga	cgc	gac			768
	Gly	Val	Thr	Thr	Tyr	Val	Asn	Arg	Arg	Gly	Ala	Ala	Arg	Gly	Arg	Asp			
				245						250				255					
	agc	aaa	atc	gag	tgg	gct	ctc	ggc	ctg	atg	aac	gat	ggc	gac	acg	att			816
	Ser	Lys	Ile	Glu	Trp	Ala	Leu	Gly	Leu	Met	Asn	Asp	Gly	Asp	Thr	Ile			
				260					265					270					
	tcc	gaa	aac	aca	aca	aac	ctt	tac	ggt	gac	ggc	aca	tac	ggc	gat	acg			864
	Ser	Glu	Asn	Thr	Thr	Asn	Leu	Tyr	Gly	Asp	Gly	Thr	Tyr	Gly	Asp	Thr			
				275			280						285						
	aaa	acg	gtt	gtt	gtc	gga	aga	gga	gag	caa	aca	gaa	aac	ttc	aca	aca			912
	Lys	Thr	Val	Val	Val	Gly	Arg	Gly	Glu	Gln	Thr	Glu	Asn	Phe	Thr	Thr			

[0476]

145

	Lys	Leu	Ala	Ala	Tyr	Ala	Leu	Glu	Arg	Phe	Arg	Gln	Leu	Asp	Gly	Val	
	305					310					315					320	
	acg	gtt	tac	ggg	ccg	gaa	gag	cgt	gcc	ggg	ctt	gtc	aca	ttt	aac	ctt	1008
	Thr	Val	Tyr	Gly	Pro	Glu	Glu	Arg	Ala	Gly	Leu	Val	Thr	Phe	Asn	Leu	
					325					330					335		
	gat	gat	gtt	cat	ccg	cac	gat	gtg	gcg	act	gtg	ctt	gat	gcg	gag	gga	1056
	Asp	Asp	Val	His	Pro	His	Asp	Val	Ala	Thr	Val	Leu	Asp	Ala	Glu	Gly	
				340					345					350			
	atc	gca	gtc	aga	gcc	ggc	cac	cat	tgt	gcg	cag	ccg	ctg	atg	aaa	tgg	1104
	Ile	Ala	Val	Arg	Ala	Gly	His	His	Cys	Ala	Gln	Pro	Leu	Met	Lys	Trp	
			355				360						365				
	ctg	gat	gta	act	gct	act	gca	aga	gca	agc	ttt	tat	ctg	tat	aat	aca	1152
	Leu	Asp	Val	Thr	Ala	Thr	Ala	Arg	Ala	Ser	Phe	Tyr	Leu	Tyr	Asn	Thr	
		370					375					380					
	gag	gaa	gag	att	gat	aag	ctt	gtg	gaa	gct	ctt	caa	aag	aca	aag	gag	1200
	Glu	Glu	Glu	Ile	Asp	Lys	Leu	Val	Glu	Ala	Leu	Gln	Lys	Thr	Lys	Glu	
	385					390					395					400	
	tat	ttt	aca	aat	gtc	ttt	taa										1221
	Tyr	Phe	Thr	Asn	Val	Phe											
					405												
	<210>	170															
	<211>	406															
	<212>	PRT															
	<213>	芽孢杆菌属															
	<400>	170															
	Met	Asn	Ile	Thr	Asp	Ile	Arg	Glu	Gln	Phe	Pro	Ile	Leu	His	Gln	Gln	
	1				5					10					15		
	Val	Asn	Gly	His	Asp	Leu	Val	Tyr	Leu	Asp	Ser	Ala	Ala	Thr	Ser	Gln	
				20					25					30			
	Lys	Pro	Arg	Ala	Val	Ile	Glu	Thr	Leu	Asp	Lys	Tyr	Tyr	Asn	Gln	Tyr	
			35				40						45				
	Asn	Ser	Asn	Val	His	Arg	Gly	Val	His	Thr	Leu	Gly	Thr	Arg	Ala	Thr	
		50					55					60					
	Asp	Gly	Tyr	Glu	Gly	Ala	Arg	Glu	Lys	Val	Arg	Lys	Phe	Ile	Asn	Ala	
	65					70					75					80	
	Lys	Ser	Met	Ala	Glu	Ile	Ile	Phe	Thr	Lys	Gly	Thr	Thr	Thr	Ser	Leu	
				85						90					95		
	Asn	Met	Val	Ala	Leu	Ser	Tyr	Ala	Arg	Ala	Asn	Leu	Lys	Pro	Gly	Asp	
				100					105					110			
[0478]	Glu	Val	Val	Ile	Thr	Tyr	Met	Glu	His	His	Ala	Asn	Ile	Ile	Pro	Trp	
			115					120					125				
	Gln	Gln	Ala	Val	Lys	Ala	Thr	Gly	Ala	Thr	Leu	Lys	Tyr	Ile	Pro	Leu	
		130						135				140					
	Gln	Glu	Asp	Gly	Thr	Ile	Ser	Leu	Glu	Asp	Val	Arg	Glu	Thr	Val	Thr	
	145					150					155					160	
	Ser	Asn	Thr	Lys	Ile	Val	Ala	Val	Ser	His	Val	Ser	Asn	Val	Leu	Gly	
				165						170					175		
	Thr	Val	Asn	Pro	Ile	Lys	Glu	Met	Ala	Lys	Ile	Ala	His	Asp	Asn	Gly	
				180					185					190			
	Ala	Val	Ile	Val	Val	Asp	Gly	Ala	Gln	Ser	Thr	Pro	His	Met	Lys	Ile	
			195					200					205				
	Asp	Val	Gln	Asp	Leu	Asp	Cys	Asp	Phe	Phe	Ala	Leu	Ser	Ser	His	Lys	
		210					215					220					
	Met	Cys	Gly	Pro	Thr	Gly	Val	Gly	Val	Leu	Tyr	Gly	Lys	Lys	Ala	Leu	
	225					230					235					240	
	Leu	Glu	Asn	Met	Glu	Pro	Ala	Glu	Phe	Gly	Gly	Glu	Met	Ile	Asp	Phe	
				245						250					255		
	Val	Gly	Leu	Tyr	Glu	Ser	Thr	Trp	Lys	Glu	Leu	Pro	Trp	Lys	Phe	Glu	
			260						265					270			
	Ala	Gly	Thr	Pro	Ile	Ile	Ala	Gly	Ala	Ile	Gly	Leu	Gly	Ala	Ala	Ile	
			275					280					285				
	Asp	Phe	Leu	Glu	Glu	Ile	Gly	Leu	Asp	Glu	Ile	Ser	Arg	His	Glu	His	
		290					295					300					
	Lys	Leu	Ala	Ala	Tyr	Ala	Leu	Glu	Arg	Phe	Arg	Gln	Leu	Asp	Gly	Val	
	305					310					315					320	
	Thr	Val	Tyr	Gly	Pro	Glu	Glu	Arg	Ala	Gly	Leu	Val	Thr	Phe	Asn	Leu	
				325						330					335		
	Asp	Asp	Val	His	Pro	His	Asp	Val	Ala	Thr	Val	Leu	Asp	Ala	Glu	Gly	
				340					345					350			
	Ile	Ala	Val	Arg	Ala	Gly	His	His	Cys	Ala	Gln	Pro	Leu	Met	Lys	Trp	
			355				360						365				
	Leu	Asp	Val	Thr	Ala	Thr	Ala	Arg	Ala	Ser	Phe	Tyr	Leu	Tyr	Asn	Thr	
		370					375					380					
	Glu	Glu	Glu	Ile	Asp	Lys	Leu	Val	Glu	Ala	Leu	Gln	Lys	Thr	Lys	Glu	
	385					390					395					400	
	Tyr	Phe	Thr	Asn	Val	Phe											
					405												
	<210>	171															

<211>	444	
<212>	DNA	
<213>	芽孢杆菌属	
<220>		
<221>	CDS	
<222>	(1) .. (444)	
<223>	iscU 操纵子中的必需基因 sufU (同义词 iscU)	
<400>	171	
atg tct ttt aat gca aac tta gat aca ttg tac aga cag gtg att atg		48
Met Ser Phe Asn Ala Asn Leu Asp Thr Leu Tyr Arg Gln Val Ile Met		
1 5 10 15		
gat cat tac aaa aac ccg cga aat aaa ggg gtt tta aat gac agc att		96
Asp His Tyr Lys Asn Pro Arg Asn Lys Gly Val Leu Asn Asp Ser Ile		
20 25 30		
gtc gtg gat atg aac aat ccg aca tgc ggc gac cgc atc aga ctg aca		144
Val Val Asp Met Asn Asn Pro Thr Cys Gly Asp Arg Ile Arg Leu Thr		
35 40 45		
atg aag ctt gac gga gac atc gtg gaa gat gcg aag ttt gaa ggg gaa		192
Met Lys Leu Asp Gly Asp Ile Val Glu Asp Ala Lys Phe Glu Gly Glu		
50 55 60		
ggc tgt tcc att tca atg gca tcc gct tcg atg atg act cag gcg att		240
Gly Cys Ser Ile Ser Met Ala Ser Ala Ser Met Met Thr Gln Ala Ile		
65 70 75 80		
aaa ggg aaa gat att gaa act gcc ctt tcc atg tct aag att ttc tcg		288
Lys Gly Lys Asp Ile Glu Thr Ala Leu Ser Met Ser Lys Ile Phe Ser		
85 90 95		
gat atg atg caa ggg aaa gag tat gat gat tct ata gat ctc ggg gat		336
Asp Met Met Gln Gly Lys Glu Tyr Asp Asp Ser Ile Asp Leu Gly Asp		
100 105 110		
att gaa gcc ctt caa ggc gtt tca aaa ttc cct gcc cgt atc aaa tgt		384
Ile Glu Ala Leu Gln Gly Val Ser Lys Phe Pro Ala Arg Ile Lys Cys		
115 120 125		
gca acc ctg tca tgg aaa gca ctt gaa aaa gga gtc gcg aaa gaa gaa		432
Ala Thr Leu Ser Trp Lys Ala Leu Glu Lys Gly Val Ala Lys Glu Glu		
130 135 140		
ggc ggc aat taa		444
Gly Gly Asn		
145		
[0479] <210>	172	
<211>	147	
<212>	PRT	
<213>	芽孢杆菌属	
<400>	172	
Met Ser Phe Asn Ala Asn Leu Asp Thr Leu Tyr Arg Gln Val Ile Met		
1 5 10 15		
Asp His Tyr Lys Asn Pro Arg Asn Lys Gly Val Leu Asn Asp Ser Ile		
20 25 30		
Val Val Asp Met Asn Asn Pro Thr Cys Gly Asp Arg Ile Arg Leu Thr		
35 40 45		
Met Lys Leu Asp Gly Asp Ile Val Glu Asp Ala Lys Phe Glu Gly Glu		
50 55 60		
Gly Cys Ser Ile Ser Met Ala Ser Ala Ser Met Met Thr Gln Ala Ile		
65 70 75 80		
Lys Gly Lys Asp Ile Glu Thr Ala Leu Ser Met Ser Lys Ile Phe Ser		
85 90 95		
Asp Met Met Gln Gly Lys Glu Tyr Asp Asp Ser Ile Asp Leu Gly Asp		
100 105 110		
Ile Glu Ala Leu Gln Gly Val Ser Lys Phe Pro Ala Arg Ile Lys Cys		
115 120 125		
Ala Thr Leu Ser Trp Lys Ala Leu Glu Lys Gly Val Ala Lys Glu Glu		
130 135 140		
Gly Gly Asn		
145		
<210>	173	
<211>	1398	
<212>	DNA	
<213>	芽孢杆菌属	
<220>		
<221>	CDS	
<222>	(1) .. (1398)	
<223>	iscU 操纵子中的必需基因 sufB	
<400>	173	
atg gct aaa aaa atg cct gat att ggt gaa tac aag tat ggt ttt cac		48
Met Ala Lys Lys Met Pro Asp Ile Gly Glu Tyr Lys Tyr Gly Phe His		
1 5 10 15		
gac aag gac gtt tcc att ttc cgt tca gag cgc gga ttg aca aaa gaa		96
Asp Lys Asp Val Ser Ile Phe Arg Ser Glu Arg Gly Leu Thr Lys Glu		

[0480]

```

      435      440      445
gca gtt gaa atg aac cgt ttg att aag ttc gaa atg gaa ggt tct atc      1392
Ala Val Glu Met Asn Arg Leu Ile Lys Phe Glu Met Glu Gly Ser Ile
      450      455      460
ggg taa
Gly
465
<210> 174
<211> 465
<212> PRT
<213> 芽孢杆菌属
<400> 174
Met Ala Lys Lys Met Pro Asp Ile Gly Glu Tyr Lys Tyr Gly Phe His
1 5 10 15
Asp Lys Asp Val Ser Ile Phe Arg Ser Glu Arg Gly Leu Thr Lys Glu
20 25 30
Ile Val Glu Glu Ile Ser Arg Met Lys Glu Glu Pro Gln Trp Met Leu
35 40 45
Asp Phe Arg Leu Lys Ser Leu Glu His Phe Tyr Asn Met Pro Met Pro
50 55 60
Gln Trp Gly Gly Asp Leu Asn Ser Leu Asn Phe Asp Glu Ile Thr Tyr
65 70 75 80
Tyr Val Lys Pro Ser Glu Arg Ser Glu Arg Ser Trp Asp Glu Val Pro
85 90 95
Glu Glu Ile Lys Gln Thr Phe Asp Lys Leu Gly Ile Pro Glu Ala Glu
100 105 110
Gln Lys Tyr Leu Ala Gly Val Ser Ala Gln Tyr Glu Ser Glu Val Val
115 120 125
Tyr His Asn Met Lys Glu Asp Leu Glu Ala Gln Gly Ile Val Phe Lys
130 135 140
Asp Thr Asp Ser Ala Leu Lys Glu Asn Glu Asp Ile Phe Arg Glu His
145 150 155 160
Trp Ala Lys Val Ile Pro Pro Thr Asp Asn Lys Phe Ala Ala Leu Asn
165 170 175
Ser Ala Val Trp Ser Gly Gly Ser Phe Ile Tyr Val Pro Lys Gly Val
180 185 190
Lys Val Glu Thr Pro Leu Gln Ala Tyr Phe Arg Ile Asn Ser Glu Asn
195 200 205
[0481] Met Gly Gln Phe Glu Arg Thr Leu Ile Ile Val Asp Glu Glu Ala Ser
210 215 220
Val His Tyr Val Glu Gly Cys Thr Ala Pro Val Tyr Thr Thr Asn Ser
225 230 235 240
Leu His Ser Ala Val Val Glu Ile Ile Val Lys Lys Gly Gly Tyr Cys
245 250 255
Arg Tyr Thr Thr Ile Gln Asn Trp Ala Asn Asn Val Tyr Asn Leu Val
260 265 270
Thr Lys Arg Thr Val Cys Glu Glu Asn Ala Thr Met Glu Trp Ile Asp
275 280 285
Gly Asn Ile Gly Ser Lys Leu Thr Met Lys Tyr Pro Ala Cys Ile Leu
290 295 300
Lys Gly Glu Gly Ala Arg Gly Met Thr Leu Ser Ile Ala Leu Ala Gly
305 310 315 320
Lys Gly Gln His Gln Asp Ala Gly Ala Lys Met Ile His Leu Ala Pro
325 330 335
Asn Thr Ser Ser Thr Ile Val Ser Lys Ser Ile Ser Lys Gln Gly Gly
340 345 350
Lys Val Thr Tyr Arg Gly Ile Val His Phe Gly Arg Lys Ala Glu Gly
355 360 365
Ala Arg Ser Asn Ile Glu Cys Asp Thr Leu Ile Met Asp Asn Lys Ser
370 375 380
Thr Ser Asp Thr Ile Pro Tyr Asn Glu Ile Leu Asn Asp Asn Ile Ser
385 390 395 400
Leu Glu His Glu Ala Lys Val Ser Lys Val Ser Glu Glu Gln Leu Phe
405 410 415
Tyr Leu Met Ser Arg Gly Ile Ser Glu Glu Glu Ala Thr Glu Met Ile
420 425 430
Val Met Gly Phe Ile Glu Pro Phe Thr Lys Glu Leu Pro Met Glu Tyr
435 440 445
Ala Val Glu Met Asn Arg Leu Ile Lys Phe Glu Met Glu Gly Ser Ile
450 455 460
Gly
465
<210> 175
<211> 480
<212> DNA
<213> 芽孢杆菌属
<220>
<221> CDS

```

	<222>	(1)..(480)	
	<223>	必需基因 accB (在 accC 操纵子中)	
	<400>	175	
	atg tta aat atc aaa gaa atc cac gag ctg att aaa gca att gac gag	48	
	Met Leu Asn Ile Lys Glu Ile His Glu Leu Ile Lys Ala Ile Asp Glu		
	1 5 10 15		
	tct aca att gac gaa ttc gta tat gaa aat gaa ggt gta tcc tta aaa	96	
	Ser Thr Ile Asp Glu Phe Val Tyr Asn Glu Gly Val Ser Leu Lys		
	20 25 30		
	ctg aaa aaa cac gaa gca ggc acg gtt caa gtc atg cag cag gca ccg	144	
	Leu Lys Lys His Glu Ala Gly Thr Val Gln Val Met Gln Gln Ala Pro		
	35 40 45		
	gca gca cct gta caa gca cag gct ccg cag gca gtt cag ccg caa gcg	192	
	Ala Ala Pro Val Gln Ala Gln Ala Pro Gln Ala Val Gln Pro Gln Ala		
	50 55 60		
	cag caa gca gcg gca cct gcc caa gaa gca cca aag caa gat gag aat	240	
	Gln Gln Ala Ala Ala Pro Ala Gln Glu Ala Pro Lys Gln Asp Glu Asn		
	65 70 75 80		
	ctg cat aaa atc act tca cca atg gta gga aca ttt tat gct tct tca	288	
	Leu His Lys Ile Thr Ser Pro Met Val Gly Thr Phe Tyr Ala Ser Ser		
	85 90 95		
	tca ccg gaa gct ggc ccg tat gta aca gcc ggt tca aaa gta aat gaa	336	
	Ser Pro Glu Ala Gly Pro Tyr Val Thr Ala Gly Ser Lys Val Asn Glu		
	100 105 110		
	aac aca gtt gtc tgc att gta gaa gcg atg aag ctt ttc aac gaa atc	384	
	Asn Thr Val Val Cys Ile Val Glu Ala Met Lys Leu Phe Asn Glu Ile		
	115 120 125		
	gaa gca gaa gtg aaa ggc gaa atc gtt gaa gta tta gta gaa aac ggc	432	
	Glu Ala Glu Val Lys Gly Glu Ile Val Glu Val Leu Val Glu Asn Gly		
	130 135 140		
	cag ctg gtc gaa tac gga caa cct cta ttt ctt gta aaa gcg gag taa	480	
	Gln Leu Val Glu Tyr Gly Gln Pro Leu Phe Leu Val Lys Ala Glu		
	145 150 155		
	<210>	176	
	<211>	159	
	<212>	PRT	
	<213>	芽孢杆菌属	
	<400>	176	
[0482]	Met Leu Asn Ile Lys Glu Ile His Glu Leu Ile Lys Ala Ile Asp Glu		
	1 5 10 15		
	Ser Thr Ile Asp Glu Phe Val Tyr Glu Asn Glu Gly Val Ser Leu Lys		
	20 25 30		
	Leu Lys Lys His Glu Ala Gly Thr Val Gln Val Met Gln Gln Ala Pro		
	35 40 45		
	Ala Ala Pro Val Gln Ala Gln Ala Pro Gln Ala Val Gln Pro Gln Ala		
	50 55 60		
	Gln Gln Ala Ala Ala Pro Ala Gln Glu Ala Pro Lys Gln Asp Glu Asn		
	65 70 75 80		
	Leu His Lys Ile Thr Ser Pro Met Val Gly Thr Phe Tyr Ala Ser Ser		
	85 90 95		
	Ser Pro Glu Ala Gly Pro Tyr Val Thr Ala Gly Ser Lys Val Asn Glu		
	100 105 110		
	Asn Thr Val Val Cys Ile Val Glu Ala Met Lys Leu Phe Asn Glu Ile		
	115 120 125		
	Glu Ala Glu Val Lys Gly Glu Ile Val Glu Val Leu Val Glu Asn Gly		
	130 135 140		
	Gln Leu Val Glu Tyr Gly Gln Pro Leu Phe Leu Val Lys Ala Glu		
	145 150 155		
	<210>	177	
	<211>	1353	
	<212>	DNA	
	<213>	芽孢杆菌属	
	<220>		
	<221>	CDS	
	<222>	(1)..(1353)	
	<223>	必需基因 accC (在 accC 操纵子中)	
	<400>	177	
	atg att aaa aag cta ttg atc gcc aac aga gga gaa att gct gtc aga	48	
	Met Ile Lys Lys Leu Leu Ile Ala Asn Arg Gly Glu Ile Ala Val Arg		
	1 5 10 15		
	atc atc aga gcc tgc aga gag ctc gga att gag act gtc gct gtt tat	96	
	Ile Ile Arg Ala Cys Arg Glu Leu Gly Ile Glu Thr Val Ala Val Tyr		
	20 25 30		
	tca gaa gct gat aaa gat gcc ctt cat gtt caa atg gcc gat gaa gct	144	
	Ser Glu Ala Asp Lys Asp Ala Leu His Val Gln Met Ala Asp Glu Ala		
	35 40 45		
	tttttgt atc gga ccg aaa gca tca aaa gac agc tat tta aac gtt aca	192	

	Phe	Cys	Ile	Gly	Pro	Lys	Ala	Ser	Lys	Asp	Ser	Tyr	Leu	Asn	Val	Thr	
	aat	att	gtg	agt	gtt	gca	aag	ctg	act	ggc	acg	gac	gcc	att	cat	cgc	240
	Asn	Ile	Val	Ser	Val	Ala	Lys	Leu	Thr	Gly	Thr	Asp	Ala	Ile	His	Pro	
	65					70				75					80		
	gga	tac	gga	ttt	tta	gct	gaa	aat	gct	gat	ttc	gct	gaa	tta	tgt	gaa	288
	Gly	Tyr	Gly	Phe	Leu	Ala	Glu	Asn	Ala	Asp	Phe	Ala	Glu	Leu	Cys	Glu	
					85					90					95		
	gaa	gtg	aat	gtc	acg	ttt	gtc	ggc	ccg	agc	gct	gac	gcc	att	tca	aaa	336
	Glu	Val	Asn	Val	Thr	Phe	Val	Gly	Pro	Ser	Ala	Asp	Ala	Ile	Ser	Lys	
				100						105				110			
	atg	gga	aca	aaa	gac	gtt	gcg	cgg	gaa	acg	atg	aaa	cag	gcc	ggc	gtg	384
	Met	Gly	Thr	Lys	Asp	Val	Ala	Arg	Glu	Thr	Met	Lys	Gln	Ala	Gly	Val	
				115					120					125			
	cca	atc	gta	ccg	ggg	tca	cag	gga	att	ata	gaa	aat	gtg	gaa	gaa	gcg	432
	Pro	Ile	Val	Pro	Gly	Ser	Gln	Gly	Ile	Ile	Glu	Asn	Val	Glu	Glu	Ala	
				130									140				
	gtt	tcg	ctt	gct	aat	gaa	att	ggg	tat	cct	gta	att	ata	aaa	gcc	acc	480
	Val	Ser	Leu	Ala	Asn	Glu	Ile	Gly	Tyr	Pro	Val	Ile	Ile	Lys	Ala	Thr	
	145					150					155				160		
	gca	ggc	gga	ggc	gga	aaa	gga	att	agg	gtt	gcc	cgc	act	gaa	gag	gaa	528
	Ala	Gly	Gly	Gly	Gly	Lys	Gly	Ile	Arg	Val	Ala	Arg	Thr	Glu	Glu	Glu	
				165						170				175			
	ctg	att	aat	ggc	att	aag	att	aca	cag	cag	gaa	gcg	gca	act	gca	ttt	576
	Leu	Ile	Asn	Gly	Ile	Lys	Ile	Thr	Gln	Gln	Glu	Ala	Ala	Thr	Ala	Phe	
				180						185				190			
	ggg	aat	cca	ggg	gta	tac	atc	gaa	aaa	tac	ata	gaa	gat	ttt	cgc	cac	624
	Gly	Asn	Pro	Gly	Val	Tyr	Ile	Glu	Lys	Tyr	Ile	Glu	Asp	Phe	Arg	His	
				195					200					205			
	gtt	gag	atc	caa	gtg	ctt	gct	gat	aac	tac	gga	aat	acg	atc	cat	ttg	672
	Val	Glu	Ile	Gln	Val	Leu	Ala	Asp	Asn	Tyr	Gly	Asn	Thr	Ile	His	Leu	
				210					215					220			
	ggc	gaa	cgc	gac	tcg	tcg	caa	aga	cgc	ctg	caa	aag	ctt	ttg	gaa		720
	Gly	Glu	Arg	Asp	Cys	Ser	Ile	Gln	Arg	Arg	Leu	Gln	Lys	Leu	Leu	Glu	
	225					230				235					240		
	gaa	tca	cca	tct	cct	gcc	ctt	gat	tca	gaa	atc	agg	gag	caa	atg	gga	768
	Glu	Ser	Pro	Ser	Pro	Ala	Leu	Asp	Ser	Glu	Ile	Arg	Glu	Gln	Met	Gly	
					245					250					255		
[0483]	gat	gca	gcg	gta	aag	gct	gcc	aaa	gcg	gtt	ggc	tat	aca	ggg	gct	gga	816
	Asp	Ala	Ala	Val	Lys	Ala	Ala	Lys	Ala	Val	Gly	Tyr	Thr	Gly	Ala	Gly	
				260					265					270			
	aca	gtt	gaa	ttt	atc	tat	gac	tac	aat	gaa	cag	cgc	tat	tac	ttc	atg	864
	Thr	Val	Glu	Phe	Ile	Tyr	Asp	Tyr	Asn	Glu	Gln	Arg	Tyr	Tyr	Phe	Met	
				275					280					285			
	gaa	atg	aac	acg	aga	atc	cag	gtg	gag	cac	cca	gtc	aca	gaa	atg	gtg	912
	Glu	Met	Asn	Thr	Arg	Ile	Gln	Val	Glu	His	Pro	Val	Thr	Glu	Met	Val	
				290					295					300			
	acg	gga	act	gac	ctg	atc	aag	gaa	caa	atc	aaa	gtg	gct	tca	gga	atg	960
	Thr	Gly	Thr	Asp	Leu	Ile	Lys	Glu	Gln	Ile	Lys	Val	Ala	Ser	Gly	Met	
	305					310								315			
	gaa	ctg	agc	ctc	aag	caa	gaa	gat	gtt	gaa	ttt	gaa	ggc	tgg	gcc	atc	1008
	Glu	Leu	Ser	Leu	Lys	Gln	Glu	Asp	Val	Glu	Phe	Glu	Gly	Trp	Ala	Ile	
				325						330					335		
	gaa	tcg	cga	atc	aac	gca	gaa	aac	cca	agc	aaa	aat	ttc	atg	ccg	tca	1056
	Glu	Cys	Arg	Ile	Asn	Ala	Glu	Asn	Pro	Ser	Lys	Asn	Phe	Met	Pro	Ser	
				340					345					350			
	cct	ggg	gaa	att	aaa	atg	tac	ctg	cct	cct	ggc	ggg	ctt	ggg	gtc	cgt	1104
	Pro	Gly	Glu	Ile	Lys	Met	Tyr	Leu	Pro	Pro	Gly	Gly	Leu	Gly	Val	Arg	
				355					360					365			
	gtc	gat	tca	gct	gcg	tac	cct	ggc	tat	tcc	atc	cca	ccg	tac	tat	gac	1152
	Val	Asp	Ser	Ala	Ala	Tyr	Pro	Gly	Tyr	Ser	Ile	Pro	Pro	Tyr	Tyr	Asp	
				370					375					380			
	agc	atg	att	gca	aag	gta	atc	aca	tac	gga	aaa	acg	aga	gat	gaa	gcg	1200
	Ser	Met	Ile	Ala	Lys	Val	Ile	Thr	Tyr	Gly	Lys	Thr	Arg	Asp	Glu	Ala	
	385					390								395			
	att	gcc	cgc	atg	aag	cgc	gca	ttg	agc	gaa	ttc	gtc	atc	gaa	ggc	att	1248
	Ile	Ala	Arg	Met	Lys	Arg	Ala	Leu	Ser	Glu	Phe	Val	Ile	Glu	Gly	Ile	
				405						410					415		
	gag	aca	aca	atc	cct	cat	tta	aaa	ctg	ctt	gaa	cac	gaa	aca	ttt		1296
	Glu	Thr	Thr	Ile	Pro	Phe	His	Leu	Lys	Leu	Glu	His	Glu	Thr	Phe		
				420					425					430			
	gtg	agc	gga	gag	ttt	aat	acg	aaa	ttt	tta	gaa	aca	tat	gat	gta	atg	1344
	Val	Ser	Gly	Glu	Phe	Asn	Thr	Lys	Phe	Leu	Glu	Thr	Tyr	Asp	Val	Met	
				435					440					445			
	ggc	tca	taa														1353
	Gly	Ser															
				450													
	<210>	178															

<211>	450	
<212>	PRT	
<213>	芽孢杆菌属	
<400>	178	
Met	Ile Lys Lys Leu Leu Ile Ala Asn Arg Gly Glu Ile Ala Val Arg	
1	5 10 15	
Ile	Ile Arg Ala Cys Arg Glu Leu Gly Ile Glu Thr Val Ala Val Tyr	
	20 25 30	
Ser	Glu Ala Asp Lys Asp Ala Leu His Val Gln Met Ala Asp Glu Ala	
	35 40 45	
Phe	Cys Ile Gly Pro Lys Ala Ser Lys Asp Ser Tyr Leu Asn Val Thr	
	50 55 60	
Asn	Ile Val Ser Val Ala Lys Leu Thr Gly Thr Asp Ala Ile His Pro	
65	70 75 80	
Gly	Tyr Gly Phe Leu Ala Glu Asn Ala Asp Phe Ala Glu Leu Cys Glu	
	85 90 95	
Glu	Val Asn Val Thr Phe Val Gly Pro Ser Ala Asp Ala Ile Ser Lys	
	100 105 110	
Met	Gly Thr Lys Asp Val Ala Arg Glu Thr Met Lys Gln Ala Gly Val	
	115 120 125	
Pro	Ile Val Pro Gly Ser Gln Gly Ile Ile Glu Asn Val Glu Glu Ala	
	130 135 140	
Val	Ser Leu Ala Asn Glu Ile Gly Tyr Pro Val Ile Ile Lys Ala Thr	
145	150 155 160	
Ala	Gly Gly Gly Gly Lys Gly Ile Arg Val Ala Arg Thr Glu Glu Glu	
	165 170 175	
Leu	Ile Asn Gly Ile Lys Ile Thr Gln Gln Glu Ala Ala Thr Ala Phe	
	180 185 190	
Gly	Asn Pro Gly Val Tyr Ile Glu Lys Tyr Ile Glu Asp Phe Arg His	
	195 200 205	
Val	Glu Ile Gln Val Leu Ala Asp Asn Tyr Gly Asn Thr Ile His Leu	
	210 215 220	
Gly	Glu Arg Asp Cys Ser Ile Gln Arg Arg Leu Gln Lys Leu Leu Glu	
225	230 235 240	
Glu	Ser Pro Ser Pro Ala Leu Asp Ser Glu Ile Arg Glu Gln Met Gly	
	245 250 255	
Asp	Ala Ala Val Lys Ala Ala Lys Ala Val Gly Tyr Thr Gly Ala Gly	
	260 265 270	
Thr	Val Glu Phe Ile Tyr Asp Tyr Asn Glu Gln Arg Tyr Tyr Phe Met	
	275 280 285	
Glu	Met Asn Thr Arg Ile Gln Val Glu His Pro Val Thr Glu Met Val	
	290 295 300	
Thr	Gly Thr Asp Leu Ile Lys Glu Gln Ile Lys Val Ala Ser Gly Met	
305	310 315 320	
Glu	Leu Ser Leu Lys Gln Glu Asp Val Glu Phe Glu Gly Trp Ala Ile	
	325 330 335	
Glu	Cys Arg Ile Asn Ala Glu Asn Pro Ser Lys Asn Phe Met Pro Ser	
	340 345 350	
Pro	Gly Glu Ile Lys Met Tyr Leu Pro Pro Gly Gly Leu Gly Val Arg	
	355 360 365	
Val	Asp Ser Ala Ala Tyr Pro Gly Tyr Ser Ile Pro Pro Tyr Tyr Asp	
	370 375 380	
Ser	Met Ile Ala Lys Val Ile Thr Tyr Gly Lys Thr Arg Asp Glu Ala	
385	390 395 400	
Ile	Ala Arg Met Lys Arg Ala Leu Ser Glu Phe Val Ile Glu Gly Ile	
	405 410 415	
Glu	Thr Thr Ile Pro Phe His Leu Lys Leu Leu Glu His Glu Thr Phe	
	420 425 430	
Val	Ser Gly Glu Phe Asn Thr Lys Phe Leu Glu Thr Tyr Asp Val Met	
	435 440 445	
Gly	Ser	
	450	
<210>	179	
<211>	408	
<212>	DNA	
<213>	芽孢杆菌属	
<220>		
<221>	CDS	
<222>	(1) .. (408)	
<223>	必需基因 yghY (在 accC 操纵子中)	
<400>	179	
atg	aaa gac aac agc ttg ctt aaa atg gat cac gaa gat acg cat ttg	48
Met	Lys Asp Asn Ser Leu Leu Lys Met Asp His Glu Asp Thr His Leu	
1	5 10 15	
ggt	aag gtt gag att gca ccg gag gtc att gag gtc att gcc ggc att	96
Gly	Lys Val Glu Ile Ala Pro Glu Val Ile Glu Val Ile Ala Gly Ile	
	20 25 30	
gca	gca tca gaa gtt gat gga gtt gcc gaa atg cgc gga aac ttt gcg	144

	Ala	Ala	Ser	Glu	Val	Asp	Gly	Val	Ala	Glu	Met	Arg	Gly	Asn	Phe	Ala	
			35				40						45				
	aca	ggc	gtc	ggt	gag	cgc	ttt	ggc	aaa	gtg	aat	cat	ggc	aaa	ggg	gtc	192
	Thr	Gly	Val	Val	Glu	Arg	Phe	Gly	Lys	Val	Asn	His	Gly	Lys	Gly	Val	
		50					55					60					
	aaa	gtc	gac	ctg	gcg	gat	gac	gga	att	acg	atc	gat	gta	tac	tgt	gtc	240
	Lys	Val	Asp	Leu	Ala	Asp	Asp	Gly	Ile	Thr	Ile	Asp	Val	Tyr	Cys	Val	
		65					70				75				80		
	ggt	aca	ttt	ggc	ggt	tcg	att	ccg	aaa	ggt	gca	gca	tcc	ggt	cag	gaa	288
	Val	Thr	Phe	Gly	Val	Ser	Ile	Pro	Lys	Val	Ala	Ala	Ser	Val	Gln	Glu	
				85						90				95			
	aac	att	cgt	caa	acc	tta	tta	aat	atg	acg	tct	ctt	tcg	att	aac	gag	336
	Asn	Ile	Arg	Gln	Thr	Leu	Leu	Asn	Met	Thr	Ser	Leu	Ser	Ile	Asn	Glu	
				100					105					110			
	atc	aat	att	cac	atc	gtc	ggc	att	caa	ttt	gac	aca	aaa	gcc	caa	gaa	384
	Ile	Asn	Ile	His	Ile	Val	Gly	Ile	Gln	Phe	Asp	Thr	Lys	Ala	Gln	Glu	
				115				120					125				
	gtc	caa	atc	gac	gaa	gaa	atg	taa									408
	Val	Gln	Ile	Asp	Glu	Glu	Met										
		130					135										
	<210>	180															
	<211>	135															
	<212>	PRT															
	<213>	芽孢杆菌属															
	<400>	180															
	Met	Lys	Asp	Asn	Ser	Leu	Leu	Lys	Met	Asp	His	Glu	Asp	Thr	His	Leu	
				5						10					15		
	Gly	Lys	Val	Glu	Ile	Ala	Pro	Glu	Val	Ile	Glu	Val	Ile	Ala	Gly	Ile	
				20					25					30			
	Ala	Ala	Ser	Glu	Val	Asp	Gly	Val	Ala	Glu	Met	Arg	Gly	Asn	Phe	Ala	
			35				40						45				
	Thr	Gly	Val	Val	Glu	Arg	Phe	Gly	Lys	Val	Asn	His	Gly	Lys	Gly	Val	
		50					55					60					
	Lys	Val	Asp	Leu	Ala	Asp	Asp	Gly	Ile	Thr	Ile	Asp	Val	Tyr	Cys	Val	
		65				70				75				80			
	Val	Thr	Phe	Gly	Val	Ser	Ile	Pro	Lys	Val	Ala	Ala	Ser	Val	Gln	Glu	
				85						90				95			
	Asn	Ile	Arg	Gln	Thr	Leu	Leu	Asn	Met	Thr	Ser	Leu	Ser	Ile	Asn	Glu	
				100					105					110			
[0485]	Ile	Asn	Ile	His	Ile	Val	Gly	Ile	Gln	Phe	Asp	Thr	Lys	Ala	Gln	Glu	
				115				120					125				
	Val	Gln	Ile	Asp	Glu	Glu	Met										
		130					135										
	<210>	181															
	<211>	1344															
	<212>	DNA															
	<213>	棒状杆菌属 (Corynebacterium)															
	<220>																
	<221>	CDS															
	<222>	(1) .. (1344)															
	<223>	棒状杆菌属必需基因 glmM															
	<400>	181															
	atg	act	cga	ctt	ttt	gga	act	gat	ggc	gtc	cgc	gga	cta	gcc	aat	gaa	48
	Met	Thr	Arg	Leu	Phe	Gly	Thr	Asp	Gly	Val	Arg	Gly	Leu	Ala	Asn	Glu	
		1			5					10				15			
	gta	ctc	acc	gca	cct	ttg	gcc	ctg	aaa	ctg	ggg	gcc	gct	gca	gct	cac	96
	Val	Leu	Thr	Ala	Pro	Leu	Ala	Leu	Lys	Leu	Gly	Ala	Ala	Ala	Ala	His	
				20					25					30			
	gta	ctt	acc	gca	gag	aaa	cgc	gta	gat	gga	cgc	cgc	ccg	ggt	gcg	atc	144
	Val	Leu	Thr	Ala	Glu	Lys	Arg	Val	Asp	Gly	Arg	Arg	Pro	Val	Ala	Ile	
				35				40					45				
	ggt	ggg	agg	gat	cct	cga	gtc	tct	gga	gag	atg	ctc	gcg	gca	gca	ctt	192
	Val	Gly	Arg	Asp	Pro	Arg	Val	Ser	Gly	Glu	Met	Leu	Ala	Ala	Ala	Leu	
		50					55					60					
	tca	gca	ggc	atg	gcc	agc	cag	ggg	ggt	gat	gtc	att	cgt	ggt	ggg	gtc	240
	Ser	Ala	Gly	Met	Ala	Ser	Gln	Gly	Val	Asp	Val	Ile	Arg	Val	Gly	Val	
		65				70				75				80			
	atc	cca	acc	cca	gct	ggt	gca	ttc	ctc	acc	gat	gat	tat	ggc	gct	gat	288
	Ile	Pro	Thr	Pro	Ala	Val	Ala	Phe	Leu	Thr	Asp	Asp	Tyr	Gly	Ala	Asp	
				85						90				95			
	atg	ggc	gtg	atg	att	tct	gca	tcc	cac	aac	cca	atg	ccg	gac	aac	gga	336
	Met	Gly	Val	Met	Ile	Ser	Ala	Ser	His	Asn	Pro	Met	Pro	Asp	Asn	Gly	
				100					105					110			
	atc	aag	ttc	ttc	tcc	gca	ggg	gga	cac	aag	ctt	cca	gac	cat	gtg	gaa	384
	Ile	Lys	Phe	Phe	Ser	Ala	Gly	Gly	His	Lys	Leu	Pro	Asp	His	Val	Glu	
				115				120					125				
	gac	gag	att	gag	cgt	gtc	atg	gac	agc	ttg	cct	gcc	gaa	ggc	cca	acc	432
	Asp	Glu	Ile	Glu	Arg	Val	Met	Asp	Ser	Leu	Pro	Ala	Glu	Gly	Pro	Thr	

130
 ggg cat gga gtt ggc cgt gtc att gaa gaa gca acc gat gca caa gac 480
 Gly His Gly Val Gly Arg Val Ile Glu Glu Ala Thr Asp Ala Gln Asp
 145
 cgc tac ctg gag cac ctg aag gaa gct gtt cct acg tca ctt gaa ggc 528
 Arg Tyr Leu Glu His Leu Lys Glu Ala Val Pro Thr Ser Leu Glu Gly
 165
 atc aag att gtt gtg gat gca gcc aat ggt gca gca agc gtt gta gct 576
 Ile Lys Ile Val Val Asp Ala Ala Asn Gly Ala Ala Ser Val Val Ala
 180
 cca aag gct tat gaa gca gca ggt gca act gta att gct att cat aac 624
 Pro Lys Ala Tyr Glu Ala Ala Gly Ala Thr Val Ile Ala Ile His Asn
 195
 aag cca gac tcc tac aac atc aac atg gat tgc ggt tcc acc cac att 672
 Lys Pro Asp Ser Tyr Asn Ile Asn Met Asp Cys Gly Ser Thr His Ile
 210
 gac cag gtg cag gcg gca gtt ctg aag cac ggt gct gac ctt gga ctc 720
 Asp Gln Val Gln Ala Ala Val Leu Lys His Gly Ala Asp Leu Gly Leu
 225
 gca cac gac ggc gat gct gac cgt tgt ttg gct gtg gac aaa gat gga 768
 Ala His Asp Gly Asp Ala Asp Arg Cys Leu Ala Val Asp Lys Asp Gly
 245
 aac ctc att gac ggt gat cag atc atg gca atg tta gcc att gcg atg 816
 Asn Leu Ile Asp Gly Asp Gln Ile Met Ala Met Leu Ala Ile Ala Met
 260
 aaa gaa aac ggc gag ctg cgc aag aac acc ctc gtg ggc acc gtc atg 864
 Lys Glu Asn Gly Glu Leu Arg Lys Asn Thr Leu Val Gly Thr Val Met
 275
 agc aac ttg gga ctg aag att gcc atg gat gaa gcc gga att acg ctg 912
 Ser Asn Leu Gly Leu Lys Ile Ala Met Asp Glu Ala Gly Ile Thr Leu
 290
 cga acc acc aag gtg ggc gac cgc tac gtg ctg gaa gac ctc aat gcg 960
 Arg Thr Thr Lys Val Gly Asp Arg Tyr Val Leu Glu Asp Leu Asn Ala
 305
 ggt gga ttc agc ctt ggt ggc gag cag tct gga cac att gtt ctt cca 1008
 Gly Gly Phe Ser Leu Gly Gly Glu Gln Ser Gly His Ile Val Leu Pro
 325
 gat cac ggc acc act ggc gat gga act ttg act ggt ctt tcc atc atg 1056
 Asp His Gly Thr Thr Gly Asp Gly Thr Leu Thr Gly Leu Ser Ile Met
 340
 gcg cgc atg gct gaa acc gga aaa tcc ttg ggc gag ttg gca caa gcg 1104
 Ala Arg Met Ala Glu Thr Gly Lys Ser Leu Gly Glu Ala Ala Gln Ala
 355
 atg acg gtt ctg cca cag gtt ctg atc aat gtg cca gtc tct gat aaa 1152
 Met Thr Val Leu Pro Gln Val Leu Ile Asn Val Pro Val Ser Asp Lys
 370
 tcc acc atc gtt tct cac cca agt gtc gtt gca gca atc gct gaa gct 1200
 Ser Thr Ile Val Ser His Pro Ser Val Val Ala Ala Ile Ala Glu Ala
 385
 gaa gcg gag ttg ggt tct acc ggt cgc gtg ctt ctt cgc gct tct ggc 1248
 Glu Ala Glu Leu Gly Ser Thr Gly Arg Val Leu Leu Arg Ala Ser Gly
 405
 acc gag gag ctt ttc cgt gtg atg gtc gaa gct ggc gat aag gaa caa 1296
 Thr Glu Glu Leu Phe Arg Val Met Val Glu Ala Gly Asp Lys Glu Gln
 420
 gca cgc cgc att gcg gga cgt ctt tct gcc gtg gtc gcg gaa gtc taa 1344
 Ala Arg Arg Ile Ala Gly Arg Leu Ser Ala Val Val Ala Glu Val
 435
 <210> 182
 <211> 447
 <212> PRT
 <213> 棒状杆菌属
 <400> 182
 Met Thr Arg Leu Phe Gly Thr Asp Gly Val Arg Gly Leu Ala Asn Glu
 1 5
 Val Leu Thr Ala Pro Leu Ala Leu Lys Leu Gly Ala Ala Ala His
 20 25
 Val Leu Thr Ala Glu Lys Arg Val Asp Gly Arg Arg Pro Val Ala Ile
 35 40
 Val Gly Arg Asp Pro Arg Val Ser Gly Glu Met Leu Ala Ala Ala Leu
 50 55
 Ser Ala Gly Met Ala Ser Gln Gly Val Asp Val Ile Arg Val Gly Val
 65 70
 Ile Pro Thr Pro Ala Val Ala Phe Leu Thr Asp Asp Tyr Gly Ala Asp
 85 90
 Met Gly Val Met Ile Ser Ala Ser His Asn Pro Met Pro Asp Asn Gly
 100 105
 Ile Lys Phe Phe Ser Ala Gly Gly His Lys Leu Pro Asp His Val Glu

[illegible]

	Ala	Ala	Ala	Pro	His	Ile	Glu	Val	Thr	Gly	Met	Phe	Thr	His	Leu	Ala	
	145					150					155					160	
	tgc	gcg	gat	gag	cca	gag	aat	ccg	gaa	act	gat	cgc	caa	att	att	gct	528
	Cys	Ala	Asp	Glu	Pro	Glu	Asn	Pro	Glu	Thr	Asp	Arg	Gln	Ile	Ile	Ala	
					165					170					175		
	ttt	cga	cgc	gcc	ctt	gcg	ctc	gcc	cgc	aag	cac	ggg	ctt	gag	tgc	ccg	576
	Phe	Arg	Arg	Ala	Leu	Ala	Leu	Ala	Arg	Lys	His	Gly	Leu	Glu	Cys	Pro	
				180					185					190			
	gtc	aac	cac	gta	tgc	aac	tca	cct	gca	ttt	ttg	act	cga	tct	gat	tta	624
	Val	Asn	His	Val	Cys	Asn	Ser	Pro	Ala	Phe	Leu	Thr	Arg	Ser	Asp	Leu	
			195					200					205				
	cac	atg	gag	atg	gtc	cga	ccg	ggt	ttg	gcc	ttt	tat	ggg	ttg	gaa	ccc	672
	His	Met	Glu	Met	Val	Arg	Pro	Gly	Leu	Ala	Phe	Tyr	Gly	Leu	Glu	Pro	
		210					215					220					
	gtg	gcg	gga	cga	gag	cat	ggt	ttg	aag	ccg	gcg	atg	acg	tgg	gag	gcg	720
	Val	Ala	Gly	Arg	Glu	His	Gly	Leu	Lys	Pro	Ala	Met	Thr	Trp	Glu	Ala	
		225				230					235				240		
	aag	gtg	agc	gtc	gta	aag	caa	att	gaa	gct	gga	caa	ggc	acc	tct	tat	768
	Lys	Val	Ser	Val	Val	Lys	Gln	Ile	Glu	Ala	Gly	Gln	Gly	Thr	Ser	Tyr	
				245						250				255			
	ggc	ctg	acc	tgg	cgc	gct	gag	gat	cgc	ggc	ttt	gtg	gct	gtg	gtg	cct	816
	Gly	Leu	Thr	Trp	Arg	Ala	Glu	Asp	Arg	Gly	Phe	Val	Ala	Val	Val	Pro	
				260					265					270			
	gcg	ggc	tat	gcc	gat	ggc	atg	ccg	cgg	cat	gcc	cag	ggg	aaa	ttc	tcc	864
	Ala	Gly	Tyr	Ala	Asp	Gly	Met	Pro	Arg	His	Ala	Gln	Gly	Lys	Phe	Ser	
			275					280					285				
	gtc	acg	att	gat	ggc	gtg	gac	tat	ccg	cag	gtt	gga	cgc	gtg	tgc	atg	912
	Val	Thr	Ile	Asp	Gly	Val	Asp	Tyr	Pro	Gln	Val	Gly	Arg	Val	Cys	Met	
		290				295					300						
	gat	cag	ttc	gtt	att	tct	ttg	ggc	gac	aat	ccg	cac	ggt	gtg	gaa	gct	960
	Asp	Gln	Phe	Val	Ile	Ser	Leu	Gly	Asp	Asn	Pro	His	Gly	Val	Glu	Ala	
		305				310				315					320		
	ggc	gcg	aag	gcc	gtg	ata	ttc	ggt	gag	aat	ggg	cat	gac	gca	act	gat	1008
	Gly	Ala	Lys	Ala	Val	Ile	Phe	Gly	Glu	Asn	Gly	His	Asp	Ala	Thr	Asp	
				325					330					335			
	ttt	gcg	gag	cgt	tta	gac	acc	att	aac	tat	gag	gta	gtg	tgc	cga	cca	1056
	Phe	Ala	Glu	Arg	Leu	Asp	Thr	Ile	Asn	Tyr	Glu	Val	Val	Cys	Arg	Pro	
				340					345					350			
[0488]	acc	ggc	cga	act	gtc	cgc	gca	tat	gtt	taa							1086
	Thr	Gly	Arg	Thr	Val	Arg	Ala	Tyr	Val								
			355					360									
	<210>	184															
	<211>	361															
	<212>	PRT															
	<213>	棒状杆菌属															
	<400>	184															
	Met	Asn	Leu	Leu	Thr	Thr	Lys	Ile	Asp	Leu	Asp	Ala	Ile	Ala	His	Asn	
	1				5					10					15		
	Thr	Arg	Val	Leu	Lys	Gln	Met	Ala	Gly	Pro	Ala	Lys	Leu	Met	Ala	Val	
				20					25					30			
	Val	Lys	Ala	Asn	Ala	Tyr	Asn	His	Gly	Val	Glu	Lys	Val	Ala	Pro	Val	
			35					40					45				
	Ile	Ala	Ala	His	Gly	Ala	Asp	Ala	Phe	Gly	Val	Ala	Thr	Leu	Ala	Glu	
		50					55					60					
	Ala	Met	Gln	Leu	Arg	Asp	Ile	Gly	Ile	Ser	Gln	Glu	Val	Leu	Cys	Trp	
	65				70					75				80			
	Ile	Trp	Thr	Pro	Glu	Gln	Asp	Phe	Arg	Ala	Ala	Ile	Asp	Arg	Asn	Ile	
				85					90					95			
	Asp	Leu	Ala	Val	Ile	Ser	Pro	Ala	His	Ala	Lys	Ala	Leu	Ile	Asp	Thr	
				100				105						110			
	Asp	Ala	Glu	His	Ile	Arg	Val	Ser	Ile	Lys	Ile	Asp	Ser	Gly	Leu	His	
		115					120						125				
	Arg	Ser	Gly	Val	Asp	Glu	Gln	Glu	Trp	Glu	Gly	Val	Phe	Ser	Ala	Leu	
		130					135					140					
	Ala	Ala	Ala	Pro	His	Ile	Glu	Val	Thr	Gly	Met	Phe	Thr	His	Leu	Ala	
	145					150				155					160		
	Cys	Ala	Asp	Glu	Pro	Glu	Asn	Pro	Glu	Thr	Asp	Arg	Gln	Ile	Ile	Ala	
				165						170				175			
	Phe	Arg	Arg	Ala	Leu	Ala	Leu	Ala	Arg	Lys	His	Gly	Leu	Glu	Cys	Pro	
				180				185					190				
	Val	Asn	His	Val	Cys	Asn	Ser	Pro	Ala	Phe	Leu	Thr	Arg	Ser	Asp	Leu	
			195					200					205				
	His	Met	Glu	Met	Val	Arg	Pro	Gly	Leu	Ala	Phe	Tyr	Gly	Leu	Glu	Pro	
		210					215					220					
	Val	Ala	Gly	Arg	Glu	His	Gly	Leu	Lys	Pro	Ala	Met	Thr	Trp	Glu	Ala	
		225				230					235				240		
	Lys	Val	Ser	Val	Val	Lys	Gln	Ile	Glu	Ala	Gly	Gln	Gly	Thr	Ser	Tyr	
				245						250					255		

	Gly	Leu	Thr	Trp	Arg	Ala	Glu	Asp	Arg	Gly	Phe	Val	Ala	Val	Val	Pro	
				260					265					270			
	Ala	Gly	Tyr	Ala	Asp	Gly	Met	Pro	Arg	His	Ala	Gln	Gly	Lys	Phe	Ser	
			275					280					285				
	Val	Thr	Ile	Asp	Gly	Val	Asp	Tyr	Pro	Gln	Val	Gly	Arg	Val	Cys	Met	
		290					295					300					
	Asp	Gln	Phe	Val	Ile	Ser	Leu	Gly	Asp	Asn	Pro	His	Gly	Val	Glu	Ala	
		305				310					315					320	
	Gly	Ala	Lys	Ala	Val	Ile	Phe	Gly	Glu	Asn	Gly	His	Asp	Ala	Thr	Asp	
				325						330					335		
	Phe	Ala	Glu	Arg	Leu	Asp	Thr	Ile	Asn	Tyr	Glu	Val	Val	Cys	Arg	Pro	
				340					345					350			
	Thr	Gly	Arg	Thr	Val	Arg	Ala	Tyr	Val								
		355						360									
	<210>	185															
	<211>	1119															
	<212>	DNA															
	<213>	米曲霉 (Aspergillus oryzae)															
	<220>																
	<221>	CDS															
	<222>	(1)..(1119)															
	<223>	曲霉属必需基因 ARG3															
	<400>	185															
	atg	act	tgc	ggg	ctg	aag	ctg	gcc	gcc	gct	cgt	tac	ggc	aac	cat	acc	48
	Met	Thr	Cys	Gly	Leu	Lys	Leu	Ala	Ala	Ala	Arg	Tyr	Gly	Asn	His	Thr	
	1			5						10				15			
	ttg	cgc	cag	aag	att	cca	ctc	aac	gct	gtg	cgc	cga	tac	act	tct	cac	96
	Leu	Arg	Gln	Lys	Ile	Pro	Leu	Asn	Ala	Val	Arg	Arg	Tyr	Thr	Ser	His	
			20					25					30				
	act	gct	act	tct	acg	act	ccc	ccc	acc	tcg	cca	ttc	gct	cct	cgc	cat	144
	Thr	Ala	Thr	Thr	Thr	Thr	Pro	Pro	Thr	Ser	Pro	Phe	Ala	Pro	Arg	His	
			35				40					45					
	ttt	ctt	tcc	atc	gct	gat	ctt	acc	tca	acg	gag	ttc	gca	acc	ctc	gtc	192
	Phe	Leu	Ser	Ile	Ala	Asp	Leu	Thr	Ser	Thr	Glu	Phe	Ala	Thr	Leu	Val	
		50				55					60						
	cgt	aat	gca	tct	tca	cac	aaa	cga	acc	atc	aag	tcg	gga	tca	atc	cct	240
	Arg	Asn	Ala	Ser	Ser	His	Lys	Arg	Thr	Ile	Lys	Ser	Gly	Ser	Ile	Pro	
	65			70						75					80		
[0489]	cag	aac	ctg	ctt	ggg	tct	atg	acc	gga	cag	act	gta	gct	atg	ctg	ttc	288
	Gln	Asn	Leu	Leu	Gly	Ser	Met	Thr	Gly	Gln	Thr	Val	Ala	Met	Leu	Phe	
				85					90					95			
	agc	aaa	cgc	agt	acc	aga	acc	agg	ata	tct	act	gaa	ggg	gct	gtg	gtg	336
	Ser	Lys	Arg	Ser	Thr	Arg	Thr	Arg	Ile	Ser	Thr	Glu	Gly	Ala	Val	Val	
			100				105						110				
	cgc	ttg	gga	gga	cat	cca	atg	ttt	ctg	ggc	aag	gat	gat	att	caa	ctc	384
	Arg	Leu	Gly	Gly	His	Pro	Met	Phe	Leu	Gly	Lys	Asp	Asp	Ile	Gln	Leu	
			115				120					125					
	ggg	gtc	aat	gag	tcc	ttg	tat	gac	tct	gcg	ggt	gta	atc	tcc	tcc	atg	432
	Gly	Val	Asn	Glu	Ser	Leu	Tyr	Asp	Ser	Ala	Val	Val	Ile	Ser	Ser	Met	
		130					135					140					
	gtc	tct	tgt	att	gta	gcc	cga	gtc	ggc	aag	cac	gct	gag	gtt	gca	gac	480
	Val	Ser	Cys	Ile	Val	Ala	Arg	Val	Gly	Lys	His	Ala	Glu	Val	Ala	Asp	
		145				150					155					160	
	ctt	gcc	aag	cat	tca	acc	gtg	cct	gtt	atc	aac	gct	ctc	tgt	gac	tcc	528
	Leu	Ala	Lys	His	Ser	Thr	Val	Pro	Val	Ile	Asn	Ala	Leu	Cys	Asp	Ser	
				165					170					175			
	ttc	cac	cct	ctc	cag	gcc	att	gcc	gat	ttt	cag	acc	atc	tac	gaa	aca	576
	Phe	His	Pro	Leu	Gln	Ala	Ile	Ala	Asp	Phe	Gln	Thr	Ile	Tyr	Glu	Thr	
				180					185					190			
	ttt	aca	cct	aag	gcc	cat	cgc	tct	gac	agt	ttg	ggg	ctg	gag	ggc	ctc	624
	Phe	Thr	Pro	Lys	Ala	His	Arg	Ser	Asp	Ser	Leu	Gly	Leu	Glu	Gly	Leu	
			195				200						205				
	aag	att	gcc	tgg	gtc	ggg	gat	gca	aac	aac	ggt	ctg	ttc	gac	atg	gca	672
	Lys	Ile	Ala	Trp	Val	Gly	Asp	Ala	Asn	Asn	Val	Leu	Phe	Asp	Met	Ala	
		210				215					220						
	att	gct	gct	acc	aag	atg	ggg	att	gat	atc	gcc	ggt	gcg	act	ccg	aag	720
	Ile	Ala	Ala	Thr	Lys	Met	Gly	Ile	Asp	Ile	Ala	Val	Ala	Thr	Pro	Lys	
		225				230					235				240		
	ggc	tat	gag	att	cct	gcc	ccg	atg	ttg	gag	ctc	atc	aag	cag	gcc	agc	768
	Gly	Tyr	Glu	Ile	Pro	Ala	Pro	Met	Leu	Glu	Leu	Ile	Lys	Gln	Ala	Ser	
				245					250					255			
	aac	ggg	gtc	tca	aag	cca	gga	aag	att	att	gag	acc	aac	ggt	cct	gag	816
	Asn	Gly	Val	Ser	Lys	Pro	Gly	Lys	Ile	Ile	Glu	Thr	Asn	Val	Pro	Glu	
				260					265					270			
	gag	gcg	ggt	aag	ggg	gcg	gat	atc	ctg	gta	act	gac	acc	tgg	ggt	tcc	864
	Glu	Ala	Val	Lys	Gly	Ala	Asp	Ile	Leu	Val	Thr	Asp	Thr	Trp	Val	Ser	
			275				280						285				
	atg	gga	cag	gag	gca	gag	tcg	atc	aag	cgg	gtg	aag	gac	ttc	gag	ggc	912

	Met	Gly	Gln	Glu	Ala	Glu	Ser	Ile	Lys	Arg	Val	Lys	Asp	Phe	Glu	Gly	
	290						295					300					
	ttc	cag	atc	act	tcg	gaa	ctt	gcc	aag	cgc	ggg	ggg	gcg	aac	gaa	ggc	960
	Phe	Gln	Ile	Thr	Ser	Glu	Leu	Ala	Lys	Arg	Gly	Gly	Ala	Asn	Glu	Gly	
	305					310					315					320	
	tgg	aag	ttc	atg	cac	tgc	ctg	cct	cgt	cac	cct	gag	gaa	gtc	agc	gat	1008
	Trp	Lys	Phe	Met	His	Cys	Leu	Pro	Arg	His	Pro	Glu	Glu	Val	Ser	Asp	
					325					330					335		
	gaa	gtt	ttc	tac	agc	cca	cgg	tcc	ttg	gtc	ttc	ccc	gaa	gcg	gag	aat	1056
	Glu	Val	Phe	Tyr	Ser	Pro	Arg	Ser	Leu	Val	Phe	Pro	Glu	Ala	Glu	Asn	
				340					345					350			
	cgg	ttg	tgg	gct	gcg	atc	tcg	gcc	atg	gag	ggc	ttt	gtt	gtc	aat	aag	1104
	Arg	Leu	Trp	Ala	Ala	Ile	Ser	Ala	Met	Glu	Gly	Phe	Val	Val	Asn	Lys	
			355					360					365				
	gga	agg	att	gag	taa												1119
	Gly	Arg	Ile	Glu													
	370																
	<210>	186															
	<211>	372															
	<212>	PRT															
	<213>	米曲霉															
	<400>	186															
	Met	Thr	Cys	Gly	Leu	Lys	Leu	Ala	Ala	Ala	Arg	Tyr	Gly	Asn	His	Thr	
	1				5					10					15		
	Leu	Arg	Gln	Lys	Ile	Pro	Leu	Asn	Ala	Val	Arg	Arg	Tyr	Thr	Ser	His	
				20					25					30			
	Thr	Ala	Thr	Ser	Thr	Thr	Pro	Pro	Thr	Ser	Pro	Phe	Ala	Pro	Arg	His	
				35					40				45				
	Phe	Leu	Ser	Ile	Ala	Asp	Leu	Thr	Ser	Thr	Glu	Phe	Ala	Thr	Leu	Val	
		50					55					60					
	Arg	Asn	Ala	Ser	Ser	His	Lys	Arg	Thr	Ile	Lys	Ser	Gly	Ser	Ile	Pro	
	65					70					75				80		
	Gln	Asn	Leu	Leu	Gly	Ser	Met	Thr	Gly	Gln	Thr	Val	Ala	Met	Leu	Phe	
				85						90					95		
	Ser	Lys	Arg	Ser	Thr	Arg	Thr	Arg	Ile	Ser	Thr	Glu	Gly	Ala	Val	Val	
				100					105					110			
	Arg	Leu	Gly	Gly	His	Pro	Met	Phe	Leu	Gly	Lys	Asp	Asp	Ile	Gln	Leu	
			115					120					125				
[0490]	Gly	Val	Asn	Glu	Ser	Leu	Tyr	Asp	Ser	Ala	Val	Val	Ile	Ser	Ser	Met	
		130					135					140					
	Val	Ser	Cys	Ile	Val	Ala	Arg	Val	Gly	Lys	His	Ala	Glu	Val	Ala	Asp	
		145				150					155				160		
	Leu	Ala	Lys	His	Ser	Thr	Val	Pro	Val	Ile	Asn	Ala	Leu	Cys	Asp	Ser	
				165						170				175			
	Phe	His	Pro	Leu	Gln	Ala	Ile	Ala	Asp	Phe	Gln	Thr	Ile	Tyr	Glu	Thr	
				180					185					190			
	Phe	Thr	Pro	Lys	Ala	His	Arg	Ser	Asp	Ser	Leu	Gly	Leu	Glu	Gly	Leu	
			195					200					205				
	Lys	Ile	Ala	Trp	Val	Gly	Asp	Ala	Asn	Asn	Val	Leu	Phe	Asp	Met	Ala	
		210					215					220					
	Ile	Ala	Ala	Thr	Lys	Met	Gly	Ile	Asp	Ile	Ala	Val	Ala	Thr	Pro	Lys	
		225				230					235				240		
	Gly	Tyr	Glu	Ile	Pro	Ala	Pro	Met	Leu	Glu	Leu	Ile	Lys	Gln	Ala	Ser	
				245						250				255			
	Asn	Gly	Val	Ser	Lys	Pro	Gly	Lys	Ile	Ile	Glu	Thr	Asn	Val	Pro	Glu	
				260				265						270			
	Glu	Ala	Val	Lys	Gly	Ala	Asp	Ile	Leu	Val	Thr	Asp	Thr	Trp	Val	Ser	
			275				280					285					
	Met	Gly	Gln	Glu	Ala	Glu	Ser	Ile	Lys	Arg	Val	Lys	Asp	Phe	Glu	Gly	
		290					295					300					
	Phe	Gln	Ile	Thr	Ser	Glu	Leu	Ala	Lys	Arg	Gly	Gly	Ala	Asn	Glu	Gly	
		305				310					315					320	
	Trp	Lys	Phe	Met	His	Cys	Leu	Pro	Arg	His	Pro	Glu	Glu	Val	Ser	Asp	
				325						330					335		
	Glu	Val	Phe	Tyr	Ser	Pro	Arg	Ser	Leu	Val	Phe	Pro	Glu	Ala	Glu	Asn	
				340				345						350			
	Arg	Leu	Trp	Ala	Ala	Ile	Ser	Ala	Met	Glu	Gly	Phe	Val	Val	Asn	Lys	
			355					360					365				
	Gly	Arg	Ile	Glu													
	370																
	<210>	187															
	<211>	2184															
	<212>	DNA															
	<213>	米曲霉															
	<220>																
	<221>	CDS															
	<222>	(1)..(2184)															
	<223>	曲霉属必需基因 adeA															

