[19] 中华人民共和国国家知识产权局

[12] 发明专利申请公布说明书

[21] 申请号 200880009662.8

[51] Int. Cl.
C12N 1/21 (2006.01)
C12N 1/19 (2006.01)
C12N 15/74 (2006.01)
C12P 7/02 (2006.01)
C12P 7/06 (2006.01)

[43] 公开日 2010 年 3 月 31 日

[11] 公开号 CN 101688175A

[22] 申请日 2008.2.8
[21] 申请号 200880009662.8
[30] 优先权
[32] 2007. 2. 9 [33] US [31] 60/900,546
[32] 2007. 2. 9 [33] US [31] 60/900,477
[32] 2007. 4. 4 [33] US [31] 60/921,927
[32] 2007. 8.17 [33] US [31] 60/956,634
[86] 国际申请 PCT/US2008/053514 2008.2.8
[87] 国际公布 WO2008/098227 英 2008.8.14
[85] 进入国家阶段日期 2009.9.24
[71] 申请人 加利福尼亚大学董事会
地址 美国加利福尼亚州
[72] 发明人 詹姆士·C·里奥 渥美正太
凯文·M·史密斯
罗·普·克莱尔·沈
安东尼·F·卡恩
迈克尔·R·康诺
[74] 专利代理机构 北京安信方达知识产权代理有限公司
代理人 阎娴斌 韩 龙

权利要求书 12 页 说明书 89 页 序列表 140 页
PCT/RO/134 表 2 页 附图 30 页

[54] 发明名称

利用重组微生物生产生物燃料

[57] 摘要

本文提供了一种用于生产生物燃料的代谢修饰微生物。 具体而言，本文提供了由合适的底物生产包括异丁醇、正丁醇、正丙醇、2 - 甲基 - 1 - 丁醇、3 - 甲基 - 1 - 丁醇和 2 - 苯基乙醇的高级醇的方法。

1.一种产生选自下组的醇类的重组微生物:

(a)正丙醇

(b)异丁醇，且每克葡萄糖的异丁醇产率约为 0.12 – 约 0.41g 异丁醇

(c)正丁醇

(d)2-甲基 1-丁醇

(e)3-甲基 1-丁醇

(f)2-苯基乙醇

其中，醇类由一种包含 2-酮酸的代谢物生成。

2.根据权利要求 1 所述的重组微生物，其中，此微生物经过 112 小时的培养，产生的乙醇低于 240mg/L。

3.根据权利要求 1 或 2 所述的重组微生物，其中，在相似条件下培养时，此微生物的醇生产状况实质上与 ATCC 登记号为 no._____或_____（分别为 SA237 或 CRS-BuOH23）的微生物醇生产状况一致。

4.根据权利要求 1 所述的重组微生物，其中，异丁醇的产率约为 0.33 – 0.36 g 异丁醇/g 葡萄糖。

5.根据权利要求 1 所述的重组微生物，其中，经过约 16 – 64 小时的培养，异丁醇产率约为 0.36 – 0.40 g 异丁醇/g 葡萄糖。

6.根据权利要求 4 或 5 所述的重组微生物，其中，每克葡萄糖的乙醇产率低于 0.0037 g/g。

7.根据权利要求 1 所述的重组微生物，其中，此微生物选自棒状杆菌属、乳杆菌属、乳球菌属、沙门氏杆菌属、肠杆菌属、肠球菌属、欧文氏菌属、泛菌属、摩根菌属、果胶杆菌属、变形菌属、沙雷氏菌属、志贺氏菌属、克雷伯氏菌属、柠檬酸杆菌属、酵母属、德克酵母属(dekkera)、克鲁维酵母属(kluyveromyces)和毕赤氏酵母属。

8.根据权利要求 1 所述的重组微生物，其中，为了产生醇类，对生物体内氨基酸的生物合成途径进行了修饰。

9.根据权利要求 1 所述的重组微生物，其中，2-酮酸选自 2-酮基丁酸、2-酮异戊酸、2-酮基戊酸、2-酮基-3-甲基戊酸、2-酮基-4-甲基-戊酸以及苯丙酮酸。