159

160

[0493]

```

Leu Leu Arg Leu Ala Gly Ile Gly Leu His Gly Gln Thr Asp Leu Leu
100 105 110
Gly Ile Thr His Thr Val Thr Arg Leu Glu His Ser Ile Gly Ala Ser
115 120 125
Leu Leu Val Arg Lys Val Gly Gly Ser Ile Gly Glu Gln Val Ala Gly
130 135 140
Leu Leu His Asp Ile Ser His Thr Val Leu Ser His Asp Val Asp Gly
145 150 155 160
Ala Leu Ser Lys Pro Gly Glu Ser Tyr His Glu Val Gln Lys Ser Arg
165 170 175
Tyr Ile Met Thr Thr Glu Leu Pro Arg Ile Leu Thr Lys His Gly Phe
180 185 190
Val Asp Leu Lys Pro Phe Asp Glu Leu Tyr Pro Leu Val Glu Arg
195 200 205
Pro Ala Pro His Leu Cys Ala Asp Arg Leu Asp Tyr Ser Leu Arg Asp
210 215 220
Ala Val Ala Phe Gly Lys Leu Ala Ile Glu Asp Ala Arg Arg Val Tyr
225 230 235 240
Asp Ser Leu Thr Ala Phe Pro Asp Ala Ser Ser Pro His Arg Leu Leu
245 250 255
Val Leu Arg Asp Ile Asp Leu Ala Leu Ala Tyr Ala Arg Ala Tyr Gly
260 265 270
Glu Cys Asp Arg Asp Val Trp Cys Asn Pro Ala His Ala Val Met Ser
275 280 285
Arg Lys Ile Gly Gln Leu Ile Gly Asp Leu Val Gln Gln Gly Ser Leu
290 295 300
Lys Glu Glu Val Leu Trp Asn Leu Ser Asp Arg Glu Phe Trp Glu Leu
305 310 315 320
Leu Lys Ser Lys Val Asp Ser Lys Gly Leu Glu Thr Ile Lys His Ile
325 330 335
Glu Ala Gly Pro His Ala Glu Asp Tyr His Arg Leu Pro Arg Gly Thr
340 345 350
Lys Ile Arg Thr Ile Asp Pro Asp Leu Leu Leu Pro Gly Ala Gly Gln
355 360 365
Pro Ser Pro Leu Ser Phe Val Lys Pro Glu Trp Ala Lys Glu Arg Gln
370 375 380
Asp Phe Ile Gln Ala Arg Gln Ala Leu Leu Ala Ser Phe Thr Val Tyr
385 390 395 400
Ile Asn Gly Gly Val Ser Trp Glu Gln Asp Ser Ile Gln Pro Gln Phe
405 410 415
Glu Asn Asn Asn Pro Ile Met Ser Glu Ala Leu Thr Asn Thr Asp Leu
420 425 430
Gln Gly Ala Leu Pro Leu Ile Ala Arg Gly Lys Val Arg Asp Leu Tyr
435 440 445
Asp Val Asp Glu Lys Thr Leu Leu Phe Val Ala Thr Asp Arg Ile Ser
450 455 460
Ala Tyr Asp Val Ile Met Glu Asn Gly Ile Pro Glu Lys Gly Ile Leu
465 470 475 480
Leu Thr Leu Cys Thr Lys Thr Trp Phe Lys Ile Leu Ser Asp Lys Ile
485 490 495
Pro Ser Leu Arg Thr His Phe Leu Thr Leu Asp Leu Pro Pro Gln Ile
500 505 510
Pro Glu Ser Leu Arg Pro Val Leu Gln Asn Arg Ser Met Gln Val Arg
515 520 525
Lys Leu Lys Ile Leu Pro Ile Glu Ala Ile Val Arg Gly Tyr Ile Thr
530 535 540
Gly Ser Ala Trp Asn Glu Tyr Lys Lys Ser Gly Thr Val His Gly Ile
545 550 555 560
Lys Val Ala Glu Gly Leu Arg Glu Ser Glu Ala Phe Pro Asp Gly Pro
565 570 575
Ile Tyr Thr Pro Ser Thr Lys Ala Glu Gln Gly Glu His Asp Glu Asn
580 585 590
Ile His Pro Asp Gln Ala Val Ala Ile Val Gly Glu Arg Tyr Ala Ser
595 600 605
Lys Ile Ala Ser Leu Ala Val Gln Leu Tyr Lys Val Ala His Glu Tyr
610 615 620
Ala Leu Thr Arg Gly Val Ile Ile Ala Asp Thr Lys Phe Glu Phe Gly
625 630 635 640
Leu Asp Pro Glu Thr Asp Glu Val Val Leu Ala Asp Glu Val Leu Thr
645 650 655
Pro Asp Ser Ser Arg Phe Trp Pro Lys Asp Ser Tyr Glu Ile Gly Arg
660 665 670
Gly Gln Gln Ser Phe Asp Lys Gln Phe Leu Arg Asp Trp Leu Thr Ser
675 680 685
Glu Gly Leu Lys Gly Lys Pro Gly Val Arg Met Thr Asp Glu Ile Ala
690 695 700
Gln Lys Thr Ser Ala Lys Tyr Arg Glu Ala Tyr Glu Arg Ile Thr Gly
705 710 715 720

```

	Glu Thr Lys Val Pro Ala Val				
		725			
<210>	189				
<211>	750				
<212>	DNA				
<213>	米曲霉				
<220>					
<221>	启动子				
<222>	(1) .. (750)				
<223>	米曲霉负荷感应启动子 p(bibA)				
<400>	189				
	ctatgtagaa tagatagtag taggggtgtc atcctccatt atctgacaag cagataccag 60				
	catgtggcca cgtgtagtga tgttgtgcaa gtcattaagc tgcagaggtc atgacttcac 120				
	agtgactaga gttgagtaac ttgagatatg gaaagtacac ggtgtactca aggtacagta 180				
	caattggagg ttgacgacgt cctttgtcca cgtatcctaa aggagcatat ggccgggggg 240				
	ctgaaaaata gggcaacttc aagaaagatt acctaacagg ggcttctagc caatgaaaca 300				
	cattttctagt atgttcgggtg gaccgcttgc agaagatcca tgaacacttt ggggcaattt 360				
	ctacggcaga atacagggcg gtggccaagg cggatgggtc attctgggtc ctgaaaggcg 420				
	atgcaccaat cacaacgcga ggagtaacct gcagagcgat ggccgtacca gccatgctgt 480				
	cctttttctct ggctcaagc atttttaaaa agctgatctc ttctctctac gtccctctgt 540				
	cagcaacttc tctcttttgc ctcaactttc ccttccttcc ccttccatt ccgtgcgctg 600				
	tggttggtt tctcttttct tctttttgct tctcttctt gaatacagga aactgtattg 660				
	aagcacaagg gatttactac caacacgtct cccgataatc acacgcgtgc cccttggcga 720				
	agaccaccac tcgatataca cataggcaca 750				
<210>	190				
<211>	750				
<212>	DNA				
<213>	米曲霉				
<220>					
<221>	启动子				
<222>	(1) .. (750)				
<223>	米曲霉负荷感应启动子 p(clxA)				
<400>	190				
	agtgtggtct gaagttttgc aaagttagtg aggaggggaa aagaacgtcg gcggggagag 60				
	gggcaacgtg aagaactcca tttttttttt ttctgatat ggataattaa ctagttagcg 120				
	gaagtagatc gcgcaatcgg tatttgcagg gatagtatgc aattacttcg aagattaaga 180				
	aaggaaacaag aggaaaaata gaaaagaatt tatttataatt catttgaaat ttacattgac 240				
	tggtattgaat tgtcttattt gaattgggtg actagcatgt ttgggagaag tgtactaata 300				
	tgaatgaatg gtatggagta ctacaaataa catacaggga cgctaagacc gtgacaacca 360				
	aagagagaaa acaaaaaaag caataaaaacg gttcttttaga atcaatggac actggacaga 420				
	acattggaga atctgccact aatcctgttt gggtcgggtg aactttaccc caaggggcaa 480				
	gcggcagcgg aagaggccag attgaactta agctaccttc ccagcttcac ataatacaacc 540				
	tccccaacac tcgtcttctc tttctctct cctctctctt ctctcttctg tcgtgcggtt 600				
	agcttgtggt acgtctctac cttgacactc ttttccctct cgaacctcta cggaggacca 660				
	attgctttag cgtttgctta tttgtttctt tctttcctaa agctactacc gagtgcgggt 720				
	tcttagtttg agagcccgtt atcattcact 750				
<210>	191				
<211>	300				
<212>	DNA				
<213>	谷氨酸棒状杆菌 (Corynebacterium glutamicum)				
<220>					
<221>	启动子				
<222>	(1) .. (300)				
<223>	谷氨酸棒状杆菌负荷感应启动子 p(groES)				
<400>	191				
	gcgaagacat cagtttggat gaaattgagc gaattaagga cggcatttcg gccgctcgta 60				
	actaccgcga tgattatcca gaggaattca acctgtggcg caacgctgta tataacctgc 120				
	gtacgcccta aagtttggct gccatgtgaa tttttagcac cctcaacagt tgagtgtctg 180				
	cactctcggg ggtagagtgc caaatagggt gtttgacaca cagttgttca cccgcgacga 240				
	cggtgtgct ggaaaccac aaccggcaca caaaaattt ttctcatgga gggattcacc 300				
<210>	192				
<211>	300				
<212>	DNA				
<213>	谷氨酸棒状杆菌				
<220>					
<221>	启动子				
<222>	(1) .. (300)				
<223>	谷氨酸棒状杆菌负荷感应启动子 p(katA)				
<400>	192				
	tttcccggat gaaatccgtg agagctgcgc gatcgccacg gccggcggct agggccagct 60				
	cggtagcgtg ggcgtcggtt cgtctttttg acttcacaat tgctcagaat accagttcag 120				
	accgggggtc ccataaagggt gtgtaggggg tgtcaaagggt ggcaaaaagt atcatacttt 180				
	tgtctttttg ctggcagttt acgtaagggc tgattatttg gtggcgcgaa gcgacattct 240				
	gatgtagact atcaatcaga cttaatccat aggtaacctt cataaaaagga aggaatgcta 300				
<210>	193				
<211>	300				

[0494]

<212>	DNA								
<213>	谷氨酸棒状杆菌								
<220>									
<221>	启动子								
<222>	(1) .. (300)								
<223>	谷氨酸棒状杆菌负荷感应启动子 p (cplX)								
<400>	193								
	ggaggtggcg	gtggtgacgac	acacaacgcg	cacgcccgat	ggttggcggt	ccatcgcgct		60	
	aaactgggga	tcacagtagc	gttctactat	ttctttgctg	agtcgatgcg	ggcgaaggcc		120	
	cctgtttgtg	gtttcaattt	cggttgogcc	gtgaataaat	tcgggattca	ttagcttaag		180	
	cgtaccgctc	tgcagagttc	acaggtaaag	cgctaagatg	gaacaaccca	ttgccaatat		240	
	tggttggttag	agttgtacgc	agtaaattct	ttcaatcgtg	gaagcgggtc	tcacagtcta		300	
<210>	194								
<211>	300								
<212>	DNA								
<213>	谷氨酸棒状杆菌								
<220>									
<221>	启动子								
<222>	(1) .. (300)								
<223>	谷氨酸棒状杆菌负荷感应启动子 p (mutM)								
<400>	194								
	tggataccaa	gttggcagag	tttgaaacct	cgctgtccac	cacgttgctg	tctgtcactg		60	
	ctgatcgttc	cgcactacgc	cgaggagctg	gagctactgg	ccgcgagctg	cgcgatgagc		120	
	aaccggcagc	gcgtgggtgaa	tatgagcggg	actacgagcg	tgattatgaa	cgaggctacg		180	
	aacgcgacga	tcgagactac	tagttttccaa	gaaaccattt	aggctttttc	acctttccgt		240	
	ctagcgatcc	tcactctgatg	tatgtagtca	cgatgaggtg	cgagtaggat	cttctctggtc		300	
<210>	195								
<211>	4694								
<212>	DNA								
<213>	合成								
<220>									
<221>	尚未归类的特征								
<222>	(1) .. (4694)								
<223>	质粒 pEG151								
<220>									
<221>	尚未归类的特征								
<222>	(88) .. (651)								
[0495]	<223>	同源区域; iscU 上游							
<220>									
<221>	启动子								
<222>	(1833) .. (2128)								
<223>	负荷成糖启动子 p (groES)								
<220>									
<221>	RBS								
<222>	(2129) .. (2134)								
<220>									
<221>	尚未归类的特征								
<222>	(2134) .. (2134)								
<223>	N 是 A、T、C 或 G								
<220>									
<221>	尚未归类的特征								
<222>	(2144) .. (2621)								
<223>	同源区域; iscU 下游								
<400>	195								
	aacaaaattc	tccagtcctc	acatcggttt	gaaaggagga	agcggaagaa	tgaagtaaga		60	
	gggatttttg	actccgaagt	aagtcctcat	tgaggaaatc	ggaaacacca	cgcatthaaga		120	
	gaaggaatat	tccgaaaaac	cattctcatt	tgttctgaaa	ggatgctttc	gtctcatctg		180	
	ccgctatttc	acgccaaatc	gtttttgogg	catccaacgc	agcactggta	caataaaaaag		240	
	caaatatatg	aagggattga	ttcattatct	gcagcagcca	ggacagactg	aacctcgatt		300	
	atacgtcaga	aatggaaaaa	gcgatgcagc	tgtcacacgc	aattggatac	gaggaatatg		360	
	gacgccggct	ggaaacaaga	atgaaagtag	agcggcagag	agagcttgat	tatgaaaaaa		420	
	gcaagcggat	ttctgcggga	gcttattgag	aaaatcaatc	ttttaaaaaa	gggtgggaat		480	
	ccactctttt	tctttttcaa	tcgctttaat	aaagtgcgtg	tctttgttaa	acttgaggaa		540	
	tgtgcaaggc	acgataaaaa	gttgtattta	cttttcatct	gttataaaaat	tagaatgaga		600	
	ataaaatgag	aataatgatt	atataatttt	ggaggtatag	atatggctgc	tagttcggat		660	
	tagatccccc	tatgcaaggg	tttattgttt	tctaaaatct	gattaccaat	tagaatgaat		720	
	atttcccaaa	tattaaataa	taaaacaaaa	aaattgaaaa	aagtgtttcc	accatttttt		780	
	caattttttt	ataatttttt	taatctgtta	tttaaatagt	ttatagttaa	atttacattt		840	
	tcattagttc	attcaatatt	ctctccaaga	taactacgaa	ctgctaacaa	aattctctcc		900	
	ctatgttcta	atggagaaga	ttcagccact	gcattttccg	caatatcttt	tggtatgatt		960	
	ttaccgcgtg	ccatagttta	aatcatacgg	cataaagtta	atatagagtt	ggtttcatca		1020	
	tcctgataat	tatctattaa	ttcctctgac	gaatccataa	tggctcttct	cacatcagaa		1080	
	aatggaatat	caggtagtaa	ttcctctaaag	tcataatttc	cgtatattct	tttatttttt		1140	
	cgttttgctt	ggtaaagcat	tatggttaaa	tctgaattta	attccttctg	aggaatgtat		1200	
	ccttgttcat	aaagctcttg	taaccattct	ccataaataa	attcttgttt	gggaggatga		1260	

[0496]

```

ttccacggta ccatcttcttg ctgaataata attgttaatt caatatatcg taagttgctt 1320
ttatctccta ttttttttga aataggtcta atttttttga taagtatttc tttactttga 1380
tctgtcaatg gttcagatac gacgactaaa aagtcaagat cactatttgg ttttagtcca 1440
ctctcaactc ctgatccaaa catgtaagta ccaataaggt tattttttta atgtttccga 1500
agtatttttt tcaactttatt aatttgttcg tatgtattca aatataatcct cctcactatt 1560
ttgattagta cctatttttat atccatagtt gtttaattaaa taaacttaat ttagttttatt 1620
tatagatttc attggcttct aaattttttta tctagataat aattattttta gtttaatttta 1680
ttctagatta tatatgatat gatctttcat ttcataaaaa ctaaaagtaag tgtaaacctt 1740
ttcattgttt taaaaatata tottgccagt cacgttacgt tattagttaa agttattata 1800
acatgtattc acgaacgaaa atcgcacgca ttaatgcgga aaactgcatt atgatataag 1860
tcaaaaataa aaaccaatac tgtttttctc aatgggtatg actcctttgt taagtgggtt 1920
tcgttcatct acagctattg taacataatc ggtacggggg tgaaaaagct aacggaaaag 1980
ggagcggaaa agaattgatg aagcgtgaaa aatttttttat cttatcactt gaaattggaa 2040
gggagattct ttattataag aatttgttta gcactcttta gtgctgagtg ctaaaattac 2100
atattcatac tattgaggag gttattttcac gagntgaaat catatggctg cttcaacatt 2160
aacgatcaaa gatcttcacg ttgaaatcga agggaaagag atcttaaagg gtgtaaacct 2220
tgaaaataaa ggtggagaat tccacgcagt aatgggcccg aacggaaactg gtaaatccac 2280
tttatcagct gctattatgg ggcacccata atatgaagta acaaaaggca gcatcacgct 2340
tgacggcaaa gatgtactgg agatggaaag ggacgagcgt gcgcaggcag gcttattcct 2400
tgcaatgcag tacccaagtg aaatcagcgg tgtgacaaat gccgacttcc ttcgctcagc 2460
aatcaacgcg cgcagagaaag aaggcgatga aatttctctc aatgaaattca tccgcaaaat 2520
ggacgaaaac atggagttcc ttgaaatgga ccttgaaatg gctcagcgt atcttaacga 2580
aggcttctca ggccggggaga aaaaacgcaa cgaaatcctt caattttatt gcaataacag 2640
gtgcttactt ttctggagtt ctttagcaaa ttttttcttt agctgaactt ctaaaattac 2700
gccatactcc tccaatccaa agctattttag aaagattact atatcctcaa acaggcggta 2760
accggcctct tcatcgggaa tgccgcgcac cttcagcatt gccggcatgt cccctggcg 2820
gacgggaaag atccagctcg aggtcgggccc gcttcttggt gcgttttcca cgttttccg 2880
ccccctgacg agcatcacia aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga 2940
ctataaagat accaggcgtt tccccctgga agctccctcg tgcgctctcc tgttccgacc 3000
ctgcgcctta ccgatacctt gtccgccttt ctcctctcgg gaagcgtggc gaagcttcat 3060
agctcacgct gtaggtatct cagttcgggtg taggtcgttc gctccaagct gggctgtgtg 3120
cacgaacccc ccgttcagcc cgaccgctgc gcttatcccg gtaactatcg tcttgagtc 3180
aaccggtaaa gacacgactt atcgcacttg gcagcagcca ctggtaacag gattagcaga 3240
gcccggtaag taggcgtgct tacagagttc ttgaaagttg tgactatggg ggcctaacta cggctacact 3300
agaaggacag tatttgggat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt 3360
gatagctctt gatccggcaa acaaacaccg gctggtagcg gtggtttttt tgtttgcaag 3420
cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg 3480
tctgacgctc agtggaacga aaactcacgt taagggattt tggtcattgag attatcaaaa 3540
aggatcttca cctagatcct tttaaatttaa aaatgaagtt ttaaatcaat ctaaaagtata 3600
tatgagtaaa cttgggtctga cagttaccaa tgcttaatca gtgaggcacc tatctcagcg 3660
atctgtctat ttctgtctat tgactccccg tgcgttagat tgcgttagat aactacgata 3720
cgggagggct taccatctgg gcaatgatac cgcgagaccc acgctcacgg 3780
gtccagatt tatcagcaat aaaccagcca gccagcgagc cagagcgagc agtggtctct 3840
gcaactttat ccgcctccat ccagtcctatt aattgttgcc gggaaagctag agtaagtagt 3900
tcgccagtta atagtttgcg caacgttggt gccattgctg caggcatcgt ggtgtcacgc 3960
tcgtcgtttg gtaggtcttc attcagctcc ggttcccaac gatcaaggcg agttacatga 4020
tccccatgt ttgtcaaaaa agcgggttagc tcttctcggtc gatcaagcg agttacatga 4080
aagttggcgg cagtggttatc actcatgggt atggcagcac tgcaataatc tcttactgtc 4140
atgccatccg taagatgctt ttctgtgact ggtgagtagt caaccaagtc attctgagaa 4200
tagtgtagtc ggcgaccgag ttgctcttgc ccggcgtaaa caccggataa taccgcgcca 4260
catagcagaa ctttaaaagt gctcatcatt ggaaaacggt cttcggggcg aaaactctca 4320
aggatcttac cgctgttgag atccagttcg atgtaaacca ctcgtgcacc caactgatct 4380
tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag gcaaaatgcc 4440
gcaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatactctt ccttttttcaa 4500
tattttgaa gcaattatca ggggttattgt ctcatgagcg gatacatatt tgaatgtatt 4560
tagaaaaata aacaaatagg ggttccgcgc acatttccc gaaaagtgc accgtacgtc 4620
taagaaacca ttattatcat gacattaacc tataaaaaata ggcgtatcac gagggccttt 4680
cgtcttcaag aatt 4694

```

<210> 196
 <211> 4697
 <212> DNA
 <213> 合成
 <220>
 <221> 尚未归类的特征
 <222> (1) .. (4697)
 <223> 质粒 pEG152
 <220>
 <221> 尚未归类的特征
 <222> (88) .. (651)
 <223> 同源区域 iscU 上游
 <220>
 <221> 启动子
 <222> (1833) .. (2131)
 <223> 负荷成瘾启动子 p (ctsR)
 <220>
 <221> rbs
 <222> (2132) .. (2137)

[0497]

tgatccccca	tggtgtgcaa	aaaagcgggt	agctccttcg	gtcctccgat	cgttgtcaga	4080
agtaagttgg	ccgcagtggt	atcactcatg	gttatggcag	cactgcataa	ttctcttact	4140
gtcatgccat	ccgtaagatg	ctttttctgtg	actgggtgagt	actcaacca	gtcattctga	4200
gaatagtgta	tgccggcgacc	gagttgctct	tgcccgcggt	caacacggga	taataccgcg	4260
ccacatagca	gaacttttaa	agtgtctatc	attggaaaac	gttcttcggg	gcgaaaactc	4320
tcaaggatct	taccgctggt	gagatccagt	tcgatgtaac	ccactcgtgc	acccaactga	4380
tcttcagcat	cttttacttt	caccagcggt	tctgggtgag	caaaaacagg	aaggcaaat	4440
gccgcaaaaa	aggggaataag	ggcgacacgg	aaatgttgaa	tactcatact	cttctctttt	4500
caatattatt	gaagcatttta	tcagggttat	tgtctcatga	gcgatacat	atttgaatgt	4560
atttagaaaa	ataaacaaat	aggggttcgg	cgcacatttc	cccgaagaat	gccacctgac	4620
gtctaagaaa	ccattatttat	catgacatta	acctataaaa	ataggcggtat	cacgaggccc	4680
tttcgtcttc	aagaatt					4697
<210>	197					
<211>	4697					
<212>	DNA					
<213>	合成					
<220>						
<221>	尚未归类的特征					
<222>	(1)..(4697)					
<223>	质粒 pEG153					
<220>						
<221>	尚未归类的特征					
<222>	(88)..(651)					
<223>	同源区域 iscU 上游					
<220>						
<221>	启动子					
<222>	(1833)..(2131)					
<223>	负荷成糖启动子 p (dnaK)					
<220>						
<221>	rhs					
<222>	(2132)..(2137)					
<223>	rhs					
<220>						
<221>	尚未归类的特征					
<222>	(2132)..(2137)					
<223>	N 是 A、T、C 或 G					
<220>						
[0498]	<221>	尚未归类的特征				
	<222>	(2147)..(2624)				
	<223>	同源区域 iscU 下游				
	<400>	197				
aacaaaaattc	tccagttcttc	acatcgggttt	gaaaggagga	agcgggaagaa	tgaagtaaga	60
gggattttttg	actccgaagt	aagtcttcat	tgagggaattc	ggaaaacacca	cgcattaaga	120
gaaggaatat	tccgaaaaaac	cattctcatt	tgttctgaaa	ggatgctttc	gtctcatctg	180
ccgctattttc	acgccaaatc	gtttttgctg	catccaacgc	agcactggta	caataaaaaag	240
caaatatatg	aagggattga	ttcattatatt	gcagcagcca	ggacagactg	aacctcgatt	300
atacgtcaga	aatggaaaaa	gcgatgcagc	tgtcacacgg	aattggatac	gaggaatatg	360
gacgccgggt	ggaaacaaga	atgaaagtag	agcggcagag	agagcttgat	tatgaaaaaa	420
gcaagcggat	ttctgcggga	gcttattgag	aaaatcaatc	ttttaaaaaa	gggtgggaat	480
ccactctttt	tctttttcaa	tcgctttaat	aaagtgtcga	tctttgttaa	acttgaggaa	540
tgtgcaagc	acgataaaaa	gttgatattta	cttttcattt	gttataaaaat	tagaatgaga	600
ataaaatgag	aataatgatt	atatattttt	ggaggtatag	atatggctgc	tagttcggat	660
tagatccccc	tatgcaaggg	tttattgttt	tctaaaaatc	gattaccaat	tagaatgaat	720
atttcccaaa	tattaaataa	taaaaacaaa	aaattgaaaa	aagtgtttcc	accatttttt	780
caattttttt	ataatttttt	taatctgtta	tttaaatagt	ttatagttaa	atttacattt	840
tcattagttcc	attcaatatt	ctctccaaga	taactacgaa	ctgctaacaa	aattctctcc	900
ctatgttcta	atggagaaga	ttcagccact	gcatttcccg	caatatcttt	tggtatgatt	960
ttaccggtgt	ccatagttaa	aatcatacgg	cataaagtta	atatagagtt	ggtttcatca	1020
tcctgataat	tatctattaa	ttcctctgac	gaatccataa	tggtctctct	cacatcagaa	1080
aatggaatat	caggtagtaa	ttcctctaa	tcataatttc	cgtatattct	tttatttttt	1140
cgttttgctt	ggtaaagcat	tatgggttaa	tctgaattta	attccttctg	aggaatgtat	1200
ccttggtcat	aaagctcttg	taaccattct	ccataaataa	attcttgttt	gggaggatga	1260
ttccacggta	ccatttcttg	ctgaataata	attgttaatt	caatatatcg	taagttgctt	1320
ttatctccta	ttttttttga	aatagggtcta	attttttgta	taagtatttc	tttactttga	1380
tctgtcaatg	gttcagatac	gacgactaaa	aagtcaagat	cactattttg	ttttagtcca	1440
ctctcaactc	ctgatccaaa	catgtaagta	ccaataaggt	tattttttta	atgtttccga	1500
agtatttttt	tcacttttatt	aatttgttcg	tatgtattca	aatatatcct	cctcactatt	1560
ttgattagta	cctatttttat	atccatagtt	gttaatttaa	taaacttaat	ttagtttatt	1620
tatagatttc	attggcttct	aaatttttta	tctagataat	aattatttta	gttaatttta	1680
ttctagatta	tatatgatat	gatctttcat	ttccataaaa	ctaaagtaag	tgtaaaccta	1740
ttcattgttt	taaaaatata	tcttgccagt	cacgttacgt	tattagttat	agttattata	1800
acatgtattc	acgaacgaaa	atcgcacgga	ttggttcaat	ccccgcacac	gggaatttcg	1860
attaaaaatcg	gaaaagaaaa	cgactatgaa	gagatggaaa	attgcagtct	gattacgggt	1920
tcttattccg	tagaccagaa	gcagatcgcc	tcaattgcga	ttatcgcccc	gacccgcagt	1980
aattattcca	gggttgtcag	cctgcttcag	catgtgactt	cggacttgtc	aaaagcatta	2040
acaagtctgt	atgatgaata	agggaaatttt	ggcaaatttt	atcgaagggc	agcacctgtc	2100

[0499]

cttctcctta	cactttgagg	gaggtgaaca	cacgagntga	aatcatatgg	ctgcttcaac	2160
attaacgatc	aaagatcttc	acgttgaaat	cgaaggga	gagatcttaa	aggggtgtaa	2220
ccttgaaata	aaaggtggag	aattccacgc	agtaattggg	ccgaacggaa	ctgggtaaatc	2280
cactttatca	gctgctatta	tggggcatcc	taaatatgaa	gtaacaaaag	gcagcatcac	2340
gcttgacggc	aaagatgtac	tggagatgga	agtggacgag	cgtgcgcagg	caggcttatt	2400
ccttgcaatg	cagtacccaa	gtgaaatcag	cgggtgtgaca	aatgccgact	tccttcgctc	2460
agcaatcaac	gcgcgcagag	aagaaggcga	tgaattttct	ctcatgaaat	tcattccgcaa	2520
aatggacgaa	aacatggagt	tccttgaaat	ggacctgaa	atggctcagc	gctatcttaa	2580
cgaaggcttc	tcaggcgggg	agaaaaaacg	caacgaaatc	cttcaatttt	attgcaataa	2640
cagggtgctta	cttttctgga	gttcttttagc	aaattttttt	attagctgaa	cttagtatta	2700
gtggccatac	tcctccaatc	caaaagctatt	tagaaagatt	actatatcct	aaacacggcg	2760
gtaaccggcc	tcctcatcgg	gaatgcgcgc	gaccttcagc	atcgccggca	tgtccccctg	2820
gcggacggga	agtatccagc	tcgaggctcg	gccgcgttgc	tggcgttttt	ccataggctc	2880
cgccccctcg	acgagcatca	caaaaaatcga	cgtcaagtc	agagggtggcg	aaacccgaca	2940
ggactataaa	gataccaggc	gtttccccct	ggaagctccc	tcgtgcgctc	tcctgtttccg	3000
accctgcccgc	ttaccggata	cctgtccgcgc	ttttccctt	cgggaagcgt	ggcgctttct	3060
catagctcac	gctgtaggta	tcctcagttcg	gtgtaggctg	ttcgtcccaa	gctgggtgtg	3120
gtgcacgaac	ccccggttca	gcccgaaccgc	tgcgccttat	ccggtaacta	tcgtcttgag	3180
tccaaccggg	taagacacga	cttatcgcca	ctggcagcag	ccactggtaa	caggattagc	3240
agagcgaggt	atgtaggcgg	tgctacagag	ttcttgaagt	gggtggcctaa	ctacgggtac	3300
actagaaggga	cagtatttgg	tatctgcgct	ctgctgaagc	cagttacctt	cggaaaaaga	3360
gttgatagct	cttgatccgg	caaacaaaacc	accgctggta	gcggtgggtt	ttttgtttgc	3420
aagcagcaga	ttacgcgcag	aaaaaaagga	tctcaagaag	atcctttgat	cttttctacg	3480
gggtctgacg	ctcagtgga	cgaaaactca	cgttaaggga	ttttggtcat	gagattatca	3540
aaaaggatct	tcacctagat	cctttttaa	taaaaatgaa	gtttttaa	aatctaaagt	3600
atatatgagt	aaacttggtc	tgacagttac	caatgcttaa	tcagtgaggc	acctatctca	3660
gcgactgttc	tatttcgttc	atccatagtt	gcctgaactc	cgtcgtgta	gataactacg	3720
atacgggagg	gcttaccatc	tggccccagt	gctgcaatga	taccgcgaga	cccacgctca	3780
ccggtctcag	atattatcagc	aataaaaccag	ccagccggaa	gggcccagcg	cagaagtggg	3840
cctgcaactt	tatccgcctc	catccagttc	attaattggt	gccgggaagc	tagagtaagt	3900
agttcgccag	ttaatagttt	gcgcgaacgtt	gttgccattg	ctgcaggcat	cgtgggtgta	3960
cgctcgtcgt	ttgggtatggc	ttoattcagc	tccggttccc	aacgatcaag	gcgagttaca	4020
tgatccccca	tggtgtgcaa	aaaagcgggt	agctccttcg	gtcctccgat	cgttgtcaga	4080
agtaagttgg	ccgcagtggt	atcactcatg	gttatggcag	cactgcataa	ttctcttact	4140
gtcatgccat	ccgtaagatg	cttttctgtg	actggtgagt	actcaaccaa	gtcattctga	4200
gaatagtgtg	tgccggcgacc	gagttgtctc	tgcccggcgt	caacacggga	taataccgcg	4260
ccacatagta	gaactttaaa	agtgtcctac	attggaatac	gttcttcggg	gcgaaaactc	4320
tcaaggatct	taccgctggt	gagatccagt	tcgatgtaac	ccactcgtgc	acccaactga	4380
tcttcagcat	cttttacttt	caccagcggt	tctgggtgag	caaaaacagg	aaggcaaaat	4440
gccgcaaaaa	aggggaataag	ggcgacacgg	aaatggtgaa	tactcatact	cttccttttt	4500
caatattatt	gaagcatttta	tcagggttat	tgtctcatga	gcggatacat	atttgaatgt	4560
atttagaata	ataaacaaat	aggggttccg	cgcacatttc	cccgaataag	gccacctgac	4620
gtctaagaaa	ccattatttat	catgacatta	acctataaaa	ataggcgtat	cacgaggccc	4680
tttcgtcttc	aagaatt					4697
<210>	198					
<211>	4697					
<212>	DNA					
<213>	合成					
<220>						
<221>	尚未归类的特征					
<222>	(1) .. (4697)					
<223>	质粒 pEG154					
<220>						
<221>	尚未归类的特征					
<222>	(88) .. (651)					
<223>	同源区域 iscU 上游					
<220>						
<221>	启动子					
<222>	(1833) .. (2131)					
<223>	负荷成瘾启动子 p(perR)					
<220>						
<221>	rb					
<222>	(2132) .. (2137)					
<223>	rb					
<220>						
<221>	尚未归类的特征					
<222>	(2137) .. (2137)					
<223>	N 是 A、T、C 或 G					
<220>						
<221>	尚未归类的特征					
<222>	(2147) .. (2624)					
<223>	同源区域 iscU 下游					
<400>	198					
aacaaaaattc	tcagttcttc	acatcggttt	gaaaggagga	agcggaagaa	tgaagtaaga	60
gggattttttg	actccgaagt	aagtcttcat	tgaggaaattc	ggaaacacca	cgcattaaga	120
gaaggaatat	tcgaaaaaac	cattctcatt	tgttctgaaa	ggatgctttc	gtctcatctg	180

[0500]

```

ccgctatttc acgccaatc gtttttgcg catccaacgc agcactggta caataaaaaa 240
caaatatatg aagggattga ttcattatct gcagcagcca ggacagactg aacctcgatt 300
atagctcaga aatggaaaaa gcgatgcagc tgcacacagg aattggatac gaggaatatg 360
gacgcggctt ggaaacaaga atgaaagtag agcgccagag agagcttgat tatgaaaaaa 420
gcaagcggat ttctgcggga gcttattgag aaaaatcaatc ttttaaaaaa ggggtgggaat 480
ccactctttt tctttttcaa tcgctttaat aaagtgcgtg tctttgttaa acttgaggaa 540
tgtgtcaaggc acgataaaaa gttgtattta cttttcattt gttataaaat tagaatgaga 600
ataaaatgag aataatgatt atatatatct ggaggtatag atatggctgc tagttcggat 660
tagatcccc ccatagttta tttattgttt tctaaaatct gattaccaat tagaatgaat 720
atttcccaaa tattaataaa taaaacaaaa aaattgaaaa aagtgtttcc accatttttt 780
caattttttt ataatttttt taatctgtta tttaaatagt ttatagttaa atttacattt 840
tcattagttc attcaatatt ctctccaaga taactacgaa ctgctaacaa aattctctcc 900
ctatgttcta atggagaaga ttcagccact gcatttcccg caatatcttt tgggtatgatt 960
ttaccctgtt aatcatacgg aatcatacgg cataaagtta atatagagt gggttccatca 1020
tcttgataat tatctattaa ttctctgcac gaatccataa tggctcttct cacatcagaa 1080
aatggataat caggtagtaa ttctcttaag tcataatttc cgtatattct tttatttttt 1140
cgttttgtct ggttaaagcat tatggttaaa cctataaata attccttctg aggaatgtat 1200
ccttgttcat aaagctcttg taaccattct ccataaataa attcttgttt gggaggatga 1260
ttccacggta ccatctcttg ctgaataata attgttaatt caatatatcg taagtgtctt 1320
ttatctccta ttttttttga aataggtcta atttttttga taagtatttc tttactttga 1380
ttctgtcaatg gttcagatac gacgactaaa aagtcacgat cactatttgg ttttagtcca 1440
ctctcaactc ctgatccaaa catgtaagta ccaataagggt tattttttta atgtttccga 1500
agtatttttt tcacttttatt aatttgttct tatgtattca aatatatcct cctcactatt 1560
ttgatttagta atcattttat atccatagtt gtttaataaa taaactttaa ttagtttatt 1620
tatagatttc attggcttct aaatttttta tctagataat aattatttta gttattttta 1680
ttctagattt tatatgatat gatctttcat ttccataaaa ctaaagttaag tgtaaaccta 1740
ttcattgttt taaaaatata tcttgccagt cagcttaccg tattagttat tgtataata 1800
acatgtattc acgaacgaaa atcgcacgta ttggaagcat ttgatgtgtg gaagctgaag 1860
aaaaactctg gaaaagagta tatgggaatc gagcgttcca cttttttaat tgataaagaa 1920
ggccggctca taaaagaatg gagaaaagta aaagttaaag accatgtagc aggaagcgtt 1980
caaacctcac aagatatgtc tgaaaaagta atctctatga gcctatgctt acttttacc 2040
tgatatcaat tgacaaatta tctagaaaag agttacacta attataaaca ttacaatgta 2100
agaatttttt ttagaagaga ggtgcacgac ggcagnttga aatcatatgg ctgcttcaac 2160
attaacgata aaagatcttc acgttgaaat cgaagggaaa gagatcttaa atgtgtaaa 2220
ccttgaaaata aaaggtggag aattccacgc agtaatgggc ccgaacggaa ctggtaaatc 2280
cactttatca gctgctatta tggggcatcc taaatatgaa gtaacaaaag gcagcatcac 2340
gcttgacggc ggaagatgac tggagatgga agtgccgagc cgtgcgcagg caggcttatt 2400
ccttgcaatg cagtacccaa gtgaaatcag cggtgtgaca aatgccgact tcttctgctc 2460
agcaatcaac gcgcgcagag aagaaggcga cgtgaaatct ctcataaata tcatccgcaa 2520
aatggacgaa aacatggagt tcttgaaat agaacctgaa atggctcagc ggtattctaa 2580
cgaaggcttc tcaggcgggg agaaaaaacg caacgaaatc cttcaatttt attagctgaa 2640
caggtgctta cttttctgga gttcttttagc aaattttttt tagaaaagatt actatatcct 2700
gtggccatcc tcttccaatc gaaatgcgcgc tagaaagatt actatatcct caaacaggcg 2760
gtaacccgac tcttcatcgc gaatgcgcgc gaccttcagc atcgccggca tgcctccctg 2820
gcggacggga agtatccagc tcgaggtcgg gccgcgttgc tggcgttttt ccataggtct 2880
gcccccctg acgagcatca caaaaaatcga cgctcaagtc agaggtggcg aaacccgaca 2940
ggactataaa gataccaggc gtttccccct ggaagctccc tctgtgcgtc tctgttcccg 3000
accctgcgcg ttaccggata cctgtccgcc ttttccctt cggaagcgt ggccgtttct 3060
catagctcac gctgtaggta tctcagttcg gtttccccct ttctccctt ttcgctccaa gctgggtgtg 3120
gtgcacgaac ccccgttca gcccgcgcgc gctgtccgcc gtgtaggctg tgcgcttatt cccgtaacta 3180
tccaaccggg taagacacga cttatcgcca ctggcagcag ccactggtaa caggattagc 3240
agagcgaggt atgtaggcgg tgctacagag ttcttgaaat ggtggcctaa ctacggctac 3300
actagaagga cagtatttgg tatctgcgct cgtctgaagt cagttacctt cggaaaaaga 3360
gttagatgct cttgatccgg caaaacaacc accgttggtg gccgttggtt ttttggttgc 3420
aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag atcctttgat cttttctacg 3480
gggtctgacg ctcagtggaa cgaaaaactca cgttaaggga ttttggctat gagattatca 3540
aaaaggatct tcacctagat cctttttaat gtaaaatgaa gttttaaatc aatctaaagt 3600
atatatgagt aaacttggtc tgacagttac caatgcttaa tcagtggagg acctatctca 3660
gcgatctgtc tatttcgttc gctgactccc gctgcaatga cctgactccc gtaaacctacg 3720
atacgggagg gcttaccatc tggccccagt gctgcaatga taccgcgaga cccacgctca 3780
ccggctccag atttatcagc ccagccggaa gggccgagcg cccagcgtca cagaagtgg 3840
cctgcaactt tatccgcctc catccagttt attaatgtgt gccgggaagc tagagtaagt 3900
agttcgccag ttaatatgtt gcgcaacggt gttgccaatt ctgcaggcat cgtgggttca 3960
cgctcgtcgt ttggtatggc aacgatcaag aacgatcaag gtcctccgat cgtgtgtaca 4020
tgatccccca tgttgtgcaa agctccttcg gttatggcag cactgcataa ttctcttact 4140
agtaagtgtg ccgcagtgtt atcactcatg gttatggcag actggtgagt actcaaccaa gtcattctga 4200
gtcatgccat ccgtaagatg cttttctgtg actggtgagt tgcctccgat caaacaggga 4260
gaatagtgtg tgccggcagc gagttgctct tctgggtgag caaaaacagg gttcttcggg 4320
ccacatagca gaactttaaa agtgctcact cgcagttaac ccaactcgtc acccaactga 4380
tcaaggatct taccgctgtt caccagcggt tctgggtgag tctggtgag caaaaacagg aaggcaaaat 4440
tcttcagcat cttttacttt ggcgcacagg tactcatact cttccttttt 4500
gccgcaaaaa gaagcattta tcaggtttat tgtctcatga cgggatacat atttgaatgt 4560
caatattatt ataaacaaat aggggttccg cgcacatttc cccgaaaagt gccacctgac 4620
atattagaaaa ccattattat catgacatta accataaaaa ataggcgtat caccaggccc 4680
gtctaagaaa aagaattt
<210> 199
<211> 4697

```

169

[0502]

gtggccatac	tcttccaatc	caaagctatt	tagaaagatt	actatatcct	caaacaggcg	2760
gtaaccggcc	tcttcatcgg	gaatgcgcgc	gaccttcagc	atcgccggca	tgtccccctg	2820
gcggacggga	agtatccagc	tcgaggtcgg	gcgcgcttgc	tggcggtttt	ccataggctc	2880
cgccccctg	acgagcatca	caaaaatcga	cgctcaagtc	agagggtggc	aaaccgcaca	2940
ggactataaa	gataccaggc	gtttccccct	ggaagctccc	tcgtgcgctc	tcctgtttccg	3000
accctgccgc	ttaccggata	cctgtccggc	ttttccctt	cgggaaagcg	ggcgctttct	3060
catagctcac	gctgtaggta	tctcagttcg	gtgtaggtcg	ttcgcctcaa	gctgggctgt	3120
gtgcacgaac	ccccggttca	gcccgaaccg	tgcgccttat	ccggtaacta	tcgtcttgag	3180
tccaaccggg	taagacacga	cttatcgcca	ctggcagcag	ccactggtaa	caggattagc	3240
agagcgaggt	atgtaggcgg	tgctacagag	ttcttgaaat	gggtggcctaa	ctacggctac	3300
actagaaggga	cagtatttgg	tatctgcgct	ctgctgaagc	cagttacctt	cggaataaga	3360
gttgatagct	cttgatccgg	caaacaaacc	accgctggta	gcggtgggtt	ttttgtttgc	3420
aagcagcaga	ttacgcgcag	aaaaaaagga	tctcaagaag	atcctttgat	cttttctacg	3480
gggtctgacg	ctcagtgga	cgaaaactca	cgtaaaggga	ttttgggtcat	gagattatca	3540
aaaaggatct	tcacctagat	cctttttaa	taaaaatgaa	gtttttaa	aatctaaagt	3600
atatatgagt	aaacttggtc	tgacagttac	caatgcttaa	tcagtggagg	acctatctca	3660
gcgactctgt	tatttcgttc	atccatagtt	gcctgactcc	ccgtcgtgta	gataactacg	3720
atcagggaag	gcttaccatc	tggtcccgag	gctgcaatga	taccgcgaga	cccacgctca	3780
ccggctccag	atattatcagc	aataaaccag	ccagccggaa	gggcccagcg	cagaagtggg	3840
cctgcaactt	tatccgcctc	catccagttc	attaattggt	gcccgggaagc	tagagtaagt	3900
agttcgccag	ttaatagttt	gcgcaacggt	gttgccattg	ctgcaggcat	cgtgggtgta	3960
cgctcgtcgt	ttgggtatggc	ttcattcagc	tccggttccc	aacgatcaag	gcgagttaca	4020
tgatccccca	tgtttgtgcaa	aaaagcgggt	agctccttcg	gtcctccgat	cgttgtcaga	4080
agtaagtggg	ccgcagtggt	atcactcatg	gttatggcag	cactgcataa	ttctcttact	4140
gtcatgccat	ccgtaagatg	cttttctgtg	actggtgagt	actcaaccaa	gtcattctga	4200
gaatagtgtg	tgccggcgacc	gagttgtctc	tgcccggcgt	caacacggga	taataccgcg	4260
ccacatagca	gaactttaaa	agtgtctcat	attggaatac	gttcttcggg	gcgaaaactc	4320
tcaaggatct	taccgctggt	gagatccagt	tcgatgtaac	ccactcgtgc	acccaactga	4380
tcttcagcat	cttttacttt	caccagcggt	tctgggtgag	caaaaacagg	aaggcaaaat	4440
gccgcaaaaa	aggggaataag	ggcgacacgg	aaatggtgaa	tactcatact	cttccttttt	4500
caatatattt	gaagcattta	tcagggttat	tgtctcatga	gcggatacat	atttgaatgt	4560
atttagaaaa	ataaacaaat	aggggttccg	cgcacatttc	cccgaataag	gccacctgac	4620
gtctaagaaa	ccattattat	catgacatta	acctataaaa	ataggcgtat	cacgaggccc	4680
tttcgtcttc	aagaatt					4697
<210>	200					
<211>	4697					
<212>	DNA					
<213>	合成					
<220>						
<221>	尚未归类的特征					
<222>	(1) .. (4697)					
<223>	质粒 pEG156					
<220>						
<221>	尚未归类的特征					
<222>	(88) .. (651)					
<223>	同源区域 iscU 上游					
<220>						
<221>	启动子					
<222>	(1833) .. (2131)					
<223>	负荷成瘾启动子 p (spxA)					
<220>						
<221>	rbs					
<222>	(2132) .. (2137)					
<223>	rbs					
<220>						
<221>	尚未归类的特征					
<222>	(2137) .. (2137)					
<223>	N 是 A、T、C 或 G					
<220>						
<221>	尚未归类的特征					
<222>	(2147) .. (2624)					
<223>	同源区域 iscU 下游					
<400>	200					
aacaaaaattc	tccagttcttc	acatcggttt	gaaaggagga	agcgggaagaa	tgaagtaaga	60
gggattttttg	actccgaagt	aagtcttcat	tgaggaaattc	ggaaacacca	cgcattaaga	120
gaagggaatat	tccgaaaaaac	cattctcatt	tgttctgaaa	ggatgctttc	gtctcatctg	180
ccgctattttc	acgccaaatc	gtttttgcgg	catccaacgc	agcactggta	caataaaaaag	240
caaatatatg	aagggtattga	ttcattattt	gcagcagcca	ggacagactg	aacctcgatt	300
atacgtcaga	aatggaaaaa	gcgatgcagc	tgtcacacgg	aattggatac	gagggaatatg	360
gacgccggct	ggaaacaaga	atgaaagtag	agcggcagag	agagcttgat	tatgaaaaaaa	420
gcaagcggat	ttctgcggga	gcttattgag	aaaatcaatc	ttttaaaaaa	gggtgggaat	480
ccactctttt	tctttttcaa	tcgctttaat	aaagtgtgta	tctttgttaa	acttgaggaa	540
tgtgcaaggg	acgataaaaa	gttgatttta	cttttcat	gttataaaa	tagaatgaga	600
ataaaatgag	aataatgatt	atatattttt	ggaggtatag	atatggctgc	tagttcggat	660
tagatcccc	tatgcaaggg	tttattgttt	tctaaaaatc	gattaccaat	tagaatgaat	720
atttcccaaa	tattaaataa	taaaacaaaa	aaattgaaaa	aagtgtttcc	accatttttt	780

[0503]

```

caattttttt ataatttttt taatctgtta tttaaatagt ttatagttaa atttacattt 840
tcattagtc attcaatatt ctctccaaga taactacgaa ctgctaacaa aattctctcc 900
ctatgttcta atggagaaga ttcagccact gcatttcccg caatatcttt tgggtatgatt 960
ttaccctgtt ccatagttta aatcatacgg cataaagtta atatagagtt ggtttcatca 1020
tcctgataat tatctattaa ttctctgcac gaatccataa tggctcttct cacatcagaa 1080
aatggaatat caggtagtaa ttctcttaag tcataatttc cgtatattct tttatttttt 1140
cgttttgctt ggtaaagcat tatggttaaa tctgaattta attccttctg aggaatgtat 1200
ccttgttcatt aaagctcttg taaccattct ccataaataa attcttggtt gggaggatga 1260
ttccacggta ccatttcttg ctgaataata attgttaatt caatatatcg taagttgctt 1320
ttatctccta ttttttttga aataggtcta attttttgta taagtatttc tttactttga 1380
tttgtcaatg ttccagatac gacgactaaa aagtcaagat cactatttgg ttttagtcca 1440
ctctcaactc ctgatccaaa catgtaagta ccaataaggt tattttttaa atgtttccga 1500
agtatttttt tcactttatt aatttgttcg tatgtattca aatatatcct cctcactatt 1560
ttgattagta cctattttat atccatagtt gtttaattaa taaacttaat ttagtttatt 1620
tatagatttc attggcttct aaatttttta tctagataat aattatttta gtttaatttta 1680
ttctagatta tatatgatat gatctttcat ttccataaaa cttaaagtaag tgtaaaccta 1740
ttcattgttt taaaaatate tcttgccagt caggttacgt tattagttat aggtattata 1800
acatgtattc acgaacgaaa atcgcacgaa ttgaaacat tcatacttat aaagatgaga 1860
aatggtacgg agaatcgtac gaaaaaacgg atgaggtctt ggaaatcatc gatgaagaaa 1920
aacaacaaaa tattttcgat cagttaaagct aattgagaat aagaacatat catgaacatt 1980
aaaaatttgc tggattgaca catttttttc ttaggatact gttcacactt acttttttat 2040
agtataatc ctgtaaagat tcttttttta gagtaatttc aatttaagaa attttaagaa 2100
gttcatccta ctaattagag gagtgaagat gacgagntga aatcatatgg ctgcttcaac 2160
atcaacgatc aaagatcttc acgttgaaat cgaagggaaa gagatcttaa aggtgtttaa 2220
ccttgaaata aaaggtggag aattccacgc agtaatgggc ccgaacggaa ctggtaaatac 2280
cactttatca gctgctatta tggggcatcc taaatatgaa gtaacaaaag gcagcatcac 2340
gcttgacggc aaagatgtac tggagatgga agtgacgag cgtgcgcagg caggcttatt 2400
ccttgcaatg cagtacccaa gtgaaatcag cggtgtgaca aatgccgact tcttcgctc 2460
agcaatcaac gcgcgcagag aagaaggcga tgaattttct ctcataaaat tcatccgcaa 2520
aatggacgaa aacatggagt tcttgaaat ggacctgaa atggctcagc gctatcttaa 2580
cgaaggcttc tcaggcgggg agaaaaaacg caacgaaatc cttcaatttt attgcaataa 2640
caggtgctta cttttctgga gttcttttagc aaattttttt attagctgaa cttagtatta 2700
gtggccatac tcttccaatc caaagctatt tagaaagatt actatatcct caaacaggcg 2760
gtaaccggcc tcttcacggc gaatgcgcgc gaccttcagc atcgccggca tggccccctg 2820
gcggacggga agtatccagc tgcaggtcgg gccgcgttgc tggcggtttt ccataggctc 2880
cgcccccttg acgagcatca caaaaaatcga cgtctcaagt agagggtggc aaacccgaca 2940
ggactataaa gataccagc gtttccccct ggaagctccc tctgtcgcctc tctgttccg 3000
accctgccgc ttaccggata cctgtccgcc tttctccctt cgggaagcgt ggcgctttct 3060
catagctcac gctgtaggtt tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt 3120
gtgcacgaac ccccggttca gcccgaccgc tgcgccttat ccggttaacta tctgtcttag 3180
tccaaccggg taagacacga cttatcgcca ctggcgagc ccactggtaa caggattatga 3240
agagcgaggt atgtaggcgg tgctacagag ttcttgaggt ggtggcctaa ctacggctac 3300
actagaagga cagtatttgg tatctgcgct ctgctgaagc cagttacctt cggaaaaaga 3360
gttgatagct cttgatccgg caaacaaaacc accgctggta gcggtggtt ttttgttgc 3420
aagcagcaga ttacgcgcag aaaaaaaggga tctcaagaag atcctttgat cttttctacg 3480
gggtctgacg tccagtggaa cgaaaaactca cgttaaaggga ttttggtcat gagattatca 3540
aaaaagatct tcacctagat ccttttaaat taaaaatgaa gttttaaatc aatctaaagt 3600
atatatgagt aaacttggtc tgacagttac caatgcttaa tcagtgaggc acctatctca 3660
gcgatctgtc tatttcgttc atccatagtt gccctgactcc cctgcgtgta gataactacg 3720
atcggggagg gcttaccatc tggccccagt gctgcaatga taccgcgaga cccacgctca 3780
ccggctccag atttatcagc aataaaaccag ccagccggaa gggccgagcg cagaagtgg 3840
cctgcaactt tatccgcctc catccagttc attaatgtt gccgggaagc tagagtaagt 3900
agttcgcag ttaatagttt gcgcaacggt gttgccattg ctgcaggcat cgtggtgtca 3960
cgctcgctcg ttggtatggc ttcattcagc tccggttccc aacgatcaag cgcagttaca 4020
tgatccccc tggtgtgcaa aaaagcgggt agctccttcg gtcctccgat cgttgtcaga 4080
agtaagttgg ccgcagtggt atcactcatg gttatggcag cactgcataa ttctcttact 4140
gtcatgccat ccgtaagatg cttttctgtg actgggtgag actcaaccaa gtcattctga 4200
gaatagtgt tgccgcgacc gagttgctct tgcgccgctg caacacggga taataccg 4260
ccacatagca gaactttaa agtgctcact attggaacac gttcttcggg gcgaaaactc 4320
tcaaggatct taccgctggt gagatccagt tccatgtaac cactcgtgc acccaactga 4380
tcttcagcat cttttacttt caccagcgtt tctgggtgag caaaaacagg aaggcaaat 4440
gccgcaaaaa aggggaataa ggcgacacgg aaatgttgaa tactcatac cttccttttt 4500
caatattatt gaagcattta tcagggttat tgcctcatga gccgatacat atttcaatgt 4560
atttgaaaaa ataacaatat aggggtttcc cgcacatttc cccgaaaagt gccacgtgac 4620
gtctaagaaa ccattattat catgacatta acctataaaa ataggcgtat cacgaggccc 4680
tttcgtcttc aagaatt 4697

```

<210> 201

<211> 4697

<212> DNA

<213> 合成

<220>

<221> 尚未归类的特征

<222> (1) .. (4697)

<223> 质粒 pEG157

<220>

<221> 尚未归类的特征

<222> (88) .. (651)

<223>	同源区域 iscU 上游						
<220>							
<221>	启动子						
<222>	(1833) .. (2131)						
<223>	负荷成糖启动子 p (sigB)						
<220>							
<221>	rbs						
<222>	(2132) .. (2137)						
<223>	rbs						
<220>							
<221>	尚未归类的特征						
<222>	(2132) .. (2137)						
<223>	N 是 A、T、C 或 G						
<220>							
<221>	尚未归类的特征						
<222>	(2147) .. (2624)						
<223>	同源区域 iscU 下游						
<400>	201						
aacaaaaattc	tccagtccttc	acatcgggttt	gaaaggaggga	agcgggaagaa	tgaagtaaga		60
gggattttttg	actccgaagt	aagtccttcac	tgagggaattc	ggaaacacca	cgcattaaga		120
gaagggaatat	tccgaaaaaac	cattctcatt	tgttctgaaa	ggatgctttc	gtctcatctg		180
ccgctattttc	acgccaaatc	gtttttgcgg	catccaacgc	agcactggta	caataaaaaag		240
caaatatatg	aagggaattga	ttcattatatt	gcagcagcca	ggacagactg	aacctcgatt		300
atacgtcaga	aatggaaaaa	gcgatgcagc	tgtcacacgg	aattggatac	gaggaatatg		360
gacgccggct	ggaaacaaga	atgaaagtag	agcggcagag	agagcttgat	tatgaaaaaa		420
gcaagcggtat	ttctgcggga	gcttattgag	aaaatcaatc	ttttaaaaaa	gggtgggaat		480
ccactctttt	tctttttcaa	tcgctttaat	aaagtgcgtga	tctttgttaa	acttgaggaa		540
tgtgcaaggc	acgataaaaa	gttgattttta	cttttcattt	gttataaaat	tagaatgaga		600
ataaaatgag	ataaatgatt	ataatatttt	ggaggtatag	atatggctgc	tagttcggat		660
tagatccccc	tatgcaaggg	tttattgttt	tctaaaatct	gattaccaat	tagaatgaat		720
atttcccaaa	tattaaataa	taaaacaaaa	aaattgaaaa	aagtgtttcc	accatttttt		780
caattttttt	ataatttttt	taatctgtta	tttaaatagt	ttatagttaa	atttacattt		840
tcatttagtcc	attcaatatt	ctctccaaga	taactacgaa	ctgctaacaa	aattctctcc		900
ctatgttcta	atggagaaga	ttcagccact	gcatttccc	caatatcttt	tggtatgatt		960
ttacccgtgt	ccatagttta	aatcacacgg	cataaagtta	atatagagtt	ggtttcatca		1020
tcctgataat	tatctattaa	ttcctctgac	gaatccataa	tggctcttct	cacatcagaa		1080
aatggaatat	caggtagtaa	ttcctctaag	tcataatttc	cgtatattct	tttatttttt		1140
cgttttgctt	ggtaaagcat	tatgggttaa	tctgaattta	attccttctg	aggaatgtat		1200
ccgttgctcat	aaagctcttg	taaccattct	ccaataataa	attcttgttt	gggaggatga		1260
ttccacggta	ccatttcttg	ctgaataata	attgttaatt	caatatatcg	taagtgtcct		1320
ttatctccta	ttttttttga	aataggtcta	attttttgta	taagtatttc	tttactttga		1380
tctgtcaatg	gttcagatac	gacgactaaa	aagtcaagat	cactatttgg	ttttagtcca		1440
ctctcaactc	ctgatccaaa	catgtaagta	ccaataaggt	tattttttta	attgtttccga		1500
agtatttttt	tcactttatt	aatttggtcg	tatgtattca	aatatatcct	cctcactatt		1560
ttgattagta	cctattttat	atccatagtt	gttaattaaa	taaacttaat	ttagttttat		1620
tatagatttc	attggcttct	aaatttttta	tctagataat	aattatttta	tttattttta		1680
ttctagatta	tatatgatat	gatctttcat	ttccataaaa	ctaaagtaag	tgtaaaccta		1740
ttcattgttt	taaaaaatct	tcttgccagt	cacgttacgt	tattagttat	agttattata		1800
acatgtattc	acgaacgaaa	atcgcactga	ttaacaattc	gatcagcatc	tggaaaagg		1860
cgatatgatt	gtttttatttt	ctgacggcgt	cacagaatgc	agaacggaaa	acggtttctt		1920
ggagcgctct	gatctgcaga	agctcattga	ggaacatatg	tgttcctctg	cgcaggaaat		1980
ggtcaaaaac	atztatgaca	gctcctccta	attgcaggat	tttcagcttc	acgatgattt		2040
tacgttaatt	gtttttgcgga	gaaagggtta	acgtctgtca	gacgagggtta	taaagcaact		2100
agtgatttga	agggaaattt	gaggtgtac	gacgagntga	aatcatatgg	ctgcttcaac		2160
cttaacgatc	aaagatcttc	acgttgaaat	cgaagggaag	gagatcttaa	agggtgtaaa		2220
ccttgaaaata	aaaggtggag	aattccacgc	agtaattggc	ccgaacggaa	ctggtaaatc		2280
cactttatca	gctgctatta	tggggcatcc	taaataatgaa	gtaacaaaag	gcagcatcac		2340
gcttgacggc	aaagatgtac	tggagatgga	agtggacgag	cgtgcgcagg	caggcttatt		2400
ccttgcaatg	cagtacccaa	gtgaaatcag	cgtgtgtaca	aatgcccact	tccttcgctc		2460
agcaatcaac	gcgcgcagag	aagaaggcga	tgaattttct	ctcatgaaat	tcacccgcaa		2520
aatggacgaa	aacatggagt	tccttgaaat	ggaccctgaa	atggctcagc	gctatcttaa		2580
cgaaggcttc	tcaggcgggg	agaaaaaacg	caacgaaatc	cttcaatttt	attgcaataa		2640
cagggtgctta	ctttttctgga	gttcttttagc	aaattttttt	attagctgaa	cttagtatta		2700
gtggccatac	tcctccaatc	caaagctatt	tagaaaagatt	actatatcct	caaacaggcg		2760
gtaaccggcc	tcttcacg	gaatgcgcgc	gaccttcage	atcgcgcgca	tgtccccctg		2820
gcggacggga	agtatccagc	tcgaggtcgg	gccgcgttgc	tggcgttttt	ccataggctc		2880
cgccccctcg	acgagcatca	caaaaatcga	cgtcaagtc	agaggtggcg	aaacccgaca		2940
ggactataaa	gataccaggc	gtttccccct	ggaagctccc	tcgtgcgctc	tcctgttccg		3000
accctgccgc	ttaccggata	cctgtccg	ttctccctt	cgggaagcgt	ggcgttttct		3060
catagctcac	gctgtaggta	tctcagttcg	gtgtaggctg	ttcgtcccaa	gctgggctgt		3120
gtgcacgaac	ccccggttca	gcccgaacgc	tgcgccttat	ccggttaacta	tcgtcttgag		3180
tccaaccogg	taagacacga	cttatcgcca	ctggcagcag	ccactggtaa	caggatttag		3240
agagcgaggt	atgtaggcgg	tgctacagag	ttcttgaaat	gggtggcctaa	ctacggctac		3300
actagaagga	cagtatttgg	tatctgcgct	ctgctgaagc	cagttacctt	cggaaaaaga		3360
gttgatagct	cttgatcgg	caaacaaaac	accgctggta	gcggtgggtt	ttttgtttgc		3420
aagcagcaga	ttacgcgcag	aaaaaaagga	tctcaagaag	atcctttgat	cttttctacg		3480
gggtctgacg	ctcagtgga	cgaaaactca	cgttaaggga	ttttggtcat	gagattatca		3540

[0504]

aaaaggatct	tcacctagat	cctttttaa	taaaaatgaa	gtttttaa	aatctaaa	3600
atatatgagt	aaacttgg	tgacagtt	caatgcttaa	tcagtgag	acctatct	3660
gcgatctgt	tatttcgt	atccatagt	gcctgact	ccgtcgtg	gataacta	3720
atacgggagg	gcttaccat	tggccccag	gctgcaat	taccgcga	cccacgct	3780
ccgggtccag	atztatcag	aataaacc	ccagccgga	gggcccga	cagaagtgg	3840
cctgcaactt	tatccgcct	catccagtc	attaattgt	gccgggaag	tagagtaa	3900
agttcgccag	ttaatagt	gcgcaacgt	gttgccatt	ctgcaggca	cgfggtgt	3960
cgctcgtcgt	ttgggtat	ttcattcag	tcoggttcc	aacgatcaa	gcgagttac	4020
tgatccccca	tgttgtgca	aaaagcgg	agctccttc	gtcctccga	cgttgtcag	4080
agtaagttgg	ccgcagtg	atcactcat	gttatggcag	cactgcata	ttctcttact	4140
gtcatgcca	ccgtaagat	cctttctgt	actggtgag	actcaacca	gtcattctga	4200
gaatagtgt	tgccgcgac	gagttgct	tgcccgcg	caacacggg	taataccgc	4260
ccacatagca	gaactttaa	agtgtctat	attggaaa	gttcttcgg	gcgaaaact	4320
tcaaggatct	taccgctgt	gagatccag	tcgatgta	ccactcgtg	acccaactga	4380
tcttcagcat	cctttacttt	caccagcgt	tctgggtgag	caaaaacagg	aaggcaaat	4440
gccgcaaaaa	aggggaata	ggcgacacg	aaatgttgaa	tactcatact	cttccttttt	4500
caatattatt	gaagcattta	tcagggttat	tgtctcatga	gcggatacat	atttgaatgt	4560
atttagaaaa	ataaacaat	aggggttccg	cgcacatttc	cccgaaga	gccacctgac	4620
gtctaagaaa	ccattattat	catgacatta	acctataaaa	ataggcgtat	cacgaggccc	4680
tttcgtcttc	agaatt					4697
<210>	202					
<211>	4697					
<212>	DNA					
<213>	合成					
<220>						
<221>	尚未归类的特征					
<222>	(1)..(4697)					
<223>	质粒 pEG158					
<220>						
<221>	尚未归类的特征					
<222>	(88)..(651)					
<223>	同源区域 iscU 上游					
<220>						
<221>	启动子					
<222>	(1833)..(2131)					
<223>	负荷成癌启动子 p(yflT)					
<220>						
<221>	rhs					
<222>	(2132)..(2137)					
<223>	rhs					
<220>						
<221>	尚未归类的特征					
<222>	(2137)..(2137)					
<223>	N 是 A、T、C 或 G					
<220>						
<221>	尚未归类的特征					
<222>	(2147)..(2624)					
<223>	同源区域 iscU 下游					
<400>	202					
aacaaaattc	tcagttcttc	acatcgggtt	gaaaggagga	agcgggaaga	tgaagtaaga	60
gggattttttg	actccgaagt	aagttcttcat	tgaggaaatt	ggaaaacacca	cgcatttaaga	120
gaagggaatat	tcggaaaaac	catttctcatt	tgttctgaaa	ggatgctttc	gtctcatctg	180
ccgctattttc	acgcctaaatc	gttttttgcgg	catccaacgc	agcactggta	caataaaaaag	240
caaatatattg	aagggattga	ttcatttattt	gcagcagcca	ggacagactg	aacctcgatt	300
atacgtcaga	aatggaaaaa	gcgatgcagc	tgtcacacgg	aattggatac	gaggaatatg	360
gacgcgcggt	ggaaacaaga	atgaaagtag	agcggcagag	agagcttgat	tatgaaaaaa	420
gcaagcggat	ttctgcggga	gcttattgag	aaaatcaatc	ttttaaaaaa	gggtgggaat	480
ccactctttt	tctttttcaa	tcgctttaat	aaagtgtctga	tctttgttaa	acttgaggaa	540
tgtgcaaggc	acgataaaaa	gttggtattta	cctttcattt	gttataaaat	tagaatgaga	600
ataaaatgag	aataatgatt	atatattttt	ggaggtatag	atatggctgc	tagttcggat	660
tagatccccc	tatgcaaggg	tttattgttt	tctaaaaatc	gattaccaat	tagaatgaat	720
attttcccaa	tattaaataa	taaaacaaaa	aaattgaaaa	aagtgtttcc	accatttttt	780
caattttttt	ataatttttt	taatctgtta	tttaaatagt	ttatagttaa	attttacattt	840
tcattagtcc	attcaatatt	ctctccaaga	taactacgaa	ctgctaacaa	aattctctcc	900
ctatgttcta	atggagaaga	ttcagccact	gcattttccg	caatatcttt	tggtatgatt	960
ttaccocgtgt	ccatagttaa	aatcatacgg	cataaagtta	atatagagtt	ggtttcatca	1020
tcttgataat	tatctattaa	ttcctctgac	gaatccataa	tggtctcttc	cacatcagaa	1080
aatggaatat	caggtagtta	ttcctctaag	tcataatttc	cgtatattct	tttatttttt	1140
cgttttgctt	ggtaaagcat	tatgggttaa	tctgaattta	attccttctg	aggaatgtat	1200
ccttggtcat	aaagctcttg	taaccattct	ccataaataa	attcttgttt	gggaggatga	1260
tttcacggtg	ccatttcttg	ctgaataata	attgttaatt	caatatatcg	taagttgctt	1320
ttatctccta	ttttttttga	aataggtcta	attttttgta	taagtatttc	tttactttga	1380
tctgtcaatg	gttcagatac	gacgactaaa	aagtaaatga	cactatttgg	ttttagtcca	1440
ctctcaactc	ctgatccaaa	catgtaagta	ccaataaggt	tattttttta	atgtttccga	1500
agtatttttt	tcactttatt	aatttgttcg	tatgtattca	aatatatcct	cctcactatt	1560
ttgattagta	cctattttat	atccatagtt	gttaatttaa	taaacttaat	ttagtttatt	1620

[0505]

[0506]

tatagatttc	attggcttct	aaatttttta	tctagataat	aattatttta	gttaatttta	1680
ttctagatta	tatatgatat	gatctttcat	ttccataaaa	ctaaagtaag	tgtaaaccta	1740
ttcattgttt	taaaaatata	tcttgccagt	caggttacgt	tattagttaa	agttattata	1800
acatgtattc	acgaacgaaa	atcgcatcga	ttgactaaat	cactcacaat	ccaagcattg	1860
tattgctgtt	ctgggtgaaa	gtattgcatt	agacatacct	cctgctcgta	cggataaagg	1920
cagcgtttca	tggtcgtgtg	ctccgtgcag	cggcttctcc	ttaattttga	tttttctgaa	1980
aataggtccc	gttccctatca	ctttaccatg	gacggaaaaac	aaatagctac	taccattcct	2040
cctgtttttc	tcttcaatgt	tctggaatct	gtttcaggta	cagacgatcg	ggatatgaaag	2100
aaatatagaa	aacatgaagg	aggaatatcg	accgagntga	aatcatatgg	ctgcttcaac	2160
attaacgatc	aaagatcttc	acgttgaaat	cgaagggaaa	gagatcttaa	aggggtgtaa	2220
ccttgaaata	aaaggtggag	aattccacgc	agtaatgggc	ccgaacggaa	ctggtaaatc	2280
cacttttatca	gctgctatta	tggggcatcc	taaatatgaa	gtaacaaaag	gcagcatcac	2340
gcttgacggc	aaagatgtac	tggagatgga	agtggacgag	cgtgcgcagg	caggcttatt	2400
cgttgcaatg	cagtacccaa	gtgaaatcag	cgggtgtgaca	aatgccgact	tccttcgctc	2460
agcaatcaac	gcgcgcagag	aagaaggcga	tgaattttct	ctcatgaaat	tcattccgca	2520
aatggacgaa	aacatggagt	tccttgaaat	ggaccctgaa	atggctcagc	gctatcttaa	2580
cgaaggcttc	tcaggcgggg	agaaaaaacg	caacgaaatc	cttcaatttt	attgcaataa	2640
cagggtgctta	cttttctgga	gttcttttagc	aaattttttt	attagctgaa	cttagtatta	2700
gtggccatac	tcctccaatc	caaagctatt	tagaaagatt	actatacctc	caaacaggcg	2760
gtaacggccc	tcttcacggc	gaatgcgcgc	gacctcagc	atcgccggca	tgtcccctcg	2820
gcggacggga	agtatccagc	tcgaggtcgg	gccgcgttgc	tggcgttttt	ccataggctc	2880
cgccccctcg	acgagcatca	caaaaaatcga	cgtccaagtc	agaggtggcg	aaaccgcaca	2940
ggactataaa	gataccaggc	gtttccccct	ggaagctccc	tcgtgcgctc	tctgtttccg	3000
accctgcgcg	ttaccggata	cctgtccgct	tttctccctt	cgggaagcgt	tccttcgctc	3060
catagctcac	gctgtaggtta	tctcagttcg	gtgtaggtcg	ttcgctccaa	gctgggctgt	3120
gtgcacgaac	ccccggttca	gcccgaaccgc	tgcgccttat	ccggtaaacta	tcgtcttgag	3180
tccaaccogg	taagacacga	cttatcgcca	ctggcagcag	ccactggtaa	caggattgaa	3240
agagcgaggt	atgtaggcgg	tgctacagag	ttcttgaagt	ggtggcctaa	ctacggctac	3300
actagaagga	cagtattttg	tatctgcgct	ctgctgaagc	cagttacctt	cggaaaaaga	3360
gttgatagct	cttgatccgg	caaaacaaacc	accgctggta	gcggtggttt	ttttgtttgc	3420
aagcagcaga	ttacgcgcag	aaaaaaagga	tctcaagaag	atcctttgat	cttttctacg	3480
gggtctgacg	ctcagtgga	cgaaaaactca	cgttaaggga	ttttggtcat	gagattatca	3540
aaaaggatct	tcacctagat	cctttttaa	taaaaatgaa	gtttttaa	aatctaaaagt	3600
atatatgagt	aaacttggtc	tgacagttac	caatgcttaa	tcagttaggc	acctatctca	3660
gcgatctgtc	tatttcgttc	atccatagtt	gcttgactcc	ccgtcgtgta	gataactacg	3720
atacgggagg	gcttaccatc	tgcccccagt	gctgcaatga	taccgcgaga	cccacgctca	3780
ccggctccag	atttatcagc	aataaaccag	ccagccggaa	gggcccagcg	cagaagtgg	3840
cctgcaactt	tatccgcctc	catccagttc	attaattgtt	gccgggaagc	tagagtaagt	3900
agttcgcag	ttaatagttt	gcgcaacgtt	gttgccattg	ctgcaggcat	cgtgggtgta	3960
cgctcgcgtg	ttggtatggc	ttcattcagc	tcogggtccc	aacgatcaag	gcgagttaca	4020
tgatccccca	tggttggtgaa	aaaagcgggt	agctccttcg	gtcctccgat	cgttggtcaga	4080
agtaagtgg	ccgcagtggt	atcactcatg	gttatggcag	cactgcataa	ttctcttact	4140
gtcatgcoat	cgttaagatg	cttttctgtg	actgggtgag	actcaacca	gtcattctga	4200
gaatagtgtg	tgccggcgacc	gagttgtctc	tgcccgccgt	caacacggga	taataccgcg	4260
ccacataagca	gaactttaaa	agtgtctatc	attggaaaaac	gttcttcggg	gcgaaaaactc	4320
tcaaggtact	taccgctgtt	gagatccagt	tcgatgtaac	ccactcgtgc	acccaactga	4380
tcttcagcat	cttttacttt	caccagcgtt	tctgggtgag	caaaaaacagg	aaggcaaaat	4440
gccgcaaaaa	agggaaataag	ggcgacacgg	aaatgttgaa	tactcatact	cttccttttt	4500
caatatattt	gaagcatttta	tcaggggttat	tgtctcatga	gcggatacat	atttgaatgt	4560
atttagaaaa	ataaacaat	aggggttccg	cgcacatttc	cccgaagagt	gccacctgac	4620
gtctaagaaa	ccattattat	catgacatta	acctataaaa	ataggcgtat	cacgaggccc	4680
tttcgtcttc	aagaattt					4697

<210> 203
 <211> 4696
 <212> DNA
 <213> 合成
 <220>
 <221> 尚未归类的特征
 <222> (1) .. (4696)
 <223> 质粒 pEG159
 <220>
 <221> 尚未归类的特征
 <222> (88) .. (676)
 <223> 同源区域 aaC 上游
 <220>
 <221> 启动子
 <222> (1858) .. (2153)
 <223> 负荷成癌启动子 p(groES)
 <220>
 <221> rbs
 <222> (2154) .. (2159)
 <223> rbs
 <220>
 <221> 尚未归类的特征
 <222> (2154) .. (2159)
 <223> N 是 A、T、C 或 G
 <220>

<221> 尚未归类的特征
<222> (2169)..(2623)
<223> 同源区域 aaC 下游
<400> 203

[0507]

aacaaaatttc	tccagtccttc	acatcgggttt	gaaaggaggga	agcgggaagaa	tgaagtaaga	60
gggattttttg	actccgaagt	aagtcttagg	gaacgaaaga	ttcatcagca	gacaaggaaa	120
catccgctga	agccagtgaa	aaaggaactg	ttgttacgga	aacagccgat	gatgattttgt	180
ttacaacgta	ccgtcttgat	ttagaagatg	ccagaagcaa	agaaagagaa	gagctgaatg	240
ccattgtgtc	aagcgatgat	gcaacagcca	aggagaaaag	cgaagcatal	gataaaatga	300
cggtctctcag	tgaagtgga	ggaacagaaa	aacagctgga	aacgctgatt	aaaacacaag	360
gttacgagga	tgcgcttggt	aatgctgaag	gagataaaat	caatattaca	gtcaaatcag	420
acaaacactc	taaatcgaag	gcgacagcca	ttatagacct	tgtggcaaaa	gaaatcaaaa	480
caatgaaaga	tgctcgctgtc	acattttgaac	cctctaaata	agaatgaggg	aaaaaagccc	540
gctaacaag	cggtcttttt	gcgttgctgt	catattagag	ttgaattcaa	aagtcogctc	600
ctgttaagatg	aacatagtat	gtacttttag	tagtaaccta	taaaaaaaga	aatttcatac	660
ataggagtgc	gattgaagtt	cggattagat	ccccctatgc	aaggggtttat	tgttttctaa	720
aatctgatta	ccaatttagaa	tgaatatttc	ccaaatatta	aataataaaa	caaaaaaatt	780
gaaaaaaagt	tttccaccat	tttttcaatt	tttttataat	tttttataat	tgttttttaa	840
atagttttata	gttaaatttta	catttttcatt	agtccattca	atattctctc	caagataact	900
acgaactgct	aacaaaattc	tctccctatg	ttctaatgga	gaagattcag	ccactgcatt	960
tcccgcata	tcttttggtta	tgatttttacc	cgtgtccata	gttaaaatca	tacggcataa	1020
agttaatata	gagttgggtt	catcatcctg	ataattatct	attaattcct	ctgacgaatc	1080
cataatggct	cttctcacat	cagaaaaatg	aatatcagg	agtaattcct	ctaagtcata	1140
atttccgctat	attcttttat	tttttcgttt	tgcttggtta	agcattatgg	ttaaatctga	1200
atttaattcc	ttctgaggaa	tgtatccttg	ttcataaagc	tcttgtaacc	attctccata	1260
aataaattct	tgtttgagg	gatgattcca	cggtagccat	tcttgctgaa	taataattgt	1320
taattcaata	tactgtaagt	tgctttttat	tctatatttt	tttgaaatag	gtcctaattt	1380
ttgtataagt	atttctttac	tttgatctgt	caatgggtta	gatacgacga	ctaaaaagtc	1440
aagatcacta	tttggttttta	gtccactctc	aactcctgat	ccaaacatgt	aagtaccaat	1500
aagggtatttt	tttaaatggt	tccgaagtat	ttttttcact	ttattaattt	tgtctgtatg	1560
attcaaatat	atcctcctca	ctattttgat	tagtacctat	tttatatcca	ttgttgttaa	1620
ttaaataaac	ttaattttagt	ttattttatag	atttcattgg	ccttctaaatt	ttttatctag	1680
ataataatta	tttttagtta	ttttattcta	gatttatatat	gatattgatct	ttcattttcca	1740
taaaactaaa	gtaagtgttaa	acctattcat	tgcttttaaaa	atatctcttg	ccagtcacgt	1800
tacgttatta	gttatagtta	ttataacatg	tattcacgaa	cgaaaatcgc	atcgattaat	1860
gcggaaact	gcatgatgat	atatgtcaaa	ataataaacc	aataactggt	tctcaaatgg	1920
tatgtactcc	tttggttaagt	gggtttctgt	catctacagc	tattgtaaca	taatcggtac	1980
gggggtgaaa	aagctaaccg	aaaaggggag	ggaaaagaat	gatgtaagcg	tgaaaaattt	2040
tttatcttat	cacttgaaat	tggaaggagg	attcttttatt	ataagaattg	tggttagact	2100
cttttagtgt	gagtgtctaaa	attacatatt	catactattg	aggaggttat	ttcacagant	2160
gaaatcatat	gttaaatatc	aaagaaatcc	acgagctgat	taaaagcaat	gacgagtcta	2220
caattgacga	attcgtatat	gaaaatgaag	gtgtatcctt	aaaaactgaa	aaacacgaag	2280
caggcacggt	tcaagtcatg	cagcaggcac	cggcagcacc	tgtacaagca	caggctccgc	2340
aggcgaatca	ccgcgaagcg	cagcaagcag	cggcacctgc	ccaagaagca	ccaagcaag	2400
atgagaatct	gcataaaatc	acttcaccaa	tggtaggaac	atttttatgct	tcttcatac	2460
cgggaactgg	cccgatgta	acagccggtt	caaaaagtaaa	tgaaaacaca	gttgtctgca	2520
ctgtagaagc	gatgaagctt	ttcaacgaaa	togaagcaga	agtgaaggcg	gaaatcgttg	2580
aagtattagt	agaaaacggc	cagctggctg	aatacggaca	accaatttta	ttgcaataac	2640
aggtgcttac	ttttctggag	ttcttttagca	aatttttttta	ttagctgaac	ttagatttag	2700
tggtccatcc	cctccaatcc	aaagctattt	agaaagatta	ctatatctct	aaacagcgcg	2760
taaccggcct	cttcactcgg	aatgcgcgcg	accttcagca	tcgcccggcat	gtccccctgg	2820
cggagcggga	gtatccagct	cgaggtcggg	ccgcgttgct	ggcggttttc	cataggtctc	2880
gccccctga	cgagcatcac	aaaaatcgac	gctcaagtca	gaggtggcga	aaccgcagac	2940
gactataaag	ataccagcgc	tttccccctg	gaagctccct	cgtgcgctct	cctggtccga	3000
ccctgccgct	taccggatac	ctgtccgcct	ttctcccttc	gggaagcgtg	gcgctttctc	3060
atagctcacg	ctgtaggtat	ctcagttcgg	tgtaggtcgt	tcgctccaag	ctgggctgtg	3120
tgcacgaacc	ccccgttcag	cccgaccgct	gcgccttate	cggtaactat	cgtcttgagt	3180
ccaaccgggt	aagacacgac	ttatcgccac	tggcagcagc	cactggtaac	aggattagca	3240
gagcgaagga	tgtaggcggt	gctacagagt	tcttgaaagt	gtggcctaac	tacggctaca	3300
ctagaaggac	agtattttgt	atctgcgctc	tgctgaagcc	agttaccttc	ggaaaaagag	3360
ttgatagctc	ttgatccggc	aaacaaacca	ccgctggtag	cggtggtttt	tttgtttgca	3420
agcagcagat	tacgcgcaga	aaaaaaggat	ctcaagaaga	tccttttgatc	ttttctacgg	3480
ggtctgacgc	tcagtggaa	gaaaactcac	gttaaagggt	tttggtcatg	agattatcaa	3540
aaaggatctt	cacctagatc	cttttaaat	aaaaatgaag	ttttaaatca	atctaaagta	3600
tatatagtat	aacttggtct	gacagttaac	aatgcttaat	cagtgaaggca	cctatctcag	3660
cgatctgtct	atttcgttca	tccatagttg	cctgactccc	cgtcgtgtag	ataactacga	3720
tacgggaagg	cttaccatct	ggccccagtg	ctgcaatgat	accgcgagac	ccacgctcac	3780
cggctccaga	tttatcagca	ataaaccagc	cagccgggaag	ggccgagcgc	agaagtggtc	3840
ctgcaacttt	atccgcctcc	atccagttca	ttaatgtttg	ccgggaagct	agagtaagta	3900
gttcgacagt	taatagtttg	cgcaacgttg	ttgcatctgc	tgcaaggcatc	gtgggtgcac	3960
gctcgtcgtt	tggtatggct	tcattcagct	ccggttccca	acgatcaagg	cgagttacat	4020
gateccccat	gttgtgcaaa	aaagcgggta	gctcctctcg	tcctccgata	gttgtcagaa	4080
gtaagtgtgc	cgcagtgtta	tcactcatgg	ttatggcagc	actgcataat	tctcttactg	4140
tcattgccatc	cgtaagatgc	ttttctgtga	ctgggtgagta	ctcaaccaag	tcattctgag	4200
aatagtgtat	gcggcgacgc	agttgtctct	gcccggcgct	aacacgggat	aataccgcgc	4260
cacatagcag	aactttaaaa	gtgctcatca	ttggaaaaacg	ttcttcgggg	cgaaaaactct	4320
caaggatctt	accgctgttg	agatccagtt	cgatgtaacc	cactcgtgca	cccaactgat	4380
cttcagcatc	ttttactttc	accagcgttt	ctgggtgagc	aaaaacagga	aggcaaaatg	4440

[0508]

```

ccgcaaaaaa ggggaataagg gcgacacgga aatgttgaat actcatactc ttccttttttc 4500
aatattattg aagcattttat cagggttatt gtctcatgag cggatacata tttgaatgta 4560
tttagaaaaa taaacaaata ggggtttccgc gcacatttcc ccgaaaagtg ccacctgacg 4620
tctaagaaac cattattatc atgacattaa cctataaaaa taggcgtatc acgaggccct 4680
ttcgtcttca agaatt
<210> 204
<211> 4699
<212> DNA
<213> 合成
<220>
<221> 尚未归类的特征
<222> (1) .. (4699)
<223> 质粒 pEG160
<220>
<221> 尚未归类的特征
<222> (88) .. (676)
<223> 同源区域 accC 上游
<220>
<221> 启动子
<222> (1858) .. (2156)
<223> 负荷成瘾启动子 p (ctsR)
<220>
<221> rbs
<222> (2157) .. (2162)
<223> rbs
<220>
<221> 尚未归类的特征
<222> (2162) .. (2162)
<223> N 是 A、T、C 或 G
<220>
<221> 尚未归类的特征
<222> (2172) .. (2626)
<223> 同源区域 accC 下游
<400> 204
aacaaaaattc tccagtcttc acatcggttt gaaaggaggga agcgggaagaa tgaagtaaga 60
gggattttttg actccgaagt aagtcttagg gaacgaaaga ttcatacagca gacaaggaaa 120
catccgctga agccagtga aaaggaactg ttgttacgga aacagccgat gatgatttgt 180
ttacaacgta ccgtcttgat ttagaagatg ccagaagcaa agaaagagaa gagctgaatg 240
ccattgtgtc aagcgatgat gcaacagcca aggagaaaag cgaagcatal gataaaatga 300
cggctctcag tgaagtggaa ggaacagaaa aacagctgga aacgctgatt aaaacacaag 360
gttacgagga tgcgcttggt aatgctgaag gagataaaat caatattaca gtcaaatcag 420
acaaacactc taaatcgaag gcgacagcca ttatagacct tgtggcaaaa gaaatcaaaa 480
caatgaaaga tgtcgctgtc acatttgaac cctctaaata agaatgaggg aaaaaagccc 540
gctaaacaag cgggcttttt gcttttagag catattagag ttgaattcaa aagtcgctc 600
ctgtaagatg aacatagtat gtacttttag tagtaaccta taaaaaaaaga aatttcatal 660
ataggagtgc gattgaagtt cggattagat cccctatgca aagggtttat tgttttctaa 720
aatctgatta ccaattagaa tgaatatctt ccaaatatta aataataaaa caaaaaaatt 780
gaaaaaagtg tttccaccat tttttcaatt tttttataat ttttttaate tgttatttaa 840
atagtttata gttaaattta cattttcatt agtccattca atattctctc caagataact 900
acgaactgct aacaaaaattc tctccctatg ttctaattga gaagattcag ccactgcatt 960
tcccgcataa tcttttggtg tgattttacc cgtgtccata gttaaaatca tacggcataa 1020
agttaatata gagttgggtt catcatcctg ataattatct attaatccct ctgacgaatc 1080
cataatggct cttctcacat cagaaaatgg aatatcaggt agtaattcct ctaagtcata 1140
atttccgtat attcttttat tttttcgttt tgcttggtaa agcattatgg ttaaattctga 1200
atttaattcc ttctgaggaa tgtatccttg ttcataaaagc tcttgtaacc attctccata 1260
aataaattct tgtttgggag gatgattcca cggtaaccatt tcttgctgaa taataattgt 1320
taattcaata tatcgtaagt tgcttttacc tctatttttt tttgaaatag gtctaatttt 1380
ttgtataagt atttctttac tttgatctgt caatggttca gatacgacga ctaaaaagtc 1440
aagatcacta tttgggtttta gtccactctc aactcctgat ccaaacatgt aagtaccaat 1500
aaggttattt tttaaatggt tccgaagtat ttttttcaat ttattaatit ttattaatit 1560
attcaaatat atcctcctca ctattttgat tagtaacctat tttatatcca tagttgttaa 1620
ttaaataaac ttaatttagt ttatttatag atttcattgg cttctaaatt ttttatctag 1680
ataataatta ttttagttaa ttttattcta gattatataat gatatgatct ttcatttcca 1740
taaaactaaa gtaagtgtaa acotattcat tgttttaaaa atatctcttg ccagtcacgt 1800
tacgttatta gttatagtta ttataacatg tattcacgaa cgaaaatcgc atcgatttgt 1860
cttcagcgcc gatggtagtc ggggggtttcc ccctgtgaga gtaggacgcc gccaaagcaag 1920
cttaaaccca gctcaatgag ctgggttttt tgatttttgg tttattggta tcataaaaatt 1980
ccacttaact gtataatata ataactttat accgaatttt aaatcagcaa tcaggtttttg 2040
tggaaccgga aaatggaaat aatgaaggat agagcgagaa agttgaaaat tctcgagaaa 2100
cggcttatag taagattaaa gtcaaatata gtcaaagtca gtaaaggagg ggggttgacga 2160
gntgaaatca tatgttaaat atcaaagaaa tccacgagct gattaaaagca attgacgagt 2220
ctacaattga cgaattcgtg tatgaaaatg aaggtgtatc cttaaaactg aaaaaacacg 2280
aagcaggcac ggttcaagtc atgcagcagg caccggcagc acctgtacaa gcacaggctc 2340
cgcaggcagt tcagccgcaa gcgcagcaag cagcggcacc tgcccaagaa gcaccaagc 2400
aagatgagaa tctgcataaa atcacttcac caatggtagg aacattttat gcttcttcat 2460
caccggaagc tggcccgatg gtaacagccg gttcaaaagt aaatgaaaac acagtttgtc 2520

```

gcattgtaga	agcgatgaag	cttttcaacg	aaatcgaagc	agaagtgaag	ggcgaaatcg	2580
ttgaagtatt	agtagaaaac	ggccagctgg	tcgaatacgg	acaaccaatt	ttattgcaat	2640
aacagggtct	tacttttctg	gagttcttta	gcaaatTTTT	ttattagctg	aacttagtat	2700
tagtggccat	actcctccaa	tccaaagcta	tttagaaaga	ttactatatc	ctcaaacagg	2760
cggtaacccg	cctcttcac	gggaatgcgc	gcgaccttca	gcacgcgcg	catgtccccc	2820
tggcggacgg	gaagtatcca	gctcgaggtc	gggcccgttt	gctggcgctt	ttccataggg	2880
tccgcccccc	tgacgagcat	cacaaaaatc	gacgctcaag	tcagaggtgg	cgaaaccgga	2940
caggactata	aagataccag	gcgtttcccc	ctggaagctc	cctcgtgcgc	tctcctgttc	3000
cgaccctgcc	gcttaccgga	tacctgtccg	cctttctccc	ttcggaagc	gtggcgcttt	3060
ctcatagctc	acgctgtagg	tatctcagtt	cgggttaggt	cgttcgctcc	aagctgggct	3120
gtgtgcacga	acccccggt	cagcccgacc	gctgcgcctt	atccggtaac	tatcgtcttg	3180
agtcacaacc	ggtaagacac	gaattatcgc	cactggcagc	agccactggg	aacaggatta	3240
gcagagcgag	gtatgtaggc	ggtgtctacg	agttcttgaa	gtgggtggcct	aactacggct	3300
acactagaag	gacagtattt	ggtatctgcg	ctctgctgaa	gccagttacc	ttcggaaaaa	3360
gagttgatag	ctcttgatcc	ggcaaacaaa	ccaccgctgg	tagcgggtgg	ttttttgttt	3420
gcaagcagca	gattacgcgc	agaaaaaaag	gatctcaaga	agatcctttg	atcttttcta	3480
cggggctctg	cgctcagtg	aacgaaaact	cacgttaagg	gatttttggtc	atgagattat	3540
caaaaagggt	cttcacctag	atccttttaa	attaaaaatg	aagtttttaa	tcaatctaaa	3600
gtatatatga	gtaaaccttg	tctgacagtt	accaatgctt	aatcagtgag	gcacctatct	3660
cagcgatctg	tctatttctg	tcatccatag	ttgcctgact	ccccgtcgtg	tagataacta	3720
cgatacggga	gggcttacca	tctggcccc	gtgctgcaat	gataccgcga	gacccacgct	3780
caccggctcc	agattttatc	gcaataaacc	agccagccgg	aagggccgag	cgcagaagtg	3840
gtcctgcaac	tttatccgcc	tccatccagt	ctattaattg	ttgcccggaa	gctagagtaa	3900
gtagtctgcc	agttaatagt	ttgcccgaac	ttgttgccat	tgctgcaggc	atcgtggtgt	3960
cacgctcgtc	gtttgggatg	gcttcattca	gctccggttc	ccaacgatca	aggcgagtta	4020
catgatcccc	catgtttgtc	aaaaaaagcg	ttagctcctt	cggctcctcc	atcgttgtca	4080
gaagaaggtt	ggcgcagtg	ttatcactca	tggttatggc	agcactgcat	aattctctta	4140
ctgtcatgcc	atccgtaaga	tgcttttctg	tgactgggtg	gtactcaacc	aagtcattct	4200
gagaatagt	tatgcggcga	ccgagttgct	cttgcccggc	gtcaacacgg	gataataccg	4260
cgccacatag	cagaacttta	aaagtgtcca	tcattggaaa	acgttcttcg	gggcgaaaa	4320
tctcaaggat	cttaccgctg	ttgagatcca	gttcgatgta	acccactcgt	gcacccaact	4380
gatcttcacg	atcttttact	ttcaccagcg	tttctgggtg	agcaaaaaa	ggaaggcaaa	4440
atgcgcgaaa	aaagggaata	agggcgacac	ggaaatgttg	aatactcata	ctcttccttt	4500
ttcaaatatta	ttgaagcatt	tatcagggtt	attgtctcat	gagcggatag	atatttgaat	4560
gtatttagaa	aaataaacia	ataggggttc	cgcgcacatt	tccccgaaaa	gtgccacctg	4620
acgtctaaga	aaccattatt	atcatgacat	taacctataa	aaataggcgt	atcacgaggg	4680
cctttcgtct	tcaagaatt					4699
<210>	205					
<211>	4699					
<212>	DNA					
<213>	合成					
<220>						
<221>	尚未归类的特征					
<222>	(1) .. (4699)					
<223>	质粒 pEG161					
<220>						
<221>	尚未归类的特征					
<222>	(88) .. (676)					
<223>	同源区域 accC 上游					
<220>						
<221>	启动子					
<222>	(1858) .. (2156)					
<223>	负荷成瘾启动子 p (dnaK)					
<220>						
<221>	rb					
<222>	(2157) .. (2162)					
<223>	rb					
<220>						
<221>	尚未归类的特征					
<222>	(2162) .. (2162)					
<223>	N 是 A、T、C 或 G					
<220>						
<221>	尚未归类的特征					
<222>	(2172) .. (2626)					
<223>	同源区域 accC 下游					
<400>	205					
aacaaaaatt	tccagttctt	acatcggttt	gaaaggagga	agcgggaagaa	tgaagtaaga	60
gggatttttt	actccgaagt	aagtcttagg	gaacgaaaga	ttcatcagca	gacaaggaaa	120
catccgctga	agccagtgaa	aaagggaact	ttgttacgga	aacagccgat	gatgattttg	180
ttacaacgta	ccgtcttgat	ttagaagatg	ccagaagcaa	agaaagagaa	gagctgaatg	240
ccattgtgtc	aagcgatgat	gcaacagcca	aggagaaaag	cgaagcatal	gataaaaatg	300
cggctctcag	tgaagtggaa	ggaacagaaa	aacagctgga	aacgctgatt	aaaacacaag	360
gttacgagga	tgcgcttggt	aatgctgaag	gagataaaat	caatattaca	gtcaaatcag	420
acaaacactc	taaatcgaag	gcgacagcca	ttatagacct	tgtggcaaaa	gaaatcaaaa	480
caatgaaaga	tgctcgctgtc	acatttgaac	cctctaaata	agaatgaggg	aaaaaagccc	540
gctaaacaag	cgggcttttt	gcgttgctgt	catattagag	ttgaattcaa	aagtcgcgtc	600

[0509]

[0510]

```

ctgtaagatg aacatagtat gtacttttag tagtaaccta taaaaaaaga aatttcatac 660
ataggagtg gattgaagtt cggatttagat cccctatgc aagggtttat tgttttctaa 720
aatctgatta ccaattagaa tgaatatttcc ccaaatatta aataataaaa caaaaaaatt 780
gaaaaaagtg tttccaccat tttttcaatt tttttataat ttttttaatc tgttatttaa 840
atagtttata gttaaattta cattttcatt agtccattca atattctctc caagataact 900
acgaactgct aacaaaaatc tctccctatg ttctaattgga gaagattcag ccactgcatt 960
tcccgcaata tcttttggtg tgattttacc cgtgtccata gttaaaatca tacggcataa 1020
agttaatata gagttggttt catcatcctg ataattatct attaatcct ctgacgaatc 1080
cataatggct cttctcacat cagaaaatgg aatatcaggt agtaattcct ctaagtcata 1140
atttcogtat attcttttat tttttcgttt tgcttggtaa agcattatgg ttaaatctga 1200
atttaattcc ttctgaggaa tgtatccttg ttcataaagg tcttgtaacc attctccata 1260
aataaattct tgtttgggag gatgattcca cgttaccatt tcttgctgaa taataattgt 1320
taattcaata tatcgtaagt tgccttttat tctatttttt tttgaaatag gtctaatttt 1380
ttgtaaatg atttctttac ttgatctgt caatggttca gatacgacga ctaaaaaagtc 1440
aagatcacta tttgggtttta gtccactctc aactcctgat aactcctgat ccaaacatgt aagtaccaat 1500
aaggttattt tttaaatggt tccgaagtat ttttttcact ttattaattt gttcgtatgt 1560
attcaaatat atcctcctca ctattttgat tagtacctat tttatatcca tagttgttaa 1620
ttaaataaac ttaatttagt ttattttatag atttcattgg cttctaaatt ttttatctag 1680
ataataatta ttttagttaa ttttattcta gattatatat gatattgatc ttcatttcca 1740
taaaactaaa gtaagtgtaa acctattcat tgttttaaaa atatctcttg ccagtcacgt 1800
tacgttatta gttatagtta ttataacatg tattcacgaa cgaaaatcgc atcgattggt 1860
tcaatccccg cacacgggaa tttcgattaa aatcggaaaa gaaaacgact atgaagagat 1920
ggaaaaattg agtctgatta cggcttctta ttccgtagac cagaagcaga tcggctcaat 1980
tgcatattc tgcccgacct gcatgaatta ttccagggtt tccagcctgc gtcagcctgc 2040
gacttcggac ttgtcaaaaag cattaacaag tctgtatgat gaataaggga attttgcaa 2100
attttatcga agggcagcac ctgtccttct ccttacactt tgagggaggt gaacacaga 2160
gntgaaatca tatgttaaat atcaaagaaa tccacgagct gattaaagca attgacaggt 2220
ctacaattga cgaattcgta aaggtgtatc aaggtgtatc cttaaaactg aaaaaacacg 2280
aagcaggcac ggttcaagtc atgcagcagg caccggcagc acctgtacaa gcacaggctc 2340
cgcaggcagt tcagccgcaa gcgccgcaac cagcggcacc tgcccaagaa gcaccaaacg 2400
aagatgagaa tctgcataaa atcacttcac caatggtagg aacattttat gcttcttcat 2460
caccggaagc tggcccgatg gtaacagccg gttcaaaagt aaatgaaaa acagttgtct 2520
gcattgtaga agcgatgaag cttttcaacg aaatcgaagc agaagtgaag ggcgaatcg 2580
ttgaagtatt agtagaaaac ggccagctgg tccaatacgg acaaccaatt ttattgcaat 2640
aacagtgctt tacttttctg gagttcttta gcaaattttt ttattagctg aacttagtat 2700
tagtggccat actcctccaa tccaaagcta tttagaanaa ttaactatct ctaaacagg 2760
cggtaaccgg cctcttcacg gggaatgcgc gcgaccttca gcatcgccgg catctcccc 2820
tgccggacgg gaagtatcca gctcgaggtc gggcccggtt gctggcgttt gctggcgttt 2880
tccgcccccc tgacgagcat cacaataatc gacgtcgaag tcagaggtgg cgaacccga 2940
caggactata agataaccag gcttttcccc ctggaagctc cctcgctgcg cctcgctgcg 3000
cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttccggaagc gtggcgcttt 3060
ctcatagctc acgctgtagg tatctcagtt cgtgtgaggt cgttcgctcc agctgggct 3120
gtgtgcagtc acccccgtt cagcccgacc gctgcgctt atccggtaac atccgtctg 3180
agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta 3240
gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct aactacgct 3300
acactagaag gacagtattt ggtatctgcg ctctgtgaa gccagttaac tccgtaaaa 3360
gagttgatag ctcttgatcc ggcaaacaaa ccaccgctgg tagcgggtgg ttttttgttt 3420
gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatccttg agatccttg 3480
cggggtctga cgtcagtggt aacgaaaact gattttgggt atgagattat 3540
caaaaaggat cttcacctag attcctttta aatttttaaa aagtttttaa atgttttaa 3600
gtatatatga gtaaaccttg accaatgctt aatcagtgag aatcagtgag 3660
cagcgatctg tctatttctg ttgctgact ccccgctgtg ccccgctgtg tagataacta 3720
cgatacggga tctggcccaa gtgctgcaat gatacgcga gatacgcga 3780
caccggctcc agatttatca gcaataaacc agccagccgg gtagagtaa 3840
gtcctgcaac tttatccgcc tccatccagt ttgctgcaat aaggccggaa gctagagtaa 3900
gtgttcgccc agttaatagt ttgocgaacg ttggtgcat tgctgcaggc atcggtgtg 3960
cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca aggcagatta 4020
catgatcccc catgttgtgc aaaaaagcgg tttagctcct cggctcctcg atcggtgca 4080
gaagtaagtt ggcgcagtg ttttatggc agcactgcat aattctctta 4140
ctgtcatgcc atccgtaaga tgactgggtg gtactcaacc aagtcattct 4200
gagaatagtg tatgcggcga cctgcccggc gtcaacacgg gataataccg 4260
cgccacatag cagaacttta aaagtgtctc tcatggaaa acgttcttcg gggcgaaaac 4320
ttcaatatta cttaccgctg ttgagatcca gttcgatgta acccactcgt gcaccact 4380
gatcttcagc atcttttact ttaccagcgg tttctgggtg agcaaaaaca ggaaggcaaa 4440
atgccgcaaa aaagggata agggcgacac aatactcata ctcttctttt 4500
ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac atatttgaat 4560
gtatttagaa aaataaacia ataggggttc atgtgtcatt tccccgaaa gtgccacctg 4620
acgtctaaga aaccattatt atcatgacat taacctataa aatataggct atcacgaggg 4680
cctttcgtct tcaagaatt

```

<210> 206
 <211> 4699
 <212> DNA
 <213> 合成
 <220>
 <221> 尚未归类的特征
 <222> (1) .. (4699)
 <223> 质粒 pEG162
 <220>

<221>	尚未归类的特征						
<222>	(88) .. (676)						
<223>	同源区域 accC 上游						
<220>							
<221>	启动子						
<222>	(1858) .. (2156)						
<223>	负荷成癮启动子 p(perR)						
<220>							
<221>	rbs						
<222>	(2157) .. (2162)						
<223>	rbs						
<220>							
<221>	尚未归类的特征						
<222>	(2162) .. (2162)						
<223>	N 是 A、T、C 或 G						
<220>							
<221>	尚未归类的特征						
<222>	(2172) .. (2626)						
<223>	同源区域 accC 下游						
<400>	206						
aacaaaattc	tccagtccttc	acatcgggttt	gaaaggaggga	agcgggaagaa	tgaagtaaga		60
gggatttttg	actccgaagt	aagtcttagg	gaacgaaaga	ttcatcagca	gacaaggaaa		120
catccgctga	agccagtgaa	aaaggaaactg	ttgttacgga	aacagccgat	gatgatttgt		180
ttacaacgta	ccgtcttgat	ttagaagatg	ccagaagcaa	agaaagagaa	gagctgaatg		240
ccattgtgtc	aagcgatgat	gcaacagcca	aggagaaaag	cgaagcatatc	gataaaatga		300
cggctctcag	tgaagtggaa	ggaacagaaa	aacagctgga	aacgctgatt	aaaacacaag		360
gttacgagga	tgcgcttggt	aatgctgaag	gagataaaaat	caatattaca	gtcaaatcag		420
acaaacactc	taaatacgaag	gcgacagcca	ttatagacct	tgtggcaaaa	gaaatcaaaa		480
caatgaaaga	tgtcgtgtgc	acatttgaac	cctctaaata	agaaatgaggg	aaaaaagccc		540
gctaacaacg	cgggctttttt	gcgttgctgt	catatttagag	ttgaattcaa	aagtcgcgtc		600
ctgtaagatg	aacatagtat	gtacttttag	tagtaaccta	taaaaaaga	aatttcatac		660
ataggagtgc	gattgaagtt	cggattagat	ccccctatgc	aagggtttat	tgttttctaa		720
aatctgatta	ccaattagaa	tgaatatattc	ccaaatatta	aataataaaa	caaaaaaatt		780
gaaaaaagtg	tttccaccat	tttttcaatt	tttttataat	ttttttaatc	tgttatitaa		840
atagtttata	gttaaaattta	cattttcatt	agtccattca	atattctctc	caagataact		900
acgaactgct	acaaaaattc	tctccctatg	ttctaattgga	gaagattcag	ccactgcatt		960
tcccgcataa	tcttttggtg	tgattttacc	cgtgtccata	gttaaaatca	tacggcataa		1020
agtttaataa	gagttgggtt	catcatcctg	ataattatct	atataattct	ctgacgaatc		1080
cataatggct	ctctccacat	cagaaaatgg	aatatcagggt	agtaattcct	ctaagtcata		1140
atttccgtat	attcttttat	tttttcgttt	tgcttggtaa	agcattatgg	ttaaatctga		1200
atttaattcc	ttctgaggaa	tgtatccttg	ttcataaagc	tcttgttaacc	attctccata		1260
aataaattct	tgtttgggag	gatgattcca	cggtaaccatt	tcttgtctgaa	taataattgt		1320
taatttcaata	tatcgttaagt	tgctttttatc	tcctatttttt	tttgaatatag	gtctaatttt		1380
ttgtataagt	atttcttttac	tttgatctgt	caatgggttca	gatacgacga	ctaaaaagtc		1440
aagatcacta	tttgggtttta	gtccactctc	aactcctgat	ccaaacatgt	aagtaccaat		1500
aactgttattt	tttaaatgtt	tcogaagtat	ttttttcaact	ttattaattt	gttcgtatgt		1560
attcaaatat	atcctcctca	ctattttgat	tagtacctat	tttataacca	tagttgttaa		1620
ttaaataaac	ttaatattagt	ttattttatag	atttcatttg	cttctaaatt	ttttatctag		1680
ataataatta	ttttagttaa	ttttatttca	gattatata	gatattgatct	ttcattttca		1740
taaaactaaa	gtaagtgtaa	acctattcat	tgtttttaaaa	atatctcttg	ccagtcacgt		1800
tacgtttata	gttatagtta	ttataacatg	tattcaacgaa	cgaataatcgc	atcgattgga		1860
agcatttgat	gtgtggaagc	tgaagaaaaa	cttcggaaaa	gagtatatgg	gaatcgagcg		1920
ttccactttt	ttaatgtata	aagaaggcgc	gtcctataaa	gaatggagaa	aagtaaaaagt		1980
aaaggaccat	gtagccgaag	cgcttcaaac	cctcaaagat	atgtctgaaa	agtaaatctc		2040
tatgagccta	tgcttacttt	taccctgata	tcaattgaca	aattatcgta	gaaagagtta		2100
cactaattat	aaacattaca	atgtaagaat	ttttttttaga	agagaggtgc	atgacggcga		2160
gntgaaatca	tatgtttaat	atcaaagaaa	tccacgagct	gattaaagca	attgacgagt		2220
ctacaattga	cgaattcgta	tatgaaaatg	aagggtgtatc	cttaaaaactg	aaaaaacacg		2280
aagcaggcac	ggttcaaagtc	atgcagcagg	caaccggcagc	acctgtacaa	gcacaggctc		2340
cgcaggcagt	tcagccgcaa	gcgcagcaag	cagcggcacc	tgcccaagaa	gcaccaaagc		2400
aagatgagaa	tctgcataaa	atcacttcac	caatggtagg	aacattttat	gcttcttcat		2460
caccggaaagc	tggcccgtat	gtaacagccg	gttcaaaaagt	aaatgaaaac	acagttgtct		2520
gcattgtaga	agcgatgaag	cttttcaacg	aaatcgaagc	agaagtgaag	ggcgaaatcg		2580
ttgaagtatt	agtagaaaac	ggccagctgg	tcgaatacgg	acaaccaatt	ttatttgcaat		2640
aacagggtgct	tacttttctg	gagttcttta	gcaaattttt	ttattagctg	aacttagtat		2700
tagtggccat	actcctccaa	tccaaagcta	tttagaaaaga	ttactatatc	ctcaaacagg		2760
cggtaaccgg	cctcttcac	gggaatgcgc	gcgaccttca	gcacgcggcg	catgtccccc		2820
tggcggacgg	gaagtatcca	gctcgaggtc	ggcgccggtt	gctggcggtt	ttccataggg		2880
tccgcccccc	tgacgagcat	cacaaaaatc	gacgctcaag	tcagaggtgg	cgaaaccgga		2940
caggactata	aagataccag	gcgtttcccc	ctggaagctc	cctcgtgcgc	tctcctgttc		3000
gcaccctgcc	gcttaaccga	tacctgtccg	cctttctccc	ttcgggaagc	gtggcgcttt		3060
ctcatagctc	acgctgtagg	tatctcagtt	cgggtgaggt	cggtcgctcc	aagctgggct		3120
gtgtgcaoga	accccccggt	cagcccgacc	gctgcgcctt	atccggtaac	tatcgtcttg		3180
agtccaaccc	ggtaagacac	gacttatcgc	cactggcagc	agccactggt	aacaggatta		3240
cgagagcgag	gtatgtaggc	ggtgctacag	agtctctgaa	gtggtggcct	aactacggct		3300
acactagaag	gacagtatct	ggtatctgcg	ctctgctgaa	gccagttacc	ttcggaaaaa		3360
gagttgatag	ctcttgatcc	ggcaaacaaa	ccaccgctgg	tagcgttggt	ttttttgttt		3420

[0511]

[0512]

```

gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta 3480
cggggtctga cgctcagtg aacgaaaact cacgttaagg gatttttggtc atgagattat 3540
caaaaaggat cttcacctag atccttttaa attaaaaatg aagtttttaa tcaatctaaa 3600
gtatatatga gtaaaacttg tctgacagtt accaatgctt aatcagtgag gcacctatct 3660
cagcgatctg tctatttcgt tcatccatag ttgacctgact ccccgctcgtg tagataacta 3720
cgattacggga gggcttacca tctggcccca gtgctgcaat gataccgcga gaccacgct 3780
caccggctcc agattttatca gcaataaacc agccagccgg aaggggccgag cgcagaagtg 3840
gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa gctagagtaa 3900
gtagttcgcc agttaaatagt ttgcgcaacg ttgttgccat tgctgcaggc atcgtggtgt 3960
cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca aggcgagtta 4020
catgatcccc catgttggtgc aaaaaagcgg tttagctcct cggctcctcg atcgttgtca 4080
gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat aattctctta 4140
ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc aagtcattct 4200
gagaatagtg tatgcggcga ccgagttgct cttgcccgcc gtcaacacgg gataataccg 4260
cgccacatag cagaacttta aaagtgtctc tcattggaaa acgttcttcg gggcgaaaac 4320
tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt gcacccaact 4380
gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaa ggaaggcaca 4440
atgccgcaaa aaaggggaata agggcgacac ggaaatggtg aatactcata ctcttccttt 4500
ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggtac atatttgaat 4560
gtatttagaa aaataaacia ataggggttc cgcgcacatt tccccgaaa gtgccacctg 4620
acgtctaaga aaccattatt atcatgacat taacctataa aaataggcgt atcacagggc 4680
cctttcgtct tcaagaatt 4699
<210> 207
<211> 4699
<212> DNA
<213> 合成
<220>
<221> 尚未归类的特征
<222> (1) .. (4699)
<223> 质粒 pEG163
<220>
<221> 尚未归类的特征
<222> (88) .. (676)
<223> 同源区域 accC 上游
<220>
<221> 启动子
<222> (1858) .. (2156)
<223> 负荷成糖启动子 p(hrcA)
<220>
<221> rbs
<222> (2157) .. (2162)
<223> rbs
<220>
<221> 尚未归类的特征
<222> (2162) .. (2162)
<223> N 是 A、T、C 或 G
<220>
<221> 尚未归类的特征
<222> (2172) .. (2626)
<223> 同源区域 accC 下游
<400> 207
aacaaaaattc tccagtccttc acatcgggttt gaaaggagga agcgggaagaa tgaagtaaga 60
gggattttttg actccgaagt aagtcttagg gaacgaaaga ttcatacagca gacaaggaaa 120
catccgctga agccagtgaa aaaggaactg ttgttacgga aacagccgat gatgatttgt 180
ttacaacgta ccgtcttgat ttagaagatg ccagaagcaa agaaagagaa gagctgaatg 240
ccattgtgtc aagcgatgat gcaacagcca aggagaaaag cgaagcatal gataaaatga 300
cggctctcag tgaagtggaa ggaacagaaa aacagctgga aacgctgatt aaaacacaag 360
gttacgagga tgcgcttggt aatgctgaag gagataaaat caatattaca gtcaaatcag 420
acaaacactc taaatcgaag cgcacagcca ttatagacct tgtggcaaaa gaaatcaaaa 480
caatgaaaga tgctcgtgtc acatttgaac cctctaaata agaattgaggg aaaaaagccc 540
gctaacaag cggttctgtg gcgttgctgt catattagag ttgaattcaa aagtcgctc 600
ctgtaagatg aacatagtat gtacttttag tagtaacct aaaaaaaga aatttcatal 660
ataggagtgc gattgaagtt cggattagat cccctatgc aagggtttat tgttttctaa 720
aatctgatta ccaattagaa tgaatatctt ccaaatatta aataataaaa caaaaaaatt 780
gaaaaaagtg tttccaccat tttttcaatt tttttataat ttttttaatc tgttatttaa 840
atagtttata gttaaaattt cattttcatt agtccattca atattctctc caagataact 900
acgaactgct aacaaaattc tctccctatg ttctaattga gaagattcag ccactgcatt 960
tcccgcata tcttttggtg tgattttacc cgtgtccata gttaaaatca tacggcataa 1020
agttaatata gagttggttt catcatcctg ataattatct attaatcct ctgacgaatc 1080
cataatggct cttctcacat cagaaaatgg aatatcaggt agtaattcct ctaagtcata 1140
atttccgtat attcttttat tttttcggtt tgcttggtaa agcattatgg ttaaatctga 1200
atttaattcc ttctgaggaa tgtatccttg ttcataaagc tcttgtaacc attctcata 1260
aataaattct tgtttgggag gatgattcca cggtaaccatt tcttgctgaa taataattgt 1320
taattcaata tatcgtaagt tgcttttata tcctattttt tttgaaatag gtctaatttt 1380
ttgtataagt atttctttac tttgatctgt caatggttca gatacgacga ctaaaaagtc 1440
aagatcacta tttggtttta gtccactctc aactcctgat ccaaacatgt aagtaaccaat 1500

```


[0513]

```

aaggttatttt tttaaatggt tccgaagtat ttttttcaact ttattaatttt gttcgtatgt 1560
attcaaatat atcctcctca ctattttgat tagtacctat tttatatcca tagttgttaa 1620
ttaaataaac ttaatttagt ttattttatag atttcattgg cttctaaatt ttttatctag 1680
ataataatta ttttagttaa ttttatttcta gattatataat gatattgatct ttcattttcca 1740
taaaactaaa gtaagtgtaa acctattcat tgttttaaaa atatctcttg ccagtcacgt 1800
tacgttatta gttatagtta ttataacatg tattcacgaa cgaaaatcgc atcgatttga 1860
ggaagaaatg ttttttaggt agcagaggtgc agcaaaaaac gattcgaga 1920
aaaatacggc cgttccctgg atggactttt cccaagcgtg ttaaaagacc tcgctgaaaa 1980
agggctgatt cacaattcgg agtctgccgt ttatttgact catcaaggga aattattagg 2040
caatgaagtt tttggcgctt ttttgggtga gttataattg acatttttct tgtggtttga 2100
tacttttgtt atagaattag cactcgctta ttgagagtgc taacagaggt gatgatgcga 2160
gntgaaatca tatgtttaa atcaaaagaaa tccacgagct gattaaagca attgacgagt 2220
ctacaattga cgaattcgtat tatgaaaatg aaggtgtatc cttaaaactg aaaaaacacg 2280
aagcaggcac gggttcaagt atgcagcagg caccggcagc acctgtacaa gcacaggctc 2340
cgcaggcagt tcagccgcaa gcgcagcaag cagcggcacc tgcccaagaa gcaccaaagc 2400
aagatgagaa tctgcataaa atcacttcac caatggtagg aacattttat gcttcttcat 2460
caccggaagc tggcccggtat gtaacagcgt gttcaaaagt aaatgaaaac acagttgtct 2520
gcattgtaga agcgatgaag cttttcaacg aaatcgaagc agaagtgaag ggcgaaatcg 2580
ttgaagtatt agtagaaaac ggccagctgg tccaatacgg acaaccaatt ttattgcaat 2640
aacaggtgct tacttttctg gagttcttta gcaaattttt ttattagctg aacttagtat 2700
tagtgccat actcctccaa tccaaagcta tttagaaaga ttactatatc ctcaaacagg 2760
cggtaaccgg cctcttcctc gggaatgcgc gcgaccttca gcatcgccgg catgtccccc 2820
tggcggcagg gaagtatcca gctcgaggtc gggccgcggt gctggcggtt ttccatagcg 2880
tcgcgccccc tgacgagcat cacaataatc gacggttagg tcagaggtgg cgaacccga 2940
caggactata aagataccag gcgtttcccc ctggaagctc cctcggtgcg tctcctgttc 3000
cgacctgccc gcttaccgga tacctgtccg cctttctccc ttccgggaagc gtggcgcttt 3060
ctcatagctc acgctgtagg tatctcagtt cgggtgaggt cgttcgctcc aagctgggct 3120
gtgtgcacga acccccgtt cagcccgacc gctgcgcctt atccggtaac tatcgtcttg 3180
agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta 3240
cgagagcgag gtatgtagc ggtgctacag agttcttgaa gtggtggcct aactacggct 3300
acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc ttccgaaaaa 3360
gagttgatag ctcttgatcc ggcaaaacaaa ccaccgctgg tagcggtggg ttttttgtt 3420
gcaagcagca gattacgcgc agaaaaaaaag gatctcaaga agatcctttg atcttttcta 3480
cggggtctga cgctcagttg aacgaaaaact cactgttaagg gatttttggtc atgagattat 3540
caaaaaaggat cttcacctag atccttttaa attaaaaatg aagtttttaa tcaatctaaa 3600
gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag gcacctatct 3660
cagcgactcg tctatttctg tcatccatag ttgcttgact ccccgctcgt tagataacta 3720
cgatacggga gggcttacca gtgctgcaat gtagctgcaat gataccgga gacccacgt 3780
caccggtccc agatttatca gcaataaacc agccagccgg aagggccgag cgcagaagtg 3840
gtcctgcaac tttatccgct cccatccagt ctattaattg ttgccgggaa gctagagtaa 3900
gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctgcaggc atcgtggtgt 3960
cacgctcgtc gtttggtagt gcttcattca gctccggttc ccaacgatca aggcagatta 4020
catgatcccc catgttgtgc ttagctcctt tttagctcct aggtctcctc atcgttgtca 4080
gaagtaagtt ggccgcagtg ttatcactca tgggttatggc agcactgcat aattctctta 4140
ctgtcatgcc atccgtaaga tgcttttctg tgactgggtg gtactcaacc aagtcattct 4200
gagaatagtg tatgcggcga cggagttgct cttgcccggc gtcaaacagg gataataccg 4260
cgacacatag cagaacttta aaagtgtcca tcattggaaa acgttcttcg gggcgaaaaa 4320
tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt gcacccaact 4380
gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaa ggaaggcaaa 4440
atgccgaaa aaagggata agggcgacac ggaatgttg aatactcata ctcttcctt 4500
ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac atatttgaat 4560
gtatttagaa aaataaacia ataggggttc cgcgcacatt tccccgaaa gtgccacctg 4620
acgtctaaga aaccattatt atcatgacat taacctataa aaatagcgt atcacgaggc 4680
cctttcgtct tcaagaatt 4699

```

<210> 208
 <211> 4699
 <212> DNA
 <213> 合成
 <220>
 <221> 尚未归类的特征
 <222> (1)..(4699)
 <223> 质粒 pEG164
 <220>
 <221> 尚未归类的特征
 <222> (88)..(676)
 <223> 同源区域 accC 上游
 <220>
 <221> 启动子
 <222> (1858)..(2156)
 <223> 负荷成糖启动子 p(spxA)
 <220>
 <221> rbs
 <222> (2157)..(2162)
 <223> rbs
 <220>
 <221> 尚未归类的特征
 <222> (2162)..(2162)

<223> N 是 A、T、C 或 G
<220>
<221> 尚未归类的特征
<222> (2172)..(2626)
<223> 同源区域 accC 下游
<400> 208