10.根据权利要求 1 所述的重组微生物，其中，与其亲本微生物相比，此微生物的乙醇生产能力较低。

11.根据权利要求 1 或 10 所述的重组微生物，其中，此微生物的乙酰-coA转化为乙醇的能力降低或受到抑制。

12.根据权利要求 1 或 10 所述的重组微生物，其中，此重组微生物的乙醇脱氢酶减少，因此，其乙醇生产能力降低。

13.根据权利要求 12 所述的重组微生物，其中，此微生物源自大肠杆菌。

14.根据权利要求 13 所述的重组微生物，其中乙醇脱氢酶为 adhE、其同源体或突变体。

15.根据权利要求 14 所述的重组微生物，此微生物包含 adhE 的缺失或敲除及其同源体或突变体的缺失或敲除。

16.根据权利要求 1 或 10 所述的重组微生物，其中，此微生物含有将丙酮酸转化为 α-酮基-异戊酸的酶的表达或增加表达。

17.根据权利要求 16 所述的重组微生物，其中，酶为 2-酮酸脱羧酶。

18.根据权利要求 1 所述的重组微生物，与其亲本微生物相比，此微生物含有 2-酮酸脱羧酶和醇脱氢酶的表达或活性的增加。

19.根据权利要求 17 或 18 所述的重组微生物，其中，2-酮酸脱羧酶选自 Pdc、Pdc1、Pdc5、Pdc6、Aro10、Thi3、Kivd、KdcA 和前述各种酶的同源体或突变体，以及与前述任何一种酶具有至少 60% 同一性且具有 2-酮酸脱羧酶活性的多肽。

20.根据权利要求 17 或 18 所述的重组微生物，其中，2-酮酸脱羧酶由与选自以下基因的核酸有至少 60% 同一性的多聚核苷酸编码: *pdc*，*pdc1*，*pdc5*，*pdc6*，*aro10*，*thi3*，*kivd*、*kdcA* 和前述任一基因的同源基因或突变体，或它们的片段，其中，这种多聚核苷酸编码一种具有 2-酮酸脱羧酶活性的多肽。

21.根据权利要求 20 所述的重组微生物，其中，2-酮酸脱羧酶由一种源自 *kivd* 基因或其同源基因的多聚核苷酸编码。

22.根据权利要求 18 所述的重组微生物，其中，醇脱氢酶选自 Adh1、Adh2、Adh3、Adh4、Adh5、Adh6、Sfa1 和前述任一酶的同源体或突变体，以及一种与前述任何一种酶有至少 60% 同一性并具有醇脱氢酶活性的多肽。

23.根据权利要求 18 所述的重组微生物，其中，醇脱氢酶由一种与选自以下的一种核酸有至少 60% 同一性的多聚核苷酸编码: *adh1*、*adh2*、*adh3*、*adh4*、

adh5、*adh6*、*sfa1* 基因和前述任一基因的同源基因，其中，这种多聚核苷酸编码一种具有 2-醇脱氢酶活性的蛋白。

24.根据权利要求 1 所述的重组微生物，其中，此重组微生物包含一个基因的一处或多处缺失或敲除，这个基因编码的酶能够催化乙酰-coA 转化为乙醇；催化丙酮酸转化为乳酸；催化延胡索酸转化为琥珀酸；催化乙酰-coA 和磷酸转化为 coA 和乙酰磷酸；催化乙酰-coA 和甲酸转化为 coA 和丙酮酸；催化乙酰-coA 的乙酰基和 3-甲基-2-氧代丁酸（2-氧代异戊酸）的缩合反应，2-异丙基苹果酸和 3-异丙基苹果酸的异构化反应；催化 α-酮酸转化为支链氨基酸，苯丙氨酸、酪氨酸、天冬氨酸或亮氨酸的合成；催化丙酮酸酯转化为乙酰-coA；催化支链氨基酸的形成；催化苏氨酸生成 α-酮丁酸的反应；催化蛋氨酸生物合成的第一步反应；以及催化苏氨酸的分解代谢。

25. 根据权利要求 1 或 24 所述的重组微生物，其中，此重组微生物包含选自下组基因的一处或多处缺失：*adhE*、*ldhA*、*frdBC*、*fnr*、*pta*、*pflB*、*leuA*、*leuB*、*leuC*、*leuD*、*ilvE*、*tyrB*、*poxB*、*ilvB*、*ilvI*、*ilvA*、*metA*、*tdh* 和前述任一基因的同源基因及自然产生的突变体。

26.根据权利要求 1 所述的重组微生物，其包含一种选自下组的基因型：

（a）选自下组基因的缺失或敲除：ΔadhE、ΔldhA、Δpta、ΔleuA、ΔleuB, ΔleuC、ΔleuD、ΔpoxB、ΔilvB、ΔilvI、ΔmetA、Δtdh 及其任何组合，并包含 kivd、ThrABC 和 adh2 的表达或增加表达，其中，此微生物可生成正丙醇。