[0514]

aacaaaattc	tccagtccttc	acatcgggttt	gaaaggaggga	agcgggaagaa	tgaagtaaga	60
gggatttttg	actccgaagt	aagtccttagg	gaacgaaaaga	ttcatcagca	gacaaggaaa	120
catccgctga	agccagtgaa	aaaggaactg	ttgttacgga	aacagccgat	gatgattttgt	180
ttacaacgta	ccgtccttgat	ttagaagatg	ccagaagcaa	agaaagagaa	gagctgaatg	240
ccattgtgtc	aagcgaatga	gcaacagcca	aggagaaaag	cgaagcatat	gataaaatga	300
cggctctcag	tgaagtggaa	ggaacagaaa	aacagctgga	aacgctgatt	aaaacacaaag	360
gttacgagga	tgcgcttgtt	aatgctgaag	gagataaaat	caatattaca	gtcaaatcag	420
acaaacactc	taaatcgaag	gcgacagcca	ttatagacct	tgtggcaaaa	gaaatcaaaa	480
caatgaaaga	tgctcgctgtc	acatttgaac	cctctaaaata	agaatgaggg	aaaaaagccc	540
gctaaacaag	cgggctttttt	gcgttgctgt	catatttagag	ttgaattcaa	aagtccgctc	600
ctgtaagatg	aacatagtat	gtacttttag	tagtaacctt	taaaaaaaga	aattttcatat	660
ataggagtgc	gattgaagtt	cggatttagat	ccccctatgc	aagggtttat	tgttttctaa	720
aatctgatta	ccaattagaa	tgaatatttc	ccaaatatta	aataataaaa	caaaaaaatt	780
gaaaaaagt	tttccaccat	tttttcaatt	tttttataat	ttttttaatc	tgttatttaa	840
atagtttata	gttaaattta	cattttcatt	agtccttcca	atattctctc	caagataaact	900
acgaactgct	aacaaaattc	tctccctatg	ttctaattgga	gaagattcag	ccactgcatt	960
tcccgcaata	gttttttgta	tgattttacc	cgtgtccata	gttaaaatca	tacggcataa	1020
agttaatata	gagttggttt	catcatcctg	ataatttatct	attaattcct	ctgacgaatc	1080
cataatggct	cttctccat	cagaaaatgg	aatatcaggt	agtaattcct	ctaagtcata	1140
atttccgtat	attcttttat	tttttcgttt	tgcttggtaa	agcattatgg	ttaaaatctga	1200
attttaatcc	ttctgaggaa	tgtatccttg	ttcataaagc	tcttgtaacc	attctccata	1260
aataaattct	tgtttgggag	gatgattcca	cgttaccatt	tcttgctgaa	taataattgt	1320
taattcaata	tatcgtaagt	tgctttttat	tcttattttt	tttgaaatag	gtctaatttt	1380
ttgtataagt	atttctttac	tttgatctgt	caatggttca	gatacgacga	ctaaaaagtc	1440
aagatcacta	tttggtttta	gtccactctc	aactcctgat	ccaaacatgt	aagtaccaat	1500
aaggttatatt	tttaaatggt	tccgaagtat	ttttttcact	ttattaattt	gttcgtatgt	1560
attcaaatat	atcctcctca	ctattttgat	tagtacctat	tttatatcca	tagttgttaa	1620
ttaaataaac	ttaatttagt	ttattttatg	atttcattgg	cttctaattt	ttttatctag	1680
ataataatta	tttttagtta	ttttattcta	gatttatatat	gatatgatct	ttcatttcca	1740
taaaactaaa	gtaagtgtaa	acctattcat	tgtttttaaaa	atatctcttg	ccagtcacgt	1800
tacgttatta	gttatagtta	ttataacatg	tattcacgaa	cgaaaatcgc	atcgattgaa	1860
aacattcata	cttataaaga	tgagaaatgg	tacggagaat	cgtacgaaaa	aacggatgag	1920
gtcttgaaaa	tcatacgatga	agaaaaacaa	aaaaatatatt	tcgatcagtt	aagctaattg	1980
agaataagaa	catatcatga	acattaaaaa	tttgctggat	tgacacattt	ttttcttagg	2040
atactgttca	cacttacttt	tttatagtat	aatacctgta	aagattcctt	tttttagagta	2100
attttcaatt	aagaataaaa	tagatgttca	tctactaat	tagaggagt	aagatgacga	2160
gntgaaatca	tatgttaaat	atcaaagaaa	tccacgagct	gattaaagca	attgacaggt	2220
ctacaattga	cgaattcgt	tatgaaaaat	aaggtgtatc	cttaaaactg	aaaaaacacg	2280
aagcaggcac	ggttcaagtc	atgcagcagg	caccggcagc	acctgtacaa	gcacaggctc	2340
cgcaggcagt	tcagccgcaa	gcgcagcaag	cagcggcacc	tgcccaagaa	gcaccaaacg	2400
aagatgagaa	tctgcataaa	atcacttcac	caatggtagg	aacattttat	gcttctctat	2460
caccggaagc	tggtcccgat	gtaacagccg	gttcaaaaag	aaatgaaaac	acagtttgtct	2520
gcattgtaga	agcgatgaag	cttttcaacg	aaatcgaagc	agaagtgaag	ggcgaaatcg	2580
ttgaagtatt	agttagaaaac	ggccagctgg	tcgaatacgg	acaaccaatt	ttattgcaat	2640
aacaggtgct	tacttttctg	gagttcttta	gcaaattttt	ttattagctg	aaacttagtat	2700
tagtggccat	actcctccaa	tccaaagcta	tttagaaaaga	ttactatata	ctcaaacagg	2760
cggtaacogg	cctcttcac	gggaatgcgc	gcgaccttca	gcacgcggg	catgtccccc	2820
tggtcggaagg	gaagtatcca	gctcgaggtc	gggcgcggtt	gctggcggtt	ttccataggc	2880
tccgcccccc	tgacgagcat	cacaaaaatc	gacgctcaag	tcagaggtgg	cgaaacccga	2940
caggactata	aagataaccag	gcgtttcccc	ctggtaagctc	cctcgtgcgc	tctctgttcc	3000
cgacctgtcc	gcttaccgga	tacctgtccg	cctttctccc	ttcgggaagc	gtggcgcttt	3060
ctcatagctc	acgctgtagg	tatctcagtt	cgggtgtagg	cgttcgctcc	aagctgggct	3120
gtgtgcacga	acccccggt	cagccccacc	gctgcgcctt	atccggtaac	tatcgtcttg	3180
agtcacaacc	ggtaagacac	gacttatcgc	cactggcagc	agccactggt	aacaggatta	3240
gcagagcgag	gtatgtaggc	ggtgctacag	agttcttgaa	gtgggtggct	aactacggct	3300
acactagaag	gacagtattt	ggtatctgcg	ctctgctgaa	gccagttacc	ttcggaaaaa	3360
gagttgatag	ctcttgatcc	ggcaaacaaa	ccaccgctgg	tagcgggtgg	ttttttgttt	3420
gcaagcagca	gattacgcgc	agaaaaaaag	gatctcaaga	agatcctttg	atcttttcta	3480
cggggtctga	cgctcagtg	aacgaaaaat	caggttaagg	gatttttggt	atgagattat	3540
caaaaaagg	cttcacctag	atccttttaa	attaaaaatg	aagtttttaa	tcaatctaaa	3600
gtatatatga	gtaaaccttg	tctgacagtt	accaatgctt	aatcagtgag	gcacctatct	3660
cagcgatctg	tctatttctg	tcatccatag	ttgcctgact	ccccgtcgtg	tagataacta	3720
cgatacggga	gggcttacca	tctggcccca	gtgctgcaat	gataccgcga	gaccacgct	3780
caccggctcc	agatttatca	gcaataaacc	agccagccgg	aagggccgag	cgcagaagtg	3840
gtcctgcaac	tttatccgcc	tccatccagt	ctattaattg	ttgccgggaa	gctagagtaa	3900
gtagttcgcc	agtttaagt	ttgcgcagc	ttgttgccat	tgctgcaggc	atcgtggtgt	3960
cacgctcgtc	gtttgggatg	gcttcattca	gctccggttc	ccaacgatca	aggcgagtta	4020
catgatcccc	catgtttgtc	aaaaaagcgg	ttagctcctt	cgttcctccg	atcgttgtca	4080
gaagttaagtt	ggcgcagtg	tttactctca	tggttatggc	agcactgcac	aattctctta	4140
ctgtcatgcc	atccgtaaga	tgcttttctg	tgactgggtg	gtactcaacc	aagtcattct	4200
gagaatagtg	tatgcggcga	cagagttgct	cttgcccgcc	gtcaacacgg	gataataccg	4260
cgcacatag	cagaacttta	aaagtgtctca	tcattggaaa	acgttctctg	ggcgaaaaac	4320

tctcaaggat	cttaccgctg	ttgagatcca	gttcgatgta	accactcgt	gcacccaact	4380
gatcttcagc	atcttttact	ttcaccagcg	tttctgggtg	agcaaaaaca	ggaaggcaaa	4440
atgccgcgaaa	aaagggaata	agggcgacac	ggaaatgttg	aatactcata	ctcttccttt	4500
ttcaatatata	ttgaagcatt	tatcagggtt	attgtctcat	gagcggatac	atatttgaat	4560
gtatttagaa	aaataaacia	atagggttct	cgcgcacatt	tccccgaaaa	gtgccacctg	4620
acgtctaaga	aaccattatt	atcatgacat	taacctataa	aaataggcgt	atcacgaggc	4680
cctttcgtct	tcaagaatt					4699
<210>	209					
<211>	4699					
<212>	DNA					
<213>	合成					
<220>						
<221>	尚未归类的特征					
<222>	(1) .. (4699)					
<223>	质粒 pEG165					
<220>						
<221>	尚未归类的特征					
<222>	(88) .. (676)					
<223>	同源区域 accC 上游					
<220>						
<221>	启动子					
<222>	(1858) .. (2156)					
<223>	负荷成糖启动子 p (sigB)					
<220>						
<221>	rbs					
<222>	(2157) .. (2162)					
<223>	rbs					
<220>						
<221>	尚未归类的特征					
<222>	(2162) .. (2162)					
<223>	N 是 A、T、C 或 G					
<220>						
<221>	尚未归类的特征					
<222>	(2172) .. (2626)					
<223>	同源区域 accC 下游					
<400>	209					
aacaaaaattc	tccagttcttc	acatcggttt	gaaaggagga	agcgggaagaa	tgaagtaaga	60
gggattttttg	actccgaagt	aagtcttagg	gaacgaaaga	ttcatcagca	gacaaggaaa	120
catccgctga	agccagtga	aaaggaaactg	ttgttacgga	aacagccgat	gatgatttgt	180
ttacaacgta	ccgtcttgat	ttagaagatg	ccagaagcaa	agaaagagaa	gagctgaatg	240
ccattgtgtc	aagcgatgat	gcaacagcca	aggagaaaag	cgaagcatal	gataaaatga	300
cggctctcag	tgaagtggaa	ggaacagaaa	aacagctgga	aacgctgatt	aaaacacaag	360
gttacgagga	tgcgcttggt	aatgctgaag	gagataaaat	caatattaca	gtcaaatcag	420
acaaacactc	taaatcgaag	gcgacagcca	ttatagacct	tgtggcaaaa	gaaatcaaaa	480
caatgaaaga	tgtcgtctgc	acatttgaac	cctctaaata	agaatgaggg	aaaaaagccc	540
gctaacaacag	cgggcttttt	gcgttgcgtg	catattagag	ttgaattcaa	aagtcgcgtc	600
ctgtaagatg	aacatagtat	gtacttttag	tagtaacct	taaaaaaaga	aatttcatac	660
ataggagtgc	gattgaagtt	cggattagat	ccccctatgc	aagggtttat	tgttttctaa	720
aatctgatta	ccaattagaa	tgaatatctc	ccaaatat	aataataaaa	caaaaaaatt	780
gaaaaaagtg	tttccaccat	tttttcaatt	tttttataat	ttttttaaat	tgttatttaa	840
atagttttata	gttaaattta	cattttcatt	agtcatttca	atattctctc	caagataaact	900
acgaactgct	aacaaaaattc	tctccctatg	ttctaattgga	gaagattcag	ccactgcatt	960
tcccgcataa	tcttttggtta	tgatttttacc	cgtgtccata	gttaaaatca	tacggcataa	1020
agtttaataa	gagttgggtt	catcatctct	ataattatct	attaattcct	ctgacgaatc	1080
cataatggct	cttctcacat	cagaaaaatgg	aatatcagggt	agtaattcct	ctaagtcata	1140
atttccgtat	attctttttat	tttttctgtt	tgtttggtaa	agcattatgg	ttaaatctga	1200
atttaattcc	ttctgaggaa	tgtatccttg	ttcataaagc	tcttgtaacc	attctccata	1260
aataaattct	tgttttgggag	gatgattcca	cggtaaccatt	tcttgctgaa	taataattgt	1320
taattcaata	tatcgtaagt	tgccttttatc	tcctattttt	tttgaaatag	gtctaatttt	1380
ttgtataagt	atttcttttac	tttgatctgt	caatgggtca	gatacgacga	ctaaaaagtc	1440
aagatcacta	tttggttttta	gtccactctc	aaactcctgat	ccaaacatgt	aagtaccaat	1500
aagggttattt	tttaaatggt	ttccgaagtat	ttttttcact	ttattaattt	gttcgtatgt	1560
attcaaatat	atcctcctca	ctattttgat	tagtaacct	tttatatcca	tagttgttaa	1620
ttaaataaac	ttaatattag	ttattttatag	atttcaattg	cttctaaatt	ttttatctag	1680
ataataatta	tttttagttta	ttttatttcta	gattatata	gatattgat	ttcattttcca	1740
taaaactaaa	gtaagtgtta	acotattcat	tgttttaaaa	atatctcttg	ccagtcacgt	1800
tacgttatta	gttatagttta	ttataacatg	tattcacgaa	cgaaaatcgc	atcgattaac	1860
aattcgatca	gcattctggaa	aagggcgata	tgattgtttt	attttctgac	ggcgtcacag	1920
aatgcagaa	ggaaaacggt	ttcttggagc	gtcctgatct	gcagaagctc	attgaggaac	1980
atatgtgttc	ctctgcgcag	gaaatgggtca	aaaacattta	tgacagcctc	ctcaaatgtc	2040
aggattttca	gcttcacgat	gatttttacgt	taattgtttt	gcggagaaa	gttttaacgtc	2100
tgtcagacga	gggtataaag	caactagtga	tttgaaaggaa	aatttgaggt	gatacgacga	2160
gntgaaatca	tatgtttaa	atcaaagaaa	tccacgagct	gattaaagca	attgacgagt	2220
ctacaattga	cgaattcgta	tatgaaaatg	aaggtgtatc	cttaaaactg	aaaaaacacg	2280
aagcaggcac	ggttcaagtc	atgcagcagg	caccggcagc	acctgtacaa	gcacaggctc	2340
cgcaggcagt	tcagccgcaa	gcgcagcaag	cagcggcacc	tgcccaagaa	gcaccaaagc	2400

[0515]

[0516]

aagatgagaa	tctgcataaa	atcactttcac	caatggtagg	aacatttttat	gcttctttcat	2460
caccggaagc	tgcccgctat	gtaacagccg	gttcaaaagt	aaatgaaaac	acagttgtct	2520
gcattgtaga	agcgatgaag	cttttcaacg	aaatcgaagc	agaagtgaac	ggcgaaatcg	2580
ttgaagtatt	agtagaaaac	ggccagctgg	tcgaatacgg	acaaccaatt	ttattgcaat	2640
aacaggtgct	tacttttctg	gagttcttta	gcaaatTTTT	ttattagctg	aacttagtat	2700
tagtggccat	actcctccaa	tccaaagcta	tttagaaaga	ttactatatc	ctcaaacagg	2760
cggtaaccgg	cctcttcate	gggaatgcgc	gcgaccttca	gcctcgccgg	catgtccccc	2820
tgccggacgg	gaagtatcca	gctcgaggtc	gggcccgggt	gctggcggtt	ttccataggg	2880
tccgcccccc	tgacgagcat	cacaaaaatc	gacgctcaag	tcagaggtgg	cgaaacccga	2940
caggactata	aagataccag	gcgtttcccc	ctggaagctc	cctcgtgcgc	tctcctgttc	3000
cgaccctgcc	gcttaccgga	tacctgtccg	cctttctccc	ttcgggaaag	gtggcgcttt	3060
ctcatagctc	acgctgtagg	tatctcagtt	cgggtgtagg	cgttcgctcc	aagctgggct	3120
gtgtgcacga	acccccggtt	cagcccgacc	gctgcgcctt	atccggtaac	tatcgtcttg	3180
agtcacaacc	ggtaagacac	gacttatcgc	cactggcagc	agccactggt	aacaggatta	3240
gcagagcgag	gtatgtaggc	ggtgctacag	agtctctgaa	gtgggtggct	aactacggct	3300
acactagaag	gacagtattt	ggtatctgcg	ctctgctgaa	gccagttacc	ttcggaaaaa	3360
gagttgtatg	ctcttgatcc	ggcaaacaaa	ccaccgctgg	tagcgggtgt	ttttttgttt	3420
gcaagcagca	gattacgcgc	agaaaaaaag	gatctcaaga	agatcctttg	atcttttcta	3480
cgggggtctga	cgctcagtg	aacgaaaact	cacgttaagg	gatttttggtc	atgagattat	3540
caaaaaggat	cttcacctag	atccttttaa	attaaaaatg	aagtttttaa	tcaatctaaa	3600
gtatatatga	gtaaaacttg	tctgacagtt	accaatgctt	aatcagttag	gcacctatct	3660
cagcgatctg	tctatttctg	tcatccatag	ttgctgact	ccccgtcgtg	tagataacta	3720
cgatacggga	gggcttacca	tctggcccca	gtgctgcaat	gataccgcga	gacccacgct	3780
caccgctgcc	gcaataaaac	gcaataaaac	agccagccgg	aagggccgag	cgcagaagtg	3840
gtcctgcaac	tttatccgcc	tccatccagt	ctattaattg	ttgccgggaa	gctagagtaa	3900
gtagttcgcc	agttaatagt	ttgcgcaacg	ttgttgccat	tgctgcaggc	atcgtggtgt	3960
caagctcgtc	gittggtatg	gcttcattca	gctccggttc	ccaacgatca	aggcgagtta	4020
catgatcccc	catgtttgtg	aaaaaagcgg	ttagctcctt	cggctcctcc	atcgttgtca	4080
gaagtaagtt	ggccgcagtg	ttatcactca	tggttatggc	agcactgcat	aattctctta	4140
ctgtcatgcc	atccgtaaga	tgcttttctg	tgactggtga	gtactcaacc	aagtcattct	4200
gataatagt	tatgcggcga	ccgagttgct	cctggccggc	gtcaacacgg	gataataacc	4260
cgccacatag	cagaacttta	aaagtgtctc	tcattggaaa	acgttcttcc	ggcggaacac	4320
tctcaaggat	cttaccgctg	ttgagatcca	gttcgatgta	acccactcgt	gcacccaact	4380
gatcttcaag	atcttttact	ttcaccagcg	ttctcgggtg	agcaaaaaaa	ggaaggcaaa	4440
atgccgcaaa	aaaggggaata	aggcgacac	ggaaatggtg	aatactcata	ctcttctctt	4500
ttcaatatata	ttgaagcatt	tatcaggggt	attgtctcat	gagcggatac	atatttgaat	4560
gtattttgaa	aaataaaaca	ataggggttc	cgcgcacatt	tccccgaaaa	gtgccacctg	4620
acgtctaaga	aaccattatt	atcatgacat	taacctataa	aaataggcgt	atcacgaggg	4680
cctttctgtc	tcaagaatt					4699
<210>	210					
<211>	4699					
<212>	DNA					
<213>	合成					
<220>						
<221>	尚未归类的特征					
<222>	(1) .. (4699)					
<223>	质粒 pEG166					
<220>						
<221>	尚未归类的特征					
<222>	(88) .. (676)					
<223>	同源区域 accC 上游					
<220>						
<221>	启动子					
<222>	(1858) .. (2156)					
<223>	负荷成瘾启动子 p(yf1T)					
<220>						
<221>	rbs					
<222>	(2157) .. (2162)					
<223>	rbs					
<220>						
<221>	尚未归类的特征					
<222>	(2162) .. (2162)					
<223>	N 是 A、T、C 或 G					
<220>						
<221>	尚未归类的特征					
<222>	(2172) .. (2626)					
<223>	同源区域 accC 下游					
<400>	210					
aacaaaaattc	tccagttcttc	acatcgggttt	gaaaggagga	agcgggaagaa	tgaagtaaga	60
gggattttttg	actccgaagt	aagtcttagg	gaacgaaaga	ttcatcagca	gacaaggaaa	120
catccgctga	agccagtga	aaaggaaactg	ttgttacgga	aacagccgat	gatgatttgt	180
ttacaacgta	ccgtcttgat	ttagaagatg	ccagaagcaa	agaaagagaa	gagctgaatg	240
ccattgtgtc	aagcgatgat	gcaacagcca	aggagaaaag	cgaagcatat	gataaaatga	300
cggctctcag	tgaagtggaa	ggaacagaaa	aacagctgga	aacgctgatt	aaaacacaa	360
gttacgagga	tgcgcttggt	aatgctgaag	gagataaaat	caatattaca	gtcaaatcag	420
acaaacactc	taaatacgaag	gcgacagcca	ttatagacct	tgtggcaaaa	gaaatcaaaa	480

[0517]

caatgaaaga	tgtcgcgtgtc	acattttgaac	cctctaaata	agaatgaggg	aaaaaagccc	540
gctaaacaag	cggttgcgtgt	gcgttgcgtgt	catattagag	ttgaattcaa	aagtcocgctc	600
ctgtaagatg	aacatagtat	gtacttttag	tagtaaccta	taaaaaaga	aatttcatac	660
ataggagtgc	gattggaagt	cggattagat	ccocctatgc	aagggtttat	tgttttctaa	720
aatctgatta	ccaattagaa	tgaatatttc	ccaaatatta	aataataaaa	caaaaaaatt	780
gaaaaaagtg	tttccaccat	tttttcaatt	tttttataat	ttttttaatc	tgttatttaa	840
atagtttata	gttaaattta	cattttcatt	agttccattca	atattctctc	caagataaact	900
acgaactgct	aacaaaattc	tctccctatg	ttctaattgga	gaagattcag	ccactgcatt	960
tcccgcataa	tcttttggtg	tgattttacc	cgtgtccata	gttaaaatca	tacggcataa	1020
agttaatata	gagttgggtt	catcatcctg	ataattatct	attaattcct	ctgacgaatc	1080
cataatggct	cttctcacat	cagaaaaatg	aataaccagt	agtaattcct	ctaagtcata	1140
atttccgtat	attcttttat	tttttcgttt	tgcttggtaa	agcattatgg	ttaaatctga	1200
atttaattcc	ttctgaggaa	tgatcccttg	ttcataaagc	tcttgtaacc	attctccata	1260
aataaattct	tgtttgaggag	gatgattcca	cggtaaccatt	tcttgctgaa	taataattgt	1320
taattcaata	tatcgtaagt	tgcttttate	tcctattttt	tttgaaatag	gtctaatttt	1380
ttgtataagt	atttcttttac	tttgatctgt	caatggttca	gatacgacga	ctaaaaagtc	1440
aagatcaacta	tttggttttta	gtccactctc	aactcctgat	ccaaacatgt	aagtaaccaat	1500
aagggttattt	tttaaatggt	tccgaagtat	tttttttact	ttattaattt	gttcgtatgt	1560
attcaaatat	atcctcctca	ctattttgat	tagtacctat	tttatatcca	tgtattgtta	1620
ttaaataaac	tttaatttagt	ttattttatag	atttcattgg	cttctaaatt	ttttatctag	1680
ataataatta	ttttagttta	ttttattcta	gattatataat	gatatgatct	ttcattttcca	1740
taaaactaaa	gtaagtgtta	acctattcat	tgtttttaaaa	atatctcttg	ccagtcacgt	1800
tacgttatta	gttatagtta	ttataacatg	tattcacgaa	cgaaaaatcg	atcgattgac	1860
taaatcactc	acaatccaag	catgtgattg	ctgtttcggg	gaaaagattt	gcatttagaca	1920
tacctcctgc	tcgtacggat	aaaggcagcg	tttcatgggtc	gtgtgctccg	tcgacggcgt	1980
tctccttaat	tttgattttt	ctgaaaaatg	gtcccgttcc	tatcacttta	ccatggacgg	2040
aaaacaaaata	gtactaccga	ttcctcctgt	ttttctcttc	aatgttctgg	aatgacctttc	2100
aggtagacag	gatcgggtat	gaaagaaata	tagaaaacat	gaaggaggaa	tatcgaccga	2160
gntgaaatca	tatgtttaat	atcaaaagaa	tcacagagct	gattaaagca	attgacgagt	2220
ctacaattga	cgaattcgtg	tatgaaaatg	aagggtgtatc	cttaaaactg	aaaaaacacg	2280
aagcaggcac	ggttcaagtc	atgcagcagg	caccggcagc	acctgtacaa	gcacaggctc	2340
cgcaggcagt	tcagccgcaa	gcgcagcaag	cagcggcacc	tgcccaagaa	gcaccaaacg	2400
aagatgagaa	ctgcataaaa	atcacctcac	caatggtagg	aacattttat	gcttctctat	2460
caccggaagc	tggcccgat	gtaacagccg	gttcaaaaagt	aaatgaaaaa	acagtgtgtc	2520
gcattgtaga	agcgatgaag	cttttcaacg	aaatcgaagc	agaagtgaag	ggcgaaatcg	2580
ttgaagtatt	agttagaaaac	ggccagctgg	tcgaatacgg	acaaccaatt	ttattgcaat	2640
aacaggtgct	tacttttctg	gagttcttta	gcaaaatttt	ttattagctg	taactgattat	2700
tagtggtcat	actcctccaa	tcocaaagcta	tttagaaaga	ttactatata	ctcaaacagg	2760
cggttaaccgg	cctcttcatc	gggaatgcgc	gcgaccttca	gcctcgcggg	catgtccccc	2820
tggcggaagg	ggaagtatcca	gctcagagtc	ggcgccggtt	gctggcggtt	ttccataggc	2880
tccgcccccc	tgacgagcat	cacaaaaate	gacgctcaag	tcagagggtg	cgaaaccoga	2940
caggactata	aagataccag	gcgtttcccc	ctggaaagctc	cctcgtgcgc	tctcctgttc	3000
cgaccctgcc	gtttaccgga	tacctgtccg	cctttctccc	ttcggaagc	gtggcgcttt	3060
ctcatagctc	acgctgtagg	tatctcagtt	cggtgtagggt	cgctcgcctc	aagctgggct	3120
gtgtgacaga	accccccggt	cagcccgacc	gctgcgcctt	atccggtaac	tatcgtcttg	3180
agttccaaacc	ggtaagacac	gacttatcgc	cactggcagc	agccaactgg	aacaggatta	3240
gcagagcgag	gtagttaggc	ggtgtacag	agttcttgaa	gtgggtggct	aactacggct	3300
acactagaag	gacagtattt	ggtatctgcg	ctctgctgaa	gccagttacc	ttcggaaaaa	3360
gagttgatag	ctcttgatcc	ggcaaacaaa	ccaccgctgg	tagcggtggg	ttttttgttt	3420
gcaagcagca	gattacgcgc	agaaaaaaaag	gatctcaaga	agatcctttg	atcttttcta	3480
cggggtctga	cgctcagtg	aacgaaaaact	cacgttaagg	gattttgggtc	atgagattat	3540
caaaaaggat	cttcacctag	atcctttttaa	attaaaaatg	aagtttttaa	tcaatctaaa	3600
gtatatatga	gtaaaacttg	tctgcaggtt	accaatgctt	aatcagtgag	gcacctatct	3660
cagcgatctg	tctatttctg	tcattccatag	ttgcctgact	ccccgctgtg	tagataacta	3720
cgatacggga	gggcttacca	tctggcccca	gtgctgcaat	gataccgcga	gacccacgct	3780
caccggtccc	agattttatca	gcaataaaacc	agccagccgg	aagggccgag	cgcagaagtg	3840
gtcctgcaac	tttatccgcc	tccatccagt	ctattaattg	ttgcggggaa	gctagagtaa	3900
gtagttccgc	agtttaatagt	ttgcgcaacg	ttgttgccat	tgctgcaggc	atcgtgggtg	3960
cacgctcgtc	gittggtagt	gcttcattca	gctcgggttc	ccaacgatca	agcgagttta	4020
catgatcccc	catgtttgtg	aaaaaagcgg	ttagctcctt	cggctcctccg	atcgttgtca	4080
gaagtaagtt	ggccgcagtg	ttatcactca	tggttatggc	agcactgcat	aattctctta	4140
ctgtcatgcc	atccgtaaga	tgcttttctg	tgactgggtga	gtactcaacc	aagtcattct	4200
gagaatagtg	tatgcggcga	ccgagttgct	cttgcccggc	gtcaaacacg	gataataccg	4260
cgccacatag	cagaacttta	aaagtgcctca	tcattggaaa	acgttcttog	gggcaaaaac	4320
tctcaaggat	cttaccgctg	ttgagatcca	gttcgatgta	acccactcgt	gcacccaact	4380
gatcttcagc	atcttttact	ttcaccagcg	ttctgggtg	agcaaaaaa	ggaaggcaaa	4440
atgccgcaaa	aaaggggaata	agggcgacac	ggaaaatgtg	aatactcata	ctcttctctt	4500
ttcaatatta	ttgaagcatt	tatcagggtt	attgtctcat	gagcggatac	atatttgaat	4560
gtatttagaa	aaataaaca	ataggggttc	cgcgcacatt	tccccgaaa	gtgccacctg	4620
acgtctaaga	aaccattatt	atcatgacat	taacctataa	aaatagggct	atcacgaggg	4680
cctttcgtct	tcaagaatt					4699

<210> 211
 <211> 783
 <212> DNA
 <213> 合成
 <220>
 <221> 尚未归类的特征
 <222> (1) .. (783)

<223> 编码 SpcR (抗性标记) 的基因		
<400> 211		
gtactaatca aaatagttag gaggatatat ttgaatacat acgaacaaat taataaagtg		60
aaaaaaatca ttccgaaaca tttaaaaaat aaccttattg gtacttacat gtttggatca		120
ggagttgaga gtggactaaa accaaatagt gatcttgact ttttagtcgt cgtatctgaa		180
ccattgacag atcaaaagtaa agaaataactt atacaaaaaa ttagacctat ttcaaaaaaa		240
ataggagata aaagcaactt acgatataat gaattaacaa ttattattca gcaagaaatg		300
gtaccgtgga atcatcctcc caaacaagaa ttattttatg gagaatgggt acaagagctt		360
tatgaacaag gatacattcc tcagaaggaa ttaaattcag atttaaccat aatgctttac		420
caagcaaaac gaaaaaataa aagaatatac ggaattatg acttagagga attactacct		480
gatattccat tttctgatgt gagaagagcc attatggatt cgtcagagga attaatagat		540
aattatcagg atgatgaaac caactctata ttaactttat gccgtatgat ttttaactatg		600
gacacgggta aaatcatacc aaaagatatt gcgggaaatg cagtggctga atcttctcca		660
ttagaacata gggagagaat tttgttagca gttcgtagtt atcttggaga gaattattgaa		720
tggactaatg aaaaatgtaaa tttaactata aactatttaa ataacagatt aaaaaaatta		780
taa		783
<210> 212		
<211> 296		
<212> DNA		
<213> 芽孢杆菌属		
<220>		
<221> 启动子		
<222> (1) .. (296)		
<223> 芽孢杆菌属负感应启动子 p (groES)		
<400> 212		
aatgcggaaa actgcatgat gatatatgtc aaaataataa accaatactg ttttctcaaa		60
tggatgttac tcttttggtt agtgggtttc gttcatctac agctattgta acataatcgg		120
tacgggggtg aaaaagctaa cggaaaaggg agcggaaaag aatgatgtaa gcgtgaaaaa		180
ttttttatct tatcacttga aattggaagg gagattcttt attataagaa ttgtgttagc		240
actctttagt gctgagtgct aaaattacat attcatacta ttgaggaggt tatcttc		296
<210> 213		
<211> 299		
<212> DNA		
<213> 芽孢杆菌属		
<220>		
<221> 启动子		
<222> (1) .. (299)		
<223> 芽孢杆菌属负感应启动子 p (ctsR)		
<400> 213		
gctcttcagc gccgatggta gtcgggggtt tccccctgtg agagtaggac gccgccaaagc		60
aagcttaaac ccagctcaat gagctgggtt ttttgatatt tggtttattg gtatcataaa		120
attccactta actgtataat ataataactt tataccgaat tttaaatcag caatcaggtt		180
ttgtggaccg ggaaaatgga aataatgaag gatagagcga gaaagttaga aattctcag		240
aaacggctta tagtaagatt aaagtcaaat atagtcaaag tcagtaaagg aggggggtg		299
<210> 214		
<211> 299		
<212> DNA		
<213> 芽孢杆菌属		
<220>		
<221> 启动子		
<222> (1) .. (299)		
<223> 芽孢杆菌属负感应启动子 p (dnaK)		
<400> 214		
ggttcaatcc ccgcacacgg gaatttcgat taaaatcgga aaagaaaacg actatgaaga		60
gatggaaaat tgcagctctga ttacggcttc ttattccgta gaccagaagc agatcggtc		120
aattgcgatt atcggcccca cccgcattga ttattccagg gttgtcagcc tgcttcagca		180
tgtgacttcg gacttgtcaa aagcattaac aagtctgtat gatgaataag ggaattttgg		240
caaattttat cgaagggcag cacctgtcct tctccttaca ctttgaggga ggtgaacac		299
<210> 215		
<211> 299		
<212> DNA		
<213> 芽孢杆菌属		
<220>		
<221> 启动子		
<222> (1) .. (299)		
<223> 芽孢杆菌属负感应启动子 p (perR)		
<400> 215		
ggaagcattt gatgtgtgga agctgaagaa aaacttcgga aaagagtata tgggaatcga		60
gcgttccact tttttaattg ataaagaagg ccggctcatt aaagaatgga gaaaagtaaa		120
agtaaaaggac catgtagccg aagcgccttca aaccctcaaa gatatgtctg aaaagtaaat		180
ctctatgagc ctatgcttac ttttaccctg atatcaattg acaaatatc gtagaaagag		240
ttacactaat tataaacatt acaatgtaag aatttttttt agaagagagg tgcattgacg		299
<210> 216		
<211> 299		
<212> DNA		
<213> 芽孢杆菌属		

```

<220>
<221> 启动子
<222> (1)..(299)
<223> 芽孢杆菌属负荷感应启动子 p(hrcA)
<400> 216
tgaggaagaa atgttttttag gtttgagaaa aacagccggt gtcagcaaaa aacgattcgc 60
agaaaaatac ggccgttccc tggatggact tttcccgaag gtgttataaa acctcgctga 120
aaaagggtcg attcacaatt cggagtcctgc cgtttatttg actcatcaag ggaaattatt 180
aggcaatgaa gtttttggcg cttttttggg tgagtataa ttgacatttt tcttgtggtt 240
tgatactttt gttatagaat tagcactcgc ttattgagag tgctaacaga ggtgatgat 299
<210> 217
<211> 299
<212> DNA
<213> 芽孢杆菌属
<220>
<221> 启动子
<222> (1)..(299)
<223> 芽孢杆菌属负荷感应启动子 p(spxA)
<400> 217
gaaaacattc atacttataa agatgagaaa tggtagcgag aatcgtagca aaaaacggat 60
gaggtcttgg aaatcatcga tgaagaaaaa caaaaaata ttttcgatca gtttaagctaa 120
ttgagaataa gaacatatca tgaacattaa aaatttgcgt gattgacaca tttttttctt 180
aggatactgt tcacacttac ttttttatag tataatacct gtaaagattc cttttttaga 240
gtaatttcaa ttttaagaata aaatagatgt tcatcctact aattagagga gtgaagatg 299
<210> 218
<211> 299
<212> DNA
<213> 芽孢杆菌属
<220>
<221> 启动子
<222> (1)..(299)
<223> 芽孢杆菌属负荷感应启动子 p(sigB)
<400> 218
aacaattcga tcagcatctg gaaaaggggc atatgattgt tttattttct gacggcgtca 60
cagaatgcag aacggaaaaa cgtttcttgg agcgtcctga tctgcagaag ctcattgagg 120
aacatatgtt ttctctctgc caggaaatgg tcaaaaaacat ttatgacagc ctcctcaaatt 180
tcaggatttt tcagcttcac gatgatttta cgttaattgt tttgcggaga aagggtttaac 240
gtctgtcaga cgagggtata aagcaactag tgatttgaag gaaaatttga ggtgatacg 299
<210> 219
<211> 299
<212> DNA
<213> 芽孢杆菌属
<220>
<221> 启动子
<222> (1)..(299)
<223> 芽孢杆菌属负荷感应启动子 p(yf1T)
<400> 219
gactaaatca ctcacaatcc aagcattgta ttgctgttct ggtgaaaagt attgcattag 60
acatacctcc tgctcgtacg gataaaggca gcgtttcatg gtcgtgtgct ccgtgcagcg 120
gcttctcctt aatttttgatt tttctgaaaa taggtcccgt tcctatcact ttaccatgga 180
cggaaaaacaa atagctacta ccattcctcc tgtttttctc ttcaatgttc tggaatctgt 240
ttcaggtaca gacgatcggg tatgaaagaa atatagaaaa catgaaggag gaatatcga 299
<210> 220
<211> 651
<212> DNA
<213> 合成
<220>
<221> CDS
<222> (1)..(651)
<223> 编码 camR 抗性标记的 chlR 基因
<400> 220
atg aac ttt aat aaa att gat tta gac aat tgg aag aga aaa gag ata 48
Met Asn Phe Asn Lys Ile Asp Leu Asp Asn Trp Lys Arg Lys Glu Ile
1 5 10 15
ttt aat cat tat ttg aac caa caa acg act ttt agt ata acc aca gaa 96
Phe Asn His Tyr Leu Asn Gln Gln Thr Thr Phe Ser Ile Thr Thr Glu
20 25 30
att gat att agt gtt tta tac cga aac ata aaa caa gaa gga tat aaa 144
Ile Asp Ile Ser Val Leu Tyr Arg Asn Ile Lys Gln Glu Gly Tyr Lys
35 40 45
ttt tac cct gca ttt att ttc tta gtg aca agg gtg ata aac tca aat 192
Phe Tyr Pro Ala Phe Ile Phe Leu Val Thr Arg Val Ile Asn Ser Asn
50 55 60
aca gct ttt aga act ggt tac aat agc gac gga gag tta ggt tat tgg 240
Thr Ala Phe Arg Thr Gly Tyr Asn Ser Asp Gly Glu Leu Gly Tyr Trp
65 70 75 80

```


	gat aag tta gag cca ctt tat aca att ttt gat ggt gta tct aaa aca	288
	Asp Lys Leu Glu Pro Leu Tyr Thr Ile Phe Asp Gly Val Ser Lys Thr	
	85 90 95	
	ttc tct ggt att tgg act cct gta aag aat gac ttc aaa gag ttt tat	336
	Phe Ser Gly Ile Trp Thr Pro Val Lys Asn Asp Phe Lys Glu Phe Tyr	
	100 105 110	
	gat tta tac ctt tct gat gta gag aaa tat aat ggt tcg ggg aaa ttg	384
	Asp Leu Tyr Leu Ser Asp Val Glu Lys Tyr Asn Gly Ser Gly Lys Leu	
	115 120 125	
	ttt ccc aaa aca cct ata cct gaa aat gct ttt tct ctt att att	432
	Phe Pro Lys Thr Pro Ile Pro Glu Asn Ala Phe Ser Leu Ser Ile Ile	
	130 135 140	
	cca tgg act tca ttt act ggg ttt aac tta aat atc aat aat agt	480
	Pro Trp Thr Ser Phe Thr Gly Phe Asn Leu Asn Ile Asn Asn Ser	
	145 150 155 160	
	aat tac ctt cta ccc att att aca gca gga aaa ttc att aat aaa ggt	528
	Asn Tyr Leu Leu Pro Ile Ile Thr Ala Gly Lys Phe Ile Asn Lys Gly	
	165 170 175	
	aat tca ata tat tta ccg cta tct tta cag gta cat cat tct gtt tgt	576
	Asn Ser Ile Tyr Leu Pro Leu Ser Leu Gln Val His His Ser Val Cys	
	180 185 190	
	gat ggt tat cat gca gga ttg ttt atg aac tct att cag gaa ttg tca	624
	Asp Gly Tyr His Ala Gly Leu Phe Met Asn Ser Ile Gln Glu Leu Ser	
	195 200 205	
	gat agg cct aat gac tgg ctt tta taa	651
	Asp Arg Pro Asn Asp Trp Leu Leu	
	210 215	
	<210> 221	
	<211> 216	
	<212> PRT	
	<213> 合成	
	<400> 221	
	Met Asn Phe Asn Lys Ile Asp Leu Asp Asn Trp Lys Arg Lys Glu Ile	
	1 5 10 15	
	Phe Asn His Tyr Leu Asn Gln Gln Thr Thr Phe Ser Ile Thr Thr Glu	
	20 25 30	
	Ile Asp Ile Ser Val Leu Tyr Arg Asn Ile Lys Gln Glu Gly Tyr Lys	
	35 40 45	
[0520]	Phe Tyr Pro Ala Phe Ile Phe Leu Val Thr Arg Val Ile Asn Ser Asn	
	50 55 60	
	Thr Ala Phe Arg Thr Gly Tyr Asn Ser Asp Gly Glu Leu Gly Tyr Trp	
	65 70 75 80	
	Asp Lys Leu Glu Pro Leu Tyr Thr Ile Phe Asp Gly Val Ser Lys Thr	
	85 90 95	
	Phe Ser Gly Ile Trp Thr Pro Val Lys Asn Asp Phe Lys Glu Phe Tyr	
	100 105 110	
	Asp Leu Tyr Leu Ser Asp Val Glu Lys Tyr Asn Gly Ser Gly Lys Leu	
	115 120 125	
	Phe Pro Lys Thr Pro Ile Pro Glu Asn Ala Phe Ser Leu Ser Ile Ile	
	130 135 140	
	Pro Trp Thr Ser Phe Thr Gly Phe Asn Leu Asn Ile Asn Asn Asn Ser	
	145 150 155 160	
	Asn Tyr Leu Leu Pro Ile Ile Thr Ala Gly Lys Phe Ile Asn Lys Gly	
	165 170 175	
	Asn Ser Ile Tyr Leu Pro Leu Ser Leu Gln Val His His Ser Val Cys	
	180 185 190	
	Asp Gly Tyr His Ala Gly Leu Phe Met Asn Ser Ile Gln Glu Leu Ser	
	195 200 205	
	Asp Arg Pro Asn Asp Trp Leu Leu	
	210 215	
	<210> 222	
	<211> 10063	
	<212> DNA	
	<213> 合成	
	<220>	
	<221> 尚未归类的特征	
	<222> (1) .. (10063)	
	<223> 质粒 pEG062	
	<220>	
	<221> 尚未归类的特征	
	<222> (1) .. (637)	
	<223> 同源区域 amyE 上游	
	<220>	
	<221> 尚未归类的特征	
	<222> (789) .. (1583)	
	<223> 抗性标记 KanR (在互补链上)	
	<220>	

<221>	终止子				
<222>	(1744) .. (1796)				
<223>	终止子 T7				
<220>					
<221>	启动子				
<222>	(1866) .. (2655)				
<223>	启动子 P3P				
<220>					
<221>	尚未归类的特征				
<222>	(2680) .. (3861)				
<223>	编码 IgA-V3 的基因				
<220>					
<221>	尚未归类的特征				
<222>	(3862) .. (3867)				
<223>	甘氨酸接头				
<220>					
<221>	尚未归类的特征				
<222>	(3868) .. (4584)				
<223>	编码 sfGFP 的基因				
<220>					
<221>	尚未归类的特征				
<222>	(4585) .. (4602)				
<223>	纯化标签 6-his				
<220>					
<221>	终止子				
<222>	(4614) .. (4682)				
<223>	终止子				
<220>					
<221>	尚未归类的特征				
<222>	(5070) .. (6262)				
<223>	同源区域 amyE 下游				
<220>					
<221>	尚未归类的特征				
<222>	(7207) .. (7989)				
<223>	编码抗性标记壮观霉素腺苷酰转移酶的 spc 基因				
<220>					
[0521] <221>	尚未归类的特征				
<222>	(8995) .. (9855)				
<223>	编码氨基青霉素抗性 β -内酰胺酶的 bla 基因 (在互补链上)				
<400>	222				
ttaacaaaat	tctccagttc	tcacatcggg	ttgaaaggag	gaagcggag	aatgaagtaa 60
gagggatttt	tgactccgaa	gtaagtcttc	aaaaaatcaa	ataaggagtg	tcaagaatgt 120
ttgcaaaacg	attcaaaacc	tctttactgc	cgttattcgc	tggattttta	ttgctgtttc 180
atattggtct	ggcaggaccg	gcggctgcga	gtgctgaaac	ggcgaacaaa	tcgaatgagc 240
ttacagcacc	gtcgatcaaa	agcggaaacca	ttcttcatgc	atggaattgg	tcgttcaata 300
cgttaaaaaca	caatatgaag	gatattcatg	atgcaggata	tacagccatt	cagacatctc 360
cgattaaacca	agtaaaaggaa	gggaatcaag	gagataaaaag	catgtcgaac	tggtactggc 420
tgtatcagcc	gacatcgtat	caaattggca	accgttactt	agggtactgaa	caagaattta 480
aagaaatgtg	tcgagccgct	gaagaatatg	gcataaagggt	cattgttgac	gcggtcatac 540
atcataccac	cagtgtattat	gccgcgattt	ccaatgaggt	taagagtatt	ccaaactgga 600
cacatggaaa	cacacaaatt	aaaaactggg	ctgatcggat	cctagaagct	tatcgaattc 660
tcattgttga	cagcttatca	tcggcaatag	ttcggataat	gagtttataa	aataaaaaaa 720
gcacctgaaa	aggtgtcttt	ttttgatggg	tttgaacttg	aagcttttta	gacatctaaa 780
tctaggtact	aaaacaattc	atccagtaaa	atataatatt	ttattttctc	ccaatcaggc 840
ttgatcccca	gtaagtcaaa	aaatagctcg	acatactgtt	cttccccgat	atcctccctg 900
atcgaccgga	cgcagaaggc	aatgtcatac	cacttgtccg	ccctgccgct	tctcccaaga 960
tcaataaagc	cacttacttt	gccatctttc	acaaagatgt	tgctgtctcc	caggtcgccg 1020
tgggaaaaga	caagttcctc	ttcgggcttt	tccgtcttta	aaaaatcata	cagctcgccg 1080
ggatctttta	atggagtgtc	ttcttcccag	ttttcgcaat	ccacatcggc	cagatcgtaa 1140
ttcagtaagt	aatccaattc	ggctaagcgg	ctgtctaaag	tattcgtata	gggacaatcc 1200
gatattgtcga	tggagtgaag	gagcctgatg	cactccgcac	acagctcgat	aatcttttca 1260
gggctttgtt	catcttcata	ctcttccgag	caaaggagcg	catcggcctc	actcatgagc 1320
agattgtctc	agccatcatg	ccgttcaaaag	tgcaggacct	ttggaacagg	cagctttcct 1380
tccagccata	gcacatgttc	cttttcccgt	tccacatcat	aggtgggtcc	tttataccgg 1440
ctgtccgtca	tttttaaata	taggttttca	ttttctccca	ccagcttata	taccttagca 1500
ggagacattc	cttccgtatc	ttttacgcag	cggatttttt	cgatcagttt	tttcaattcc 1560
ggtgatattc	tcatttttagc	cattttattat	ttccttccct	ttttctacag	tattttaaaga 1620
taccccaaga	agctaattat	aacaagacga	actccaattc	actgttcctt	gcattctaaa 1680
accttaataa	ccagaaaaaca	gacccaggtc	tttttcaaag	ttgttttgaa	agttggcccg 1740
catcccccc	agcataaccc	cttggggcct	ctaaacgggt	cttgaggggg	tttttggagc 1800
tcaaacaatt	aatgtgagtt	agctcactca	ttaggcaccc	caggctttac	actttatgag 1860
tctattcttt	tggagaagaaa	tatagggaag	atgggtattg	ttaaaaattc	tgaatattta 1920
tacaatatca	tatgtttcac	attgaaatat	gctgtccaga	ctgtccgctg	tgtaaaaaat 1980
aggaataaag	gggggttggg	attatttttac	tgatattgtaa	aatataattt	gtataagaaa 2040
attcgaaacg	taagatgaaa	ccttagataa	aagtgccttt	tttgttgcaa	ttgaagaatt 2100

[0522]

attaatgtta	agcttaatta	aagataatat	ctttgaattg	taacgcccc	caaaagtaag	2160
aactacaaaa	aaagaatacg	ttatatagaa	atatgtttga	accttcttca	gattacaaat	2220
atattcggac	ggactctacc	tcaaagtctt	atctaactat	agaatgacat	acaagcacaa	2280
ccttgaaaat	ttgaaaatat	aactaccaat	gaacttggtc	atgtgaatta	tcgctgtatt	2340
taattttctc	aattcaatat	ataatatgcc	aatacattgt	tacaagtaga	aattaagaca	2400
cccttgatag	ccttactata	cctaacatga	tgtagtatta	aatgaatatg	taaatatatt	2460
tatgataaga	agcgacttat	ttataatcat	tacatatttt	tctattggaa	tgattaaagt	2520
tccaatagaa	tagtgtataa	attattttatc	ttgaaaggag	ggatgcctaa	aaacgaagaa	2580
cattaaaaac	atatattttgc	accgtctaat	ggattttatga	aaaatcattt	tatcagtttg	2640
aaaattatgt	attatgatta	actaataagg	aggacaaaac	tgatccaaaa	acgcaaacgc	2700
acagtttagct	ttagactggg	tctgatgtgc	acactgctgt	ttgtttcact	gccgattaca	2760
aaaacaagcg	cacaagttca	actgcaagaa	tcaggcggag	gcctggttca	agcaggcgga	2820
tcactgagac	tgatcatgcgc	agcatcaggc	cttacatttg	atacatatgc	aatgggctgg	2880
tttagacaag	caccgggaaa	aaaacgcgaa	tatgttgacg	caatttcacg	gacaggcatc	2940
agcacatatt	atgcggtat	tgcaaaagga	cgctttacaa	ttagccgtga	taatgcgaaa	3000
aacacgctgt	atctgcaaat	ggattcactg	aaaccgggaag	atacagcagt	ctattattgc	3060
gcagcaaaaa	aatcactgaa	tggtccggca	ccgtgggatt	attggggaca	aggcacacaa	3120
gttacagttt	catcagatcc	gtgtccgcaa	tgctgcaaac	cgctcactgtc	actgcaaccg	3180
cctgcatggg	cggatctgct	gctgggctca	aatgcacac	tgacatgcac	actgtcagga	3240
ctgaaaaaat	ctgaaggcgt	ttcattttaca	tggaaccgtg	caggcggaaa	agatgcagtt	3300
caagcatcac	cgacaagaga	ttcatgcggc	tgctattcag	tttcaagcat	tcttccggga	3360
tgcgctgata	cgtggaataa	aggcgaaaca	ttttcatgca	cagcagcaca	ttcagaactg	3420
aaatcagcac	tgacagcgac	aattacaaaa	ccgaaagtca	atacatttcg	tccgcaagtt	3480
catctgtctc	gcctcccgag	cgaagaactg	gcactgtatg	aactggttac	acttatactg	3540
ctgggttagag	gcttttcacc	gaaagatggt	cttggttagat	ggctgcaagg	cggacaagaa	3600
ctgcgagag	ataaatatct	ggtttgggaa	tcactgccgg	aaaccgggaca	agcaattccg	3660
acatatgcag	ttacatcagt	tcttagagtt	gatggcggaag	attggaacaa	ggcgagcata	3720
tttagctgca	tggttgccca	tgaagcactg	ccgctggcat	ttacacaaaa	aacaattgat	3780
agactggcag	gcaaacggac	acatgttaat	gtttcagttg	ttatggcaga	agcagaaggc	3840
atctgctata	aagatgaact	gggtggcatg	cgtaaaggcg	aagagctggt	cactggtttc	3900
gtcactattc	tggtggaact	ggatgggtgat	gtcaacgggtc	ataagttttc	cgtgcgtggc	3960
gagggtgaag	gtgacgcaac	taatggtaaa	ctgacgcgtga	agttcatctg	tactactggt	4020
aaactgccgg	tacccttgcc	gactctggta	acgacgcgtga	cttatgggtg	tcagtgcctt	4080
gctcgttatc	gcgaccacat	gaagcagcat	gacttcttca	agtcgcccat	gccgaaggcc	4140
tatgtgcagg	aacgcacgat	ttccttttaag	gatgacggca	cgtacaaaac	gcgtgcggaa	4200
gtgaaatttg	aaggcgatac	cctggtaaac	cgcattgagc	tgaaaggcat	tgacttttaa	4260
tgagaaatga	atatcctggg	ccataagctg	gaatacaatt	ttaacagcca	caatgtttac	4320
atcacgcgcg	ataaacaaaa	aaatggcatt	aaagcgaatt	ttaaaaatcg	ccacaacgtg	4380
gaggatggca	gcgtgcagct	ggctgatcac	taccagcaaa	acactccaat	cgggtgatgtt	4440
cctgtttcgc	tgccagacaa	tcactatctg	agcacgcaaa	gcgttctgtc	tgaaagctcg	4500
aacgagaaac	gcgatcacat	ggttctgctg	gagttcgtaa	ccgcagcggg	catcacgcat	4560
ggatgtgagt	aactgtacaa	ataacaccac	catcatcacc	attaaaagag	catgagagga	4620
cggattttct	gaaggaaatc	cgttttttta	ttttcccgct	cttataaatt	cttttgatta	4680
caaagctgtt	ctttttataat	atgagataat	gccgactgta	ctttttacag	tcgggtttct	4740
aatgtcacta	acctgccccg	ttagttgaag	aaggttttta	tattacagct	ccagatccct	4800
tacgccggac	gcactcgtgg	cggcatcacc	ggcgccacag	gtgcgggttg	tgggcgcctat	4860
atcgccgaca	tcaccgatgg	ggaagatcgg	gctcgccact	tcgggctcat	gagcgcctgt	4920
ttcggcgtgg	gtatgggtgg	aggccccgtg	gccgggggac	tgttggggcg	catctccttg	4980
catgcaccat	tccttgccgc	ggcgggtgct	aacgggcctca	acctactact	gggtgtgctt	5040
ctaagtgcag	agtcgcataa	gggagagcgt	cgcactggat	gagcagatgat	gatctccgtt	5100
taggctgggc	ggtgatagct	tctcgttcag	gcagtaacgc	tcttttcttt	tccagacctg	5160
agggagcgcg	aaatggtgtg	aggttcccgg	ggaaaagcca	aataggcgat	cgcgggagtg	5220
ctttatttga	agatcaggct	atcactcgcg	tcacatgatt	tcacaatgtg	atggctggac	5280
agcctgagga	actctcgaac	ccgaatggaa	acaaccagat	atttatgaat	cagcgcggct	5340
cacatggcgt	tgtgctggca	aatgcaggtt	catcctctgt	ctctatcaat	acggcaacaa	5400
aattgcctga	tggcaggtat	gacaataaag	ctggagcggg	ttcatttcaa	gtgaacgatg	5460
gtaaaactga	aggcacgatc	aatgccaggt	ctgtatgctg	gctttatcct	gtgatatttg	5520
caaaagcgcc	tcattgttttc	cttgagaatt	acaaaacagg	tgtaaacacat	tctttcaatg	5580
atcaactgac	gattaccttg	cgtgcagatg	cgaatacaac	aaaagccggt	tatcaaatca	5640
ataatggacc	agacgacagg	cgttttaagga	tgagatcaaa	ttcacaaatcg	gaaaaggaga	5700
tccaatttgg	caaaacatac	accatcatgt	taaaaggaaac	gaacagtgat	ggtgtaacga	5760
ggaccgagaa	atacagtttt	gttaaaagag	atccagcgtc	ggccaaaacc	atcggtatc	5820
aaaatccgaa	tcatttgagc	caggtaaatg	cttatatcta	taaacatgat	gggagccgag	5880
taattgaatt	gaccggatct	tgccctggaa	aaccaatgac	taaaaatgca	gacggaattt	5940
acacgctgac	gctgcctgcg	gacacggata	caaccaacgc	aaaagtgatt	tttaataatg	6000
gcagcgccca	agtgcctggg	cagaatcagc	ctggctttga	ttacgtgcta	aatgggttat	6060
ataatgactc	gggtttaagc	ggttctcttc	ccacttgagg	gcaaggctag	acgggactta	6120
ccgaaagaaa	ccatcaatga	tggtttcttt	ttgtttcata	aatcagacaa	aacttttctc	6180
ttgcaaaagt	ttgtgaagtg	ttgcacaata	taaattgtgaa	atacttcaca	aacaaaaaga	6240
catcaaaagag	aaacataacc	tgcaaggatg	ctgatattgt	ctgcatttgc	gccggagcaa	6300
acaaaaaac	tggtgagaca	cgcttgaat	tagtagaaaa	gaacttgaag	attttcaaag	6360
gcactgcttag	tgaagtcatg	gcgagcggat	ttgacggcat	tttcttagtc	gcgacgcgag	6420
gctggatggg	cttccccatt	atgattcttc	tcgcttccgg	cggcatcggg	atgcccgctg	6480
tgacggccat	gctgtccagg	caggtagatg	acgacctca	gggacagctt	caaggatcgc	6540
tcgcgctctc	taccagccta	acttcgatca	ctggagcgtg	gatcgtcacg	gcgatttatg	6600
ccgcctcggc	gagcacatgg	aacgggttgg	catgtattgt	aggcgccgcc	ctataccttg	6660
tctgcctccc	cgcggttgcg	cgcggtgcat	ggagccgggc	cacctactga	agtggatttc	6720
tttaagagct	cctttaactt	cctcaccagt	agttgtatcg	gtaccataag	tagaagcagc	6780

[0523]

aacccaagta	gctttaccag	catccgggttc	aaccagcata	gtaagaatct	tactggacat	6840
cggcagttct	tcgaacagtg	cgccaactac	cagctctttc	tgcagttcat	tcagggcacc	6900
ggagaacctg	cgtgcaatcc	atcttgttca	atcatgcgaa	acgatcccta	tcctgtctct	6960
tgatccatgg	attacgcgtt	aaccggggcc	cgcggaatga	tatgatcaga	tcttaaggcc	7020
taggtctaga	ggatcgatct	gtataataaa	gaataattat	taatctgtag	acaaattgtg	7080
aaaggatgta	cttaaacgct	aacgggtcagc	tttattgaac	agtaatttaa	gtatatgtcc	7140
aatctagggt	aagtaaatgt	agtatcaata	taaaccttat	atgaacataa	tcaacgaggt	7200
gaaatcatga	gcaatttcat	taacggaaaa	ataccaaatc	aagcgattca	aacattaaaa	7260
atcgtaaaag	atatttttgg	aagttcaata	gttggaatga	atctatttgg	ttcagcagta	7320
aatgggtggt	tacgcattaa	cagcgatgta	gatgttctag	tcgtcgtgaa	tcatagttta	7380
cctcaattaa	ctcgaaaaaa	actaacagaa	agactaatga	ctatatcagg	aaagattgga	7440
aatacggatt	ctgttagacc	acttgaagtt	acggttataa	ataggagtga	agttgtccct	7500
tggcaatc	ctccaaaaag	agaatttata	tacggtgagt	ggctcagggg	tgaatttgag	7560
aatggacaaa	ttcaggaacc	aagctatgat	cctgatttgg	ctattgtttt	agcacaagca	7620
agaaagaata	gtattttctct	atttgggtcct	gattcttcaa	gtatacttgt	ctccgtacct	7680
ttgacagata	ttcgaagagc	aattaaggat	tctttgccag	aactaattga	ggggataaaa	7740
ggatgatgag	gtaatgtaat	tttaacccta	gctcgaaatg	ggcaaacagt	gactactggt	7800
gaaattacct	cgaaagatgt	cgctgcagaa	tgggctatac	ctcttttacc	taaagagcat	7860
gtaactttac	tggatatagc	tagaaaaggc	tatcggggag	agtgtgatga	taagtgggaa	7920
ggactatatt	caaagggtgaa	agcactcgtt	aagtatatga	aaaattctat	agaaaacttct	7980
ctcaatttag	ctaattttat	tgcataaaca	ggtgcttact	tttctggagt	ttcttagcaa	8040
atttttttat	tagctgaact	tagtattagt	ggccatactc	ctccaatcca	aagctattta	8100
gaaagattac	tatatcctca	aacaggcggt	aaccggcctc	ttcatcggga	atgcgcgcga	8160
ccttcagcat	cgccggcatg	tccccctggc	ggacgggaag	tatccagctc	gaggtcgggc	8220
cgcggttctg	gcgtttttcc	ataggctccg	ccccctgac	gagcatcaca	aaaatcgacg	8280
ctcaagttag	aggtggcgaa	acccgcacag	actataaaga	taccaggcgt	ttccccctgg	8340
aatgctccctc	gtgcgctctc	ctgttccgac	cctgcgcgtt	accggatacc	tgctccgctt	8400
tctcccttcg	ggaagcgtgg	cgctttctca	tagctcacgc	tgtaggtatc	tcagttcggg	8460
gtaggtcgtt	cgctccaagc	tgggctgtgt	gcacgaaccc	cccgttcagc	ccgaccgctg	8520
cgcttatacc	ggttaactatc	gtcttgagtc	caaccgggta	agacacgact	tatcgccact	8580
ggcagcagtc	actggtaaca	ggattagcag	agcgaggtat	gtaggcggtg	ctacagagtt	8640
cttgaagtgg	tggcctaact	acggctacac	tagaaggaca	gtatttggta	tctgcgctct	8700
gctgaagcca	gttaccttcg	gaaaaagagt	tgatagctct	tgatccggca	aacaaaccac	8760
cgctggtagc	ggtgggtttt	ttgtttgcaa	gcagcagatt	acgcgcagaa	aaaaaggatc	8820
tcaagaagat	cctttgatct	tttctacggg	gtctgacgct	cagtggaaag	aaaactcacg	8880
ttaagggatt	ttgggtcatga	gattatcaaa	aaggatcttc	acctagatcc	ttttaaatta	8940
aaaatgaagt	tttaaatcaa	tctaaagtat	atatgagtaa	acttgggtctg	acagttacca	9000
atgcttaatc	agtgaaggac	ctatctcagc	gatctgtcta	tttcgttcat	ccatagttgc	9060
ctgactcccc	gtcgtgtaga	taactacgat	acgggagggc	ttaccatctg	gccccagtg	9120
tgcaatgata	ccgcgagacc	cacgctcacc	ggctccagat	ttatcagcaa	ttaaccagcc	9180
agccggaagg	gcccagcgca	gaagtgggtc	tgcaacttta	tccgcctcca	tccagttcat	9240
taattgttgc	cgggaaagcta	gagtaagtag	ttcgccagtt	aatagtttgc	gcaacgttgt	9300
tgccattgct	gcaggcatcg	tggtgtcacg	ctcgtcgttt	ggtatggctt	cattcagctc	9360
cggttcccaa	cgatcaaggc	gagttacatg	atccccatg	ttgtgcaaaa	aagcggttag	9420
ctccttcggg	cctccgatcg	ttgtcagaag	taagtggg	gcagtgttat	cactcatggg	9480
tatggcagca	ctgcataatt	ctcttactgt	catgccatcc	gtaagatgct	tttctgtgac	9540
tggtgagtac	tcaaccaagt	cattctgaga	atagtgtatg	cggcgaccga	gttgctcttg	9600
cccggcgctca	acacgggata	ataccgcgcc	acatagcaga	acttttaaaag	tgctcatcat	9660
tggaaaacgt	tcttcggggc	gaaaactctc	aaggatctta	ccgctgttga	gatccagttc	9720
gatgtaaccc	actcgtgcac	ccaactgate	ttcagcatct	ttacttttca	ccagcgtttc	9780
tgggtgagca	aaaacaggaa	ggcaaaatgc	cgcaaaaaag	ggaataaggg	cgacacggaa	9840
atggtgaata	ctcatactct	tcttttttca	atattattga	agcattttatc	agggttattg	9900
tctcatgagc	ggatacatat	ttgaatgtat	ttagaaaaat	aaacaaatag	gggttccgcg	9960
cacatttccc	cgaaaagtgc	cacctgacgt	ctaagaaacc	attattatca	tgacattaac	10020
ctataaaaat	aggcgtatca	cgaggccctt	tcgtcttcaa	gaa		10063
<210>	223					
<211>	20					
<212>	DNA					
<213>	合成					
<220>						
<221>	引物 结合					
<222>	(1) . (20)					
<223>	引物 E257					
<400>	223					
ctgagtc	aat	ttcatcaagg				20
<210>	224					
<211>	22					
<212>	DNA					
<213>	合成					
<220>						
<221>	引物 结合					
<222>	(1) . (22)					
<223>	引物 E258					
<400>	224					
cagaagaaaa	cattctcgaa	gc				22
<210>	225					
<211>	20					

<212> DNA
 <213> 合成
 <220>
 <221> 引物 结合
 <222> (1) .. (20)
 <223> 引物 E261
 <400> 225
 gaggtcgcaa ctgaaaaagc 20
 <210> 226
 <211> 20
 <212> DNA
 <213> 合成
 <220>
 <221> 引物 结合
 <222> (1) .. (20)
 <223> 引物 E262
 <400> 226
 atgttcagtc tccttactcc 20
 <210> 227
 <211> 30
 <212> DNA
 <213> 合成
 <220>
 <221> 引物 结合
 <222> (1) .. (30)
 <223> 引物 E372
 <400> 227
 attataacau gtattcacga acgaaaatcg 30
 <210> 228
 <211> 35
 <212> DNA
 <213> 合成
 <220>
 <221> 引物 结合
 <222> (1) .. (35)
 <223> 引物 E373
 <400> 228
 atcagatttu agaaaacaat aaacccttgc atagg 35
 <210> 229
 <211> 20
 <212> DNA
 <213> 合成
 <220>
 <221> 引物 结合
 <222> (1) .. (20)
 <223> 引物 E374
 <400> 229
 aaaatctgau acggggcagg 20
 <210> 230
 <211> 37
 <212> DNA
 <213> 合成
 <220>
 <221> 引物 结合
 <222> (1) .. (37)
 <223> 引物 E375
 <400> 230
 atgttataau acccattagt tcaacaaacg aaaattg 37
 <210> 231
 <211> 264
 <212> PRT
 <213> 合成
 <220>
 <221> 肽
 <222> (1) .. (264)
 <223> KanR
 <400> 231
 Met Ile Glu Gln Asp Gly Leu His Ala Gly Ser Pro Ala Ala Trp Val
 1 5 10 15
 Glu Arg Leu Phe Gly Tyr Asp Trp Ala Gln Gln Thr Ile Gly Cys Ser
 20 25 30
 Asp Ala Ala Val Phe Arg Leu Ser Ala Gln Gly Arg Pro Val Leu Phe
 35 40 45
 Val Lys Thr Asp Leu Ser Gly Ala Leu Asn Glu Leu Gln Asp Glu Ala
 50 55 60

Ala Arg Leu Ser Trp Leu Ala Thr Thr Gly Val Pro Cys Ala Ala Val
 65 70 75 80
 Leu Asp Val Val Thr Glu Ala Gly Arg Asp Trp Leu Leu Gly Glu
 85 90 95
 Val Pro Gly Gln Asp Leu Leu Ser Ser His Leu Ala Pro Ala Glu Lys
 100 105 110
 Val Ser Ile Met Ala Asp Ala Met Arg Arg Leu His Thr Leu Asp Pro
 115 120 125
 Ala Thr Cys Pro Phe Asp His Gln Ala Lys His Arg Ile Glu Arg Ala
 130 135 140
 Arg Thr Arg Met Glu Ala Gly Leu Val Asp Gln Asp Asp Leu Asp Glu
 145 150 155 160
 Glu His Gln Gly Leu Ala Pro Ala Glu Leu Phe Ala Arg Leu Lys Ala
 165 170 175
 Arg Met Pro Asp Gly Glu Asp Leu Val Val Thr His Gly Asp Ala Cys
 180 185 190
 Leu Pro Asn Ile Met Val Glu Asn Gly Arg Phe Ser Gly Phe Ile Asp
 195 200 205
 Cys Gly Arg Leu Gly Val Ala Asp Arg Tyr Gln Asp Ile Ala Leu Ala
 210 215 220
 Thr Arg Asp Ile Ala Glu Glu Leu Gly Gly Glu Trp Ala Asp Arg Phe
 225 230 235 240
 Leu Val Leu Tyr Gly Ile Ala Ala Pro Asp Ser Gln Arg Ile Ala Phe
 245 250 255
 Tyr Arg Leu Leu Asp Glu Phe Phe
 260

<210> 232
 <211> 2734
 <212> DNA
 <213> 合成
 <220>
 <221> 尚未归类的特征
 <222> (1)..(2734)
 <223> 包含 P(gsh2) 启动子、cia1 必需基因和 TIS1 翻译起始序列的整合片段 INT1。
 <220>
 <221> CDS
 <222> (1024)..(1398)
 <223> blaShe (用于选择染色体整合的抗性基因)
 <220>
 <221> 启动子
 <222> (1724)..(2234)
 <223> P13_pGSH2
 <220>
 <221> 启动子
 <222> (2229)..(2234)
 <223> TIS1
 <220>
 <221> 基因
 <222> (2235)..(2734)
 <223> CIA1 基因的第一部分, 用于通过同源重组整合
 <400> 232

tggtcctcgt ctatttggtg tttaaaagtg atagtatact tcttgaactc gctgggggtta 60
 agaaattcca agtcttcacg ttctacaaag aagtagttgt acgagaacaa tagccataca 120
 gtaccaccta tgacgaccag actaggaata aatctcaacc atgggtttct cgcgtttcta 180
 atgacctttt ctgctttcaa ctccgatggg gaaggtggag gctgggtgtc ataagtgacc 240
 ctttggttgg gtcttggaat gtattgtttt ttcagttgaa tcccttgcaa cggttgaagg 300
 ccgccagtea acttgggggc gtttttctta gcagggactg cataatttct aacgagccgt 360
 agtattctct ggttcacgga tcatagaact gaggggaagg ggaagtgtag aaatgagcag 420
 aaatggtagg ggatttggag ttctctctaa gcagcgaaca gaaaattcaa gcaccctaga 480
 agcagatccc attaatgcaa tccgccccat caaccatcaa ctacaagatc tatcacaac 540
 acaccacac accatagctt caaaatgttt ctactcctt tttactcttc cagattttct 600
 cggactccgc gcacgcgct accacttcaa aacaccaag cacagcatac taaattttcc 660
 ctctttcttc ctctaggggtg tcgttaatta cccgtactaa aggtttggaa aagaaaaaag 720
 agaccgcctc gtttcttttt ctctgctgaa aaaggcaata aaaattttta tcacgtttct 780
 ttttcttgaa attttttttt ttagtttttt tctctttcag tgacctccat tgatatttaa 840
 gttaataaac ggtcttcaat ttctcaagtt tcagtttcat ttttcttggt ctattacaac 900
 tttttttact tcttggtcat tagaaagaaa gcatagcaat ctaatctaag gggcgggtgt 960
 gacaattaat catcggcata gtatatcggc atagtataat acgacaaggt gaggaactaa 1020
 acc atg gcc aag ttg acc agt gcc gtt ccg gtg ctc acc gcg cgc gac 1068
 Met Ala Lys Leu Thr Ser Ala Val Pro Val Leu Thr Ala Arg Asp
 1 5 10 15
 gtc gct gga gcg gtc gag ttc tgg acc gac cgg ctc ggg ttc tcc cgg 1116
 Val Ala Gly Ala Val Glu Phe Trp Thr Asp Arg Leu Gly Phe Ser Arg
 20 25 30
 gac ttc gtg gag gac gac ttc gct ggt gtg gtc cgg gac gac gtg acc 1164
 Asp Phe Val Glu Asp Asp Phe Ala Gly Val Val Arg Asp Asp Val Thr
 35 40 45

ctg ttc atc agc gcg gtc cag gac cag gtg gtg ccg gac aac acc ttg	1212
Leu Phe Ile Ser Ala Val Gln Asp Gln Val Val Pro Asp Asn Thr Leu	
50 55 60	
gct tgg gtg tgg gtg cgc ggt ttg gac gag ctg tac gcc gag tgg tcg	1260
Ala Trp Val Trp Val Arg Gly Leu Asp Glu Leu Tyr Ala Glu Trp Ser	
65 70 75	
gag gtc gtg tcc acg aac ttc ccg gac gcc tcc ggt cca gcc atg acc	1308
Glu Val Val Ser Thr Asn Phe Arg Asp Ala Ser Gly Pro Ala Met Thr	
80 85 90 95	
gag atc ggc gag cag ccg tgg ggg agg gag ttc gcc ctg cgc gat ccg	1356
Glu Ile Gly Glu Gln Pro Trp Gly Arg Glu Phe Ala Leu Arg Asp ccg	
100 105 110	
gct ggc aac tgc gtt cac ttc gtg gcc gag gag cag gac tga	1398
Ala Gly Asn Cys Val His Phe Val Ala Glu Glu Gln Asp	
115 120	
cacgtccgac ggcgccccac ggggtcccagg cctcggagat ccgtccccct tttcctttgt	1458
cgatatcatg taattagtta tgtaacgctt acattcacgc cctcccccca catccgctct	1518
aaccgaaaag gaaggagtta gacaacctga agtctagggtc cctattttatt tttttatagt	1578
tatgttagta ttaagaacgt tattttatatt tcaaattttt ctttttttttc tgtacagacg	1638
cggttacgca tgtaacatta tactgaaaac ctgtcttgag aagggttttg gacgctcgaa	1698
ggctttaatt tgcaagctat caactctggc ccgaggaagt cagcgaaaaga ggatacttta	1758
ccccagagg agaataataa attgatgaca atgcagctca agccatgaaa gactcaatgc	1818
tgtacaagat gtcattttac agattcgggtg aactttttccc ctctggagat gccatagacc	1878
gtgtcagagg tcaaagattg agtcgctctt acgctgagtc tatagatctt aacattgttg	1938
aagaagtatt cacaagtgag aactggctgg ttctgtctta caagctcaaa gagccagaca	1998
acttggaag aagcctactg accctgaagg ataacgagaa aaagttaggcc actaaaaaag	2058
gtcgagatt gaggtcaac aagaaaccct cattggatct acgggtgtag tgtagttgtg	2118
tatttttagat agacagagct ctgggggttag aatgataaca tattgcgtgc ttatgtaatg	2178
atgttcacct ttctgtaggtc tcaaaaagtta gcctagtctt agcccgtaac tgatatatgt	2238
ctctatcttt gggtcacaca ttacacaggtc acacagataa atgctgggct acttccattc	2298
actctaaact tcttttattg gcaacagctc ccggggacag aacttgtcgg gtctataact	2358
tggaacgaa gagactgatt gtcgtactcg acgatgattc tcattcaaaag actctttcca	2418
cgctcgaatg gaagccaact ggggaattcc ctctgtggc aatcggctcg tttgattcta	2478
caatctctat ttgggggaat gaagaagctg tcctggaaga cgaggataac tggacactga	2538
tggtcattat cgaaggacat gagaacgaga ttaaaggggg gagttggagc catgatggag	2598
tgtatctggc atcatgctca cgagacaagt ccatttggat atgggaggct gatgataaca	2658
acgaagagtt tgagtgcgtc tttgtgtgac aggaacacag ccaagatgtc aagcatgtaa	2718
catggcatca gcatga	2734
<210> 233	
<211> 124	
<212> PRT	
<213> 合成	
<400> 233	
Met Ala Lys Leu Thr Ser Ala Val Pro Val Leu Thr Ala Arg Asp Val	
1 5 10 15	
Ala Gly Ala Val Glu Phe Trp Thr Asp Arg Leu Gly Phe Ser Arg Asp	
20 25 30	
Phe Val Glu Asp Asp Phe Ala Gly Val Val Arg Asp Asp Val Thr Leu	
35 40 45	
Phe Ile Ser Ala Val Gln Asp Gln Val Val Pro Asp Asn Thr Leu Ala	
50 55 60	
Trp Val Trp Val Arg Gly Leu Asp Glu Leu Tyr Ala Glu Trp Ser Glu	
65 70 75 80	
Val Val Ser Thr Asn Phe Arg Asp Ala Ser Gly Pro Ala Met Thr Glu	
85 90 95	
Ile Gly Glu Gln Pro Trp Gly Arg Glu Phe Ala Leu Arg Asp Pro Ala	
100 105 110	
Gly Asn Cys Val His Phe Val Ala Glu Glu Gln Asp	
115 120	
<210> 234	
<211> 2734	
<212> DNA	
<213> 合成	
<220>	
<221> 尚未归类的特征	
<222> (1) .. (2734)	
<223> 包含 P(gsh2) 启动子、cia1 必需基因和 TIS3 翻译起始序列的整合片段 INT3。	
<220>	
<221> 启动子	
<222> (1724) .. (2234)	
<223> P13_pGSH2	
<220>	
<221> 启动子	
<222> (2229) .. (2234)	
<223> TIS3	
<220>	
<221> 基因	

[0526]

<222> (2235) .. (2734)
<223> CIA1 基因的第一部分, 用于通过同源重组整合
<400> 234

tggctcctcgt	ctatattgttg	tttaaaagt	atagtataact	tcttgaactc	gctgggggta	60
agaaattcca	agtcttcac	ttctacaaag	aagtagttgt	acgagaacaa	tagccataca	120
gtaccaccta	tgacgaccag	actaggaata	aatctcaacc	atgggtttct	ccgctttcta	180
atgacctttt	ctgctttcaa	ctccgatggg	gaaggtggag	gctggtgttc	ataagtgacc	240
ctttgggttg	gtcttggaat	gtattgtttt	ttcagttgaa	tcccttgcaa	cggttgaagg	300
ccgccagtca	acttgggggc	gtttttctta	gcagggactg	cataatttct	aacgagccgt	360
agtattctct	ggttcacgga	tcatagaact	gaggggaagg	ggaagtgtag	aaatgagcag	420
aaatggtagg	ggatttgga	ttctctctaa	gcagcgaaac	gaaaattcaa	gcaccctaga	480
agcagatccc	attaatgcaa	tccgccccat	caaccatcaa	ctacaagatc	tatcacaaac	540
acaccacac	accatagctt	caaaatgttt	ctactccttt	tttactcttc	cagattttct	600
cggaactcgc	gcacgcgcgt	accacttcaa	aacacccaag	cacagcatac	taaattttcc	660
ctctttcttc	ctctaggggtg	tcgttaatta	cccgtactaa	aggtttggaa	aagaaaaaag	720
agaccgcctc	gtttcttttt	cttcgtcgaa	aaaggcaata	aaaattttta	tcacgtttct	780
ttttctgaa	attttttttt	ttagtttttt	ttcttttcag	tgacctccat	tgatatttaa	840
gttaataaac	ggtcttcaat	ttctcaagtt	tcagtttcat	ttttcttgtt	ctattacac	900
tttttttaact	tcctgttcat	tagaaagaaa	gcatagcaat	ctaactaag	ggcgcggtgt	960
gacaattaat	catcggcata	gtatatcggc	atagtataat	acgacaagg	gaggaaacta	1020
accatggcca	agttgaccag	tgccgttccg	gtgctcaccg	cgcgcgacgt	cgctgggacg	1080
gtcgagttct	ggaccgaccg	gctcgggttc	tcocgggact	tcgtggagga	cgacttcgct	1140
ggtgtggtcc	gggacgacgt	gacctgttc	atcagcgcg	tcaggaacca	ggtggtgccc	1200
gacaacacct	tggtttgggt	gtgggtgcgc	ggtttggacg	agctgtacgc	cgagtgttgc	1260
gaggtcgtgt	ccacgaactt	ccgggacgcc	tcgggtccag	ccatgaccga	gatcgggcag	1320
cagccgtggg	ggagggagtt	cgccctgcgc	gatccggctg	gcaactgcgt	tcacttcgtg	1380
gccgaggagc	aggactgaca	cgcccgacgg	cgcccgacgg	gtcccaggcc	tcggagatcc	1440
gtcccccttt	tcctttgtcg	atatcatgta	attagttatg	tcacgcttac	attcacgccc	1500
tcoccccaata	tcgcctctaa	ccgaaaagga	aggagttaga	caacctgaag	tctaggtccc	1560
tattttatttt	tttatagtta	tgttagtatt	aagaacgtta	tttatatttc	aaatttttct	1620
ttttttctcg	tacagacgcg	tgtaacgatg	taacattata	ctgaaaacct	tgcttgagaa	1680
ggtttttgga	cgctcgaagg	ctttaatttg	caagctatca	actctggccc	gaggaaagta	1740
gcgaaagagg	atactttacc	cccagaggag	aatataaaaat	tgatgacaat	gcagctcaag	1800
ccatgaaaga	ctcaatgctg	tacaagatgt	cattttacag	attcggtgaa	cttttcccct	1860
ctggagatgc	catagaccgt	gtcagagggtc	aaagattgag	tcgctcttac	gctgagtcta	1920
tagatcttaa	cattgttgaa	gaagtattca	caagtgaaga	ctggctgggt	cgtctttaca	1980
agctcaaaga	ccagacaac	ttgggaagaa	gcctactgac	cctgaaggat	aacgagaaaa	2040
agttggccac	taaaaaaggt	cgagatttga	gggtcaacaa	gaaacctca	ttggatctac	2100
gggtgtagt	tagttgtgta	ttttagatag	acagagctct	ggggttagaa	tgataacata	2160
ttgcgtgctt	atgtaatgat	gttcaccttt	cgtaggtctc	aaaagttagc	ctagtcttag	2220
cccgtaacac	gttcacgtct	ctatcttttg	ttcacacatt	cacaggtcac	acagataaat	2280
gctgggtcac	ttccattcac	tctaaacttc	ctttattggc	aacagtctcc	ggggacagaa	2340
cttgctgggt	ctataacttg	gaaacgaaga	gactgattgt	cgtaactcgac	gatgattctc	2400
attcaaaagc	tctttccagc	gtcgaatgga	agccaaactg	ggaattccct	tcgttggaac	2460
tcggctcgtt	tgattctaca	atctctatct	gggggaatga	agaagctgtc	ctggaagacg	2520
aggataactg	gacactgatg	gccattatcg	aaggacatga	gaacgagatt	aaaggggtga	2580
gttgagcca	tgatggagtg	tatctggcat	catgctcacg	agacaagtcc	atttggtat	2640
gggaggtcga	tgataacaac	gaagagtttg	agtgcgtctt	tggtgtacag	gaacacagcc	2700
aagatgtcaa	gcattgtaaca	tgccatcagc	atga			2734

<210> 235
<211> 2734
<212> DNA
<213> 合成
<220>
<221> 尚未归类的特征
<222> (1) .. (2734)
<223> 包含 P(gsh2) 启动子、cia1 必需基因和 TIS5 翻译起始序列的整合片段 INT5。
<220>
<221> 启动子
<222> (1724) .. (2234)
<223> P13_pGSH2
<220>
<221> 启动子
<222> (2229) .. (2234)
<223> TIS5
<220>
<221> 基因
<222> (2235) .. (2734)
<223> CIA1 基因的第一部分, 用于通过同源重组整合
<400> 235

tggctcctcgt	ctatattgttg	tttaaaagt	atagtataact	tcttgaactc	gctgggggta	60
agaaattcca	agtcttcac	ttctacaaag	aagtagttgt	acgagaacaa	tagccataca	120
gtaccaccta	tgacgaccag	actaggaata	aatctcaacc	atgggtttct	ccgctttcta	180
atgacctttt	ctgctttcaa	ctccgatggg	gaaggtggag	gctggtgttc	ataagtgacc	240
ctttgggttg	gtcttggaat	gtattgtttt	ttcagttgaa	tcccttgcaa	cggttgaagg	300
ccgccagtca	acttgggggc	gtttttctta	gcagggactg	cataatttct	aacgagccgt	360
agtattctct	ggttcacgga	tcatagaact	gaggggaagg	ggaagtgtag	aaatgagcag	420

aaatggtagg	ggattttggag	ttctctcttaa	gcagcgaaca	gaaaattcaa	gcaccctaga	480
agcagatccc	attaatgcaa	tccgccccat	caaccatcaa	ctacaagatc	tatcacaaac	540
acacccacac	accatagcct	caaaaatgttt	ctactccttt	tttactcttc	cagattttct	600
cggactccgc	gcacgcgcgt	accacttcaa	aacacccaag	cacagcatac	taaaattttcc	660
ctctttcttc	ctctaggggtg	tcgttaatta	cccgactacta	aggttttggaa	aagaaaaaag	720
agaccgcctc	gtttctttttt	cttcgtcgaa	aaaggcaata	aaaattttta	tcacgtttct	780
ttttcttgaa	attttttttt	ttagtttttt	tctctttcag	tgacctccat	tgatatttaa	840
gttaataaac	ggtcttcaat	ttctcaagtt	tcagtttcat	ttttcttggt	ctattacaac	900
tttttttaact	tcttggttcat	tagaaagaaa	gcatagcaat	ctaataaag	gggagggtgt	960
gacaattaat	catcggcata	gtatatcggc	atagtataat	acgacaagg	gaggaaactaa	1020
accatggcca	agttgaccag	tgccgttccg	gtgctcaccg	cgcgcgacgt	cgctggagcg	1080
gtcagattct	ggaccgaccg	gctcgggttc	tcccgggact	tcgtggaggga	cgacttcgct	1140
ggtgtggtcc	gggacgacgt	gacctgttc	atcagcgcg	tccaggacca	ggtggtgccc	1200
gacaacacct	tggcttgggt	gtgggtgcgc	ggtttggacg	agctgtacgc	cgatgggtcg	1260
gaggctcgtg	ccacgaactt	ccgggacgcc	tccggtccag	ccatgaccga	gatcggcgag	1320
cagccgtggg	ggaggagggt	cgccctgcgc	gatccggctg	gcaactcgct	tcacttcgtg	1380
gccgaggagc	aggactgaca	cgctccgacg	cggcccaagg	gtcccaggcc	tcggagatcc	1440
gtcccccttt	tcctttgtcg	atatcatgta	attagttatg	tcacgcttac	attcacgccc	1500
tcccccaca	tcgcctctaa	ccgaaaagga	aggagttaga	caacctgaag	cttaggtccc	1560
tatttatttt	tttatagtta	tgttagtatt	aagaacgtta	tttataattc	aaatttttct	1620
ttttttctcg	tacagacgcg	tgtagcatg	taacattata	ctgaaaacct	tgcttgagaa	1680
ggttttggga	cgctcgaagg	ctttaatttg	caagctatca	actctggccc	gagggaagta	1740
cgcaaaagag	atactttacc	cccagaggag	aatataaaat	tgatgacaat	gcagctcaag	1800
ccatgaaaga	ctcaatgctg	tacaagatgt	cattttacag	attcggtgaa	cttttcccc	1860
ctggagatgc	catagaccgt	gtcagagggt	aaagattgag	tcgctcttac	gctgagtcta	1920
tagatcttaa	cattgttgaa	gaagtattca	caagttagaa	ctggctgggt	cgtctttaca	1980
agctcaaaag	gccagacaac	ttgggaagaa	gcctactgac	cctgaaggat	aacgagaaaa	2040
agttggccac	taaaaaaggt	cgagatttga	gggtcaacaa	gaaacctca	ttggatctac	2100
gggtgtagt	tagttgtgta	tttttagatag	acagagctct	ggggttagaa	tgataacata	2160
ttgcgtgctt	atgtaatgat	gttcaccttt	cgtaggcttc	aaaagttagc	ctagtcttag	2220
cccgtaaacta	ggttatgtct	ctatcttttg	ttcacacatt	cacaggtcac	acagataaat	2280
gctgggttac	ttccattcac	tctaaacttc	ctttattggc	aacagtctcc	ggggacagaa	2340
cttgctgggt	ctataacttg	gaaacgaaga	gactgattgt	cgtactcgac	gatgattctc	2400
attcaaaagc	tctttccagc	gtcgaatgga	agccaaactgg	ggaattccct	tcgttggaac	2460
tcggctcgtt	tgattctaca	atctctatct	gggggaatga	agaagctgtc	ctggaagacg	2520
aggataactg	gacactgatg	gccattatcg	aaggacatga	gaacgagatt	aaaggggtga	2580
gttgagcca	tgatggagtg	tatctggcat	catgctcacg	agacaagtcc	atttggtat	2640
gggagcgtga	tgataacaac	gaagagtttg	agtgcgtctt	tgttgtacag	gaacacagcc	2700
aaagatgtcaa	gcattgtaaca	tgccatcagc	atga			2734
<210>	236					
<211>	2734					
<212>	DNA					
<213>	合成					
<220>						
<221>	尚未归类的特征					
<222>	(1) .. (2734)					
<223>	包含 P(gsh2) 启动子、cia1 必需基因和 TIS9 翻译起始序列的整合片段 INT9。					
<220>						
<221>	启动子					
<222>	(1724) .. (2234)					
<223>	P13_pGSH					
<220>						
<221>	启动子					
<222>	(2229) .. (2234)					
<223>	TIS9					
<220>						
<221>	基因					
<222>	(2235) .. (2734)					
<223>	CIA1 基因的第一部分，用于通过同源重组整合					
<400>	236					
tggtcctcgt	ctattttgttg	tttaaaagtg	atagtatact	tcttgaactc	gctgggggtta	60
agaaaattcca	agtcttcatc	ttctacaaag	aagttagttgt	acgagaacaa	tagccataca	120
gtaccaccta	tgacgaccag	actaggaata	aatctcaacc	atgggtttct	ccgcttttcta	180
atgacctttt	ctgcttttcaa	ctccgatggt	gaagggtggag	gctgggtgtc	ataagtgacc	240
cttttggttg	gtcttggaat	gtattgtttt	ttcagttgaa	tcccttgcaa	cggttgaagg	300
ccgccagtca	acttgggggc	gtttttctta	gcagggaactg	cataatttct	aacgagccgt	360
agtattctct	ggttcatgga	tcatagaact	gagggggaagg	ggaagtgtag	aaatgagcag	420
aaatggttag	ggattttggag	ttctctctaa	gcagcgaaca	gaaaattcaa	gcaccctaga	480
agcagatccc	attaatgcaa	tccgccccat	caaccatcaa	ctacaagatc	tatcacaaac	540
acacccacac	accatagcct	caaaaatgttt	ctactccttt	tttactcttc	cagattttct	600
cggactccgc	gcacgcgcgt	accacttcaa	aacacccaag	cacagcatac	taaaattttcc	660
ctctttcttc	ctctagggtg	tcgttaatta	cccgactacta	aggttttggaa	aagaaaaaag	720
agaccgcctc	gtttcttttt	cttcgtcgaa	aaaggcaata	aaaattttta	tcacgtttct	780
ttttcttgaa	attttttttt	ttagtttttt	tctctttcag	tgacctccat	tgatatttaa	840
gttaataaac	ggtcttcaat	ttctcaagtt	tcagtttcat	ttttcttggt	ctattacaac	900
tttttttaact	tcttggttcat	tagaaagaaa	gcatagcaat	ctaataaag	gggagggtgt	960
gacaattaat	catcggcata	gtatatcggc	atagtataat	acgacaagg	gaggaaactaa	1020

	accatggcca	agttgaccag	tgccgttccg	gtgctcaccg	cgcgcgacgt	cgctggagcg	1080
	gtcgagttct	ggaccgaccg	gctcgggttc	tcccgggact	tcgtggagga	cgacttcgct	1140
	ggtgtggtcc	gggacgacgt	gacctgttc	atcagcgcg	tccaggacca	ggtggtgccc	1200
	gacaacacct	tggttgggt	gtgggtgccc	ggtttggacg	agctgtacgc	cgagtggctg	1260
	gaggtcgtgt	ccacgaactt	ccgggacgcc	tccgggtccag	ccatgaccga	gatcggcgag	1320
	cagccgtggg	ggagggagtt	cgccctgccc	gatccggctg	gcaactgcgt	tcacttcgtg	1380
	gccgaggagc	aggactgaca	cgcccgacgg	cgcccccagg	gtcccaggcc	tcggagatcc	1440
	gtcccccttt	tcctttgtcg	atatcatgta	attagttatg	tcacgcttac	attcacgccc	1500
	tcctcccccaca	tcgcgtctaa	ccgaaaaggga	aggagttaga	caacctgaag	tcagggtccc	1560
	tattttatttt	tttatagtta	tgtagtatt	aagaacgtta	tttatatttc	aaatttttct	1620
	tttttttctg	tacagacgcg	tgtagcgtg	taacattata	ctgaaaacct	tgcttgagaa	1680
	ggttttggga	cgctcgaagg	ctttaatttg	caagctatca	actctggccc	gaggaagtca	1740
	gcgaaagagg	atactttacc	cccagaggag	aataataaat	tgatgacaat	gcagctcaag	1800
	ccatgaaaga	ctcaatgctg	tacaagatgt	cattttacag	attcgggtgaa	cttttcccct	1860
	ctggagatgc	catagaccgt	gtcagaggtc	aaagattgag	tcgctcttac	gctgagtcta	1920
	tagatcttaa	cattgttgaa	gaagtattca	caagttagaa	ctggctgggt	cgctcttaca	1980
	agctcaaga	gccagacaac	ttgggaagaa	gcctactgac	cctgaaggat	aacgagaaaa	2040
	agttggccac	taaaaaaggt	cgagatttga	gggtcaacaa	gaaacctca	ttggatctac	2100
	gggtgtagt	tagttgtgta	ttttagatag	acagagctct	gggttagaa	tgataacata	2160
	ttgcgtgctt	atgtaatgat	gttcaccttt	cgtaggctct	aaaagtttag	ctagtcttag	2220
	cccgtaacaa	aacaatgtct	ctatctttgg	ttcacacatt	cacaggtcac	acagataaat	2280
	gctgggctac	ttccattcac	tctaaacttc	ctttattggc	aacagtctcc	ggggacagaa	2340
	cttgcgggt	ctataacttg	gaaacgaaga	gactgattgt	cgtactcgac	gatgattctc	2400
	attcaagac	tccttccagc	gtcgaatgga	agccaactgg	ggaattcccct	tcgttggcaa	2460
	tcggctcgtt	tgattctaca	atctctatct	gggggaatga	agaagctgtc	ctggaagacg	2520
	aggataactg	gacactgatg	gccattatcg	aaggacatga	gaacgagatt	aaaggggtga	2580
	gttgaggcca	tgatggagtg	tatctggcat	catgctcag	agacaagtcc	atttggatat	2640
	ggaggagctga	tgataacaac	gaagagtttg	agtgcgtctt	tgttgtacag	gaacacagcc	2700
	aagatgtcaa	gcattgtaaca	tgcatcagc	atga			2734
	<210>	237					
	<211>	19					
	<212>	DNA					
	<213>	合成					
	<220>						
	<221>	引物 结合					
	<222>	(1) .. (19)					
	<223>	引物 E251					
	<400>	237					
[0529]	acgcgtgtac	gcattgtaac					19
	<210>	238					
	<211>	20					
	<212>	DNA					
	<213>	合成					
	<220>						
	<221>	引物 结合					
	<222>	(1) .. (20)					
	<223>	引物 E627					
	<400>	238					
	cataactgga	gcttgctagg					20
	<210>	239					
	<211>	750					
	<212>	DNA					
	<213>	米曲霉					
	<220>						
	<221>	启动子					
	<222>	(1) .. (750)					
	<223>	米曲霉负感应启动子 p(agsA)					
	<400>	239					
	ggtgtttctt	aatataggca	cacggatccc	gccatcgatc	ggcgccctgt	tcgagtatgg	60
	ccatcctcac	acattgtcgg	gtcccgctgg	cgctgectca	acatatatttt	aaatggggaa	120
	accataagca	aagagactcc	atcggggcgca	ctacacatct	cgctctctct	ggacctctgt	180
	gatcgggact	gggggtggatg	gcaaaccgtc	aagttgccc	cgctctctct	cgctgttgcc	240
	gcgtcacgcc	atataggcga	ctctcccaat	catggaaagc	ctgcgttcca	gcgatattca	300
	acgacgtggt	ggcttggatt	gcagcttcga	ccagatcaca	gcaatagaat	agcgatgatt	360
	ccagccgtgg	gccttttaggg	tagaaaccat	actgggctag	gaaacatgcc	gtatcaagtg	420
	agaaaacaccg	ctaccccacc	acggctgaaa	agtagatagt	ggatgggtcc	atgtcatcta	480
	tcagcaggcg	aggagtgggt	ctatagtaga	ggttgtaagt	gtcgtggtag	aggtccctgg	540
	cccattttta	tcctccatct	tcccgccctg	atggttccaa	gccttcttta	cgaaaagtgt	600
	tgacagctgt	ttttcgagtc	gacaggctgc	tcgttccctc	gtttgatttc	cctcgctttt	660
	gaccacgccc	agggcatgta	ggattttatc	cgtttgagat	tatcgagagc	ctaagtcgcg	720
	cgagttgtc	ggtagcaggc	cactgccggg				750
	<210>	240					
	<211>	1200					
	<212>	DNA					
	<213>	米曲霉					
	<220>						

	<221> CDS																	
	<222> (1)..(1200)																	
	<223> ERG10 曲霉属必需基因 (mRNA																	
	<400> 240																	
	atg	tct	tct	ctt	ccg	ccg	gtc	tat	att	gtt	tcc	tct	gcc	cgc	acc	cca	48	
Met	Ser	Ser	Leu	Pro	Pro	Val	Tyr	Ile	Ile	Val	Ser	Ser	Ala	Arg	Thr	Pro		
1				5						10					15			
	gtc	ggc	tct	ttc	ttg	ggg	tcg	ctc	tca	agt	ctc	act	gcc	ccg	cag	tta	96	
Val	Gly	Ser	Phe	Leu	Gly	Ser	Leu	Ser	Ser	Ser	Leu	Thr	Ala	Pro	Gln	Leu		
			20					25						30				
	ggc	tct	cat	gct	att	aaa	gct	gcg	ctc	agc	aaa	gcg	gat	gga	atc	aag	144	
Gly	Ser	His	Ala	Ile	Lys	Ala	Ala	Leu	Ser	Lys	Ala	Asp	Gly	Ile	Lys			
		35				40						45						
	ccg	tct	gat	atc	cag	gag	gtc	ttc	ttt	ggc	aat	gtc	atc	tcc	gca	aac	192	
Pro	Ser	Asp	Ile	Gln	Glu	Val	Phe	Phe	Phe	Gly	Asn	Val	Ile	Ser	Ala	Asn		
		50				55						60						
	gtt	gga	caa	aat	cct	gct	aga	cag	tgt	gct	ctc	ggc	gct	ggt	ctc	aat	240	
Val	Gly	Gln	Asn	Pro	Ala	Arg	Gln	Cys	Ala	Leu	Gly	Ala	Gly	Leu	Asn			
		65			70				75					80				
	gaa	tca	act	gtc	tgt	act	acg	gtt	aat	aag	gtg	tgc	gcg	tct	ggc	ttg	288	
Glu	Ser	Thr	Val	Cys	Thr	Thr	Val	Asn	Asn	Lys	Val	Cys	Ala	Ser	Gly	Leu		
				85					90						95			
	aaa	gcg	gtt	att	ctc	ggt	gca	cag	acc	atc	atg	act	ggc	aat	gcg	gat	336	
Lys	Ala	Val	Ile	Leu	Gly	Ala	Gln	Thr	Thr	Ile	Met	Thr	Gly	Asn	Ala	Asp		
			100					105						110				
	att	gtc	gta	gca	ggc	ggt	gct	gaa	tcc	atg	tct	aac	gcc	cct	cat	tac	384	
Ile	Val	Val	Ala	Gly	Gly	Ala	Glu	Ser	Met	Met	Ser	Asn	Ala	Pro	His	Tyr		
		115				120							125					
	ctt	cca	aac	ctt	cgc	gtc	ggt	gcg	aaa	tac	ggc	aac	cag	agt	ctg	gtg	432	
Leu	Pro	Asn	Leu	Arg	Val	Gly	Gly	Ala	Lys	Tyr	Gly	Asn	Gln	Ser	Leu	Val		
		130				135						140						
	gac	ggt	att	atg	aag	gat	ggc	ttg	aca	gac	gca	gga	aag	cag	gaa	ctc	480	
Asp	Gly	Ile	Met	Lys	Asp	Gly	Leu	Thr	Thr	Asp	Ala	Gly	Lys	Gln	Glu	Leu		
		145			150					155						160		
	atg	ggc	ttg	caa	gcc	gag	gag	tgt	gct	cag	gat	cat	ggc	ttt	agc	agg	528	
Met	Gly	Leu	Gln	Ala	Glu	Glu	Cys	Ala	Ala	Gln	Asp	His	Gly	Phe	Ser	Arg		
				165				170							175			
	gaa	caa	cag	gat	gat	tat	gcc	att	cgc	act	tac	gaa	aaa	gca	cag	gcg	576	
Glu	Gln	Gln	Asp	Asp	Tyr	Ala	Ile	Ile	Arg	Thr	Tyr	Glu	Lys	Ala	Gln	Ala		
			180					185						190				
	gct	caa	aag	gct	ggc	ctt	ttt	gac	gaa	gaa	att	gcg	cct	att	gaa	ctt	624	
Ala	Gln	Lys	Ala	Gly	Leu	Phe	Asp	Glu	Glu	Glu	Ile	Ala	Pro	Ile	Glu	Leu		
		195					200						205					
	cct	ggc	ttt	agg	ggc	aag	cca	ggt	gtg	act	gtg	tca	caa	gac	gaa	gaa	672	
Pro	Gly	Phe	Arg	Gly	Lys	Pro	Gly	Val	Thr	Thr	Val	Ser	Gln	Asp	Glu	Glu		
		210				215						220						
	cca	aag	aat	ctt	aac	ccg	gat	aag	ctt	cga	gct	atc	aag	cct	gca	ttt	720	
Pro	Lys	Asn	Leu	Asn	Pro	Asp	Lys	Leu	Arg	Ala	Ala	Ile	Lys	Pro	Ala	Phe		
		225			230					235					240			
	atc	ccc	gga	tcc	ggc	acg	gtc	aca	gcc	ccg	aat	tcc	tca	cct	ctt	aac	768	
Ile	Pro	Gly	Ser	Gly	Thr	Val	Thr	Ala	Ala	Pro	Asn	Ser	Ser	Pro	Leu	Asn		
				245						250					255			
	gac	ggt	gct	gct	gct	gtt	atc	ctc	gtc	tca	gaa	gct	aaa	ctg	aaa	gag	816	
Asp	Gly	Ala	Ala	Ala	Val	Ile	Leu	Val	Val	Ser	Glu	Ala	Lys	Leu	Lys	Glu		
			260					265						270				
	ctt	aac	cta	aag	cct	gtt	gca	aag	att	ctt	ggc	tgg	gga	gat	gcc	gcc	864	
Leu	Asn	Leu	Lys	Pro	Val	Ala	Lys	Lys	Ile	Leu	Gly	Trp	Gly	Asp	Ala	Ala		
		275				280							285					
	cag	cag	cca	agc	aaa	ttc	aca	act	gcc	cca	gct	cta	gca	att	ccc	aag	912	
Gln	Gln	Pro	Ser	Lys	Phe	Thr	Thr	Ala	Ala	Pro	Ala	Leu	Ala	Ile	Pro	Lys		
		290				295						300						
	gcc	ctc	agc	cat	gca	ggt	gtg	gct	cag	gat	gct	gtt	gat	gcg	ttc	gag	960	
Ala	Leu	Ser	His	Ala	Gly	Val	Ala	Gln	Asp	Ala	Ala	Val	Asp	Ala	Phe	Glu		
		305			310				315						320			
	att	aac	gaa	gcg	ttc	agc	gta	gtt	gct	ctg	gcc	aat	atg	aaa	ctc	ctg	1008	
Ile	Asn	Glu	Ala	Phe	Ser	Val	Val	Ala	Ala	Leu	Ala	Asn	Met	Lys	Leu	Leu		
				325					330						335			
	ggg	ttg	gct	gaa	gat	aaa	gtc	aac	atc	cat	ggt	ggt	gca	gtg	gct	atc	1056	
Gly	Leu	Ala	Glu	Asp	Lys	Val	Val	Asn	Ile	His	Gly	Gly	Ala	Val	Ala	Ile		
			340					345						350				
	ggt	cat	cct	atc	ggc	gcc	agc	ggt	gct	cgt	atc	ttg	act	aca	ttg	ctc	1104	
Gly	His	Pro	Ile	Gly	Ala	Ser	Gly	Gly	Ala	Arg	Ile	Leu	Thr	Thr	Leu	Leu		
		355				360							365					
	ggt	gta	ttg	aaa	gcg	aga	aag	ggt	aag	att	ggt	tgt	gcc	ggg	att	tgt	1152	
Gly	Val	Leu	Lys	Ala	Arg	Lys	Gly	Gly	Lys	Ile	Gly	Cys	Ala	Gly	Ile	Cys		
		370				375						380						
	aat	gga	gga	ggt	ggt	gct	agc	gct	att	gtt	gtc	gaa	tct	ctc	gtc	tga	1200	
Asn	Gly	Gly	Gly	Gly	Ala	Ser	Ala	Ile	Ile	Val	Val	Glu	Ser	Leu	Val			

[0530]

199

[0532]	Leu 65	Ile	Asp	Glu	Cys	Lys 70	Pro	Asp	Leu	Thr	Lys 75	Ser	Phe	Glu	Ser	His 80	
	gct	ggt	att	tct	cac	act	cgt	tgg	gcc	act	cat	ggg	acg	cct	tcc	cgc	288
	Ala	Gly	Ile	Ser	His 85	Thr	Arg	Trp	Ala	Thr 90	His	Gly	Thr	Pro	Ser 95	Arg	
	ttg	aac	tgc	cac	ccc	cac	agg	tcc	gat	ctg	aac	tgg	gaa	ttt	tca	ggt	336
	Leu	Asn	Cys	His	Pro	His	Arg	Ser	Asp 105	Leu	Asn	Trp	Glu	Phe 110	Ser	Val	
	gtc	cac	aat	ggt	atc	att	acc	aac	tac	aaa	gag	ctc	aag	gct	ttg	ttg	384
	Val	His	Asn 115	Gly	Ile	Ile	Thr	Asn	Tyr 120	Lys	Glu	Leu	Lys 125	Ala	Leu	Leu	
	gag	agc	aag	gga	ttc	cgc	ttt	gag	act	gac	aca	gac	acc	gaa	tgt	att	432
	Glu	Ser 130	Lys	Gly	Phe	Arg	Phe	Glu	Thr	Asp	Thr	Asp	Thr	Glu	Cys	Ile	
	gca	aag	ctt	acc	aaa	tat	ctc	tat	gac	caa	cag	ccc	gat	atc	gaa	ttc	480
	Ala	Lys 145	Leu	Thr	Lys	Tyr 150	Leu	Tyr	Asp	Gln	Gln	Pro	Asp	Ile	Glu	Phe 160	
	aca	gtt	ttg	gct	aag	gct	gtt	gtg	aag	gag	ctt	gag	ggc	gcg	ttt	ggg	528
	Thr	Val	Leu	Ala	Lys 165	Ala	Val	Val	Lys	Glu 170	Leu	Glu	Gly	Ala	Phe 175	Gly	
	ctt	ctc	atc	aaa	tcc	gta	cat	tat	ccc	cat	gag	gtc	att	gcg	gct	cgc	576
	Leu	Leu	Ile	Lys 180	Ser	Val	His	Tyr	Pro 185	His	Glu	Val	Ile	Ala 190	Ala	Arg	
	aag	ggg	tcg	ccg	ctt	gtt	att	ggc	gtg	aga	acg	tcg	aag	aag	atg	aag	624
	Lys	Gly 195	Ser	Pro	Leu	Val	Ile	Gly	Val 200	Arg	Thr	Ser	Lys 205	Lys	Met	Lys	
	gtg	gat	ttc	gtg	gac	gtc	gaa	tac	tca	gaa	gat	gga	gcc	ctt	cct	gca	672
	Val	Asp 210	Phe	Val	Asp	Val	Glu	Tyr	Ser	Glu	Asp	Gly	Ala	Leu	Pro	Ala	
	gag	cag	gcc	tcc	cag	aac	gtg	gcc	atc	aag	aag	tct	gcg	acc	ggc	ctc	720
	Glu	Gln 225	Ala	Ser	Gln	Asn 230	Val	Ala	Ile	Lys	Lys	Ser	Ala	Thr	Gly	Leu 240	
	ctt	gcc	cca	cct	gac	aaa	tcc	ctt	ttg	cat	agg	tca	cag	tct	cgc	gct	768
	Leu	Ala	Pro	Pro	Asp 245	Lys	Ser	Leu	Leu	His 250	Arg	Ser	Gln	Ser	Arg	Ala	
	ttc	ctt	tcc	gat	ggc	atc	ccc	cag	cca	gct	gag	ttt	ttc	ttg	tct		816
	Phe	Leu	Ser	Asp 260	Asp	Gly	Ile	Pro	Gln 265	Pro	Ala	Glu	Phe	Phe 270	Leu	Ser	
	tct	gac	cca	tct	gca	att	gtc	gag	cat	acg	aag	aag	gtc	ctc	tat	ctc	864
	Ser	Asp 275	Pro	Ser	Ala	Ile	Val	Glu	His 280	Thr	Lys	Lys	Val 285	Leu	Tyr	Leu	
	gaa	gat	gat	gat	atc	gct	cat	gtg	cat	gag	gga	cag	tta	aac	att	cat	912
	Glu	Asp 290	Asp	Asp	Ile	Ala	His 295	Val	His	Glu	Gly	Gln	Leu	Asn	Ile	His	
	cgc	ctc	aca	aaa	gat	gat	ggc	act	tct	aat	gtt	cgc	gct	att	cag	aca	960
	Arg	Leu 305	Thr	Lys	Asp	Asp 310	Gly	Thr	Ser	Asn	Val	Arg	Ala	Ile	Gln	Thr 320	
	att	gaa	ctc	gaa	ttg	cag	gag	att	atg	aag	ggc	aag	ttt	gac	cac	ttc	1008
	Ile	Glu	Leu	Glu	Leu 325	Gln	Glu	Ile	Met	Lys 330	Gly	Lys	Phe	Asp	His 335	Phe	
	atg	caa	aag	gaa	att	ttc	gag	caa	ccc	gag	tct	gtg	atc	aat	acc	atg	1056
	Met	Gln	Lys	Glu	Ile	Phe	Glu	Gln	Pro 345	Glu	Ser	Val	Ile	Asn 350	Thr	Met	
	aga	gga	cga	ctg	gat	gtt	gca	aac	aag	caa	gtc	aca	ctc	ggt	ggc	ctg	1104
	Arg	Gly 355	Arg	Leu	Asp	Val	Ala	Asn 360	Lys	Gln	Val	Thr	Leu	Gly 365	Gly	Leu	
	cgg	cag	tac	att	tct	act	atc	cgc	cgc	tgc	aga	aga	atc	ata	ttt	gtt	1152
	Arg	Gln 370	Tyr	Ile	Ser	Thr	Ile	Arg	Arg	Cys	Arg	Arg	Ile	Ile	Phe	Val	
	gcc	tgt	gga	act	agt	tac	cat	tca	tgc	atg	gct	gtg	cgt	gga	gtc	ttt	1200
	Ala	Cys 385	Gly	Thr	Ser	Tyr 390	His	Ser	Cys	Met	Ala	Val	Arg	Gly	Val	Phe 400	
	gaa	gag	ctt	act	gag	att	cct	att	tct	gtt	gaa	ctt	gct	tcc	gat	ttc	1248
	Glu	Glu	Leu	Thr	Glu 405	Ile	Pro	Ile	Ser	Val	Glu	Leu	Ala	Ser	Asp	Phe	
	ctg	gat	aga	cag	gct	cca	gta	ttc	cgt	gat	gac	acg	tgt	gtt	ttc	gtt	1296
	Leu	Asp	Arg	Gln	Ala	Pro	Val	Phe	Arg 425	Asp	Asp	Thr	Cys	Val	Phe	Val	
	tct	cag	tct	ggc	gag	acc	gct	gac	tcc	ctt	atg	gcc	ttg	cgc	tac	tgt	1344
	Ser	Gln 435	Ser	Gly	Glu	Thr	Ala	Asp 440	Ser	Leu	Met	Ala	Leu	Arg	Tyr	Cys	
	ctc	gaa	cgc	gga	gca	ttg	acc	gtt	ggc	att	gtt	aat	gtt	gta	ggg	tcc	1392
	Leu	Glu 450	Arg	Gly	Ala	Leu	Thr	Val	Gly 455	Ile	Val	Asn	Val	Val	Gly	Ser	
	tct	atc	tca	ctt	ctc	act	cac	tgc	ggc	gtg	cac	atc	aac	gct	gga	ccc	1440
	Ser	Ile 465	Ser	Leu	Leu	Thr 470	His	Cys	Gly	Val	His 475	Ile	Asn	Ala	Gly	Pro 480	
	gaa	att	ggg	gta	gct	tct	acc	aag	gcc	tac	acg	tcc	cag	ttc	gtt	gct	1488

201

[0534]

Phe Leu Ser Asp Asp Gly Ile Pro Gln Pro Ala Glu Phe Phe Leu Ser
 260 265 270
 Ser Asp Pro Ser Ala Ile Val Glu His Thr Lys Lys Val Leu Tyr Leu
 275 280 285
 Glu Asp Asp Asp Ile Ala His Val His Glu Gly Gln Leu Asn Ile His
 290 295 300
 Arg Leu Thr Lys Asp Asp Gly Thr Ser Asn Val Arg Ala Ile Gln Thr
 305 310 315
 Ile Glu Leu Glu Leu Gln Glu Ile Met Lys Gly Lys Phe Asp His Phe
 320 325 330 335
 Met Gln Lys Glu Ile Phe Glu Gln Pro Glu Ser Val Ile Asn Thr Met
 340 345 350
 Arg Gly Arg Leu Asp Val Ala Asn Lys Gln Val Thr Leu Gly Gly Leu
 355 360 365
 Arg Gln Tyr Ile Ser Thr Ile Arg Arg Cys Arg Arg Ile Ile Phe Val
 370 375 380
 Ala Cys Gly Thr Ser Tyr His Ser Cys Met Ala Val Arg Gly Val Phe
 385 390 395 400
 Glu Glu Leu Thr Glu Ile Pro Ile Ser Val Glu Leu Ala Ser Asp Phe
 405 410 415
 Leu Asp Arg Gln Ala Pro Val Phe Arg Asp Asp Thr Cys Val Phe Val
 420 425 430
 Ser Gln Ser Gly Glu Thr Ala Asp Ser Leu Met Ala Leu Arg Tyr Cys
 435 440 445
 Leu Glu Arg Gly Ala Leu Thr Val Gly Ile Val Asn Val Val Gly Ser
 450 455 460
 Ser Ile Ser Leu Leu Thr His Cys Gly Val His Ile Asn Ala Gly Pro
 465 470 475 480
 Glu Ile Gly Val Ala Ser Thr Lys Ala Tyr Thr Ser Gln Phe Val Ala
 485 490 495
 Met Val Met Phe Ala Leu Ser Leu Ser Glu Asp Arg Ala Ser Lys Gln
 500 505 510
 Lys Arg Arg Glu Glu Ile Met Glu Gly Leu Ser Lys Val Ser Glu Gln
 515 520 525
 Phe Lys Glu Ile Leu Lys Leu Asn Glu Pro Ile Lys Gln Leu Cys Ala
 530 535 540
 Asn Phe Lys Asn Gln Lys Ser Leu Leu Leu Leu Gly Arg Gly Gly Gln
 545 550 555 560
 Phe Pro Thr Ala Leu Glu Gly Ala Leu Lys Ile Lys Glu Ile Ser Tyr
 565 570 575
 Leu His Cys Glu Ala Val Met Ser Gly Glu Leu Lys His Gly Val Leu
 580 585 590
 Ala Leu Val Asp Glu Asn Leu Pro Ile Ile Met Ile Leu Thr Arg Asp
 595 600 605
 Asn Leu Phe Thr Lys Ser Leu Asn Ala Tyr Gln Gln Val Ile Ala Arg
 610 615 620
 Gly Gly Arg Pro Ile Val Ile Cys Asn Ser Asp Asp Pro Glu Phe Ser
 625 630 635 640
 Ser Ala Gln Thr Val Lys Ile Glu Val Pro Lys Thr Val Asp Cys Leu
 645 650 655
 Gln Gly Leu Leu Asn Val Ile Pro Leu Gln Leu Ile Ala Tyr Trp Leu
 660 665 670
 Ala Val Ser Glu Gly Leu Asn Val Asp Phe Pro Arg Asn Leu Ala Lys
 675 680 685
 Ser Val Thr Val Glu
 690

<210> 244

<211> 1347

<212> DNA

<213> 米曲霉

<220>

<221> CDS

<222> (1)..(1347)

<223> TUB1 曲霉属必需基因 (mRNA)

<400> 244

atg	aga	gaa	ggt	att	agt	ttg	aac	gtt	ggt	cag	gct	ggt	tgc	cag	atc	48
Met	Arg	Glu	Val	Ile	Ser	Leu	Asn	Val	Gly	Gln	Ala	Gly	Cys	Gln	Ile	
1			5					10					15			
gcc	aat	tct	tgc	tgg	gag	ctc	tac	tgt	ctt	gag	cac	ggc	atc	cag	ccc	96
Ala	Asn	Ser	Cys	Trp	Glu	Leu	Tyr	Cys	Leu	Glu	His	Gly	Ile	Gln	Pro	
			20					25					30			
gat	ggt	tac	ttg	acc	gag	gaa	cgc	aag	aag	gaa	gac	cct	gac	cat	ggt	144
Asp	Gly	Tyr	Leu	Thr	Glu	Glu	Arg	Lys	Lys	Glu	Asp	Pro	Asp	His	Gly	
			35				40					45				
ttc	agc	acc	ttc	ttc	tcc	gaa	act	ggc	cag	ggc	aag	tat	gtt	cct	cgt	192
Phe	Ser	Thr	Phe	Phe	Ser	Glu	Thr	Gly	Gln	Gly	Lys	Tyr	Val	Pro	Arg	
			50				55					60				
acc	atc	tac	gcc	gat	ctg	gag	ccc	aat	gtt	gtc	gat	gag	gtc	cgc	act	240

	Thr 65	Ile	Tyr	Ala	Asp	Leu 70	Glu	Pro	Asn	Val	Val 75	Asp	Glu	Val	Arg	Thr 80	
	ggc	acc	tac	cgt	acc	ctt	ttc	cac	ccc	gag	aac	atg	atc	acc	ggc	aag	288
	Gly	Thr	Tyr	Arg	Thr 85	Leu	Phe	His	Pro	Glu	Asn	Met	Ile	Thr	Gly 95	Lys	
	gag	gat	gcc	tgc	aac	aac	tat	gcc	cgt	ggt	cac	tac	acc	gtt	ggc	aag	336
	Glu	Asp	Ala	Ser	Asn	Asn	Tyr	Ala	Arg 105	Gly	His	Tyr	Thr	Val 110	Gly	Lys	
	gag	atg	atc	gac	cag	gtc	ctc	gac	aag	gtt	cgc	cgt	gtg	gcc	gac	aac	384
	Glu	Met	Ile	Asp	Gln	Val	Leu	Asp	Lys 120	Val	Arg	Arg	Val	Ala 125	Asp	Asn	
	tgc	gct	ggg	ctc	cag	ggc	ttc	ctc	gtc	ttc	cac	tct	ttc	ggg	ggg	ggg	432
	Cys	Ala	Gly	Leu	Gln	Gly	Phe	Leu	Val	Phe	His	Ser	Phe	Gly	Gly	Gly	
	act	ggg	tcc	ggg	ttc	ggg	gct	ctc	ctg	atg	gag	cgt	ctg	tct	gtg	gac	480
	Thr	Gly	Ser	Gly	Phe	Gly 150	Ala	Leu	Leu	Met	Glu	Arg	Leu	Ser	Val	Asp 160	
	tac	ggc	aag	aag	tcc	aag	ctg	gag	ttc	tgc	gtc	tac	cct	gcc	ccc	cag	528
	Tyr	Gly	Lys	Lys	Ser 165	Lys	Leu	Glu	Phe	Cys 170	Val	Tyr	Pro	Ala 175	Pro	Gln	
	aat	gcc	acc	tcc	gtc	gtt	gag	ccc	tac	aac	tcc	atc	ctg	act	acc	cac	576
	Asn	Ala	Thr	Ser	Val 180	Val	Glu	Pro	Tyr 185	Asn	Ser	Ile	Leu	Thr 190	Thr	His	
	acc	acc	ctt	gag	cac	tcc	gac	tgc	agt	ttc	atg	gtt	gac	aac	gaa	gcc	624
	Thr	Thr	Leu	Glu	His	Ser	Asp	Cys 200	Ser	Phe	Met	Val	Asp 205	Asn	Glu	Ala	
	atc	tac	gac	atc	tgc	cgc	cgc	aac	ctt	ggc	atc	gag	cgt	ccc	agc	tat	672
	Ile	Tyr	Asp	Ile	Cys	Arg	Arg	Asn	Leu	Gly	Ile	Glu	Arg	Pro	Ser	Tyr	
	gag	aac	ctg	aac	cgc	ctg	att	gct	cag	gtt	gtc	tcc	tcc	atc	acc	gcc	720
	Glu	Asn	Leu	Asn	Arg	Leu	Ile	Ala	Gln	Val	Val	Ser	Ser	Ile	Thr	Ala 240	
	tcc	ctg	cgt	ttc	gat	ggg	tcc	ctg	aac	gtg	gat	ctc	aac	gag	ttc	cag	768
	Ser	Leu	Arg	Phe	Asp 245	Gly	Ser	Leu	Asn	Val 250	Asp	Leu	Asn	Glu	Phe 255	Gln	
	acc	aac	ctg	gtt	ccc	tac	ccc	cgt	att	cac	tcc	cct	ctc	gtt	gcc	tat	816
	Thr	Asn	Leu	Val	Pro 260	Tyr	Pro	Arg	Ile 265	His	Phe	Pro	Leu	Val 270	Ala	Tyr	
[0535]	gct	ccc	gtc	atc	tcc	gcg	gcc	aag	gcc	tcc	cac	gag	gcc	aac	tcc	gtc	864
	Ala	Pro	Val 275	Ile	Ser	Ala	Ala	Lys 280	Ala	Ser	His	Glu	Ala 285	Asn	Ser	Val	
	aac	gag	atc	acc	atg	tct	tgc	ttc	gag	ccc	aac	aac	cag	atg	gtc	aag	912
	Asn	Glu	Ile	Thr	Met	Ser	Cys 295	Phe	Glu	Pro	Asn	Asn	Gln	Met	Val	Lys	
	tgt	gac	ccc	cgc	aat	ggc	aag	tac	atg	gct	act	tgc	ttg	ctg	tac	cgt	960
	Cys	Asp	Pro	Arg	Asn	Gly 310	Lys	Tyr	Met	Ala	Thr	Cys	Leu	Leu	Tyr	Arg 320	
	ggg	gat	gtt	gtg	ccc	aag	gag	acc	cac	gcc	gcc	gtt	gct	acc	ctc	aag	1008
	Gly	Asp	Val	Val	Pro 325	Lys	Glu	Thr	His	Ala 330	Ala	Val	Ala	Thr	Leu 335	Lys	
	acc	aag	cgt	acc	atc	cag	ttc	gtc	gac	tgg	tgc	cct	act	ggg	ttc	aag	1056
	Thr	Lys	Arg	Thr 340	Ile	Gln	Phe	Val	Asp 345	Trp	Cys	Pro	Thr	Gly 350	Phe	Lys	
	atc	ggg	atc	tgc	tac	cag	ccc	cct	aag	cag	gtt	ccc	aac	ggg	gac	ctt	1104
	Ile	Gly	Ile 355	Cys	Tyr	Gln	Pro	Pro 360	Lys	Gln	Val	Pro	Asn 365	Gly	Asp	Leu	
	gcc	aac	ctc	agc	cgt	gct	gtc	tgc	atg	ctg	tct	aac	acc	acc	gcc	atc	1152
	Ala	Asn	Leu	Ser	Arg	Ala	Val 375	Cys	Met	Leu	Ser	Asn	Thr	Thr	Ala	Ile	
	tcc	gag	gcc	tgg	tcc	gct	ctc	gac	cac	aag	ttc	gat	ctc	atg	tac	tcc	1200
	Ser	Glu	Ala	Trp	Ser	Ala 390	Leu	Asp	His	Lys	Phe	Asp	Leu	Met	Tyr	Ser 400	
	aag	cgt	gct	ttc	gtt	cac	tgg	tat	gtt	gga	gag	ggg	atg	gag	gag	ggg	1248
	Lys	Arg	Ala	Phe	Val 405	His	Trp	Tyr	Val	Gly 410	Glu	Gly	Met	Glu	Glu	Gly	
	gaa	ttc	tcc	gag	gcc	cgt	gag	gac	ctg	gct	gcc	ctc	gag	cgc	gac	tac	1296
	Glu	Phe	Ser	Glu	Ala	Arg	Glu	Asp	Leu 425	Ala	Ala	Leu	Glu	Arg	Asp	Tyr	
	gag	gag	gtc	gcc	agc	gac	tgc	ctg	gag	gag	ggt	gag	gct	gag	tac		1344
	Glu	Glu	Val 435	Ala	Ser	Asp	Ser	Leu	Glu 440	Glu	Glu	Val	Glu	Ala	Glu	Tyr	
	taa																1347
	<210>	245															
	<211>	448															
	<212>	PRT															
	<213>	米曲霉															
	<400>	245															
	Met	Arg	Glu	Val	Ile	Ser	Leu	Asn	Val	Gly	Gln	Ala	Gly	Cys	Gln	Ile	