（b）选自下组基因的缺失或敲除：ΔadhE、ΔldhA、ΔfrdB、ΔfrdC、Δfnr、Δpta、ΔpflB、ΔleuA、ΔilvE、ΔpoxB、ΔilvA 及其任何组合，并包含 kivd、ThrABC 和 adh2 的表达或增加表达，其中此微生物可生成异丁醇。

（c）选自下组基因的缺失或敲除：ΔadhE、ΔldhA、Δpta、ΔpoxB、ΔilvB、ΔilvI、ΔmetA、Δtdh 及其任何组合，并包含 kivd、ThrABC 和 adh2 的表达或增加表达，其中此微生物可生成正丁醇。

（d）选自下组基因的缺失或敲除：ΔilvB、ΔilvI、ΔmetA、Δtdh 及其任何组合，并包含 kivd、ThrABC 和 adh2 的表达或增加表达，其中此微生物可生成 2-甲基-1-丁醇。

（e）选自下组基因的缺失或敲除：ΔadhE、ΔldhA、ΔfrdB、ΔfrdC、Δfnr、Δpta、ΔpflB、ΔilvE、ΔtyrB 及其任何组合，并包含 kivd、ThrABC 和 adh2 的表达或增加表达，其中此微生物可生成 3-甲基-1-丁醇。

（f）选自下组基因的缺失或敲除：ΔadhE、ΔldhA、ΔfrdB、ΔfrdC、Δfnr、Δpta、ΔpflB、ΔleuA、ΔilvE、ΔpoxB、ΔilvA 及其任何组合，并包含 kivd、ThrABC 和 adh2 的表达或增加表达，其中此微生物可生成 2-苯基乙醇。

27.根据权利要求 26 的重组微生物，其中，ThrABC 是一种抗反馈抑制 ThrA*。

28.一种重组微生物，与其亲本微生物相比，含有以下酶的表达增加或活性增加：

(a) 乙酰羟酸合成酶；

(b) 乙酰羟酸还原异构酶；

(c) 二羟酸脱水酶；

(d) 2-酮酸脱羧酶；

(e) 醇脱氢酶；

并且其中此重组微生物含有至少一个编码以下一种酶的基因敲除或破坏：乙醇脱氢酶、乳酸脱氢酶、延胡索酸还原酶、磷酸乙酰基转移酶、甲酸乙酰基转移酶及其任何组合；

其中，此重组微生物产生异丁醇。

29.一种重组微生物，与其亲本微生物相比，含有以下酶的表达增加或活性增加：

(a) 乙酰乳酸合成酶；

(b) 乙酰羟酸还原异构酶；

(c) 二羟酸脱水酶；

(d) 2-酮酸脱羧酶；和

(e) 醇脱氢酶；

此重组微生物含有至少一个编码以下一种酶的基因敲除或破坏：乙醇脱氢酶、乳酸脱氢酶、延胡索酸还原酶、磷酸乙酰基转移酶、甲酸乙酰基转移酶及其任何组合；

其中，此重组微生物产生异丁醇。

30.根据权利要求 28 或 29 所述的重组微生物，与亲本微生物相比，其还含有丙酮酸通量的增加。

31.根据权利要求 28 或 29 所述的重组微生物，其中，所述敲除或破坏包括 *adhE*、 *ldhA*、*frdBC*、*fnr*、*pflB* 或 *pta* 基因、其任何组合或其任何同源基因或其任何自然产生的突变基因的缺失或表达破坏。

32.根据权利要求 28 或 29 所述的重组微生物，其中，所述敲除或破坏包括 *adhE*、*ldhA*、*frdBC* 基因或其任何同源基因或任何自然产生的突变基因的缺失或表达破坏。

33.根据权利要求 28 或 29 所述的重组微生物，其中，所述敲除或破坏包括选自下组基因的缺失或破坏：*adhE*、*ldhA*、*frdBC*、*pta*、*fnr* 、其任何组合，及其任何同源基因或任何自然产生的突变基因。

34.根据权利要求 33 所述的重组微生物，包含 *adhE*、*ldhA*、*frdBC* 和 *pta*，及其同源基因或突变基因的缺失或破坏。

35.根据权利要求 33 所述的重组微生物，包含 *adhE*、*ldhA*、*frdBC*、*pta* 和 *fnr*，及其同源基因或突变基因的缺失或破坏。

36.根据权利要求 33 所述的重组微生物，包含 *adhE*、*ldhA*、*frdBC* 和 *fnr*，及其同源基因或突变基因的缺失或破坏。

37.根据权利要求 28 或 29 所述的重组微生物，包含 *adhE*、*ldhA*、*frdBC*、*pta* 和 *pflB* 的缺失或破坏。

38.根据权利要求 28 或 29 所述的重组微生物，其中，微生物为大肠杆菌。

39.根据权利要求 38 所述的重组微生物，其中，敲除包含缺失此重组微生物基因组的一部分，约为大肠杆菌基因组的第 1,397,551 –第 1,439,877 个核苷酸。