[0536]

```

1      5      10      15
Ala Asn Ser Cys Trp Glu Leu Tyr Cys Leu Glu His Gly Ile Gln Pro
20      25      30
Asp Gly Tyr Leu Thr Glu Glu Arg Lys Lys Glu Asp Pro Asp His Gly
35      40      45
Phe Ser Thr Phe Phe Ser Glu Thr Gly Gln Gly Lys Tyr Val Pro Arg
50      55      60
Thr Ile Tyr Ala Asp Leu Glu Pro Asn Val Val Asp Glu Val Arg Thr
65      70      75      80
Gly Thr Tyr Arg Thr Leu Phe His Pro Glu Asn Met Ile Thr Gly Lys
85      90      95
Glu Asp Ala Ser Asn Asn Tyr Ala Arg Gly His Tyr Thr Val Gly Lys
100      105      110
Glu Met Ile Asp Gln Val Leu Asp Lys Val Arg Arg Val Ala Asp Asn
115      120      125
Cys Ala Gly Leu Gln Gly Phe Leu Val Phe His Ser Phe Gly Gly Gly
130      135      140
Thr Gly Ser Gly Phe Gly Ala Leu Leu Met Glu Arg Leu Ser Val Asp
145      150      155      160
Tyr Gly Lys Lys Ser Lys Leu Glu Phe Cys Val Tyr Pro Ala Pro Gln
165      170      175
Asn Ala Thr Ser Val Val Glu Pro Tyr Asn Ser Ile Leu Thr Thr His
180      185      190
Thr Thr Leu Glu His Ser Asp Cys Ser Phe Met Val Asp Asn Glu Ala
195      200      205
Ile Tyr Asp Ile Cys Arg Arg Asn Leu Gly Ile Glu Arg Pro Ser Tyr
210      215      220
Glu Asn Leu Asn Arg Leu Ile Ala Gln Val Val Ser Ser Ile Thr Ala
225      230      235
Ser Leu Arg Phe Asp Gly Ser Leu Asn Val Asp Leu Asn Glu Phe Gln
245      250      255
Thr Asn Leu Val Pro Tyr Pro Arg Ile His Phe Pro Leu Val Ala Tyr
260      265      270
Ala Pro Val Ile Ser Ala Ala Lys Ala Ser His Glu Ala Asn Ser Val
275      280      285
Asn Glu Ile Thr Met Ser Cys Phe Glu Pro Asn Asn Gln Met Val Lys
290      295      300
Cys Asp Pro Arg Asn Gly Lys Tyr Met Ala Thr Cys Leu Leu Tyr Arg
305      310      315
Gly Asp Val Val Pro Lys Glu Thr His Ala Ala Val Ala Thr Leu Lys
320      325      330
Thr Lys Arg Thr Ile Gln Phe Val Asp Trp Cys Pro Thr Gly Phe Lys
335      340      345
Ile Gly Ile Cys Tyr Gln Pro Pro Lys Gln Val Pro Asn Gly Asp Leu
350      355      360
Ala Asn Leu Ser Arg Ala Val Cys Met Leu Ser Asn Thr Thr Ala Ile
365      370      375
Ser Glu Ala Trp Ser Ala Leu Asp His Lys Phe Asp Leu Met Tyr Ser
380      385      390
Lys Arg Ala Phe Val His Trp Tyr Val Gly Glu Gly Met Glu Glu Gly
395      400      405
Glu Phe Ser Glu Ala Arg Glu Asp Leu Ala Ala Leu Glu Arg Asp Tyr
410      415      420
Glu Glu Val Ala Ser Asp Ser Leu Glu Glu Glu Val Glu Ala Glu Tyr
425      430      435
<210> 246
<211> 5509
<212> DNA
<213> 合成
<220>
<221> 基因
<222> (1970)..(2868)
<223> pyrG
<220>
<221> 启动子
<222> (3260)..(4009)
<223> agsA
<220>
<221> 启动子
<222> (4004)..(4009)
<223> TIS
<220>
<221> 基因
<222> (4010)..(5474)
<223> 必需基因 ERG10
<220>
<221> 外显子

```

[0537]

```

<222> (4010) .. (4074)
<220>
<221> 外显子
<222> (4142) .. (4191)
<220>
<221> 外显子
<222> (4249) .. (4324)
<220>
<221> 外显子
<222> (4398) .. (4887)
<220>
<221> 外显子
<222> (4956) .. (5474)
<400> 246
aacgcgacgt tggccttttct tgggcttagt cgcatttttt cgccttgctc aggcgaagaca 60
ctttcaatga cagtgttgac ttcttcaatc acatgcttga gcttgaaata ggcgtcgttg 120
ggtggttatct tcagtttcgag gatcagccga tccattttgt tcacaacccat ggtcaacgga 180
aggccctcta gcacagcgtg cttgatgatt tgctccgtat tcgattgtac tccttccaca 240
atatcaacga ctatcaccac gccatcaacc agtcgtgatg aggcagctac ttcatcaaca 300
aaattcacat gccccgggtg atcaagaatg ttgaacaaat gagacttccc cttagtcccc 360
tgaagaacta gggtcatagg agctgattta atagaaaggc ctgcctctct ctctagaaaag 420
tgaatatcgg tataacgaag ctgtacttcc ttctttttac ctgctcgtct ctctagccgt 480
tcagtcaaatt catgctgttg cataacaagg atatccatga acgcggtctt cccatgatgt 540
aaatgaccca caagcgcaat gtctcttatt tgttcaggaa aattcagaag atcagtcagt 600
aattcccttg aaaagtgaac cggggggagc tcagtctctt caatggcgaa cttcttttgc 660
tgtacggggg caacgatagg ctcggaagag ggctgtgcgt cttcttcttg aactaatgtc 720
tctacctctg tcccgtagac ctggtgtgct ctagggtagt attgtttgtc ttctgttaga 780
atgacggcat ttgagggggc ttctgtaatt agatatatca gcaccgacta aatccagggg 840
cttttaagaga gtgccaccgc tccacctcca tagactgctg cccatataat tcttcatctt 900
cttcctcccc aaacgcctcg ttgaaggcga aagcccagag cttcacttct tctgtgttggc 960
cgtcttctac tgaatctgca gcctctccaa tgtagtgtgc aaacctgtgt ggtcttgatt 1020
agaattttga caattagagc gttagattgg atgcaactta ctcatcgtag agatcatcca 1080
tcgtgcggtg ggaaagggtg ttcaagctgc caaaaacttg tcaccagtga gctgtcgcct 1140
ctctttttgt cccaagctgc tgaagtcctg ctaatggact caccgagtgg tccgtgcgtc 1200
gaccttacca ttttggtagc cccgcgcgcg gtagtgcacg tcaggagggg gtctttaat 1260
gactttgatt tctcaccacc tcttctctgt tcagagatac ctactgcta cgtctggctg 1320
aataactcaa ctctttcttg ttcccttctt tccccctttt aatcccacca atcttcttta 1380
ttaccggcaa gttctgtcta caagggtgct ctttaccaca caccgaacct ctttaaacgc 1440
ttcacgttac agaagctttt ggtttacgac tcttctaate agctatcgta ctttaaacgc 1500
agctgattaa atattctaga cccaagccgc tgctggaatt gacattatta tggccgatag 1560
ggttgggctt attgctatgt cctgaaagg atatcaaaa gaggcaaaaa gccaggcata 1620
atccccgctt ggacggtacc ctaaggatag gccctaatt tatctacatg tgactgcctc 1680
gatgtgtttg gtcaaaatga ggcattgtgg tcaccccaca ggcggagaaa cgtgtggcta 1740
gtgcatgaca gtccccctca tagattcaat ttaatttttc gcggaatttg tctgtcagtt 1800
tgtatctaca tttcattcca tatatcaaga gtagtagttt ggacatcctg attattttgt 1860
ctaattactg aaaactcgaa gtactaacct actaataagc cagtttcaac cactaagtgc 1920
tcattttatac aatatttgca gaaccccgcg ctacccctcc atcgccaaca tgtcttccaa 1980
ctcgcaattg acctacagcg cacgcgctag caagcaccce aatgcgctcg taaagaagct 2040
cctcgagggt gccgaggcca agaaaaacca tgtaaccggt tccgcccagc tagaacccac 2100
caaagagctg ctggatttgg ctgaccgtat ggcacccggg gatgccactt acatgtgatc 2160
tagtaattgt taatgggtga ttatataaca ggactcgggt cgtacattgc cgtgatcaaa 2220
actcacatcg atatctctc cgatttcoagc gaagaaaacca tcaccggtct tcaccgctct 2280
gcagagaagg acaatttctc catcttcgaa gatcgcaagt tcacgatata cggaaacaca 2340
gtccaaaagg agtaccatgg cggcactctg cgtatctctg agtgggcca catcatcaac 2400
tgcagtattc tgcccgtgta gggatatcgt gaggctctgg cccagactgc ttccggcgag 2460
gacttccccct acggctccga gaggggcctt ttgatccttg cggagatgac ctccaaggga 2520
tctttggcta ccggtcaata tactacttct tctgttgact atgctcgga gtataagaag 2580
tttgtgatgg gattcgtctc gacacgtcac cttggcgagg ttcatctga agttagctcg 2640
ccttcggagg aggaagattt tctcgtcttc acgacaggtg tcaacctctc ctggaagggt 2700
gacaagctgg gacagcagta ccaaaactct gagtcggctg ttggacggcg tgcgactttt 2760
attattgctg gccgtggaat ttatgctgct cctgatcccc tggaggcggc gaagcagtag 2820
cagaaggagg gatgggatgc atacctgaag cgtgttggtg cgcaataagt agtgggtgat 2880
acgtactcct tttatggcag tatgtcgcaa gtagatgcg atttataaat tcagcactcg 2940
aaatgactac tactatgtgt ctacgacaga taccctctcc gtacgaataa gacacctgcc 3000
tcgatatatg gacaaattca aaatcagggt caagggtcat gtttcaaagt cacaacaatc 3060
tccaacatag acgagaattt gtaccggagt gtctgaagg gtctgaagg gcagctggag atttgtctat 3120
tttcttagaa tggggatatc ctaatgtaca gtcgcactct atcgtaaaaa caatcacaat 3180
tatatacaag atttcccacc accccctact ctaaacggcg acaattatcc atcgactcag 3240
agcctagcca ccatttggtg ttgtttctta atataggcac acggatcccc ccatcgatcg 3300
gcgcccgtgt cgagtatggc catctcaca cattgtcggg ctccgctggc gctgcctcaa 3360
catattttta aatggggaaa ccataagcaa agagactcca tcgggcgcac tacacatctc 3420
gcctctctct gacctgtgg atcgggactg ggtgtgatgg caaacgctca agttgcccgc 3480
gcttcttcgc gctgttgctg cgtcacgcca tataggcgac tctcccaatc atggaaagcc 3540
tgcttccag cgatattcaa cgacgtgggt gcttggattg cagcttcgac cagatcacag 3600
caatagaata cggatgattc cagccgtggg cctttagggt agaaaaccata ctgggctagg 3660
aaacatgccg tatcaagtga gaaacaccgc cggctgaaaa gtagatagtg 3720
gatggttcca tgtcatctat cagcaggcga ggagtgggtc tatagtagag gttgtaattg 3780

```

	tcgtggtaga ggtccctggc ccatttttat ccccatcctt cccggcctga tggttccaag	3840
	ccttctttac gaaaagtgtt gcagctgctt ttctgagtcg acaggctcgt cgttccttcg	3900
	tttgatttcc ctgcgttttg accacgcccc gggcatgtag gatttattcc gtttgagatt	3960
	atcgagagcc taagtgcgcg gcagttgtcg gtagcaggcc actgcggg atg tct tct	4018
	Met Ser Ser	
	ctt ccg ccg gtc tat att gtt tcc tct gcc cgc acc cca gtc ggc tct	4066
	Leu Pro Pro Val Tyr Ile Val Ser Ser Ala Arg Thr Pro Val Gly Ser	
	5 10 15	
	ttc ttg gg gtgagattgg ccttgtgcct tgtctatacg tacaatattg	4114
	Phe Leu Gly	
	20	
	acttacactg actatccgct tccacag g tcg ctc tca agt ctc act gcc ccg	4166
	Ser Leu Ser Ser Leu Thr Ala Pro	
	25 30	
	cag tta ggc tct cat gct att aaa g gtatgacagc tctcgagtgg	4211
	Gln Leu Gly Ser His Ala Ile Lys	
	35	
	cggtacgtgt tgatatacta acaaactcga aattcag ct gcg ctc agc aaa gcg	4265
	Ala Ala Leu Ser Lys Ala	
	40	
	gat gga atc aag ccg tct gat atc cag gag gtc ttc ttt ggc aat gtc	4313
	Asp Gly Ile Lys Pro Ser Asp Ile Gln Glu Val Phe Phe Gly Asn Val	
	45 50 55 60	
	atc tcc gca aa gtgagtcaac gtccaccctt ataccttata gccaacatcg	4364
	Ile Ser Ala Asn	
	aaggcacact tacattctcg ttctatggat cag c gtt gga caa aat cct gct	4416
	Val Gly Gln Asn Pro Ala	
	65 70	
	aga cag tgt gct ctc ggc gct ggt ctc aat gaa tca act gtc tgt act	4464
	Arg Gln Cys Ala Leu Gly Ala Gly Leu Asn Glu Ser Thr Val Cys Thr	
	75 80 85	
	acg gtt aat aag gtg tgc gcg tct ggc ttg aaa gcg gtt att ctc ggt	4512
	Thr Val Asn Lys Val Cys Ala Ser Gly Leu Lys Ala Val Ile Leu Gly	
	90 95 100	
	gca cag acc atc atg act ggc aat gcg gat att gtc gta gca ggc ggt	4560
	Ala Gln Thr Ile Met Thr Gly Asn Ala Asp Ile Val Val Ala Gly Gly	
	105 110 115	
[0538]	gct gaa tcc atg tct aac gcc cct cat tac ctt cca aac ctt cgc gtc	4608
	Ala Glu Ser Met Ser Asn Ala Pro His Tyr Leu Pro Asn Leu Arg Val	
	120 125 130	
	ggt gcg aaa tac ggc aac cag agt ctg gtg gac ggt att atg aag gat	4656
	Gly Ala Lys Tyr Gly Asn Gln Ser Leu Val Asp Gly Ile Met Lys Asp	
	135 140 145 150	
	ggc ttg aca gac gca gga aag cag gaa ctc atg ggc ttg caa gcc gag	4704
	Gly Leu Thr Asp Ala Gly Lys Gln Glu Leu Met Gly Leu Gln Ala Glu	
	155 160 165	
	gag tgt gct cag gat cat ggc ttt agc agg gaa caa cag gat gat tat	4752
	Glu Cys Ala Gln Asp His Gly Phe Ser Arg Glu Gln Gln Asp Asp Tyr	
	170 175 180	
	gcc att cgc act tac gaa aaa gca cag gcg gct caa aag gct ggc ctt	4800
	Ala Ile Arg Thr Tyr Glu Lys Ala Gln Ala Ala Gln Lys Ala Gly Leu	
	185 190 195	
	ttt gac gaa gaa att gcg cct att gaa ctt cct ggc ttt agg ggc aag	4848
	Phe Asp Glu Glu Ile Ala Pro Ile Glu Leu Pro Gly Phe Arg Gly Lys	
	200 205 210	
	cca ggt gtg act gtg tca caa gac gaa gaa cca aag aat gtgaggaaga	4897
	Pro Gly Val Thr Val Ser Gln Asp Glu Glu Pro Lys Asn	
	215 220 225	
	tattactttc tacagatttt ttaagtcgtc ggccactgct aatatactca ttgcgtag	4955
	ctt aac ccg gat aag ctt cga gct atc aag cct gca ttt atc ccc gga	5003
	Leu Asn Pro Asp Lys Leu Arg Ala Ile Lys Pro Ala Phe Ile Pro Gly	
	230 235 240	
	tcc ggc acg gtc aca gcc ccg aat tcc tca cct ctt aac gac ggt gct	5051
	Ser Gly Thr Val Thr Ala Pro Asn Ser Ser Pro Leu Asn Asp Gly Ala	
	245 250 255	
	gct gct gtt atc ctc gtc tca gaa gct aaa ctg aaa gag ctt aac cta	5099
	Ala Ala Val Ile Leu Val Ser Glu Ala Lys Leu Lys Glu Leu Asn Leu	
	260 265 270 275	
	aag cct gtt gca aag att ctt ggc tgg gga gat gcc gcc cag cag cca	5147
	Lys Pro Val Ala Lys Ile Leu Gly Trp Gly Asp Ala Ala Gln Gln Pro	
	280 285 290	
	agc aaa ttc aca act gcc cca gct cta gca att ccc aag gcc ctc agc	5195
	Ser Lys Phe Thr Thr Ala Pro Ala Leu Ala Ile Pro Lys Ala Leu Ser	
	295 300 305	
	cat gca ggt gtg gct cag gat gct gtt gat gcg ttc gag att aac gaa	5243
	His Ala Gly Val Ala Gln Asp Ala Val Asp Ala Phe Glu Ile Asn Glu	

	310	315	320	
gcg ttc agc gta gtt gct ctg gcc aat atg aaa ctc ctg ggg ttg gct				5291
Ala Phe Ser Val Val Ala Leu Ala Asn Met Lys Leu Leu Gly Leu Ala				
325	330	335		
gaa gat aaa gtc aac atc cat ggt ggt gca gtg gct atc ggt cat cct				5339
Glu Asp Lys Val Asn Ile His Gly Gly Ala Val Ala Ile Gly His Pro				
340	345	350	355	
atc ggc gcc agc ggt gct cgt atc ttg act aca ttg ctc ggt gta ttg				5387
Ile Gly Ala Ser Gly Ala Arg Ile Leu Thr Thr Leu Leu Gly Val Leu				
360	365	370		
aaa gcg aga aag ggt aag att ggt tgt gcc ggg att tgt aat gga gga				5435
Lys Ala Arg Lys Gly Lys Ile Gly Cys Ala Gly Ile Cys Asn Gly Gly				
375	380	385		
ggt ggt gct agc gct att gtt gtc gaa tct ctc gtc tga gtcattgccga				5484
Gly Gly Ala Ser Ala Ile Val Val Glu Ser Leu Val				
390	395			
gtcgcccttg atcttggtat ggaat				5509
<210> 247				
<211> 5509				
<212> DNA				
<213> 合成				
<220>				
<221> 基因				
<222> (1970)..(2868)				
<223> pyrG				
<220>				
<221> 启动子				
<222> (3260)..(4009)				
<223> agsA 启动子				
<220>				
<221> 启动子				
<222> (4004)..(4009)				
<223> TIS				
<220>				
<221> 基因				
<222> (4010)..(5509)				
<223> 必需基因 FKS2				
<220>				
<221> 外显子				
<222> (4010)..(4014)				
<220>				
<221> 外显子				
<222> (4072)..(4142)				
<220>				
<221> 外显子				
<222> (4229)..(4460)				
<220>				
<221> 外显子				
<222> (4534)..(5509)				
<400> 247				
ataaaaaattt cgccaccgaa gaagtatcaa ggaaactgtc tagcgagagg gactggactg				60
gaacacttga gtattcggta tccagttaag gtgatttaag ccaccccttc tcccttcaat				120
ccctacccttc cttcaaaaag gggaagaaca gtagctaaag gtagttgcca tcaagaactc				180
atacttaaaa acacgaggta agaggccgt aatagtaatt gggtgcaaga gaaaaggggg				240
gttatecttc tgagcatcac agttttgatt acgggaagaa ccagaatatg aacgtggcaa				300
cagcggtgtg cggtagcaga ccaggccttt attggatacc ggaggagtcc ctctttcgtg				360
tgtttcttac ccctttttaca gtacaattac aacctacata atgtcactga tagatattta				420
tacgtgtgccc agaagatact tatgggttag tcaatgtaaa tagcttctaa ttttgattca				480
tgattgaggg gtctaaaatt ggactgaacc taatcctact caagtaggag ttagatgacg				540
atggtgagtg tgaaatgccca ggattaagggt tagcatatca gccgcacagg aactaccgag				600
gggggtggagg ctaacagccc tctttttttc tctccctttt tctcattttt cctctctctt				660
ctcccgatga tgacgatttg tagattcata agttatagaa aaatattttgt tctatgcaca				720
cgctaatagg acatagtagt gcttattgta tgtatatgcc tgggtgtttaa agaagtagac				780
ttcttgagtt ccgtacggag cactatttta caatacgaat aaaaaaaatc catggagcag				840
ggaagggaag gatatagcta gctagggtccc taatagtcga tatgttattt ttgatctgtt				900
agtggagcca ctgacctgtt gaattttttg taggtaaact tagcaacatg tccatgccat				960
aaaagggtct caaccgagag tcacgagtga aacataccag taggtccaac aaactgtgta				1020
agggtgcagaa aaggcacatt attggccagc tccatgttga atgaatggct gcgcccgtga				1080
gccggtcccat agtagcttcc ctgggtctctg tgtgggccga tctcacaaca taaaacatag				1140
ctcttccccc tctttccaac agattgtcga tctactcaat ttcccttttg ctatcaactt				1200
ctccccctcc tcgttctctt ctctcactcg ccacacggcc ctttaagaat atcacaagta				1260
atattctagc tggattatcc aaagtgtgtt ttatgttgtc cttccatggt atcttccaat				1320
tcgtttgttg acagtatatt cccttccaga gtgcactttg aagaaacccc actaccgttt				1380
acgtgacttg ctttcaatat ccggaggcaa catttgttat cgaagcatta caacgcctag				1440
tcattctcatt ctacaccttt ccacattttg actcgtctct tctcttgaca taccttaacc				1500
agctgattaa atattctaga cccaagccgc tgctggaatt gacattatta tggccgatag				1560

[0539]

[0540]

```

ggttgggctt attgctatgt cctgaaaggg atatcaaaag caggcaaaaa gccaggcata 1620
atccccgcgt ggacgggtacc ctaaggatag gccctaactt tatctacatg tgactgcac 1680
gatgtgtttg gtcaaaaatga ggcatgtggc tcaccccaca ggcgagaaaa cgtgtggcta 1740
gtgcattgaca gtcccccctcca tagattcaat ttaatttttc gcggaatttg tegtgcagtt 1800
tgtatctaca tttcattcca tatatcaaga gtttagtagtt ggacatcctg attattttgt 1860
ctaattactg aaaactcgaa gtactaacct actaataaagc cagtttcaac cactaagtgc 1920
tcattttatac aatatttgca gaaccccgcg ctacccctcc atcgccaaca tgtcttccaa 1980
gtcgcaattg acctacagcg cacgcgttag caagcacccc aatgcgctcg taaagaagct 2040
cttcgaggtt gccgaggcca agaaaaccaa tgtcacctgt tccgcccagc tgacaaccac 2100
caaagagctg ctggatttgg ctgaccgtat gcgcaccggg gatgccactt acatgtgatc 2160
tagtaatggt taatggtgga ttatataaca ggactcggtc cgtacattgc cgtgatcaaa 2220
actcacatcg atatcctctc cgatttcagc gaagaaaacca tcaccgggtc gaaggccctt 2280
gcagagaagc acaatttcct catcttcgaa gatcgcaagt tcactgatat cggaaacaca 2340
gtccaaaagc agtaccatgg cggcactctg cgtatctctg agtgggcccc catcatcaac 2400
tgcagtattc tgcccgttga gggatcgtc gaggtctctg cccagactgc ttccggccgag 2460
gacttccctt acgggtccga gaggggcctt ttgatccttg cggagatgac ctccaaggga 2520
tcctttggta cgggtcaata tactacttct tctgttgact atgctcgaa agtataagaag 2580
tttgtgatgg gattcgtctc gacacgtcac cttggcgagg ttcagtctga agttagctcg 2640
ccttcggagg aggaagattt tgtcgtcttc acgacagggt tcaacctctc ctogaagggt 2700
gacaagctgg gacagcagta ccaaactcct gactcggctg ttggacgcgg tgccgacttt 2760
attattgctg gccgtggaat ttatgctgct cctgatcccc tggaggcgcc gaagcagtac 2820
cagaaggagg gatgggatgc atacctgaag cgtgttggtg cgcaataagt agtgggtgat 2880
acgtactcct tttatggcag tatgtcgcaa gtatgatgag atttataaat tcagcactcg 2940
aaatgactac tactatgtgt ctacgacaga taccctctcc gtacgaataa gacgactgcc 3000
tcgatatacg gacaaattca aaatcagggt caagggtcat gtttcaaagt cacaacaatc 3060
tccaacatag acgagaattt gtaccggagt gtctgaaggt gcagctggag attgggtctat 3120
tttcttagag tggggtatca ctaatgtaca gtcggtcact atcgtaaaaa caatcaaat 3180
tatatacaag atttcccacc accccctact ctaacacggc acaattatcc atcgagtcag 3240
agcctagcca ccatttgggtg ttgtttctta atataggcac acggatcccc ccatcgatcg 3300
gcgcccgttt cagtatggc catctcaca cattgtcggg ctccgctggc gctgcctcaa 3360
cataatttta aatggggaaa ccataagcaa agagactcca tcgggcgcac tacacatctc 3420
gcctcttctg gacctgtgg atcgggactg ggggtgatgg caaacctgca agttgccgcc 3480
gcttcttcgc gctgttgccg cgtcacgcca tataggcgac tctcccaatc atggaaagcc 3540
tgctgttcca cgatattcaa cgactgggtg gcttggattg cagcttcgac cagatcacag 3600
caatagaata gcgatgatcc cagccgtggg cctttagggt agaaaccata ctgggctagg 3660
aaacatgccg tatcaagtga gaaacaccgc taccacacca cggctgaaaa gtagatagtg 3720
gatggttcca tgtcatctat cagcaggcga ggaagtgggt tatagtagag gttgtaattg 3780
tcgtggtaga ggtccctggc ccatttttat ccccatcctt cccggcctga tggttccaag 3840
ccttctttac gaaaagtttt gcagctgctt ttctcagtcg acaggtcgtc cgttccttcg 3900
tttgatttcc ctgccttttg accacgcccc gggcatgtag gatttattcc gtttgagatt 3960
atcgagagcc taagtgcgcg gcagttgtcg gtagcaggcc actgcccggg atg tg 4014
Met Cys
1
gtacgttccc gaagaccag gcttactgaa tgtttgatgt aactaattga cccctag c 4072
ggc atc ttc ggc tac att aac tac ctc gtc gag agg gac cgc aag ttt 4120
Gly Ile Phe Gly Tyr Ile Asn Tyr Leu Val Glu Arg Asp Arg Lys Phe
5 10 15
att ctt gac act ctt ctc aat g gttggcattc ctgttgtttt caccctcact 4172
Ile Leu Asp Thr Leu Leu Asn
20 25
ggaccttggt tactctgggt ttgggagtggt ttgtgtattc ctctgacata ttgaag ga 4230
Gly
ctc tcc cgt ctt gaa tac cga ggc tac gac tct gcg ggt ctt gcc ata 4278
Leu Ser Arg Leu Glu Tyr Arg Gly Tyr Asp Ser Ala Gly Leu Ala Ile
30 35 40
gat ggc gat aag aag aat gaa gtt tgt gct ttc aag gaa gtg ggc aag 4326
Asp Gly Asp Lys Lys Asn Glu Val Cys Ala Phe Lys Glu Val Gly Lys
45 50 55
gtt gct aag ctt cga gag ctt att gat gag tgc aag cct gac ttg acc 4374
Val Ala Lys Leu Arg Glu Leu Ile Asp Glu Cys Lys Pro Asp Leu Thr
60 65 70
aaa tcc ttt gaa tct cat gct ggt att tct cac act cgt tgg gcc act 4422
Lys Ser Phe Glu Ser His Ala Gly Ile Ser His Thr Arg Trp Ala Thr
75 80 85 90
cat ggg acg cct tcc cgc ttg aac tgc cac ccc cac ag gtgagtcoga 4470
His Gly Thr Pro Ser Arg Leu Asn Cys His Pro His Arg
95 100
tggtccacgg tttaatcagc tcgactagga tcaggcttgt ttgctaattt tccctggcagc 4530
cag g tcc gat ctg aac tgg gaa ttt tca gtt gtc cac aat ggt atc att 4579
Ser Asp Leu Asn Trp Glu Phe Ser Val Val His Asn Gly Ile Ile
105 110 115
acc aac tac aaa gag ctc aag gct ttg ttg gag agc aag gga ttc cgc 4627
Thr Asn Tyr Lys Glu Leu Lys Ala Leu Leu Glu Ser Lys Gly Phe Arg
120 125 130
ttt gag act gac aca gac acc gaa tgt att gca aag ctt acc aaa tat 4675
Phe Glu Thr Asp Thr Asp Thr Glu Cys Ile Ala Lys Leu Thr Lys Tyr
135 140 145 150

```


	ctc tat gac caa cag ccc gat atc gaa ttc aca gtt ttg gct aag gct	4723
	Leu Tyr Asp Gln Gln Pro Asp Ile Glu Phe Thr Val Leu Ala Lys Ala	
	155 160 165	
	gtt gtg aag gag ctt gag ggc gcg ttt ggg ctt ctc atc aaa tcc gta	4771
	Val Val Lys Glu Leu Glu Gly Ala Phe Gly Leu Leu Ile Lys Ser Val	
	170 175 180	
	cat tat ccc cat gag gtc att gcg gct cgc aag ggg tgc ccg ctt gtt	4819
	His Tyr Pro His Glu Val Ile Ala Ala Arg Lys Gly Ser Pro Leu Val	
	185 190 195	
	att ggc gtg aga acg tgc aag aag atg aag gtg gat ttc gtg gac gtc	4867
	Ile Gly Val Arg Thr Ser Lys Lys Met Lys Val Asp Phe Val Asp Val	
	200 205 210	
	gaa tac tca gaa gat gga gcc ctt cct gca gag cag gcc tcc cag aac	4915
	Glu Tyr Ser Glu Asp Gly Ala Leu Pro Ala Glu Gln Ala Ser Gln Asn	
	215 220 225 230	
	gtg gcc atc aag aag tct gcg acc ggc ctc ctt gcc cca cct gac aaa	4963
	Val Ala Ile Lys Lys Ser Ala Thr Gly Leu Ala Pro Pro Asp Lys	
	235 240 245	
	tcc ctt ttg cat agg tca cag tct cgc gct ttc ctt tcc gat gat ggc	5011
	Ser Leu Leu His Arg Ser Gln Ser Arg Ala Phe Leu Ser Asp Asp Gly	
	250 255 260	
	atc ccc cag cca gct gag ttt ttc ttg tct tct gac cca tct gca att	5059
	Ile Pro Gln Pro Ala Glu Phe Phe Leu Ser Ser Asp Pro Ser Ala Ile	
	265 270 275	
	gtc gag cat acg aag aag gtc ctc tat ctc gaa gat gat gat atc gct	5107
	Val Glu His Thr Lys Lys Val Leu Tyr Leu Glu Asp Asp Asp Ile Ala	
	280 285 290	
	cat gtg cat gag gga cag tta aac att cat cgc ctc aca aaa gat gat	5155
	His Val His Glu Gly Gln Leu Asn Ile His Arg Leu Thr Lys Asp Asp	
	295 300 305 310	
	ggg act tct aat gtt cgc gct att cag aca att gaa ctc gaa ttg cag	5203
	Gly Thr Ser Asn Val Arg Ala Ile Gln Thr Ile Glu Leu Glu Leu Gln	
	315 320 325	
	gag att atg aag ggc aag ttt gac cac ttc atg caa aag gaa att ttc	5251
	Glu Ile Met Lys Gly Lys Phe Asp His Phe Met Gln Lys Glu Ile Phe	
	330 335 340	
	gag caa ccc gag tct gtg atc aat acc atg aga gga cga ctg gat gtt	5299
	Glu Gln Pro Glu Ser Val Ile Asn Thr Met Arg Gly Arg Leu Asp Val	
	345 350 355	
[0541]	gca aac aag caa gtc aca ctc ggt ggc ctg cgg cag tac att tct act	5347
	Ala Asn Lys Gln Val Thr Leu Gly Gly Leu Arg Gln Tyr Ile Ser Thr	
	360 365 370	
	atc cgc cgc tgc aga aga atc ata ttt gtt gcc tgt gga act agt tac	5395
	Ile Arg Arg Cys Arg Arg Ile Ile Phe Val Ala Cys Gly Thr Ser Tyr	
	375 380 385 390	
	cat tca tgc atg gct gtg cgt gga gtc ttt gaa gag ctt act gag att	5443
	His Ser Cys Met Ala Val Arg Gly Val Phe Glu Glu Leu Thr Glu Ile	
	395 400 405	
	cct att tct gtt gaa ctt gct tcc gat ttc ctg gat aga cag gct cca	5491
	Pro Ile Ser Val Glu Leu Ala Ser Asp Phe Leu Asp Arg Gln Ala Pro	
	410 415 420	
	gta ttc cgt gat gac acg	5509
	Val Phe Arg Asp Asp Thr	
	425	
	<210> 248	
	<211> 5509	
	<212> DNA	
	<213> 合成	
	<220>	
	<221> 基因	
	<222> (1970) .. (2868)	
	<223> pyrG	
	<220>	
	<221> 启动子	
	<222> (3260) .. (4009)	
	<223> agsA	
	<220>	
	<221> 启动子	
	<222> (4004) .. (4009)	
	<223> TIS	
	<220>	
	<221> 基因	
	<222> (4010) .. (5509)	
	<223> TUB1 必需基因	
	<220>	
	<221> 外显子	
	<222> (4010) .. (4034)	

<220>									
<221>	外显子								
<222>	(4124) .. (4164)								
<220>									
<221>	外显子								
<222>	(4239) .. (4265)								
<220>									
<221>	外显子								
<222>	(4337) .. (4412)								
<220>									
<221>	外显子								
<222>	(4473) .. (4527)								
<220>									
<221>	外显子								
<222>	(5481) .. (5509)								
<400>	248								
cggggaagatg	ttccatagaa	agggatcaag	gatgatatcc	gctgcaaccg	ggggaaacatt				60
ctccattagt	gctttgttaa	gagaaaaagta	gtccgcgcgc	gcagtaacac	gaaggaaaaag				120
caccgaatca	gacgtaggtg	cttgatcatg	aaaagaggaa	gatttgcgag	gaacagcggt				180
agactgaggg	gaggcgagat	gggcttggga	atagaggcta	atcgaggaga	agagtgcggg				240
atcattcaatg	gcttgtggaa	gagtatgcgt	tgtcagggtg	aaggcggtcg	gttgctttga				300
ctacttctga	gctaattttc	tttggtcgcc	ttgctgggca	ctggtagagg	gtatatgagg				360
acgtacagtt	gccaaccaac	acacccgcac	ctcataccgc	tggccggggg	tcaagttctc				420
cagaaaaaac	catgagtctg	taccagggaa	ctcgttggtc	ggaaatgtag	cgtaattctc				480
ggttcgaaga	acgggactag	aggaggagag	gcgatctaga	cctaggtcat	cgagggtagg				540
atcgaccgtg	gggatgatga	ggggttgcgg	agcgatgaag	atcgttttct	cgacattggc				600
gatagcgggg	aatgcgaggg	tgcagagtaa	gagtagata	tggagtagca	tcatgacaga				660
gagttcgcta	tgttcaattg	acgctgctgc	tagacggtea	tggttctagt	ccacacgcag				720
ttgagtatgg	taggctata	gtatgtatgt	agtaaagtga	gtagatctga	tactaactgt				780
tagattccta	agtgaagtcc	ttacatgaaa	ttgattcaat	gtcaatacga	tcttaccata				840
aattaggcca	ttgtgcgcag	caatgttgg	gactagaaat	agagcactgt	gtgtccgaga				900
cttgttctat	gtgagtattt	tcaatagtag	ttactatcgc	aagctgtacg	gatagtcaaa				960
tacaatgcag	ctgcacgaag	agatgccgtc	attgacacgt	gacgcaccct	tgcctgaac				1020
taagcccaaa	taagtggctg	agtcacgcgt	gggctggaag	ggcggtcacg	gcacgcgaca				1080
agccaatcag	aaatgggatt	gtgggcgacg	cgaccaacg	gatggatcaa	caatgaaaaag				1140
gtggaagtcc	tgttttggg	ttgacagcgc	atcgcaagac	cacaacaaat	tgaccgatca				1200
ttcgggtctt	gggcaattgt	ttaattgate	cccttagtct	atttttctcc	togtcaatca				1260
ccccaaaggt	ttgttctctt	ccccatttc	ctccccacgg	ttcccccttc	cctttctttt				1320
ggttcttttc	ccaggtggtg	tttttttcgt	ttacttaccc	ctgaaaacttc	cttacagatc				1380
agggcatttt	tccccctcta	aagtccttcc	acattccacc	cactccgttt	cacacacccc				1440
cctccttcca	atcttaattc	ctttgtcact	ccgacattcg	tcccttttag	actcgacaag				1500
agctgattaa	atattctaga	cccaagccgc	tgtggaatt	gacattatta	tggccgatag				1560
ggttgggctt	attgctatgt	ccctgaaagg	atatcaaaa	caggcaaaaa	gccaggcata				1620
atccccgcgt	ggacggatcc	ctaaggatag	gccctaatct	tatctacatg	tgactgcac				1680
gatgtgtttg	gtcaaaatga	ggcatgtggc	tcaccacaca	ggcggagaaa	cgtgtggcta				1740
gtgcataaca	tgtcccctcca	tagattcaat	ttaatttttc	gcggaattg	togtgcagtt				1800
tgtatctaca	tttcattcca	tatatcaaga	gttagtagtt	ggacatcctg	attattttgt				1860
ctaattactg	aaaactcgaa	gtactaacct	actaataage	cagtttcaac	cactaagtgc				1920
tcatttatac	aatatttgca	gaaccccgcg	ctaccctcc	atcgccaaca	tgtcttccaa				1980
gtcgcaattg	acctacagcg	cacgcgctag	caagcacccc	aatgcgctcg	taaagaagct				2040
cttcgaggtt	gccgaggcca	agaaaaacaa	tgtcaccggt	tccgccgacg	tgacaaccac				2100
caaagctgg	ctggatttgg	ctgacggtat	gcgcaccggg	gatgccactt	acatgtgatc				2160
tagtaatggt	taatgggtga	ttatataaca	ggactcggtc	cgtacattgc	cgtgatcaaa				2220
actcacatcg	atatcctctc	cgatttcagc	gaagaaacca	tcaccggtct	gaaggccctt				2280
gcagagaagc	acaatttctc	catcttcgaa	gatcgatat	tcacgatata	cggaacacaa				2340
gtccaaaaag	agtaccatgg	cggcactctg	cgtatctctg	agtgggcca	catcatcaac				2400
tgcagtatcc	tgcgccgtga	gggtatcgtc	gaggctctgg	cccagactgc	ttcggccgag				2460
gacttcccct	acgggtccga	gaggggcctt	ttgatccttg	cggagatgac	ctccaaggga				2520
tcttttgcta	ccggtcaata	tactacttct	tctgttgact	atgctcggaa	gtataagaag				2580
tttgtgatgg	gattcgtctc	gacacgtcac	cttggcgagg	ttcagtcctga	agttagctcg				2640
ccttcggagg	aggaagattt	tgtcgtcttc	acgacagggt	tcaacctctc	ctcgaagggt				2700
gacaagctgg	gacagcagta	ccaaaactct	gagtcggctg	ttggacgcgg	tgcgcacttt				2760
attattgctg	gccgtggaat	ttatgctgct	cctgatcccg	tggaggcggc	gaagcagtac				2820
cagaaggagg	gatgggatgc	atacctgaag	cgtgttgggt	cgcaataagt	agtggtggat				2880
acgtactcct	tttatggcag	tatgtcgcaa	gtatgatgcg	atttataaat	tcagcactcg				2940
aaatgactac	tactatgtgt	ctacgacaga	tacctctctc	gtacgaataa	gacacctgcc				3000
tcgatatatg	gacaaaattca	aaatcagggt	caagggtcat	gtttcaaaat	cacaacaatc				3060
tccaacatag	acgagaattt	gtaccggagt	gtctgaagg	gcagctggag	attggtctat				3120
tttcttagag	tggggtatca	ctaattgtaca	gtcggtcact	atcgtacaaa	caatcacaat				3180
tatatacaag	atttcccacc	accccctact	ctaacacggc	acaattatcc	atcgagtcag				3240
agcctagcca	ccatttgggtg	ttgtttctta	atataggcac	acggatcccg	ccatcgatcg				3300
gcgccctgtt	cgagtatggc	catcctcaca	cattgtcggg	ctccgctggc	gctgcctcaa				3360
catattttta	aatggggaaa	ccataagcaa	agagactcca	tggggcgcac	tacacatctc				3420
gcctcttctg	gacctgtgg	atcgggactg	gggtggatgg	caaaccgtca	agttgccgcc				3480
gcttcttcgc	gctgttgcgg	cgtaacgcca	tataggcgac	tctcccaatc	atggaaagcc				3540
tgcgttccag	cgatattcaa	cgacgtgggtg	gcttggattg	cagcttcgac	cagatcacag				3600
caatagaata	gcgatgatcc	cagccgtggg	cctttagggg	agaaaccata	ctgggctagg				3660

[0542]

	aaacatgccg	tatcaagtga	gaaacaccgc	tacccaccca	cggctgaaaa	gtagatagt	3720
	gatgggtcca	tgatcatctat	cagcaggcga	ggagtgggtc	tatagtagag	gttgtaatgg	3780
	tcgtggtaga	ggccccctggc	ccattttttat	ccccatcctt	cccggcctga	tggttccaag	3840
	ccttcctttac	gaaaagtgttt	gcagctgctt	tttcgagtcg	acaggtcgct	cgttccttcg	3900
	tttgattttcc	ctcgtctttt	accacgcccc	gggcatgtag	gattttattcc	gtttgagatt	3960
	atcgagagcc	taagtgcgcg	gcagttgtcg	gtagcaggcc	actgccggg	atg aga gaa	4018
						Met Arg Glu	
						1	
	ggt att agt ttg aac g	gtatgtgttt	tctcaatctt	ttcctcacac	cgctccttc		4074
	Val Ile Ser Leu Asn						
	5						
	accggagtcg	gtctgttgac	cggtactaat	tggttttttta	cggtgtgcag	tt ggt cag	4131
						Val Gly Gln	
						10	
	gct ggt tgc cag atc gcc aat tct tgc tgg gag	gtacggagat	ggaactatac				4184
	Ala Gly Cys	Gln Ile Ala Asn Ser	Cys Trp Glu				
	15	20					
	tagaagatag	aggtgttttg	aatgtgacta	ttgaaactga	cattatgggt	ctag ctc	4241
						Leu	
	tac tgt ctt gag cac ggc atc cag	gtacgtggcc	acgatgacaa	tccatccatc			4295
	Tyr Cys Leu Glu His Gly Ile Gln						
	25	30					
	ttcatctcca	tcgggcacac	agactaacgc	gtatacgata	g ccc gat ggt tac ttg		4351
					Pro Asp Gly Tyr Leu		
					35		
	acc gag gaa cgc aag aag gaa gac cct gac cat ggt ttc agc acc ttc						4399
	Thr Glu Glu Arg Lys Lys Glu Asp Pro Asp His Gly Phe Ser Thr Phe						
	40	45	50				
	ttc tcc gaa act g	gtacgtcgta	ttgacttggt	atcttgccgc	cgagccctct		4452
	Phe Ser Glu Thr						
	55						
	aacaggctat	gggtgtcttag	gc cag ggc aag tat gtt cct cgt acc atc tac				4504
			Gly Gln Gly Lys Tyr Val Pro Arg Thr Ile Tyr				
			60	65			
	gcc gat ctg gag ccc aat gtt gt	cgatgaggtc	cgcaactggca	cctaccgtac			4557
	Ala Asp Leu Glu Pro Asn Val Val						
	70	75					
[0543]	ccttttccac	cccgagaaca	tgatcacccg	caaggaggat	gcctcgaaca	actatgcccg	4617
	tggtcactac	accgtttggca	aggagatgat	cgaccaggtc	ctcgacaagg	ttcgccgtgt	4677
	ggccgacaac	tgcgttggtc	tccagggtct	cctcgtcttc	cactcttttcg	gtgggtggtag	4737
	tggttccggt	ttcgttgctc	tctgatgga	gcgtctgtct	gtggactacg	gcaagaagtc	4797
	caagctggag	ttctgcgtct	accctgcccc	ccagaatgcc	acctccgtcg	ttgagcccta	4857
	caactccatc	ctgactaccc	acaccacct	tgagcactcc	gactgcagtt	tcatgggtga	4917
	caacgaagcc	atctacgaca	tctgccgcgc	caaccttgcc	atcgagcgtc	ccagctatga	4977
	gaacctgaac	cgcttgattg	ctcaggttgt	ctcctccatc	accgctccc	tgcttttcga	5037
	tggttccctg	aacgtggatc	tcaacgagtt	ccagaccaac	ctggttccct	acccccgtat	5097
	tcacttccct	ctcgttgctc	atgctcccg	catctccgcg	gccaaggcct	cccacgaggc	5157
	caactccgtc	aacgagatca	ccatgtcttg	cttcgagccc	aacaaccaga	tggtcaagtg	5217
	tgacccccgc	aatggcaagt	acatggctac	ttgcttgctg	taccgtgggtg	atgtttgtgc	5277
	caaggagacc	cacgccgcgc	ttgctaccct	caagaccaag	cgtaccatcc	agttcgtcga	5337
	ctgggtgccct	actggtttca	agatcggtat	ctgctaccag	ccccctaagc	aggttcccaa	5397
	cggtgacctt	gccaacctca	gccgtgctgt	gtaagtccc	tatctaccac	tcgtgcaaga	5457
	gataaaaagc	tgaccagtta	tag c tgc atg ctg tct	aac acc acc gcc atc t			5509
			Cys Met Leu Ser	Asn Thr Thr Ala Ile			
				80			
<210>	249						
<211>	11018						
<212>	DNA						
<213>	合成						
<220>							
<221>	基因						
<222>	(1970) .. (2868)						
<223>	pyrG						
<220>							
<221>	启动子						
<222>	(3260) .. (4009)						
<223>	bipA 启动子						
<220>							
<221>	启动子						
<222>	(4004) .. (4009)						
<223>	TIS						
<220>							
<221>	基因						
<222>	(4010) .. (5474)						
<223>	ERG10 必需基因						
<220>							

<221>	外显子								
<222>	(4010) .. (4074)								
<220>									
<221>	外显子								
<222>	(4142) .. (4191)								
<220>									
<221>	外显子								
<222>	(4249) .. (4324)								
<220>									
<221>	外显子								
<222>	(4398) .. (4887)								
<220>									
<221>	外显子								
<222>	(4956) .. (5474)								
<400>	249								
aacgcgacgt	tgccctttttc	tgggcttagt	cgccattttt	cgcccttgctc	aggcaagaca		60		
ctttcaatga	cagtgttgac	ttcttcaatc	acatgcctga	gcttgaaata	ggcgctgttg		120		
ggtgggtatct	tcagtttcgag	gatcagccga	tccattttgt	tcacaacccat	ggtcaacgga		180		
aggccctcta	gcacagcgtg	cttgatgatt	tgctccgtat	tcgattgtac	tccttccaca		240		
atatcaacga	ctatcaccac	gccatcaacc	agtcgtgatg	aggcagctac	ttcatcaaca		300		
aaattcacat	gccccggggt	atcaagaatg	ttgaacaaat	gagacttccc	cttagtgccc		360		
tgaagaacta	ggctcatagg	agctgattta	atagaaaggc	ctcgctctct	ctctagaaag		420		
tgaatatcgg	tataacgaag	ctgtacttcc	tttcttttac	ctgctcgctt	ctctagccgt		480		
tcagtcaaat	catgcgtttg	cataacaagc	atatccatga	acgcggtctt	cccatgatgt		540		
aaatgaccga	caagcgcaat	gttccttatt	tggtcaggaa	aattcagaag	atcagtcattg		600		
aatctccctg	aaaagtgaac	ccgggggagc	tcagtcctct	caatggcgaa	cttcttttgc		660		
tgtacggggg	caacgatagg	ctcggagaga	ggctgtgcgt	cttcttcttg	aactaatgtc		720		
tctacctctg	tcccgtagac	ctgttgtgcg	ctagggtagt	attgtttgtc	ttcgtgtaga		780		
atgacggcat	ttgagggggc	ttctgtaatg	agatatatca	gcaccgacta	aatccagggg		840		
ctttaagaga	gtgcccaccg	tccacctcca	tagactgctg	cccatataat	tcttcatctt		900		
cttccctcccc	aaacgcctcg	ttgaaggcga	aagcccgagg	cttcacttcg	ctcgtgttggc		960		
cgtcttccatc	tgaatctgca	gcctctccaa	tgtagttgcc	aaacctgtgt	ggctcttgatt		1020		
agaatttgca	caattagagc	gttagattgg	atgcaactta	ctcatcgtag	agatcatcca		1080		
tcgtgcggtg	ggaaaggttg	ttcaagctgc	caaaaacttg	tcaccagtga	gctgtcgcat		1140		
ctctttttgt	cccaagctgc	tgaagtcctg	ctaattgga	caccgagtgg	tcogtgcgtc		1200		
gaccttaccg	ttttggtagc	cccgcgcgcg	ggatgccacg	tccaggaggg	gtcctttaa		1260		
gacttttgatt	tctcaccacc	tcttctcgtg	tcagagatac	ctcactgcta	tcgtcggctg		1320		
aataactcaa	ctcttttcttg	tttcccttct	tccccctttt	aatcccacca	atcttcttta		1380		
ttaccggcaa	gttctgtcta	cacgggtgcc	ctttaccaca	caccgaaccc	ctttaaaccg		1440		
ttcacgctac	agaagctttt	ggtttacgac	tcttctaate	agctatcgta	ggccgcgat		1500		
agctgattaa	atattctaga	cccaagccgc	tgctggaatt	gacattatta	tgcccgatag		1560		
ggttgggctt	attgctatgt	ccctgaaagg	atatcaaaa	caggcaaaaa	gccaggcata		1620		
atcccccgct	ggacgggtacc	ctaaggatag	gccctaactc	tatctacatg	tgactgcate		1680		
gatgtgtttg	gtcaaaatga	ggcatgtggc	tcaccccaca	ggcggagaaa	cgtgtggcta		1740		
gtgcattgaca	gtcccctcca	tagattcaat	ttaattttct	gcggaattg	tcgtgcagtt		1800		
tgtatctaca	tttcattcca	tatatcaaga	gttagtagtt	ggacatcctg	attattttgt		1860		
ctaattactg	aaaactcgaa	gtactaacct	actaataage	cagtttcaac	cactaagtgc		1920		
tcattttatac	aatattttgca	gaaccccgcg	ctacccctcc	atcgccaaca	tgtcttccaa		1980		
tcgcgaattg	acctacagcg	cacgcgctag	caagcacccc	aatgcgctcg	taagaaagct		2040		
cttcgaggtt	gccgaggcca	agaaaaccaa	tgtaaccgtt	tccgcgcagc	tgacaaccac		2100		
caaagagctg	ctggattttg	ctgaccgtat	gcgcacccgg	gatgccactt	acatgtgatc		2160		
tagtaatggg	taatgggtga	ttatataaca	ggactcggtc	cgtaacattg	cgatgacaaa		2220		
actcacatcg	atatcctctc	cgatttcagc	gaagaaacca	tcaccggtct	gaaggccctt		2280		
gcagagaagc	acaattttct	catcttccga	gatcgcaagt	tcacgatgat	cggaaacaca		2340		
gtccaaaagc	agtaccatgg	cggcactctg	cgtactctcg	agtgggcccc	catcatcaac		2400		
tgcagtattc	tgcccgggtga	gggtatcgtc	gaggctctgg	cccagactgc	ttcggccgag		2460		
gacttccctc	acggctccga	gaggggcctt	ttgatccttg	cggagatgac	ctccaaggga		2520		
tctttggcta	ccgggtcaata	tactacttct	tctgttgact	atgctcggaa	gtataagaag		2580		
tttgtgatgg	gattcgtctc	gacacgtcac	cttgccgagg	ttcagtcctga	agttagctcg		2640		
ccttcggagg	aggaagattt	tgctcgtctc	acgacagggt	tcaacctctc	ctcgaagggt		2700		
gacaagctgg	gacagcagta	ccaaactcct	gagtcggctg	ttggacgcgg	tgccgacttt		2760		
attattgctg	gocgtggaat	ttatgctgct	cctgatcccg	tgagggcggc	gaagcagtac		2820		
cagaaggagg	gatgggatgc	atacctgaag	cgtgttggtg	cgcaataagt	agtggtggat		2880		
acgtactcct	tttatggcag	tatgtcgcaa	gtatgatgcg	atttataaat	tcagcactcg		2940		
aaatgactac	tactatgtgt	ctacgacaga	tacctctctc	gtacgaataa	gacacctgcc		3000		
tcgatatatg	gacaaaattca	aaatcagggt	caagggtcat	gtttcaaaagt	cacaacaatc		3060		
tccaacatag	acgagaattt	gtaccggagt	gtctgaaggt	gcagctggag	attgtctaat		3120		
ttcttagag	tgggggtatca	ctaattgtaca	gtcggtcact	atcgtaaaaa	caatcacatc		3180		
tatatacaag	atttcccacc	accccctact	ctaacaacgg	acaattatcc	atcgagtcag		3240		
agcctagcca	ccattttggtc	tatgtagaat	agatagtagt	aggggtgtca	tcctccatta		3300		
tctgacaagc	agataccagc	atgtggccac	gtgtagtgtg	gttgtgcaag	tcattaagct		3360		
gcagaggtca	tgacttcaca	gtgactagag	ttgagtaact	tgagatatgg	aaagtacacg		3420		
gtgtactcaa	ggtacagtac	aattggagggt	tgaogacgct	ctttgtccac	gtatcctaaa		3480		
ggagcatatg	gccggggggc	tgaaaaaatag	ggcaacttca	agaaaagatta	cctaacaggg		3540		
gcttctagcc	aatgaaacac	atttctagta	tggtcgggtg	accgcttgca	gaagatccat		3600		
gaacactttg	gggcaatttc	tacggcgaga	tacaggcgcg	tggccaaggc	ggatgggtca		3660		
ttctgggtccc	tgaaggcgca	tgcaccaatc	acaacgcgag	gagtaacctg	cagagcgatg		3720		

[0544]

	gccgtaccag	ccatgctgtc	cttttctctg	gcctcaagca	ttttaaaaaa	gctgatctct	3780
	tcctctcaag	tccttcgttc	agcaacttct	ctcttttgcc	tcaactttcc	cttccttccc	3840
	ccttccattc	cgtcgcctgt	ggttgggttc	tccttttctt	ccttttgctt	ctcttttctg	3900
	aatacaggaa	actgtattga	agcacaaggg	attttactacc	aacacgtctc	ccgataatca	3960
	cacgcgtgcc	ccttggcgaa	gaccaccact	cgatatacac	ataggcaca	atg tct tct	4018
						Met Ser Ser	
	ctt ccg ccg gtc tat att gtt tcc tct gcc cgc acc cca gtc ggc tct						4066
	Leu Pro Pro Val Tyr Ile Val Ser Ser Ala Arg Thr Pro Val Gly Ser						
	5 10 15						
	ttc ttg gg gtgagattgg ccttgtgcct tgtctatacg tacaatattg						4114
	Phe Leu Gly						
	20						
	acttacactg actatccgct tccacag g tcg ctc tca agt ctc act gcc ccg						4166
				Ser Leu Ser Ser Leu Thr Ala Pro			
				25 30			
	cag tta ggc tct cat gct att aaa g gtatgacagc tctcgagtgg						4211
	Gln Leu Gly Ser His Ala Ile Lys						
	35						
	cggtacgtgt tgatatacta acaaactcga aattcag ct gcg ctc agc aaa gcg						4265
				Ala Ala Leu Ser Lys Ala			
				40			
	gat gga atc aag ccg tct gat atc cag gag gtc ttc ttt ggc aat gtc						4313
	Asp Gly Ile Lys Pro Ser Asp Ile Gln Glu Val Phe Phe Gly Asn Val						
	45 50 55 60						
	atc tcc gca aa gtgagtcaac gtccaccctt ataccttata gccaacatcg						4364
	Ile Ser Ala Asn						
	aaggcacact tacattctcg ttctatggat cag c gtt gga caa aat cct gct						4416
				Val Gly Gln Asn Pro Ala			
				65 70			
	aga cag tgt gct ctc ggc gct ggt ctc aat gaa tca act gtc tgt act						4464
	Arg Gln Cys Ala Leu Gly Ala Gly Leu Asn Glu Ser Thr Val Cys Thr						
	75 80 85						
	acg gtt aat aag gtg tgc gcg tct ggc ttg aaa gcg gtt att ctc ggt						4512
	Thr Val Asn Lys Val Cys Ala Ser Gly Leu Lys Ala Val Ile Leu Gly						
	90 95 100						
	gca cag acc atc atg act ggc aat gcg gat att gtc gta gca ggc ggt						4560
	Ala Gln Thr Ile Met Thr Gly Asn Ala Asp Ile Val Val Ala Gly Gly						
	105 110 115						
	gct gaa tcc atg tct aac gcc cct cat tac ctt cca aac ctt cgc gtc						4608
	Ala Glu Ser Met Ser Asn Ala Pro His Tyr Leu Pro Asn Leu Arg Val						
	120 125 130						
	ggt gcg aaa tac ggc aac cag agt ctg gtg gac ggt att atg aag gat						4656
	Gly Ala Lys Tyr Gly Asn Gln Ser Leu Val Asp Gly Ile Met Lys Asp						
	135 140 145 150						
	ggc ttg aca gac gca gga aag cag gaa ctc atg ggc ttg caa gcc gag						4704
	Gly Leu Thr Asp Ala Gly Lys Gln Glu Leu Met Gly Leu Gln Ala Glu						
	155 160 165						
	gag tgt gct cag gat cat ggc ttt agc agg gaa caa cag gat gat tat						4752
	Glu Cys Ala Gln Asp His Gly Phe Ser Arg Glu Gln Gln Asp Asp Tyr						
	170 175 180						
	gcc att cgc act tac gaa aaa gca cag gcg gct caa aag gct ggc ctt						4800
	Ala Ile Arg Thr Tyr Glu Lys Ala Gln Ala Ala Gln Lys Ala Gly Leu						
	185 190 195						
	ttt gac gaa gaa att gcg cct att gaa ctt cct ggc ttt agg ggc aag						4848
	Phe Asp Glu Glu Ile Ala Pro Ile Glu Leu Pro Gly Phe Arg Gly Lys						
	200 205 210						
	cca ggt gtg act gtg tca caa gac gaa gaa cca aag aat gtgaggaaga						4897
	Pro Gly Val Thr Val Ser Gln Asp Glu Glu Pro Lys Asn						
	215 220 225						
	tattactttc tacagatttt ttaagtcgtc ggccactgct aatatactca ttgcgtag						4955
	ctt aac ccg gat aag ctt cga gct atc aag cct gca ttt atc ccc gga						5003
	Leu Asn Pro Asp Lys Leu Arg Ala Ile Lys Pro Ala Phe Ile Pro Gly						
	230 235 240						
	tcc ggc acg gtc aca gcc ccg aat tcc tca cct ctt aac gac ggt gct						5051
	Ser Gly Thr Val Thr Ala Pro Asn Ser Ser Pro Leu Asn Asp Gly Ala						
	245 250 255						
	gct gct gtt atc ctc gtc tca gaa gct aaa ctg aaa gag ctt aac cta						5099
	Ala Ala Val Ile Leu Val Ser Glu Ala Lys Leu Lys Glu Leu Asn Leu						
	260 265 270 275						
	aag cct gtt gca aag att ctt ggc tgg gga gat gcc gcc cag cag cca						5147
	Lys Pro Val Ala Lys Ile Leu Gly Trp Gly Asp Ala Ala Gln Gln Pro						
	280 285 290						
	agc aaa ttc aca act gcc cca gct cta gca att ccc aag gcc ctc agc						5195
	Ser Lys Phe Thr Thr Ala Pro Ala Leu Ala Ile Pro Lys Ala Leu Ser						
	295 300 305						
	cat gca ggt gtg gct cag gat gct gtt gat gcg ttc gag att aac gaa						5243

His	Ala	Gly	Val	Ala	Gln	Asp	Ala	Val	Asp	Ala	Phe	Glu	Ile	Asn	Glu	
		310					315					320				
gcg	ttc	agc	gta	ggt	gct	ctg	gcc	aat	atg	aaa	ctc	ctg	ggg	ttg	gct	5291
Ala	Phe	Ser	Val	Val	Ala	Leu	Ala	Asn	Met	Lys	Leu	Leu	Gly	Leu	Ala	
		325					330					335				
gaa	gat	aaa	gtc	aac	atc	cat	ggg	ggg	gca	gtg	gct	atc	ggg	cat	cct	5339
Glu	Asp	Lys	Val	Asn	Ile	His	Gly	Gly	Ala	Val	Ala	Ile	Gly	His	Pro	
		340					345				350				355	
atc	ggc	gcc	agc	ggg	gct	cgt	atc	ttg	act	aca	ttg	ctc	ggg	gta	ttg	5387
Ile	Gly	Ala	Ser	Gly	Ala	Arg	Ile	Leu	Thr	Thr	Leu	Leu	Gly	Val	Leu	
				360					365					370		
aaa	gcg	aga	aag	ggg	aag	att	ggg	tgt	gcc	ggg	att	tgt	aat	gga	gga	5435
Lys	Ala	Arg	Lys	Gly	Lys	Ile	Gly	Cys	Ala	Gly	Ile	Cys	Asn	Gly	Gly	
			375					380					385			
ggg	ggg	gct	agc	gct	att	ggt	gtc	gaa	tct	ctc	gtc	tga	gtcatg	gccga		5484
Gly	Gly	Ala	Ser	Ala	Ile	Val	Val	Glu	Ser	Leu	Val					
		390					395									
gtcgcgccttg	atcttgggtat	ggaataacgc	gacgttgccc	ttttctgggc	ttagtcgcca											5544
tttttcgcct	tgtccaggca	agacactttc	aatgacagt	ttgacttctt	caatcacatg											5604
cttgagcttg	aaataggcgt	cgttgggtgg	tatcttcagt	tcgaggatca	gccgatccat											5664
tttggttcaca	accatgggtca	acggaaggcc	ctctagcaca	gcgtgcttga	tgatttgctc											5724
cgtattcgat	tgtactcctt	ccacaatatc	aacgactatc	accacgccat	caaccagtcg											5784
tgatgaggca	gctacttcat	caacaaaatt	cacatgcccc	gggggtatcaa	gaatgttgaa											5844
caaatgagac	ttccccttag	tgcccctgaag	aactaggctc	ataggagctg	atttaataaga											5904
aaggcctcgc	tctctctcta	gaaagtgaat	atcgttataa	cgaagctgta	cttcctttct											5964
tttacctgct	cgcttctcta	gccgttcagt	caaactatgc	gtttgcataa	caagcatatc											6024
catgaacgcg	gtcttcccat	gatgtaaatg	accgacaagc	gcaatgttcc	ttatttggtc											6084
agggaaaattc	agaagatcag	tcatgaattc	ccttgaaaag	tgaaccgggg	ggagctcagt											6144
ctcttcaatg	gcgaacttct	tttgctgtac	gggggcaacg	ataggctcgg	agagaggctg											6204
tgctgtctct	tcttgaacta	atgtctctac	ctctgtcccg	tagacctgtt	gtgcgtcagg											6264
gtagtattgt	ttgtcttcgt	gtagaatgac	ggcatttgag	ggggcttctg	taatgagata											6324
tatcagcacc	gactaaatcc	aggggcttta	agagagtgc	caccgtccac	ctccatagac											6384
tgctgcccc	ataattcttc	atcttcttcc	tcccaaacg	cctcgttgaa	ggcgaaagcc											6444
cgaggcttca	cttcgtctg	ttggcctgtg	tcatctgaat	ctgcagcctc	tccaatgtag											6504
ttgccaaaacc	tgtgtgggtct	tgattagaat	ttgcaacaat	agagcgttag	attggatgca											6564
acttactcat	cgtagagatc	atccatcgtg	cggtgggaaa	gggtggttcaa	gctgccaaaa											6624
acttctcacc	agttagctgt	cgcactctct	ttgttcccaa	gctgctgaag	tcctgtctaat											6684
ggactcaccg	agtgggtccg	gcgtcgaccc	taccattttg	gtagccccc	cgcgcggtatg											6744
ccacgtccag	gaggggtccct	taaatgaact	tgattttctca	ccacctcttc	tcgtgtcaga											6804
gacactccac	tgctatcgtc	ggctgaataa	ctcaactctt	tcttgtttcc	tccttctccc											6864
cctttaatcc	caccaatctt	ctttattacc	ggcaagtctt	gtctacacgg	gtgcccttta											6924
ccacacaccg	aaccccttta	aacgcttcac	gctacagaag	cttttggttt	acgactcttc											6984
taatcgccta	tcgtaggccg	ccgatagctg	attaaatatt	ctagacccaa	cgcgctcgtg											7044
gaattgacat	tattatggcc	gatagggttg	ggcttattgc	tatgtccctg	aaaggatatc											7104
aaaagcaggc	aaaaagccag	gcataatccc	cgctggagcg	gtaccctaag	gataggccct											7164
aatcttatct	acatgtgact	gcacatcgtg	gtttgggtcaa	aatgaggcat	gtggctcacc											7224
ccacagcggg	agaaacgtgt	ggctagtgcg	tgacagtccc	ctccatagat	tcaatttaat											7284
ttttcgcggc	aattgtcgtg	cagtttggtat	ctacatttca	ttccatatat	caagagttag											7344
tagttggaca	tcttgattat	tttgtctaat	tactgaaaac	tcgaagtact	aacctactaa											7404
taagccagtt	tcaaccaata	agtgtctcatt	tatacataat	ttgcagaacc	ccgcgtctacc											7464
cctccatcgc	caacatgtct	tccaagtcgc	aattgacctc	cagcgcacgc	gctagcaagc											7524
accccaatgc	gctcgtaaaag	aagctctctc	aggttgccga	ggccaaagaa	accaatgtca											7584
ccgtttccgc	cgacgtgaca	accaccaaa	agctgtgga	tttggctgac	ccgatgtcca											7644
ccgggggatgc	cacttacatg	tgatctagta	atgggttaatg	gtggattata	taacaggact											7704
cggtccgtac	attgccgtga	tcaaaaactca	catcgatata	ctctccgatt	tcagcgaaga											7764
aaccatcacc	ggctggaagg	cccttcgaga	gaagcacaat	ttcctcatct	tcgaagtctg											7824
caagttcatc	gatatcgga	acacagtcga	aaagcagtac	catggcggca	ctctgcgtat											7884
ctctgagtgg	gcccacatca	tcaactgcag	tattctgccc	ggtaggggta	tcgtcgaggc											7944
tctggcccag	actgcttcgg	ccgaggactt	cccctacggc	tccgagaggg	gccttttgat											8004
ccttgccggag	atgacctcca	agggatcttt	ggctaccggg	caatatacta	ctctctctgt											8064
tgactatgct	cggaagtata	agaagtttgt	gatgggattc	gtctcgacac	gtcaccttgg											8124
cgaggttcag	tctgaagtta	gctcgccttc	ggaggaggaa	gattttgtcg	tcttcacgac											8184
aggtgtcaac	ctctcctcga	aggggtgaca	gctgggacag	cagtaccaaa	ctccttgatc											8244
ggctgttgga	cgcggtgcgc	actttattat	tgctggccgt	ggaatttatg	ctgctcctga											8304
tcccgtggag	gcggcggaagc	agtaccagaa	ggagggatgg	gatgcatacc	tgaagcgtgt											8364
tggtgcgcaa	taagtagtgg	tgataacgta	ctccttttat	ggcagtatgt	cgcaagtatg											8424
atgcgattta	taaattcagc	actcgaaatg	actactacta	tgtgtctacg	acagataccc											8484
tctccgtatc	aaatagacac	ctgcctcgat	atatggacaa	attcaaaaatc	aggggtcaagg											8544
gtcatgtttc	aaagtacaaa	caatctccaa	catagacgag	aatttgtacc	ggagtgtctg											8604
aaggtgcagc	tgagatttgg	tctattttct	tagagtgggg	tatcactaat	gtacagtcgg											8664
tcactatcgt	acaacaatc	acaattatat	acaagatttc	ccaccacccc	ctactctaac											8724
acggcacaa	tatccatcga	gtcagagcct	agccaccatt	tggtctatgt	agaatataga											8784
gtagtagggg	gtcatcctc	cattatctga	caagcagata	ccagcatgtg	gccacgtgta											8844
gtgatgttgt	gcaagtcatt	aagctgcaga	ggctcatgact	tcacagtgc	tagagttag											8904
taacttgaga	tatggaaagt	acacgggtga	ctcaaggtac	agtacaattg	gaggttgacg											8964
acgtcccttg	tccacgtatc	ctaaaggagc	atatggccgg	ggggctgaaa	aatagggcaa											9024
cttcaagaaa	gattacctaa	caggggcttc	tagccaatga	aacacatttc	tagtatgttc											9084
gggtggaccgc	ttgcagaaga	tccatgaaca	ctttggggca	atttctacgg	cagaatacag											9144

[0546]

	ggcgggtggcc	aaggcggatg	ggtcattctg	gtccctgaaa	ggcgatgcac	caatcacaac	9204
	gcgaggagta	acctgcagag	cgatggccgt	accagccatg	ctgtcccttt	ctctggcctc	9264
	aagcatttta	aaaaagctga	tctcttccctc	tcacgtccct	cgttcagcaa	cttctctctt	9324
	ttgcctcaac	tttcccttcc	ttccccccttc	cattccgtcg	cctgtgggtg	ggttctcttt	9384
	ttcttctctt	tgttctctct	tcttgaatac	aggaaactgt	attgaagcac	aagggattta	9444
	ctaccaacac	gtctcccgat	aatcacacgc	gtgccccttg	gcgaagacca	ccactcgata	9504
	tacacatagg	cacaatgtct	tctcttccgc	cggtctatat	tgtttctctc	gcccgcaccc	9564
	cagtcggctc	tttcttgggg	tgagattggc	cttgtgcctt	gtctatacgt	acaatattga	9624
	cttacactga	ctatccgctt	ccacaggctcg	ctctcaagtc	tcactgcccc	gcagtttaggc	9684
	tctcatgcta	ttaaaggtat	gacagctctc	gagtggcggt	acgtgttgat	atactaacia	9744
	actcgaaatt	cagctgcgct	cagcaaaagcg	gatggaaatca	agccgtctga	tatccaggag	9804
	gtcttctttg	gcaatgtcat	ctccgcaaag	tgagtcaacg	tccaccctta	taccttatag	9864
	ccaacatcga	aggcacactt	acattctcgt	tctatggatc	agcgttggac	aaaatcctgc	9924
	tagacagtgt	gctctcggcg	ctgggtctcaa	tgaatcaact	gtctgtacta	cggttaataa	9984
	ggtgtgcgcg	tctggcctga	aagcggttat	tctcggtgca	cagaccatca	tgactggcaa	10044
	tgccgatat	gtcgtagcag	gcgggtgctga	atccatgtct	aacgcccctc	attaccttcc	10104
	aaacctctgc	gtcgggtgcga	aatacggcaa	ccagagtctg	gtggacggta	ttatgaagga	10164
	tggtctgaca	gacgcaggaa	agcaggaact	catgggcttg	caagccgagg	agtgtgctca	10224
	ggatcatggc	tttagcaggg	aacaacagga	tgattatgcc	attcgcactt	acgaaaaagc	10284
	acaggcggct	caaaaggctg	gcctttttga	cgaagaaatt	gcgcctattg	aacttccctg	10344
	cttttaggggc	aagccagggtg	tgactgtgtc	acaagacgaa	gaaccaaaga	atgtgaggaa	10404
	gatattactt	tctacagatt	ttttaagtcg	tcggccactg	ctaataact	cattgcgtag	10464
	cttaaccggg	ataagcttcc	agctatcaag	cctgcattta	tccccggatc	cggcacggtc	10524
	acagcccaga	attcttcacc	tcttaacgac	ggtgctgctg	ctgttatcct	cgtctcagaa	10584
	gctaaactga	aagagcttaa	cctaaagcct	gttgcaaaga	ttcttggctg	gggagatgcc	10644
	gccccagcagc	caagcaaat	cacaactgcc	ccagctctag	caattcccaa	ggccctcagc	10704
	catcgagggtg	tggtcagga	tgctgttgat	gcgttcgaga	ttaacgaagc	gttcagcgta	10764
	gttgctctgg	ccaatatgaa	actcctgggg	ttggctgaag	ataaagtcaa	catccatggt	10824
	ggtgcagtg	ctatcgggtca	tctatcggc	gccagcggtg	ctcgtatctt	gactacattg	10884
	ctcgggtgat	tgaaagcgag	aaagggttaag	attggtgtg	cggggatttg	taatggagga	10944
	ggtggtgcta	gcgctattgt	tgctgaatct	ctcgtctgag	tcatgccgag	tcgccttga	11004
	tcttggtatg	gaat					11018
	<210>	250					
	<211>	5509					
	<212>	DNA					
	<213>	合成					
	<220>						
	<221>	基因					
[0547]	<222>	(1970) .. (2868)					
	<223>	pyrG					
	<220>						
	<221>	启动子					
	<222>	(3260) .. (4009)					
	<223>	bipA 启动子					
	<220>						
	<221>	启动子					
	<222>	(4004) .. (4009)					
	<223>	TIS					
	<220>						
	<221>	基因					
	<222>	(4010) .. (5509)					
	<223>	PFS2 必需基因					
	<220>						
	<221>	外显子					
	<222>	(4010) .. (4014)					
	<220>						
	<221>	外显子					
	<222>	(4072) .. (4142)					
	<220>						
	<221>	外显子					
	<222>	(4229) .. (4460)					
	<220>						
	<221>	外显子					
	<222>	(4534) .. (5509)					
	<400>	250					
	ataaaaaattt	cgccaccgaa	gaagtatcaa	ggaaactgtc	tagcgagagg	gactggactg	60
	gaacacttga	gtattcggta	tccagttaag	gtgatttaag	ccacccttc	tcccttcaat	120
	ccctaccccc	cttcaaaaag	gggaagaaca	gtagctaaag	gtagttgcca	tcaagaactc	180
	atacttaaaa	acacgaggta	agaggcccg	aatagtaatt	ggttgcaaga	gaaaaggggg	240
	gttatecttc	tgagcatcac	agttttgatt	acgggaagaa	ccagaatatg	aacgtggcaa	300
	cagcggctgt	cggtaccaga	ccaggccctt	attggatacc	ggaggagtcc	ctcttctgtg	360
	tgtttctctac	ccctttttaca	gtacaattac	aacctacata	atgtcactga	tagatattta	420
	tacgtgtgcc	agaagatact	tatgggttag	tcaatgtaaa	tagcttctaa	ttttgattca	480
	tgattgaggg	gtctaaaatt	ggactgaacc	taatcctact	caagtaggag	ttagatgacg	540
	atgggtgagt	tgaaatgcc	ggattaagg	tagcatatca	gccgcacagg	aactaccgag	600
	gggggtggagg	ctaacagccc	tctttttttc	tttccctttt	tcctcatttt	cctctctctt	660

[0548]	ctccccgatga	tgacgatttg	tagattcata	agttatagaa	aaatatttgt	tctatgcaca	720
	cgctaataagg	acatagtaact	gcttatttgta	tgtatatgcc	tggtgtttta	agaagtagac	780
	ttcttgagtt	ccgtacggag	cactattttta	caatacgaat	aaaaaaaaatc	catggagcag	840
	ggaaggggaag	gatatagcta	gctaggtccc	taatagtoga	tatgtttattt	ttgatctggt	900
	agtggagcca	ctgacctgtt	gaatttttgg	taggttaaact	tagcaacatg	tccatgccat	960
	aaaagggtctc	caaccgagag	tcacgagtgga	aacataccag	taggtccaac	aaactgtgta	1020
	aggtgcagaa	aaggcacatt	attggccagc	tccatgttga	atgaatggct	gcccctgga	1080
	gccggtccat	agtagcttcc	ctgggtctctg	tgtgggcccga	tctcacaca	tataacatag	1140
	ctcttccccc	tccttccaac	agattgtcga	tctactcaat	ttcccccttg	ctatcaactt	1200
	ctccccctcc	tcgcttccct	ctctcactcg	ccacaccgcc	ctttaagaat	atcacaagta	1260
	atattctagc	tggtattatcc	aaagtgtgtt	ttatgtttgc	cttccatggg	atcttccaat	1320
	tcgtttgttg	acagtatatt	cccttccaga	gtgcactttg	aagaaacccc	actaccgttt	1380
	acgtgacttg	ctttcaatat	ccggaggcaa	catttgttat	cgaagcatta	caacgcctag	1440
	tcactctcatt	ctacaccttt	ccacattttg	actcgctctc	tctcttgaca	taccttaacc	1500
	agctgattaa	atattctaga	cccaagccgc	tgctggaatt	gacattatta	tggccgatag	1560
	gggtgggctt	attgctatgt	ccctgaaagg	atatcaaaag	caggcaaaaa	gccaggcata	1620
	atccccgctt	ggacggtaac	ctaaggatag	gcctaatct	tatctacatg	ttcgccctac	1680
	gatgtgtttg	gtcaaaatga	ggcatgtggc	tcaccccaca	ggcggagaaa	cgtgtggcta	1740
	gtgcatgaca	gtcccctcca	tagattcaat	ttattttttc	gcggaatttg	tcgtgcagtt	1800
	tgtattctaca	tttcattcca	tatatcaaga	gttagtagtt	ggacatcctg	attattttgt	1860
	ctaattactg	aaaactcgaa	gtactaacct	actaaataag	cagtttcaac	cactaagtgc	1920
	tcattttatac	aatattttgca	gaaccccgcg	ctaccctcc	atcgccaaca	tgtcttccaa	1980
	gtcgcaattg	acctacagcg	cacgcgctag	caagcaccce	aatgcgctcg	taaagaagct	2040
	cttcgaggtt	gccgaggcca	agaaaaacca	tgtcacgctt	tcgcgcgacg	tgacaaccac	2100
	caaagagctg	ctggattttg	ctgaccgtat	gcgcacccgg	gatgccactt	acatgtgatc	2160
	tagtaatggt	taatgggtga	ttatataaca	ggactcggtc	cgtacattgc	cgtgatcaaa	2220
	actcacatct	cgattttcagc	cgattttcagc	gaagaaacca	tcaccggtct	tcaggccact	2280
	gcagagaagc	acaattttcct	catcttccga	gatcgcaagt	tcacgatgat	cggaaacaca	2340
	gtccaaaagc	agtaccatgg	cggcactctg	cgtatctctg	agtggggcca	catcatcaac	2400
	tgccgtattc	tgcccgggtga	gggtatcgtc	gaggtctctg	cccagactgc	ttcgccgag	2460
	gacttccctt	acggctccga	gaggggcctt	ttgatccttg	cggagatgac	ctccaaaggga	2520
	tctttggcta	ccggtcaata	tactacttct	tctgttgact	atgctcggaa	gtataagaag	2580
	tttgtgatgg	gattcgtctc	gacacgtcac	cttggcgagg	ttcagtcga	agttagctcg	2640
	ccttcggagg	aggaagattt	tgtcgtcttc	acgacaggtg	tcaacctctc	ctcgaagggt	2700
	gacaagctgg	gacagcagta	ccaaactcct	gagtcggctg	ttggacgcgg	tgccgacttt	2760
	attattgctg	gccgtggaat	ttatgctgct	cctgatcccg	tggaggcggc	gaagcagtag	2820
	cagaaggagg	gtgggtagtc	atacctgaag	cgtgttggtg	cgcaataaag	agtggtggat	2880
	acgtactcct	tttatggcag	tatgtcgcaa	gtatgatgcg	atttataaat	tcagcactcg	2940
	aaatgactac	tactatgtgt	ctacgcagaa	tacctctccc	gtacgaataa	gacacctgcc	3000
	tcgatatatg	gacaaattca	aaatcagggt	caagggtcat	gtttcaaaag	cacaacaatc	3060
	tccaacatag	acgagaattt	gtaccggagt	gtctgaagg	gcagctggag	attgggtctat	3120
	ttctcttagag	tgggggtatca	ctaattgtaca	gtcggtcact	atcgtacaaa	caatcacaat	3180
	tatatacaag	atttcccacc	accccctact	ctaaactatcc	acaattatcc	atcgagtcag	3240
	agcctagcca	ccattttggtc	tatgtagaat	agatagtagt	aggggtgtca	tcctccatta	3300
	tctgacaagc	agataccagc	atgtggccac	gtgtagtgt	gtttgtgcaag	tcattaagct	3360
	gcagaggtea	tgactttcaca	gtgactagag	ttgagtaact	tgagatatgg	aaagtagacg	3420
	gtgtactcaa	ggtacagtag	aattggagg	tgacgacgtc	ctttgtccac	gtatcctaaa	3480
	ggagcatatg	gccggggggc	tgaaaaatag	ggcaacttca	agaaagatta	cctaacaggg	3540
	gcttctagcc	aatgaaacac	atttctagta	tgttcgggtg	accgcttgca	gaagatccat	3600
	gaacactttg	gggcaatttc	tacggcagaa	tacagggcgg	tggccaaggc	gtatcggtca	3660
	ttctggctcc	tgaaggcgca	tgcaccaatc	acaacgcgag	gagtaacctg	cagagcgatg	3720
	gccgtacacg	ccatgctgtc	cttttctctg	gcctcaagca	ttttaaaaaa	gctgatctct	3780
	tcctctcacg	tccttcgttc	agcaacttct	ctcttttgcc	tcaactttcc	cttccctccc	3840
	ccttccattc	cgtcgctgtg	ggttgggttc	tctttttctt	cctttttgctt	ctctttcttg	3900
	aatacaggaa	actgtattga	agcacaagg	atttactacc	aacacgtctc	ccgataatca	3960
	cacgcgtgcc	ccttggcgaa	gaccaccact	cgatatacac	ataggcaca	atg tg	4014
						Met Cys	
						1	
	gtacgttccc	gaagaccag	gcttactgaa	tgtttgatgt	aactaattga	cccctag c	4072
	ggc atc ttc	ggc tac att	aac tac ctc	gtc gag agg	gac cgc aag	ttt	4120
	Gly Ile Phe	Gly Tyr Ile	Asn Tyr Leu	Val Glu Arg	Asp Arg Lys	Phe	
		5	10	15			
	att ctt gac	act ctt ctc	aat g gtt	ggcattc	ctgttgtttt	cacctcact	4172
	Ile Leu Asp	Thr Leu Leu	Asn				
		20	25				
	ggaccttgtt	tactctgggt	ttgggagtg	ttgtgtattc	ctctgacata	ttgaag ga	4230
						Gly	
	ctc tcc cgt	ctt gaa tac	cga ggc tac	gac tct gcg	ggt ctt gcc	ata	4278
	Leu Ser Arg	Leu Glu Tyr	Arg Gly Tyr	Asp Ser Ala	Gly Leu Ala	Ile	
		30	35	40			
	gat ggc gat	aag aag aat	gaa gtt tgt	gct ttc aag	gaa gta	ggc aag	4326
	Asp Gly Asp	Lys Lys Asn	Glu Val Cys	Ala Phe Lys	Glu Val Gly	Lys	
		45	50	55			
	ggt gct aag	ctt cga gag	ctt att gat	gag tgc aag	cct gac ttg	acc	4374
	Val Ala Lys	Leu Arg Glu	Leu Ile Asp	Glu Cys Lys	Pro Asp Leu	Thr	
		60	65	70			
	aaa tcc ttt	gaa tct cat	gct ggt att	tct cac act	cgt tgg gcc	act	4422

	Lys Ser Phe Glu Ser His Ala Gly Ile Ser His Thr Arg Trp Ala Thr	
	75 80 85 90	
	cat ggg acg cct tcc cgc ttg aac tgc cac ccc cac ag gtgagtcgga	4470
	His Gly Thr Pro Ser Arg Leu Asn Cys His Pro His Arg	
	95 100	
	tggtccacgg ttttaatcagc tcgactagga tcaggcttgt ttgctaattt tcctggcagc	4530
	cag g tcc gat ctg aac tgg gaa ttt tca gtt gtc cac aat ggt atc att	4579
	Ser Asp Leu Asn Trp Glu Phe Ser Val Val His Asn Gly Ile Ile	
	105 110 115	
	acc aac tac aaa gag ctc aag gct ttg ttg gag agc aag gga ttc cgc	4627
	Thr Asn Tyr Lys Glu Leu Lys Ala Leu Leu Glu Ser Lys Gly Phe Arg	
	120 125 130	
	ttt gag act gac aca gac acc gaa tgt att gca aag ctt acc aaa tat	4675
	Phe Glu Thr Asp Thr Asp Thr Glu Cys Ile Ala Lys Leu Thr Lys Tyr	
	135 140 145	
	ctc tat gac caa cag ccc gat atc gaa ttc aca gtt ttg gct aag gct	4723
	Leu Tyr Asp Gln Pro Asp Ile Glu Phe Thr Val Leu Ala Lys Ala	
	155 160 165	
	ggt gtg aag gag ctt gag ggc gcg ttt ggg ctt ctc atc aaa tcc gta	4771
	Val Val Lys Glu Leu Glu Gly Ala Phe Gly Leu Leu Ile Lys Ser Val	
	170 175 180	
	cat tat ccc cat gag gtc att gcg gct cgc aag ggg tcg ccg ctt gtt	4819
	His Tyr Pro His Glu Val Ile Ala Ala Arg Lys Gly Ser Pro Leu Val	
	185 190 195	
	att ggc gtg aga acg tcg aag aag atg aag gtg gat ttc gtg gac gtc	4867
	Ile Gly Val Arg Thr Ser Lys Lys Met Lys Val Asp Phe Val Asp Val	
	200 205 210	
	gaa tac tca gaa gat gga gcc ctt cct gca gag cag gcc tcc cag aac	4915
	Glu Tyr Ser Glu Asp Gly Ala Leu Pro Ala Glu Gln Ala Ser Gln Asn	
	215 220 225 230	
	gtg gcc atc aag aag tct gcg acc ggc ctc ctt gcc cca cct gac aaa	4963
	Val Ala Ile Lys Lys Ser Ala Thr Gly Leu Leu Ala Pro Pro Asp Lys	
	235 240 245	
	tcc ctt ttg cat agg tca cag tct cgc gct ttc ctt tcc gat gat ggc	5011
	Ser Leu Leu His Arg Ser Gln Ser Arg Ala Phe Leu Ser Asp Asp Gly	
	250 255 260	
	atc ccc cag cca gct gag ttt ttc ttg tct tct gac cca tct gca att	5059
	Ile Pro Gln Pro Ala Glu Phe Phe Leu Ser Ser Asp Pro Ser Ala Ile	
	265 270 275	
[0549]	gtc gag cat acg aag aag gtc ctc tat ctc gaa gat gat gat atc gct	5107
	Val Glu His Thr Lys Lys Val Leu Tyr Leu Glu Asp Asp Asp Ile Ala	
	280 285 290	
	cat gtg cat gag gga cag tta aac att cat cgc ctc aca aaa gat gat	5155
	His Val His Glu Gly Gln Leu Asn Ile His Arg Leu Thr Lys Asp Asp	
	295 300 305 310	
	ggt act tct aat gtt cgc gct att cag aca att gaa ctc gaa ttg cag	5203
	Gly Thr Ser Asn Val Arg Ala Ile Gln Thr Ile Glu Leu Glu Leu Gln	
	315 320 325	
	gag att atg aag ggc aag ttt gac cac ttc atg caa aag gaa att ttc	5251
	Glu Ile Met Lys Gly Lys Phe Asp His Phe Met Gln Lys Glu Ile Phe	
	330 335 340	
	gag caa ccc gag tct gtg atc aat acc atg aga gga cga ctg gat gtt	5299
	Glu Gln Pro Glu Ser Val Ile Asn Thr Met Arg Gly Arg Leu Asp Val	
	345 350 355	
	gca aac aag caa gtc aca ctc ggt ggc ctg cgg cag tac att tct act	5347
	Ala Asn Lys Gln Val Thr Leu Gly Gly Leu Arg Gln Tyr Ile Ser Thr	
	360 365 370	
	atc cgc cgc tgc aga aga atc ata ttt gtt gcc tgt gga act agt tac	5395
	Ile Arg Arg Cys Arg Arg Ile Ile Phe Val Ala Cys Gly Thr Ser Tyr	
	375 380 385 390	
	cat tca tgc atg gct gtg cgt gga gtc ttt gaa gag ctt act gag att	5443
	His Ser Cys Met Ala Val Arg Gly Val Phe Glu Glu Leu Thr Glu Ile	
	395 400 405	
	cct att tct gtt gaa ctt gct tcc gat ttc ctg gat aga cag gct cca	5491
	Pro Ile Ser Val Glu Leu Ala Ser Asp Phe Leu Asp Arg Gln Ala Pro	
	410 415 420	
	gta ttc cgt gat gac acg	5509
	Val Phe Arg Asp Asp Thr	
	425	
	<210> 251	
	<211> 5509	
	<212> DNA	
	<213> 合成	
	<220>	
	<221> 基因	
	<222> (1970) .. (2868)	
	<223> pyrG	
	<220>	

<221>	启动子								
<222>	(3260) .. (4009)								
<223>	bipA 启动子								
<220>									
<221>	启动子								
<222>	(4004) .. (4009)								
<223>	TIS								
<220>									
<221>	基因								
<222>	(4010) .. (5509)								
<223>	TUB1 必需基因								
<220>									
<221>	外显子								
<222>	(4010) .. (4034)								
<220>									
<221>	外显子								
<222>	(4124) .. (4164)								
<220>									
<221>	外显子								
<222>	(4239) .. (4265)								
<220>									
<221>	外显子								
<222>	(4337) .. (4412)								
<220>									
<221>	外显子								
<222>	(4473) .. (5427)								
<220>									
<221>	外显子								
<222>	(5481) .. (5509)								
<400>	251								
cgggaagatg	ttccatagaa	agggatcaag	gatgatatcc	gctgcaaccg	ggggaacatt		60		
ctccattagt	gctttgttaa	gagaaaagta	gtccgccgcc	gcagtaacac	gaaggaaaaag		120		
caccgaatca	gacgtaggtg	cttgatcatg	aaaagaggaa	gattttgcgag	gaacagcggtt		180		
agactgaggg	gagggcgagat	gggcttgagg	atagaggcta	atcgaggaga	agagtgcggg		240		
atcatcaatg	gcttgtggaa	gagtatgcgt	tgtcagggta	aaggcggtcg	gttgctttga		300		
ctacttgtca	gctaattttc	tttggtcgcc	ttgctgggca	ctggtagagg	gtatatgagg		360		
acgtacagtt	gccaaccaac	acacccgcac	ctcataccgc	tggccggggg	tcaagttctc		420		
cagaaaaaac	catgagtcgt	taccagggaa	ctcgttggtc	ggaaatgtag	cgtaattctc		480		
ggttcgaaga	acgggactag	aggaggagag	gcgatctaga	cctagggtcat	cgagggtagg		540		
atcgaccgtg	gggatgatga	ggggttgcgg	agcgatgaag	atcgttttct	cgacattggc		600		
gatagcgggg	aatgcgaggg	tgcagagtaa	gagtagatat	tggagtagca	tcatgacaga		660		
gagttcgcta	tgtttcaattg	acgctgctgc	tagacggtca	tggttctagt	ccacacgcag		720		
ttgagtatgg	taggctatat	gtatgtatgt	agtaaatgag	gtagatctga	tactaaactgt		780		
tagattctca	agtgaagtcc	ttacatgaaa	ttgattcaat	gtcaatacga	tcttaccata		840		
aatttaggcca	ttgtgcgcag	caatgttggt	gactagaaat	agagcactgt	gtgtccgaga		900		
cttgttctcat	gtgagtatatt	tcaatagtag	ttactatcgc	aagctgtacg	gatagtcaaa		960		
tacaatgcag	ctgcacgaag	agatgccgtc	attgacacgt	gacgcaccct	tgcctgaac		1020		
taagcccaaa	taagtggctg	agtcagcggt	gggctggaag	ggcggtcacg	gcacgcgaca		1080		
agccaatcag	aaatgggatt	gtgggcgacg	cgaccaaacg	gatggatcaa	caatgaaaag		1140		
gtggaagtcc	tgccttttggg	ttgacagcgc	atcgcaagac	cacaacaaat	tgaccgatca		1200		
ttcgggtctt	gggcaattgt	ttaattgata	cccttagtct	atctttctcc	tcgtcaatca		1260		
ccccaaagggt	ttgtttctcct	cccccaatttc	ctcccaccgg	ttccccttct	ccctttctttt		1320		
gggtcttttc	ccaggttggtg	tttttttctg	ttacttaccc	ctgaaacttc	cttacagatc		1380		
agggcaatttt	tcccctctca	aagtccttcc	acattccaac	cactccgttt	cacacacccc		1440		
cctccttcca	atcttaattc	ctttgtcact	ccgacattcg	tcccttttag	actcgacaag		1500		
agctgattaa	atattctaga	cccaagccgc	tgttggaatt	gacattatta	tggccgatag		1560		
ggttgggctt	attgctatgt	ccctgaaagg	atatcaaaag	caggcaaaaa	gccaggcata		1620		
atcccccggt	ggacgggtacc	ctaaggatag	gccctaattct	tatctacatg	tgactgcata		1680		
gatgtgtttg	gtcaaaatga	ggcatgtggc	tcaccccaca	ggcgagagaa	cgtgtggcta		1740		
gtgcatgaca	gtcccctcca	tagattcaat	ttaatttttc	gcggaattg	tcgtgcagtt		1800		
tgtatctaca	tttcattcca	tatatcaaga	gttagtagtt	ggacatccgt	attattttgt		1860		
ctaattactg	aaaactcgaa	gtactaacct	actaataagc	cagtttcaac	cactaagtgc		1920		
tcattttatac	aatatattgca	gaacccccgc	ctacccctcc	atcgccaaca	tgtcttccaa		1980		
tgcgcaattg	acctacagcg	cacgcgctag	caagcacccc	aatgcgctcg	taaaagaagt		2040		
cttcgaggtt	gccgaggcca	agaaaaaccac	tgtcacccgt	tccgcgcgac	tgacaaccac		2100		
caaagagctg	ctggatttgg	ctgaccgtat	gcgcaccggg	gatgccactt	acatgtgatc		2160		
tagtaattgg	taatgggtgga	ttatataaca	ggactcggtc	cgtacattgc	cgtgatcaaa		2220		
actcacatcg	atatcctctc	cgatttccagc	gaagaaaacca	tcaccggtct	gaaggccctt		2280		
gcagagaagc	acaatttctc	catcttcgaa	gatcgcaagt	tcacatgat	cggaacacac		2340		
gtccaaaagc	agtaccatgg	cggaactctg	cgtatctctg	agtgggcccc	catcatcaac		2400		
tgcagtatct	tgcccgggtga	gggtatcgtc	gaggtctctg	cccagactgc	ttcgcccgag		2460		
gacttcccct	acgggtccga	gaggggcctt	ttgatccttg	cggagatgac	ctccaaggga		2520		
tctttggcta	ccgggtcaata	tactacttct	tctgttgact	atgctcgga	gtataagaag		2580		
tttgtgatgg	gattcgtctc	gacacgtcac	cttgccgagg	ttcagtcctga	agttagctcg		2640		
ccttcggagg	aggaagattt	tgtcgtcttc	acgacaggtg	tcaacctctc	ctcgaagggt		2700		
gacaagctgg	gacagcagta	ccaaactcct	gagtcggctg	ttggacgcgg	tgccgacttt		2760		

attattgctg	gccgtggaat	ttatgctgct	cctgatcccg	tggaggcggc	gaagcagtag	2820
cagaaggagg	gatgggatgc	atacctgaag	cgtgttggtg	cgcaataagt	agtgggtggat	2880
acgtactcct	tttatggcag	tatgtcgcaa	gtatgatgcg	atttataaat	tcagcactcg	2940
aaatgactac	tactatgtgt	ctacgacaga	taccctctcc	gtacgaataa	gacacctgcc	3000
tcgatatatg	gacaaattca	aaatcagggt	caagggtcat	gtttcaaagt	cacaacaatc	3060
tccaacatag	acgagaattt	gtaccggagt	gtctgaagg	gcagctggag	attgggtctat	3120
tttcttagag	tgggggtaca	ctaattgtaca	gtcggtcact	atcgtacaaa	caatcacaa	3180
tatatacaag	atttcccacc	accccctact	ctaaccacggc	acaattatcc	atcgagtcag	3240
agcctagcca	ccattttggtc	tatgtagaat	agatagtagt	aggggtgtca	tcctccatta	3300
tctgacaagc	agataccagc	atgtggccac	gtgtagtgt	gtttgtgcaag	tcattaagct	3360
gcagagggtca	tgacttcaca	gtgactagag	ttgagtaact	tgagatatgg	aaagtacacg	3420
gtgtactcaa	ggtacagtag	aattggagggt	tgacgacgtc	cittgtccac	gtatcctaaa	3480
ggagcatatg	gccggggggc	tgaaaaatag	ggcaacttca	agaaagatta	cctaacagggt	3540
gcttctagcc	aatgaaacac	atttctagta	tgctcgggtg	accgcttgca	gaagatccat	3600
gaacactttg	gggcaatttc	tacggcagaa	tacaggcggg	tggccaaggc	ggatgggtca	3660
ttctggtccc	tgaaaggcga	tgacccaatc	acaacgcgag	gagtaacctg	cagagcgatg	3720
gccgtaccag	ccatgctgtc	citttctctg	gcctcaagca	ttttaaaaaa	gctgatctct	3780
tcctctcacg	tccttcgttc	agcaacttct	ctcttttgcc	tcaactttcc	cttctctccc	3840
ccttccattc	cgctgcctgt	ggttgggttc	tccttttctt	ccttttgctt	ctcttctctg	3900
aatacaggaa	actgtattga	agcacaaggg	atttactacc	aacacgtctc	ccgataatca	3960
cacgcgtgcc	ccttggcgaa	gaccaccact	cgatatacac	ataggcaca	atg aga gaa	4018
					Met Arg Glu	
					1	
ggt att agt ttg aac g	gtatgtgttt	tctcaatctt	ttcctcacac	cgccctcttc		4074
Val Ile Ser Leu Asn						
					5	
accggagtcg	gtctgttgac	cgttactaat	tgttttttta	cgtgtgcag	tt ggt cag	4131
					Val Gly Gln	
					10	
gct ggt tgc cag atc gcc aat tct tgc tgg gag	gtacggagat	ggaactatac				4184
Ala Gly Cys Gln Ile Ala Asn Ser Cys Trp Glu						
					15	
tagaagatag	aggtgtttgg	aatgtgacta	ttgaaactga	cattatgggt	ctag ctc	4241
					Leu	
tac tgt ctt gag cac ggc atc cag	gtacgtggcc	acgatgacaa	tcctccatc			4295
Tyr Cys Leu Glu His Gly Ile Gln						
					25	
ttcatctcca	tcgggcacac	agactaacgg	gtatacgata	g ccc gat ggt tac ttg		4351
				Pro Asp Gly Tyr Leu		
					35	
acc gag gaa cgc aag aag gaa gac cct gac cat ggt ttc agc acc ttc						4399
Thr Glu Glu Arg Lys Lys Glu Asp Pro Asp His Gly Phe Ser Thr Phe						
					40	
ttc tcc gaa act g	gtacgtcgta	ttgacttgtt	atcttgccgc	cgagccctct		4452
Phe Ser Glu Thr						
					55	
aacaggctat	ggtgtcttag	gc cag ggc aag tat gtt cct cgt acc atc tac				4504
		Gly Gln Gly Lys Tyr Val Pro Arg Thr Ile Tyr				
					60	
gcc gat ctg gag ccc aat gtt gtc gat gag gtc cgc act ggc acc tac						4552
Ala Asp Leu Glu Pro Asn Val Val Asp Glu Val Arg Thr Gly Thr Tyr						
					70	
cgt acc ctt ttc cac ccc gag aac atg atc acc ggc aag gag gat gcc						4600
Arg Thr Leu Phe His Pro Glu Asn Met Ile Thr Gly Lys Glu Asp Ala						
					85	
tcg aac aac tat gcc cgt ggt cac tac acc gtt ggc aag gag atg atc						4648
Ser Asn Asn Tyr Ala Arg Gly His Tyr Thr Val Gly Lys Glu Met Ile						
					100	
gac cag gtc ctc gac aag gtt cgc cgt gtg gcc gac aac tgc gct ggt						4696
Asp Gln Val Leu Asp Lys Val Arg Arg Val Ala Asp Asn Cys Ala Gly						
					120	
ctc cag ggc ttc ctc gtc ttc cac tct ttc ggt ggt ggt act ggt tcc						4744
Leu Gln Gly Phe Leu Val Phe His Ser Phe Gly Gly Gly Thr Gly Ser						
					135	
ggt ttc ggt gct ctc ctg atg gag cgt ctg tct gtg gac tac ggc aag						4792
Gly Phe Gly Ala Leu Leu Met Glu Arg Leu Ser Val Asp Tyr Gly Lys						
					150	
aag tcc aag ctg gag ttc tgc gtc tac cct gcc ccc cag aat gcc acc						4840
Lys Ser Lys Leu Glu Phe Cys Val Tyr Pro Ala Pro Gln Asn Ala Thr						
					165	
tcc gtc gtt gag ccc tac aac tcc atc ctg act acc cac acc acc ctt						4888
Ser Val Val Glu Pro Tyr Asn Ser Ile Leu Thr Thr His Thr Thr Leu						
					180	
gag cac tcc gac tgc agt ttc atg gtt gac aac gaa gcc atc tac gac						4936
Glu His Ser Asp Cys Ser Phe Met Val Asp Asn Glu Ala Ile Tyr Asp						
					200	
atc tgc cgc cgc aac ctt ggc atc gag cgt ccc agc tat gag aac ctg						4984

```

Ile Cys Arg Arg Asn Leu Gly Ile Glu Arg Pro Ser Tyr Glu Asn Leu
      215      220      225
aac cgc ctg att gct cag gtt gtc tcc tcc atc acc gcc tcc ctg cgt      5032
Asn Arg Leu Ile Ala Gln Val Val Ser Ser Ile Thr Ala Ser Leu Arg
      230      235      240
ttc gat ggt tcc ctg aac gtg gat ctc aac gag ttc cag acc aac ctg      5080
Phe Asp Gly Ser Leu Asn Val Asp Leu Asn Glu Phe Gln Thr Asn Leu
      245      250      255
gtt ccc tac ccc cgt att cac ttc cct ctc gtt gcc tat gct ccc gtc      5128
Val Pro Tyr Pro Arg Ile His Phe Pro Leu Val Ala Tyr Ala Pro Val
      260      265      270      275
atc tcc gcg gcc aag gcc tcc cac gag gcc aac tcc gtc aac gag atc      5176
Ile Ser Ala Ala Lys Ala Ser His Glu Ala Asn Ser Val Asn Glu Ile
      280      285      290
acc atg tct tgc ttc gag ccc aac aac cag atg gtc aag tgt gac ccc      5224
Thr Met Ser Cys Phe Glu Pro Asn Asn Gln Met Val Lys Cys Asp Pro
      295      300      305
cgc aat ggc aag tac atg gct act tgc ttg ctg tac cgt ggt gat gtt      5272
Arg Asn Gly Lys Tyr Met Ala Thr Cys Leu Leu Tyr Arg Gly Asp Val
      310      315      320
gtg ccc aag gag acc cac gcc gcc gtt gct acc ctc aag acc aag cgt      5320
Val Pro Lys Glu Thr His Ala Ala Val Ala Thr Leu Lys Thr Lys Arg
      325      330      335
acc atc cag ttc gtc gac tgg tgc cct act ggt ttc aag atc ggt atc      5368
Thr Ile Gln Phe Val Asp Trp Cys Pro Thr Gly Phe Lys Ile Gly Ile
      340      345      350      355
tgc tac cag ccc cct aag cag gtt ccc aac ggt gac ctt gcc aac ctc      5416
Cys Tyr Gln Pro Pro Lys Gln Val Pro Asn Gly Asp Leu Ala Asn Leu
      360      365      370
agc cgt gct gt gtaagttccc tatctaccac tcgtgcaaga gataaaaagc      5467
Ser Arg Ala Val
      375
tgaccagtta tag c tgc atg ctg tct aac acc acc gcc atc t      5509
      Cys Met Leu Ser Asn Thr Thr Ala Ile
      380

```

[0552]

```

<210> 252
<211> 5509
<212> DNA
<213> 合成
<220>
<221> 基因
<222> (1970)..(2868)
<223> pyrG
<220>
<221> 启动子
<222> (3260)..(4009)
<223> clxA 启动子
<220>
<221> 启动子
<222> (4004)..(4009)
<223> TIS
<220>
<221> 基因
<222> (4010)..(5474)
<223> ERG10 必需基因
<220>
<221> 外显子
<222> (4010)..(4074)
<220>
<221> 外显子
<222> (4142)..(4191)
<220>
<221> 外显子
<222> (4249)..(4324)
<220>
<221> 外显子
<222> (4398)..(4887)
<220>
<221> 外显子
<222> (4956)..(5474)
<400> 252
aacgcgacgt tgccttttct tgggcttagt cgccattttt cgccttgtcc aggcaagaca      60
ctttcaatga cagtgttgac ttcttcaatc acatgcttga gcttgaaata ggcgtcgttg      120
ggtgggtatc tcagttcgag gatcagccga tccattttgt tcacaacccat gggtcaacgga      180
aggccctcta gcacagcgtg cttgatgatt tgcctcgtat tcgattgtac tccttccaca      240
atatcaacga ctatcaccac gccatcaacc agtcgtgatg aggcagctac ttcattcaaca      300

```


[0553]

```

aaattcacat gccccggggt atcaagaatg ttgaacaaat gagacttccc cttagtgtccc 360
tgaagaacta ggctcatagg agctgattta atagaagggc ctgcgtctct ctctagaaaag 420
tgaatatcgg tataacgaag ctgtacttcc ttctttttac ctgcgtcgtt ctctagccgt 480
tcaggtcaaat catgcgtttg cataacaagc atatccatga acgcggtott cccatgatgt 540
aaatgaccga caagcgcaat gtcccttatt tgttcaggaa aattcagaag atcagtcattg 600
aattcccttg aaaagtgaac ccggggggagc tcagttctctt caatggcgaa cttctttttgc 660
tgtacggggg caacgatagg ctgggagaga ggtctgtcgt cttcttcttg aactaatgtc 720
tctacctctg tcccgtagac ctgttgtgag ctagggtagt attgtttgtc ttcgtgtaga 780
atgacggcat ttgagggggc ttctgtaatg agatatatca gcaccgacta aatccagggg 840
cttttaagaga gtgcccaccc tccacctcca tagactgtcg cccatataat tcttcatctt 900
cttctctccc aaacgcctcg ttgaaggcga aagcccaggg cttcacttcg tcgtgtttggc 960
cgtcttctac tgaatctgca gctcttccaa tgtagtgtgc aaacctgtgt ggtcttgatt 1020
agaatttgca caattagagc gttagattgg atgcaactta ctcactgtag agatcatcca 1080
tcgtgcggtg ggaaagggtg ttcaagctgc caaaaacttg tcaccagtga gctgtgcgat 114
ctctttttgt cccaagctgc tgaagtcctg ctaatggact caccgagtgg tccgtgcgtc 1200
gaccctacca ttttggtagc ccgcgcgcgc ggtagccacg tccaggaggg gtccttaaat 1260
gacttttgatt tctcaccacc tcttctcgtg tcagagatac ctcactgcta tcgtcggctg 1320
aataactcaa ctctttcttg tttcccttct tccccctttt aatcccacca atcttcttta 1380
ttaccggcaa ttctctgtct caccgggtgc ctttaccaca caccgaacct ctttaaacgc 1440
ttcacgctac agaagctttt ggtttacgac tcttctaate agctatcgta ggccgcgat 1500
agctgattaa atattctaga cccaagccgc tgcctgaatt tgcttgaatt gacattatta tggccgatag 1560
ggttggtgtt attgctatgt ccctgaaagg atatcaaaaag caggcaaaaa gccaggcata 1620
atcccccgct ggacggtagc ctaaggatag gccctaatct tatctacatg tgactgcctc 1680
gatgtgtttg gtcaaaaatga ggcatgtggc tcaccccaca cggcgagaaa cgtgtggcta 1740
gtgcatgaca gtccccctcca tagattcaat ttaatttttc gcggcaattg tcgtgcagtt 1800
tgtatctaca tttcattcca tatatcaaga gttagttagt ggacatcctg attattttgt 1860
ctaattagtg aaaactcgaa gtactaacct actaataaag cagtttcaac cactaatgtc 1920
tcatttatac aatatttgca gaaccccgcg ctacccctcc atcgccaaca tgtcttccaa 1980
gtcgcatttg acctacagcg caccgcgtag caagcacccc aatgcgctcg taaagaagt 2040
cttcgaggtt gccgaggcca agaaaaccaa tgtcagctgt tccgcccagc tgacaaccac 2100
caaagagctg ctggatttgg ctgaccgtat gcgcaccggg gatgccactt acatgtgatc 2160
tagtaatggt taatggtgga ttatataaca ggactcggtc cgtacattgc cgtgatcaaa 2220
actcacatcg atatctctc ogatttcagc gaagaaacca tcaccgtctt gaaggccctt 2280
gcagagaagc acaatttctt catcttcgaa gatcgcaagt tcatcgatat cggaaacaca 2340
gtccaaaagc agtaccatgg cggcactctg cgtatctctg agtgggcca catcatcaac 2400
tgcatatttc tgcccgtgga ggggtatcgt gaggtctctg cccagactgc ttcggccgag 2460
gacttccctt accgtccga gaggggcctt ttgactcttg cggagatgac ctccaaggga 2520
tctttggcta ccggtcaata tactacttct tctgttgact atgctcggaa gtataagaag 2580
tttgtgatgg gattcgtctc gacacgtcac cttggcgagg ttccagtctg agttagctcg 2640
ccttcggagg aggaagattt tctcgtcttc acgacagggt tcaacctctc ctogaagggt 2700
gacaagctgg gacagcagta ccaaactcct gagtcggctg ttggacggcg tgcgcacttt 2760
attattgctg gccgtggaat ttatgctgct cctgatcccg tggaggcggc gaagcagtag 2820
cagaaggagg gatgggatgc atacctgaag cgtgttggtg cgcaataagt cgtgggtgat 2880
acgtactcct tttatggcag tatgtcgcaa gttatgtgag atttataaat tcagcactcg 2940
aaatgactac tactatgtgt ctacgacaga taccctctcc gtacgaataa gacacctgac 3000
tcgatatatg gacaaattca aaatcagggt caagggtcat gtttcaaagt cacaacaatc 3060
tccaacatag acgagaattt gtaccggagt gcttgaagg gtagctggag atttgtctat 3120
tttcttagag tgggggtatc ctaatgtaca gtcggtcact atcgtacaaa caatcacaa 3180
tatatacaag atttcccacc accccctact ctaaacaggc acaattatcc atcgagtcag 3240
agccttagca ccatttggtg gtgtgtgtct aagttttgca aagttagtga agaggggaaa 3300
agaacgtcgg cggggagagg ggcaacgtga agaactccat tttttttttt tctggatatg 3360
gataattaac tagtagcggg aagtagatcg cgcaatcggt attttgcaagg atagtatgca 3420
attacttcga agattaagaa ggaaacaaga ggaataaagt aaaagaattt atttatattc 3480
atltgaaatt tacattgact ggattgaatt gtcttatttg aattgggtga ctagcatgtt 3540
tgggagaagt gtactaatat gaatgaatgg tatggagtac tacaataaac atacagggac 3600
gctaagaccg tgacaaccaa agagagaaaa caaaaaagc aataaaaagg ttctttagaa 3660
tcaatggaca ctggacagaa cattggagaa tctgccacta atcctgtttg gtcggttgga 3720
actttacccc aaggggcaag cggcagcgga agaggccaga ttgaacttaa gctaccttcc 3780
cagcttcata taatcaacct ccccaacact cgtcttctct ttctctctct ctctctcttc 3840
ttccttctgt cgtgcggtta gcttgtggtg cgtctcaacc ttgacactct tttcccctc 3900
gaacctctac ggaggaccaa ttgcttgatc gtttgcttat ttgtttctt ctttctctaa 3960
gctactaccg agtcgggctt cttagtttga gagcccgta tcattcctc atg tct tct 4018
Met Ser Ser
1
ctt ccg ccg gtc tat att gtt tcc tct gcc cgc acc cca gtc ggc tct 4066
Leu Pro Pro Val Tyr Ile Val Ser Ser Ala Arg Thr Pro Val Gly Ser
5 10 15
ttc ttg gg gtgagattgg ccttgtgcct tgtctatacg tacaatattg 4114
Phe Leu Gly
20
acttacactg actatccgct toacag g tgc etc tca agt etc act gcc ccg 4166
Ser Leu Ser Ser Leu Thr Ala Pro
25 30
cag tta ggc tct cat gct att aaa g gtatgacagc tctcagtggt 4211
Gln Leu Gly Ser His Ala Ile Lys
35
cggtagctgt tgatatacta acaaaactcga aattcag ct gcg etc agc aaa gcg 4265
Ala Ala Leu Ser Lys Ala

```

222

223

attattgctg	gccgtggaat	ttatgctgct	cctgatcccg	tggaggcggc	gaagcagtac	2820
cagaaggagg	gatgggatgc	atacctgaag	cgtgttggtg	cgcaataagt	agtgggtggat	2880
acgtactcct	tttatggcag	tatgtcgcaa	gtatgatgcg	atttataaat	tcagcactcg	2940
aaatgactac	tactatgtgt	ctacgacaga	tacctctcc	gtacgaataa	gacacctgcc	3000
tcgatatatg	gacaaattca	aaatcagggg	caagggtcat	gtttcaaagt	cacaacaatc	3060
tccaacatag	acgagaattt	gtaccggagt	gtctgaagg	gcagctggag	attgggtctat	3120
tttcttagag	tgggggtatca	ctaattgtaca	gtcggtcact	atcgtacaaa	caatcacaa	3180
tatatacaag	atttcccacc	acccccctact	ctaaccacggc	acaattatcc	atcgagtcag	3240
agcctagcca	ccatttggta	gtgtgggtctg	aagttttgca	aagtgaagtga	ggaggggaaa	3300
agaacgtcgg	cggggagagg	ggcaacgtga	agaactccat	tttttttttt	tctggatatg	3360
gataattaac	tagtagcggg	aagtagatcg	cgcaatcggg	atttgcaagg	atagtatgca	3420
attacttcga	agattaagaa	aggaacaaga	ggaaaaatag	aaaagaattt	atttatattc	3480
atttgaaatt	tacattgact	ggattgaatt	gtcttatttg	aattgggtga	ctagcatgtt	3540
tgggagaagt	gtactaatat	gaatgaatgg	tatggagtag	tacaaaatac	atacagggac	3600
gctaagaccg	tgacaaccaa	agagagaaaa	caaaaaaagc	aataaaaacg	ttcttttagaa	3660
tcaatggaca	ctggacagaa	cattggagaa	tctgccacta	atcctgtttg	ggtcggtgga	3720
actttacccc	aaggggcaag	cggcagcggg	agagccaga	ttgaacttaa	gctaccttcc	3780
cagcttcata	taatcaacct	ccccaacact	cgtcttctct	ttctctctct	ctctctctct	3840
ttccttctgt	cgtgcggtta	gcttgttgga	cgtctcacc	ttgacactct	tttccccctc	3900
gaacctctac	ggaggaccaa	ttgcttgatc	gtttgcttat	ttgtttcttt	cttctctaaa	3960
gctactaccg	agtcgggctt	cttagtttga	gagcccggtta	tcattcctc	atg tct	4014
					Met Cys	
					1	
gtacgttccc	gaagacccag	gcttactgaa	tggttgatgt	aactaattga	cccctag c	4072
ggc atc ttc	ggc tac att	aac tac ctc	gtc gag agg	gac cgc aag	ttt	4120
Gly Ile Phe	Gly Tyr Ile	Asn Tyr Leu	Val Glu Arg	Asp Arg Lys	Phe	
	5	10	15			
att ctt gac	act ctt ctc	aat g gtt	ggcattc ctgtt	gtttt caccct	cact	4172
Ile Leu Asp	Thr Leu Leu	Asn				
	20	25				
ggaccttggt	tactctgggt	ttgggagtgt	ttgtgtattc	ctctgacata	ttgaag ga	4230
					Gly	
ctc tcc cgt	ctt gaa tac	cga ggc tac	gac tct gcg	ggt ctt gcc	ata	4278
Leu Ser Arg	Leu Glu Tyr	Arg Gly Tyr	Asp Ser Ala	Gly Leu Ala	Ile	
	30	35	40			
gat ggc gat	aag aag aat	gaa gtt tgt	gct ttc aag	gaa gtg ggc	aag	4326
Asp Gly Asp	Lys Lys Asn	Glu Val Cys	Ala Phe Lys	Glu Val Gly	Lys	
	45	50	55			
gtt gct aag	ctt cga gag	ctt att gat	gag tgc aag	cct gac ttg	acc	4374
Val Ala Lys	Leu Arg Glu	Leu Ile Asp	Glu Cys Lys	Pro Asp Leu	Thr	
	60	65	70			
aaa tcc ttt	gaa tct cat	gct ggt att	tct cac act	cgt tgg gcc	act	4422
Lys Ser Phe	Glu Ser His	Ala Gly Ile	Ser His Thr	Arg Trp Ala	Thr	
	75	80	85	90		
cat ggg acg	cct tcc ttg	aac tgc cac	ccc cac ag	gtgagtcga		4470
His Gly Thr	Pro Ser Arg	Leu Asn Cys	His Pro His	Arg		
	95	100				
tggtccacgg	tttaatcagc	tcgactagga	tcaggcttgt	ttgctaattt	tctgggcacg	4530
cag g tcc	gat ctg aac	tgg gaa ttt	tca gtt gtc	cac aat ggt	atc att	4579
	Ser Asp Leu	Asn Trp Glu	Phe Ser Val	Val His Asn	Gly Ile Ile	
	105	110	115			
acc aac tac	aaa gag ctc	aag gct ttg	ttg gag agc	aag gga ttc	cgc	4627
Thr Asn Tyr	Lys Glu Leu	Lys Ala Leu	Leu Glu Ser	Lys Gly Phe	Arg	
	120	125	130			
ttt gag act	gac aca gac	acc gaa tgt	att gca aag	ctt acc aaa	tat	4675
Phe Glu Thr	Asp Thr Asp	Thr Glu Cys	Ile Ala Lys	Leu Thr Lys	Tyr	
	135	140	145	150		
ctc tat gac	caa cag ccc	gat atc gaa	ttc aca gtt	ttg gct aag	gct	4723
Leu Tyr Asp	Gln Gln Pro	Asp Ile Glu	Phe Thr Val	Leu Ala Lys	Ala	
	155	160	165			
gtt gtg aag	gag ctt gag	ggc gcg ttt	ggg ctt ctc	atc aaa tcc	gta	4771
Val Val Lys	Glu Leu Glu	Gly Ala Phe	Gly Leu Ile	Lys Ser Val		
	170	175	180			
cat tat ccc	cat gag gtc	att gcg gct	cgc aag ggg	tcg ccg ctt	gtt	4819
His Tyr Pro	His Glu Val	Ile Ala Arg	Lys Gly Ser	Pro Leu Val		
	185	190	195			
att ggc gtg	aga acg tcg	aag aag atg	aag gtg gat	ttc gtg gac	gtc	4867
Ile Gly Val	Arg Thr Ser	Lys Met Lys	Val Asp Phe	Val Asp Val		
	200	205	210			
gaa tac tca	gaa gat gga	gcc ctt cct	gca gag cag	gcc tcc cag	aac	4915
Glu Tyr Ser	Glu Asp Gly	Ala Leu Pro	Ala Glu Gln	Ala Ser Gln	Asn	
	215	220	225	230		
gtg gcc atc	aag aag tct	gcg acc ggc	ctc ctt gcc	cca cct gac	aaa	4963
Val Ala Ile	Lys Lys Ser	Ala Thr Gly	Leu Leu Ala	Pro Pro Asp	Lys	
	235	240	245			
tcc ctt ttg	cat agg tca	cag tct cgc	gct ttc ctt	tcc gat gat	ggc	5011
Ser Leu Leu	His Arg Ser	Gln Ser Arg	Ala Phe Leu	Ser Asp Asp	Gly	

		250		255		260		
	atc ccc cag cca gct gag ttt ttc ttg tct tct gac cca tct gca att							5059
	Ile Pro Gln Pro Ala Glu Phe Phe Leu Ser Ser Asp Pro Ser Ala Ile							
	265			270				
	gtc gag cat acg aag aag gtc ctc tat ctc gaa gat gat gat atc gct							5107
	Val Glu His Thr Lys Lys Val Leu Tyr Leu Glu Asp Asp Asp Ile Ala							
	280			285				
	cat gtg cat gag gga cag tta aac att cat cgc ctc aca aaa gat gat							5155
	His Val His Glu Gly Gln Leu Asn Ile His Arg Leu Thr Lys Asp Asp							
	295			300				310
	ggt act tct aat gtt cgc gct att cag aca att gaa ctc gaa ttg cag							5203
	Gly Thr Ser Asn Val Arg Ala Ile Gln Thr Ile Glu Leu Glu Leu Gln							
	315			320				325
	gag att atg aag ggc aag ttt gac cac ttc atg caa aag gaa att ttc							5251
	Glu Ile Met Lys Gly Lys Phe Asp His Phe Met Gln Lys Glu Ile Phe							
	330			335				340
	gag caa ccc gag tct gtg atc aat acc atg aga gga cga ctg gat gtt							5299
	Glu Gln Pro Glu Ser Val Ile Asn Thr Met Arg Gly Arg Leu Asp Val							
	345			350				355
	gca aac aag caa gtc aca ctc ggt ggc ctg cgg cag tac att tct act							5347
	Ala Asn Lys Gln Val Thr Leu Gly Gly Leu Arg Gln Tyr Ile Ser Thr							
	360			365				370
	atc cgc cgc tgc aga aga atc ata ttt gtt gcc tgt gga act agt tac							5395
	Ile Arg Arg Cys Arg Arg Ile Ile Phe Val Ala Cys Gly Thr Ser Tyr							
	375			380				385
	cat tca tgc atg gct gtg cgt gga gtc ttt gaa gag ctt act gag att							5443
	His Ser Cys Met Ala Val Arg Gly Val Phe Glu Glu Leu Thr Glu Ile							
	395			400				405
	cct att tct gtt gaa ctt gct tcc gat ttc ctg gat aga cag gct cca							5491
	Pro Ile Ser Val Glu Leu Ala Ser Asp Phe Leu Asp Arg Gln Ala Pro							
	410			415				420
	gta ttc cgt gat gac acg							5509
	Val Phe Arg Asp Asp Thr							
	425							
	<210> 254							
	<211> 5509							
	<212> DNA							
	<213> 合成							
[0557]	<220>							
	<221> 基因							
	<222> (1970) .. (2868)							
	<223> pyrG							
	<220>							
	<221> 启动子							
	<222> (3260) .. (4009)							
	<223> clxA 启动子							
	<220>							
	<221> 启动子							
	<222> (4004) .. (4009)							
	<223> TIS							
	<220>							
	<221> 基因							
	<222> (4010) .. (5509)							
	<223> TUB1 必需基因							
	<220>							
	<221> 外显子							
	<222> (4010) .. (4034)							
	<220>							
	<221> 外显子							
	<222> (4124) .. (4164)							
	<220>							
	<221> 外显子							
	<222> (4239) .. (4265)							
	<220>							
	<221> 外显子							
	<222> (4337) .. (4412)							
	<220>							
	<221> 外显子							
	<222> (4473) .. (5427)							
	<220>							
	<221> 外显子							
	<222> (5481) .. (5509)							
	<400> 254							
	cggaagatg ttccatagaa agggatcaag gatgatatcc gctgcaacog ggggaacatt							60
	ctccattagt gctttgttaa gagaaaagta gtccgccgcc gcagtaacac gaaggaaaag							120
	caccgaatca gacgtaggtg cttgatcatg aaaagaggaa gatttgcgag gaacagcgtt							180

[0558]

agactgaggg	gagggcgagat	gggcttgagg	atagaggcta	atcgaggaga	agagtgcggg	240
atcatcaatg	gcttgtggaa	gagtatgcgt	tgtcagggta	aaggcgggctg	gttgcttttga	300
ctacttgta	gctaatttttc	tttgggtcgcc	ttgctggggca	ctggtagagg	gtatatgagg	360
acgtacagtt	gccaaccaac	acaccgcac	ctcataccgc	tggccggggg	tcaagttctc	420
cagaaaatac	catgagtcctg	taccaggga	ctcggttggtc	ggaaatgtag	cgtaatttct	480
ggttcgaaga	acgggactag	aggaggagag	gcgatctaga	cctagggtcat	cgagggtagg	540
atcgaccgtg	gggatgatga	gggggttcgg	agcgatgaag	atcgttttct	cgacattggc	600
gatagcgggg	aatgcgaggg	tgcaagta	gagtagatat	tggagtagca	tcacacagca	660
gagttcgcta	tggtcaattg	acgctgctgc	tagacgggtca	tggttctagt	ccacacgcag	720
ttgagtatgg	taggttatat	gtatgtatgt	agtaaatgag	gtagatctga	tactaactgt	780
tagattcccta	agtgaagtcc	ttacatgaaa	ttgattcaat	gtcaatacga	tcttaccata	840
aattaggcca	ttgtgcgcag	caatgttgg	gactagaaat	agagcactgt	gtgtccgaga	900
cttgtttctat	gtgagtatatt	tcaatagtac	ttactatcgc	aagctgtacg	gatagtcaaa	960
tacaatgcag	ctgcacgaag	agatgcgctg	attgacacgt	gatcgacct	tgccttgaac	1020
taagcccaaa	taagtggctg	agtcagcgg	gggctggaag	ggcgggtcacg	gcatcggaca	1080
agccaatcag	aaatgggatt	gtgggcgacg	cgacccaacg	gatggatcaa	caatgaaaag	1140
gtggaagtcc	tgctttttggg	ttgacagcgc	atcgcaagac	cacaacaaat	tgaccgatca	1200
ttcgggtctt	gggcaattgt	ttaatgtatc	cccttagtct	atttttctcc	tcgtcaatca	1260
ccccagggt	ttgttctctc	ccccatttc	ctcccaccgg	ttcccccttc	cctttctttt	1320
ggttcttttc	ccagggtggg	ttttttctgt	ttacttacct	ctgaaacttc	cttacagatc	1380
agggcatttt	tccccctcta	aagtcccttc	acattccacc	cactccgttt	cacacacccc	1440
cctccttcca	atcttaatttc	ctttgtcact	ccgacattcg	tcccccttag	actcgacaag	1500
agctgattaa	atattctaga	cccaagccgc	tgctggaatt	gacattatta	tggccgatag	1560
ggttgggctt	attgctatgt	ccctgaaagg	atatcaaaag	caggcaaaaa	gccagcata	1620
atccccgcgt	ggacgggtacc	ctaaggatag	gcccataatc	tatctacatg	tgactgcac	1680
gatgtgtttg	gtcaaaatga	ggcatgtggc	tcacccacac	ggcggagaaa	cgtgtggcta	1740
gtgcatgaca	gtccccctcca	tagattcaat	ttaatttttc	gcggcaattg	tcgtgcagtt	1800
tgtattcaca	tttcattcca	tatatcaaga	gttagtagtt	ggacatcctg	attattttgt	1860
ctaattactg	aaaactcgaa	gtactaacct	actaataagc	cagtttcaac	cactaagtgc	1920
tcatttatatc	aatattttgca	gaaccccgcg	ctacccctcc	atcgccaaca	tgtcttccaa	1980
gtcgcaattg	acctacagcg	caacgcctag	caagcaccac	aatgcgctcg	taagagagct	2040
cttcgaggtt	gccgaggcca	agaaaaccac	tgtcacccgt	tccgcccagc	tgacaaccac	2100
caaagagctg	ctggattttg	ctgaccgtat	gcgcaccggg	gatgccactt	acatgtgatc	2160
tagtaattgg	taatgggtgga	ttatataaca	ggactcggtc	cgtacattgc	cgtgatcaaa	2220
actcacattc	atatcctctc	cgatttcagc	gaagaaacca	tcaccggctc	gaaggccctt	2280
gcagagaagc	acaatttctc	catcttcgaa	gatcgcaagt	tcacgatgat	cggaacacac	2340
gtccaaaagc	agtaccatgg	cggcactctg	cgtatctctg	agtggggcca	catcatcaac	2400
tgcaagtattc	tgcccggtga	gggtatcgtg	gaggetctgg	cccagactgc	tccggccgag	2460
gacttcccc	acggctccga	gaggggcctt	ttgatccttg	cggagatgac	ctccaaggga	2520
tttttggcta	ccgggtcaata	tactactctc	tctgttgact	atgctcggaa	gtataagaag	2580
tttgtgatgg	gattcgtctc	gacacgtcac	cttgccgagg	ttcagtcctg	agttagctcg	2640
ccttcggagg	aggaagattt	tgtcgtcttc	acgacagggt	tcaaacctctc	ctcgaagggt	2700
gacaagctgg	gacagcagta	ccaaaactct	gagtcggctg	ttggacgcgg	tgccgacttt	2760
attattgctg	gccgtggaat	ttatgctgct	cctgatcccg	tggaggcggc	gaagcagtag	2820
cagaaggagg	gattgggatgc	atacctgaa	cgtgttggtg	cgcaataagt	agtgctggat	2880
acgtactcct	tttatggcag	tatgtcgcaa	gtatgatgcg	atttataaat	tcagcactcg	2940
aaatgactac	tactatgtgt	ctacgacaga	tacctctctc	gtacgaataa	gacacctgcc	3000
tcgatataatg	gacaaaattca	aaatcagggt	caagggtcat	gtttcaaagt	cacaacactc	3060
tccaacatag	acgagaatttt	gtaccggagt	gtctgaagg	gcagctggag	attgggtctat	3120
tttcttagag	tgggggtatca	ctaattgtaca	gtcggctact	atcgtaaaaa	caatcacact	3180
tatatcaaac	atttcccacc	acccctact	ctaaccggc	acaattatcc	atcgagtcag	3240
agcctagcca	ccattttggta	gtgtggtctg	aagtttttgc	aagttagtga	ggaggggaaa	3300
agaacgtcgg	cgggggagagg	ggcaacgtga	agaactccat	tttttttttt	tctggatatg	3360
gataattaac	tagtagcggg	aagtagatcg	cgcaatcgg	atttgcaagg	atagtatgca	3420
attacttcga	agattaagaa	aggaacaaga	ggaaaaatag	aaaagaattt	atttatattc	3480
atttgaaatt	tacattgact	ggattgaatt	gtcttatttt	aattgggtga	ctagcatgtt	3540
tgggagaagt	gtactaatat	gaatgaatgg	tatggagtac	tacaaataac	atacagggac	3600
gctaagaccg	tgacaaccaa	agagagaaaa	caaaaaaagc	aataaaaacg	ttcttttagaa	3660
tcaatggaca	ctggacagaa	cattggagaa	tctgccacta	atcctgtttg	ggtcgggtgga	3720
actttacccc	aaggggcaag	cggcagcgga	agaggccaga	ttgaacttaa	gctaccttcc	3780
cagcttcata	taatcaacct	ccccaacact	cgtcttctct	ttctctctct	ctctctctct	3840
ttccttctgt	cgtgcgggta	gcttgtggta	cgtctctacc	ttgacactct	tttccccctc	3900
gaacctctac	ggaggaccaa	ttgcttgatc	gittgcttat	ttgtttcttt	ctttcctaaa	3960
gctactaccg	agtcgggctt	cttagtttga	gagcccggtta	tcatttcac	atg aga gaa	4018
					Met Arg Glu	
					1	
gtt att agt ttg aac g	gtatgtgttt	tctcaatctt	ttcctcacac	cgctctcttc		4074
Val Ile Ser Leu Asn						
5						
accggagtcg	gtctgttgac	cgttactaat	tgttttttta	cgtgtgcag	tt ggt cag	4131
					Val Gly Gln	
					10	
gct ggt tgc cag atc gcc aat tct tgc tgg gag	gtacggagat	ggaactatac				4184
Ala Gly Cys	Gln Ile Ala Asn Ser	Cys Trp Glu				
15		20				
tagaagatag	aggtgttttg	aatgtgacta	ttgaaactga	cattatgggt	ctag ctc	4241
					Leu	

	tac tgt ctt gag cac ggc atc cag gtacgtggcc acgatgacaa tccatccatc	4295
	Tyr Cys Leu Glu His Gly Ile Gln	
	25 30	
	ttcatctcca tcgggcacac agactaaccg gtatacgata g ccc gat ggt tac ttg	4351
	Pro Asp Gly Tyr Leu	
	35	
	acc gag gaa cgc aag aag gaa gac cct gac cat ggt ttc agc acc ttc	4399
	Thr Glu Glu Arg Lys Lys Glu Asp Pro Asp His Gly Phe Ser Thr Phe	
	40 45 50	
	ttc tcc gaa act g gtacgtcgta ttgacttggt atcttgccgc cgagccctct	4452
	Phe Ser Glu Thr	
	55	
	aacaggctat ggtgtcttag gc cag ggc aag tat gtt cct cgt acc atc tac	4504
	Gly Gln Gly Lys Tyr Val Pro Arg Thr Ile Tyr	
	60 65	
	gcc gat ctg gag ccc aat gtt gtc gat gag gtc cgc act ggc acc tac	4552
	Ala Asp Leu Glu Pro Asn Val Val Asp Glu Val Arg Thr Gly Thr Tyr	
	70 75 80	
	cgt acc ctt ttc cac ccc gag aac atg atc acc ggc aag gag gat gcc	4600
	Arg Thr Leu Phe His Pro Glu Asn Met Ile Thr Gly Lys Glu Asp Ala	
	85 90 95	
	tcg aac aac tat gcc cgt ggt cac tac acc gtt ggc aag gag atg atc	4648
	Ser Asn Asn Tyr Ala Arg Gly His Tyr Thr Val Gly Lys Glu Met Ile	
	100 105 110 115	
	gac cag gtc ctc gac aag gtt cgc cgt gtg gcc gac aac tgc gct ggt	4696
	Asp Gln Val Leu Asp Lys Val Arg Arg Val Ala Asp Asn Cys Ala Gly	
	120 125 130	
	ctc cag ggc ttc ctc gtc ttc cac tct ttc ggt ggt ggt act ggt tcc	4744
	Leu Gln Gly Phe Leu Val Phe His Ser Phe Gly Gly Gly Thr Gly Ser	
	135 140 145	
	ggt ttc ggt gct ctc ctg atg gag cgt ctg tct gtg gac tac ggc aag	4792
	Gly Phe Gly Ala Leu Leu Met Glu Arg Leu Ser Val Asp Tyr Gly Lys	
	150 155 160	
	aag tcc aag ctg gag ttc tgc gtc tac cct gcc ccc cag aat gcc acc	4840
	Lys Ser Lys Leu Glu Phe Cys Val Tyr Pro Ala Pro Gln Asn Ala Thr	
	165 170 175	
	tcc gtc gtt gag ccc tac aac tcc atc ctg act acc cac acc acc ctt	4888
	Ser Val Val Glu Pro Tyr Asn Ser Ile Leu Thr Thr His Thr Thr Leu	
	180 185 190 195	
[0559]	gag cac tcc gac tgc agt ttc atg gtt gac aac gaa gcc atc tac gac	4936
	Glu His Ser Asp Cys Ser Phe Met Val Asp Asn Glu Ala Ile Tyr Asp	
	200 205 210	
	atc tgc cgc cgc aac ctt ggc atc gag cgt ccc agc tat gag aac ctg	4984
	Ile Cys Arg Arg Asn Leu Gly Ile Glu Arg Pro Ser Tyr Glu Asn Leu	
	215 220 225	
	aac cgc ctg att gct cag gtt gtc tcc atc acc gcc tcc ctg cgt	5032
	Asn Arg Leu Ile Ala Gln Val Val Ser Ser Ile Thr Ala Ser Leu Arg	
	230 235 240	
	ttc gat ggt tcc ctg aac gtg gat ctc aac gag ttc cag acc aac ctg	5080
	Phe Asp Gly Ser Leu Asn Val Asp Leu Asn Glu Phe Gln Thr Asn Leu	
	245 250 255	
	ggt ccc tac ccc cgt att cac ttc cct ctc gtt gcc tat gct ccc gtc	5128
	Val Pro Tyr Pro Arg Ile His Phe Pro Leu Val Ala Tyr Ala Pro Val	
	260 265 270 275	
	atc tcc gcg gcc aag gcc tcc cac gag gcc aac tcc gtc aac gag atc	5176
	Ile Ser Ala Ala Lys Ala Ser His Glu Ala Asn Ser Val Asn Glu Ile	
	280 285 290	
	acc atg tct tgc ttc gag ccc aac aac cag atg gtc aag tgt gac ccc	5224
	Thr Met Ser Cys Phe Glu Pro Asn Asn Gln Met Val Lys Cys Asp Pro	
	295 300 305	
	cgc aat ggc aag tac atg gct act tgc ttg ctg tac cgt ggt gat gtt	5272
	Arg Asn Gly Lys Tyr Met Ala Thr Cys Leu Leu Tyr Arg Gly Asp Val	
	310 315 320	
	gtg ccc aag gag acc cac gcc gcc gtt gct acc ctc aag acc aag cgt	5320
	Val Pro Lys Glu Thr His Ala Ala Val Ala Thr Leu Lys Thr Lys Arg	
	325 330 335	
	acc atc cag ttc gtc gac tgg tgc cct act ggt ttc aag atc ggt atc	5368
	Thr Ile Gln Phe Val Asp Trp Cys Pro Thr Gly Phe Lys Ile Gly Ile	
	340 345 350 355	
	tgc tac cag ccc cct aag cag gtt ccc aac ggt gac ctt gcc aac ctc	5416
	Cys Tyr Gln Pro Pro Lys Gln Val Pro Asn Gly Asp Leu Ala Asn Leu	
	360 365 370	
	agc cgt gct gt gtaagtcccc tatctaccac tcgtgcaaga gataaaaagc	5467
	Ser Arg Ala Val	
	375	
	tgaccagtta tag c tgc atg ctg tct aac acc acc gcc atc t	5509
	Cys Met Leu Ser Asn Thr Thr Ala Ile	
	380。	

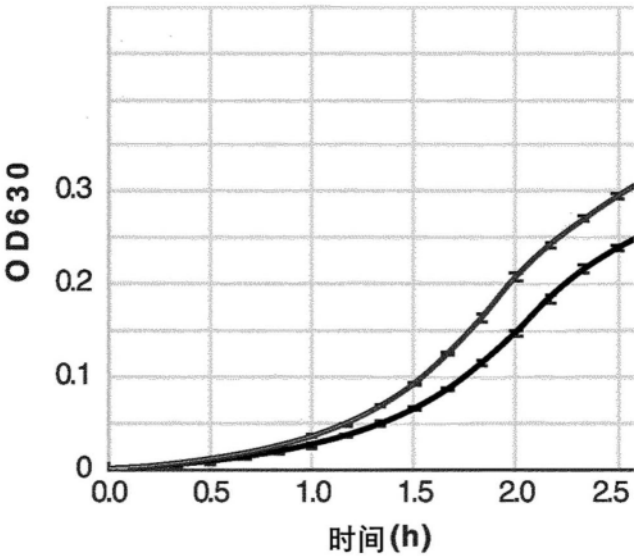


图1

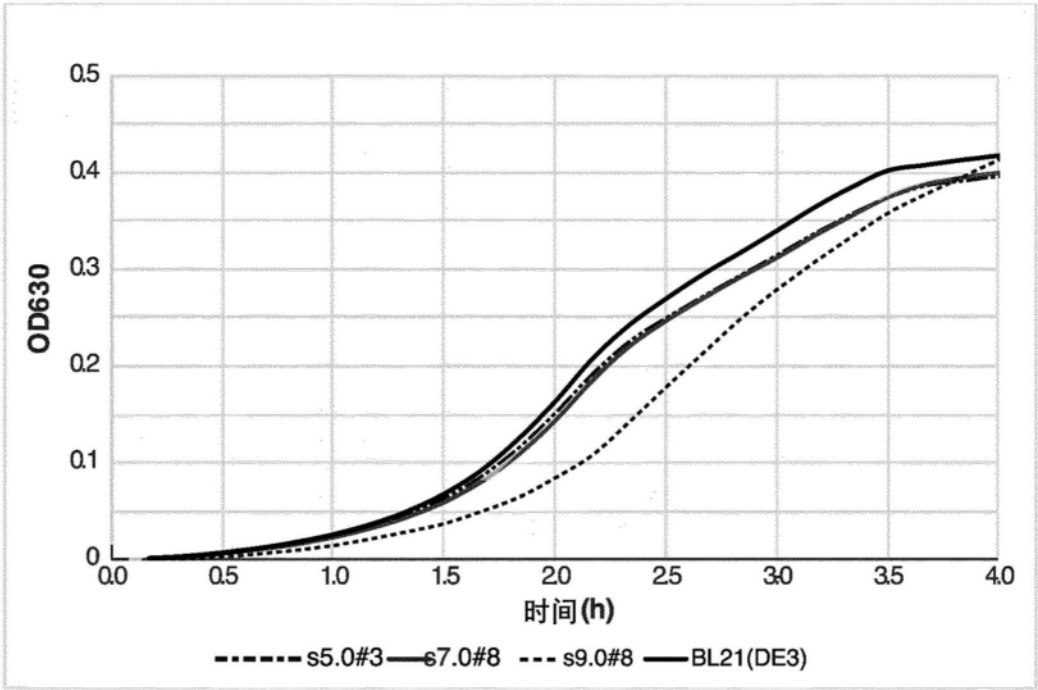


图2

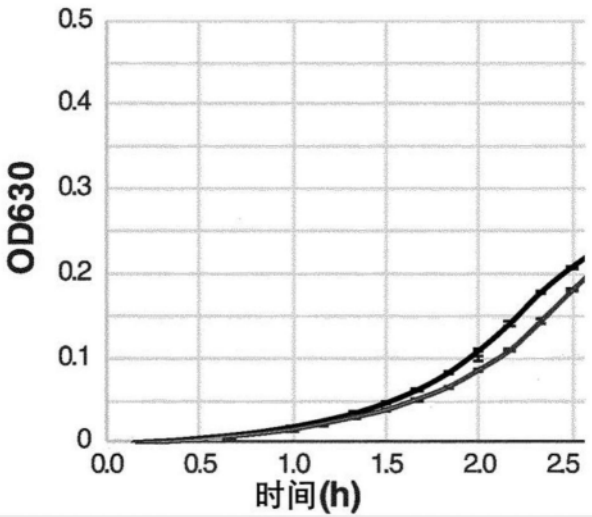


图3

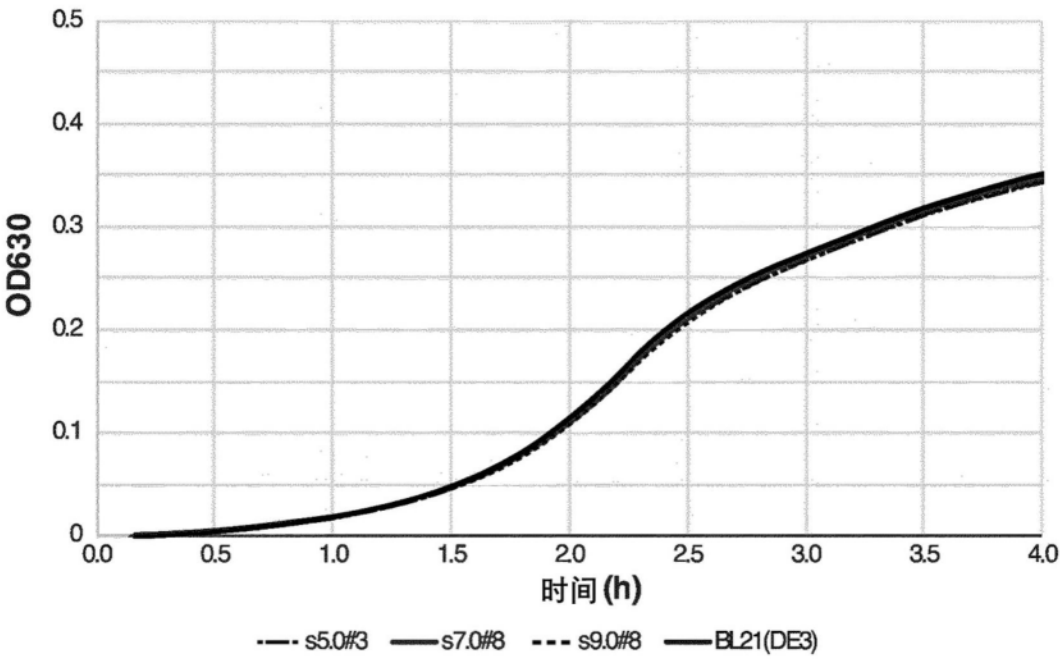


图4

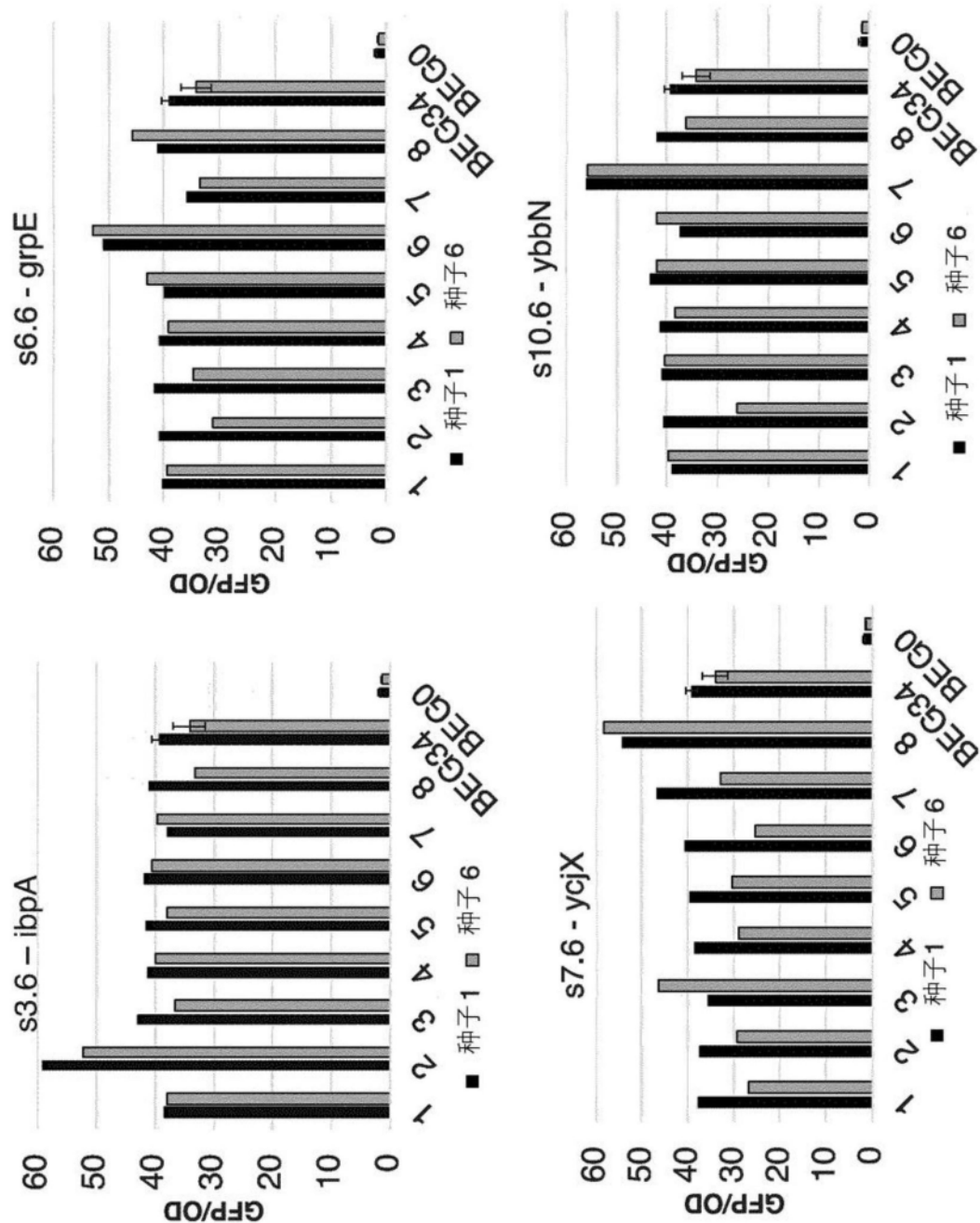


图5

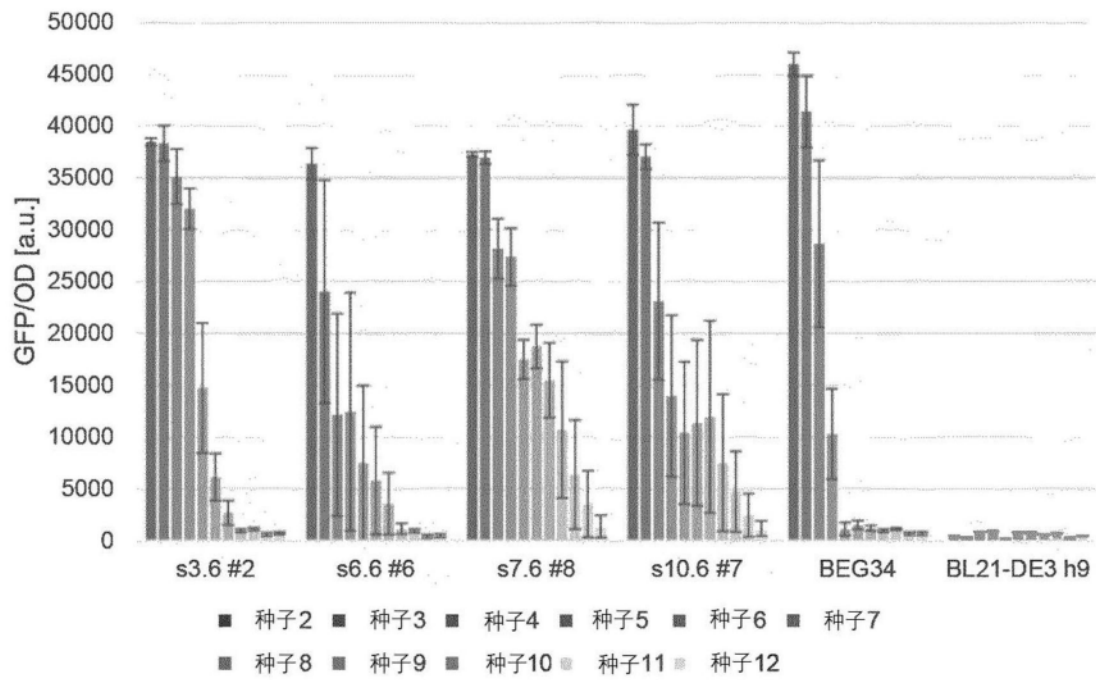


图6

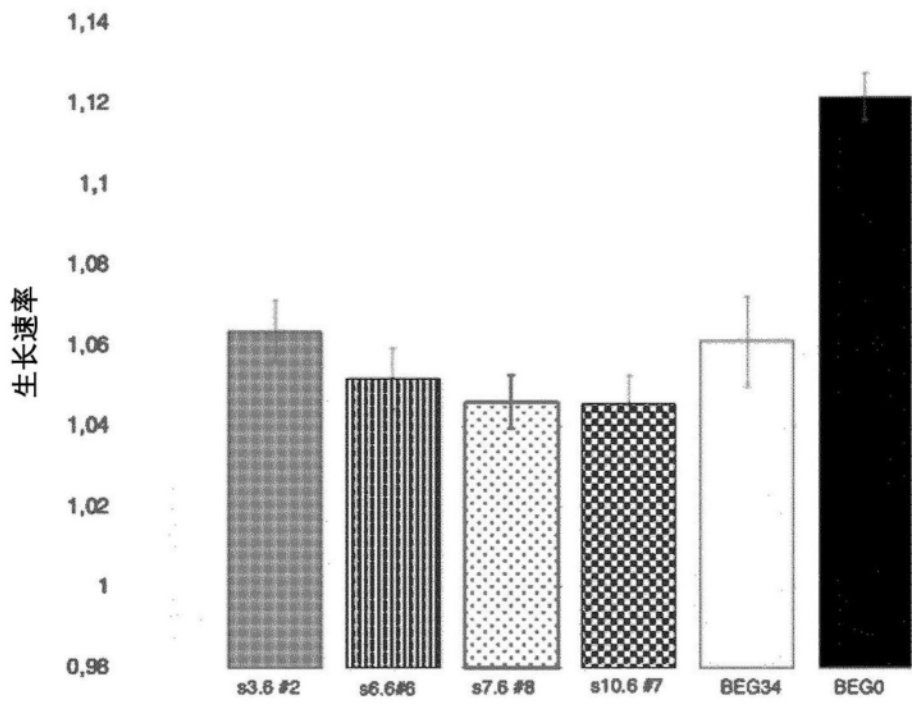


图7

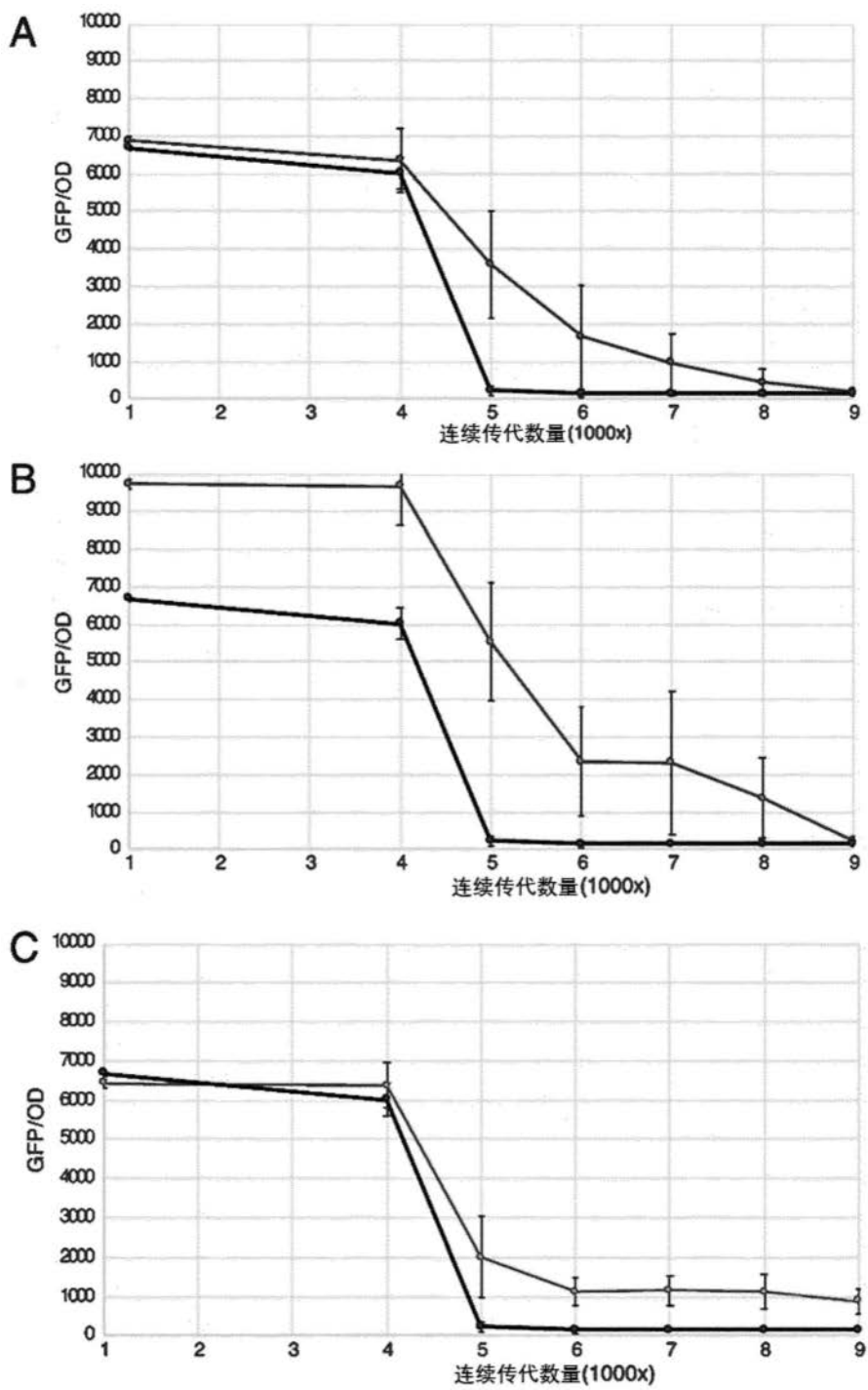


图8

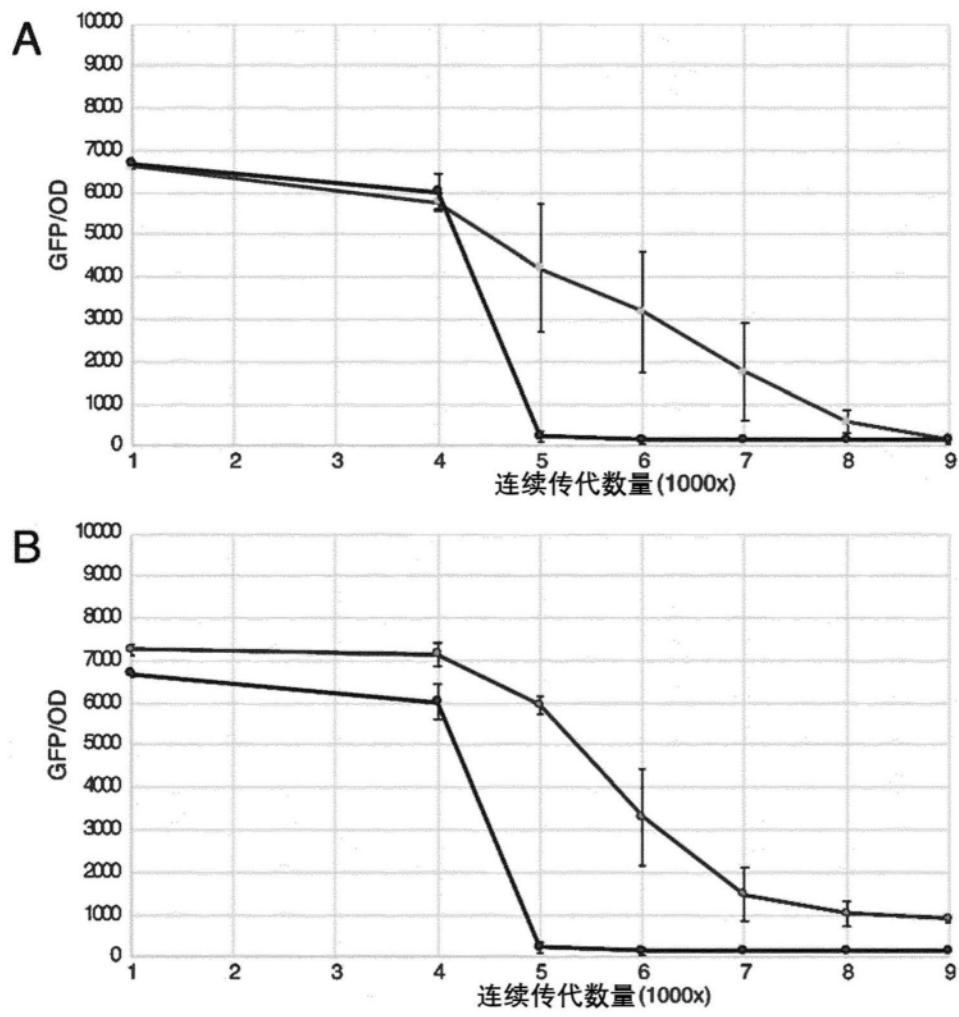


图9

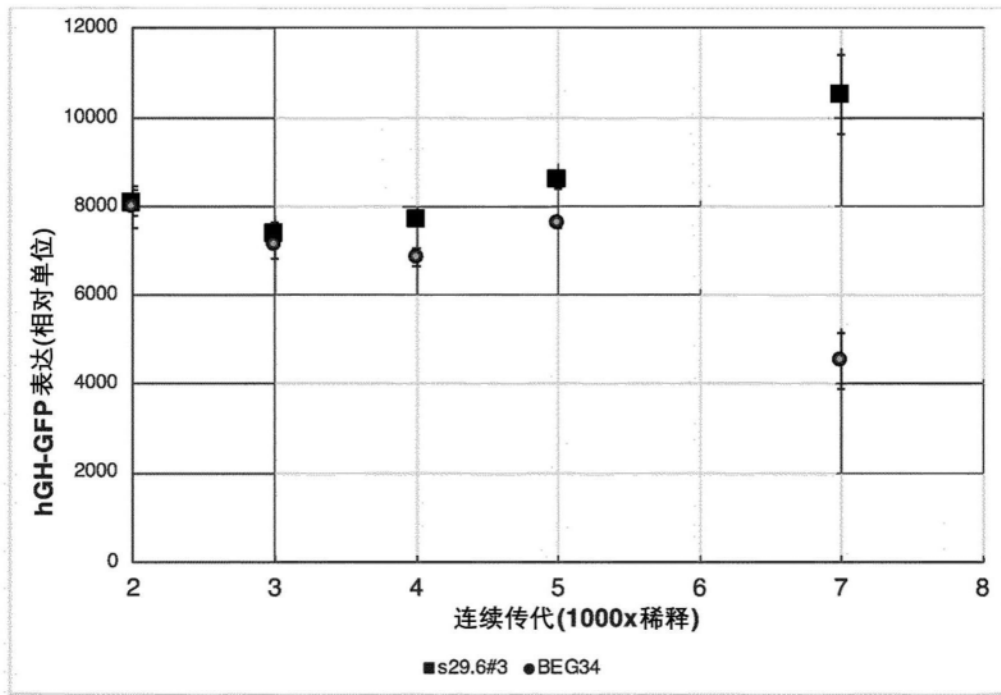


图10

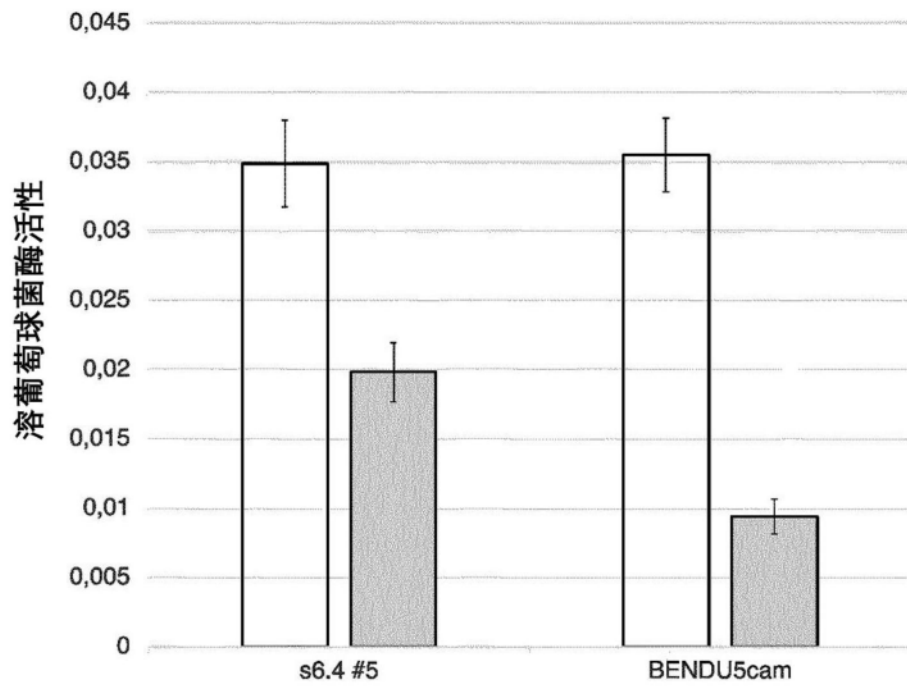


图11

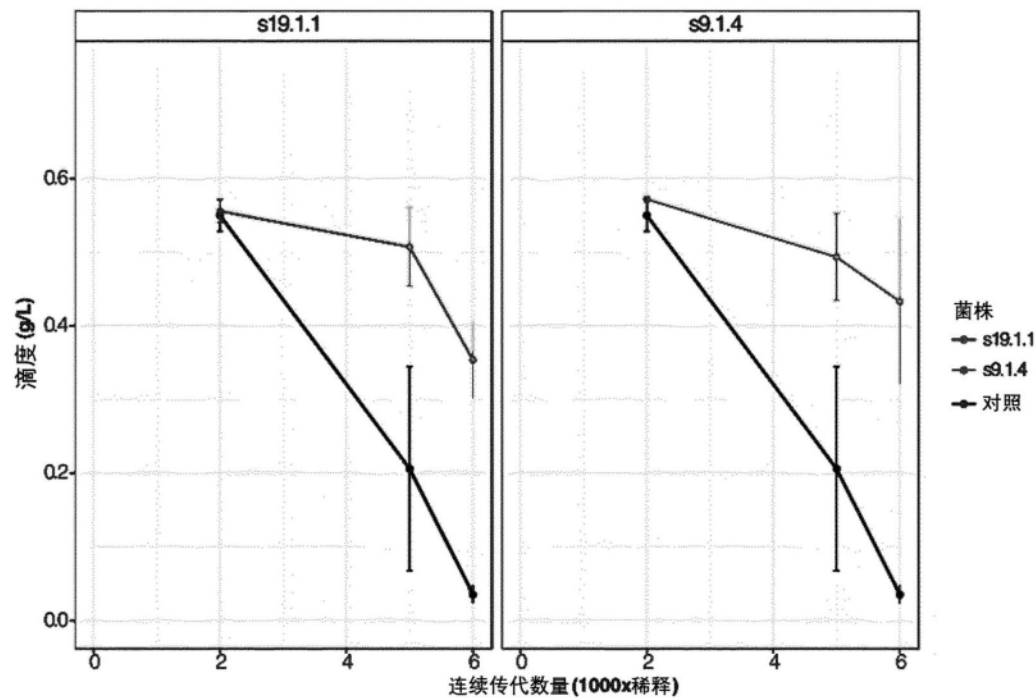


图12

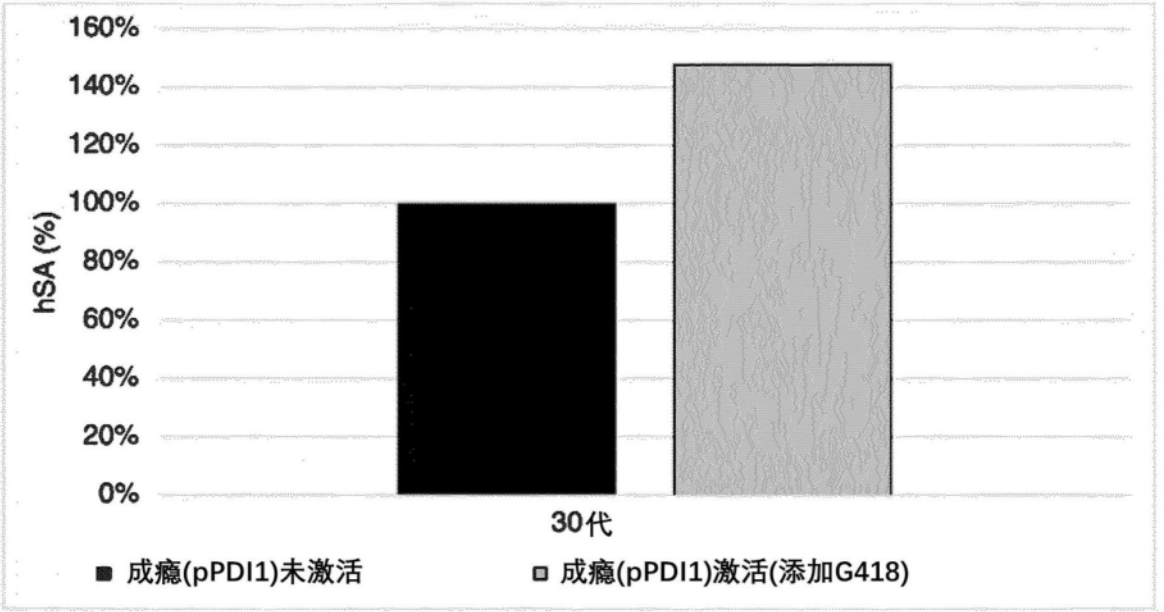


图13

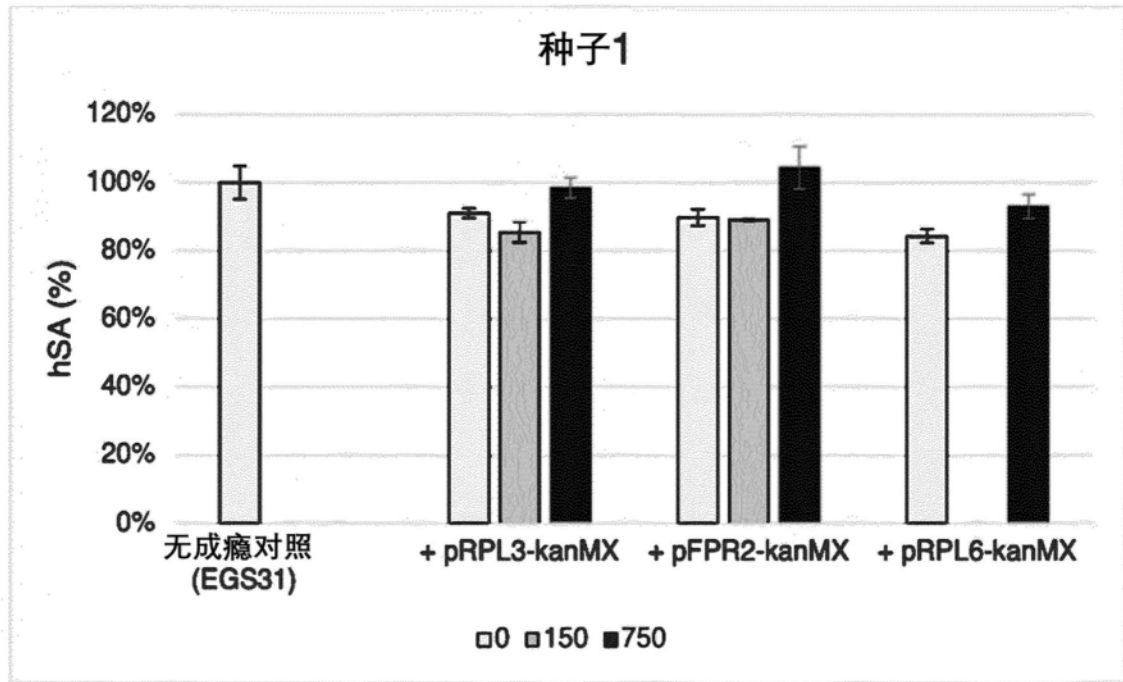


图14A

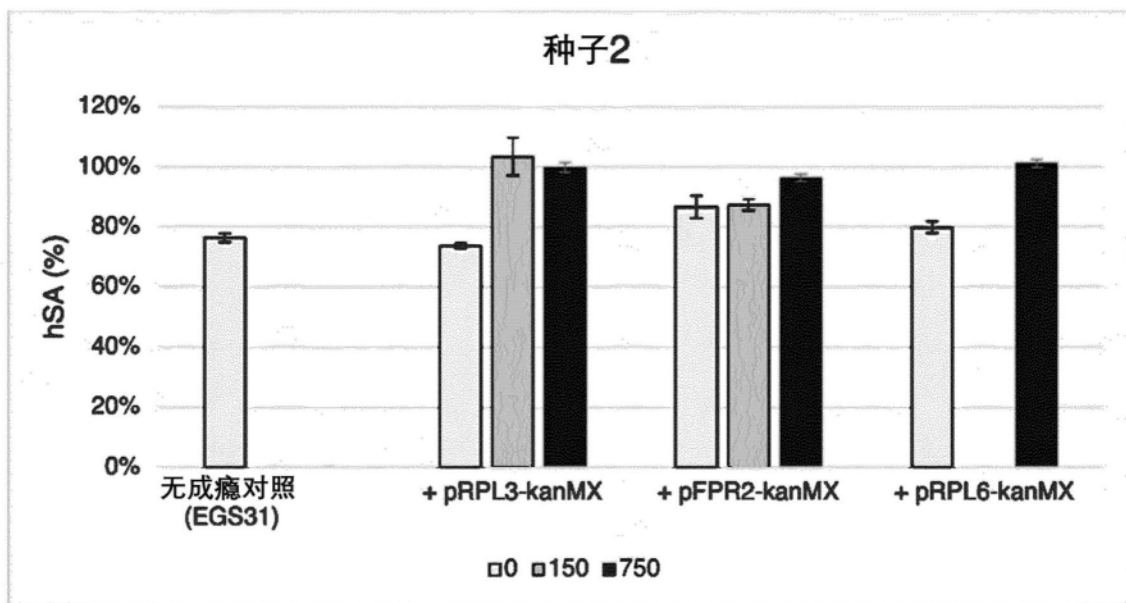


图14B

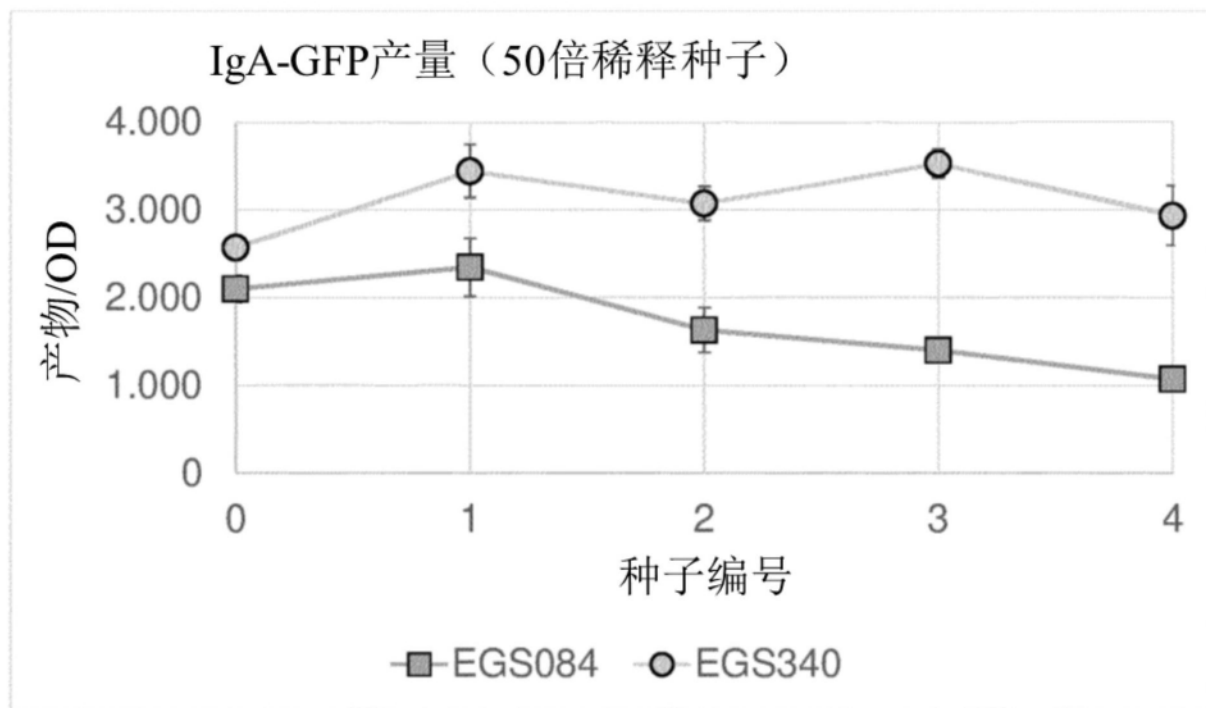


图15

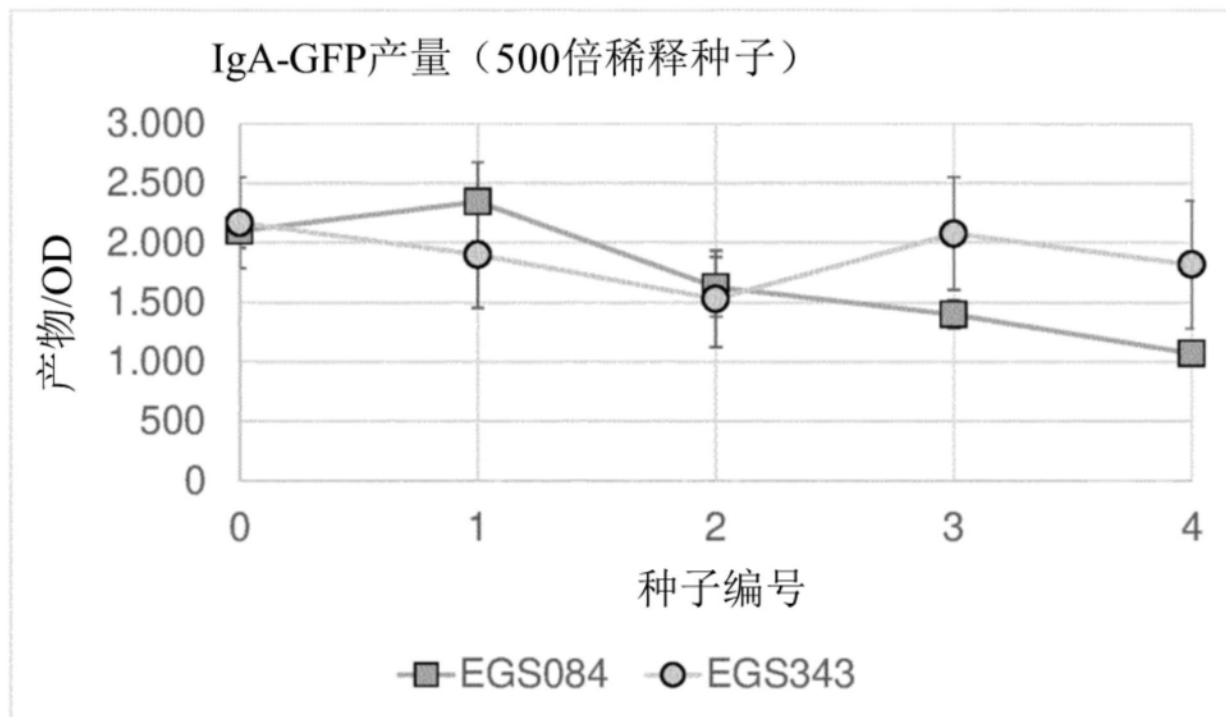


图16

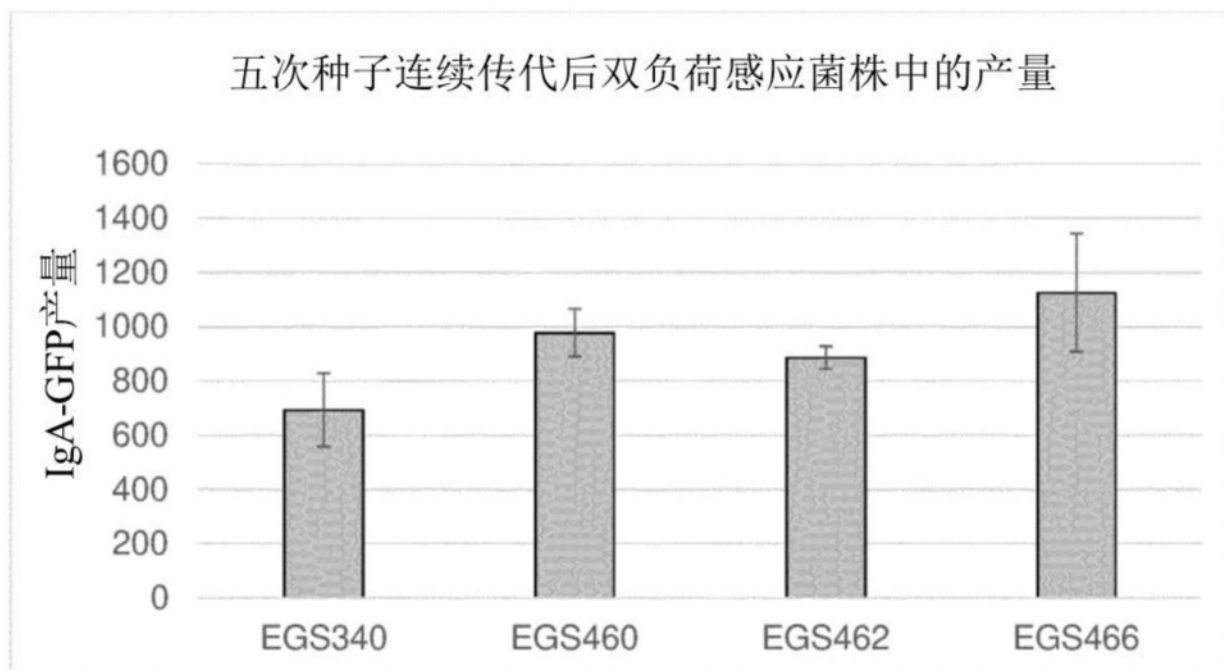


图17

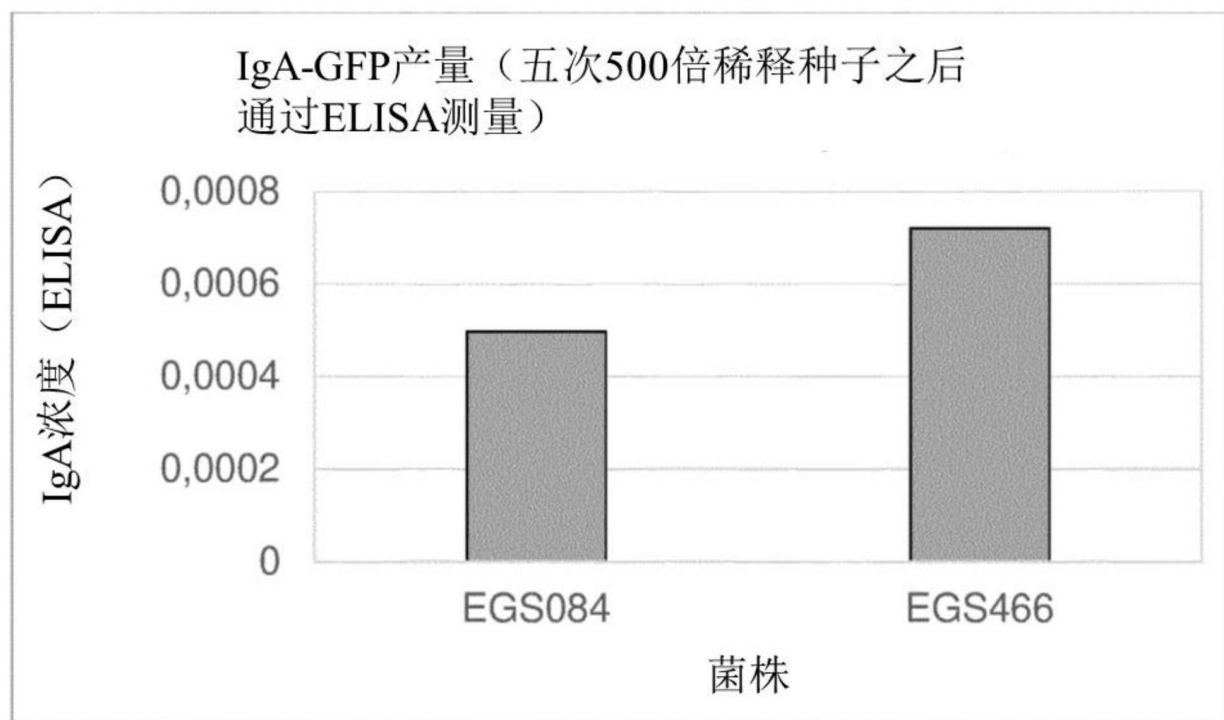


图18

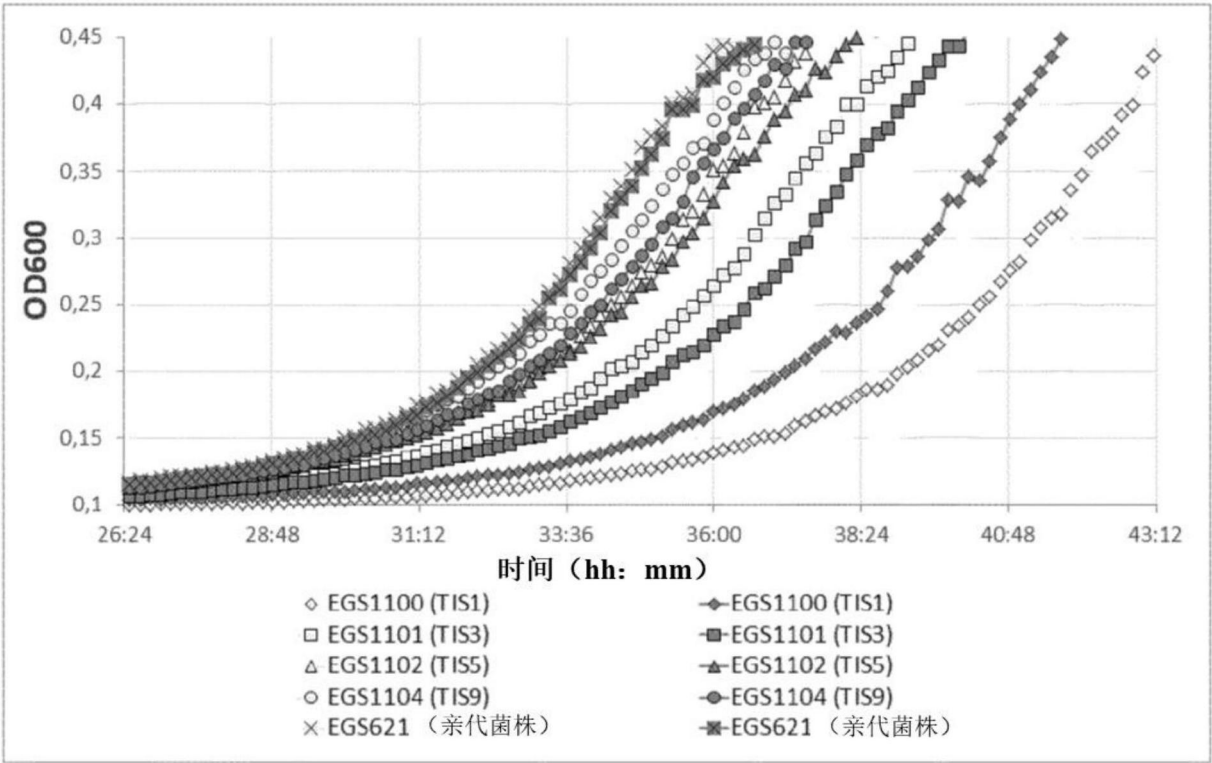


图19