40.根据权利要求 28 所述的重组微生物，其中，乙酰羟酸合成酶由一种源自 *ilvIH* 操纵子或其同源体的多聚核苷酸编码。

41.根据权利要求 40 所述的重组微生物，其中，*ilvIH* 操纵子的 *ilvI* 基因编码乙酰羟酸合成酶大亚基多肽。

42.根据权利要求 40 所述的重组微生物，其中，*ilvIH* 操纵子的 *ilvH* 基因编码乙酰羟酸合成酶小亚基多肽。

43.根据权利要求 29 所述的重组微生物，其中乙酰羟酸合成酶是来自枯草芽孢杆菌的 *alsS*。

44.根据权利要求 28 或 29 所述的重组微生物，其中，乙酰羟酸还原异构酶由一种源自 *ilvC* 基因或其同源基因的多聚核苷酸编码。

45.根据权利要求 28 或 29 所述的重组微生物，其中，二羟酸脱水酶由一种源自 *ilvD* 基因或其同源基因的多聚核苷酸编码。

46.根据权利要求 28 或 29 所述的重组微生物，其中，2-酮酸脱羧酶由一种源自 *kivd* 基因或其同源基因的、或 *ARO10* 基因、或其同源基因的多聚核苷酸编码。

47.根据权利要求 28 或 29 所述的重组微生物，其中，醇脱氢酶由一种源自 *ADH2* 基因或其同源基因的多聚核苷酸编码。

48.根据权利要求 28 或 29 所述的重组微生物，包含 ATCC 登记号为 no.(名称为 SA237)的一种表型。

49.根据权利要求 28 或 29 所述的重组微生物，包含 ATCC 登记号为 no.(名称为 CRS-BuOH23)的基因组。

50.一种重组微生物，与亲本微生物相比，含有以下酶的表达增加或活性增加：

(a)2-异丙基苹果酸合成酶；

(b)β-异丙基苹果酸脱氢酶；

(c)异丙基苹果异构酶；

(d)苏氨酸脱水酶；

其中，此重组微生物产生正丁醇。

51.根据权利要求 50 所述的重组微生物，与亲本微生物相比，此微生物还可含有 2-酮异戊酸、2-酮基-3-甲基戊酸或 2-酮基-4-甲基戊酸或其任何组合的水平减少。

52.根据权利要求 50 所述的重组微生物，与其亲本微生物相比，此微生物还含有 *ilvD* 基因缺失或表达破坏。

53.根据权利要求 50 所述的重组微生物，其中，2-异丙基苹果酸合成酶由一种源自 *leuA* 基因、其同源基因或其自然产生的突变基因的多聚核苷酸编码。

54.根据权利要求 50 所述的重组微生物，其中，β-异丙基苹果酸合成酶由一种源自 *leuB* 基因、其同源基因或其自然产生的突变基因的多聚核苷酸编码。

55.根据权利要求 50 所述的重组微生物，其中，异丙基苹果酸异构酶由一种源自 *leuCD* 操纵子或其同源体的多聚核苷酸编码。

56.根据权利要求 55 所述的重组微生物，其中，*leuCD* 操纵子的 *leuC* 基因编码异丙基苹果酸异构酶大亚基多肽。

57.根据权利要求55所述的重组微生物，其中，*leuCD*操纵子的*leuD*基因编码异丙基苹果酸异构酶小亚基多肽。

58.根据权利要求50所述的重组微生物，其中，苏氨酸脱水酶由一种源自*ilvA*基因或其同源基因的多聚核苷酸编码。

59.根据权利要求50所述的重组微生物，与其亲本微生物相比，此重组微生物还含有以下酶的表达或活性的增加：磷酸烯醇式丙酮酸羧化酶、丙酮酸羧化酶、天冬氨酸转氨酶、高丝氨酸脱氢酶、天冬氨酸-半醛脱氢酶、高丝氨酸激酶、苏氨酸合成酶、L-丝氨酸脱水酶或苏氨酸脱水酶或其任何组合。

60.根据权利要求50到59中任一项所述的的重组微生物，其中，磷酸烯醇式丙酮酸羧化酶、丙酮酸羧化酶、天冬氨酸转氨酶、高丝氨酸脱氢酶、天冬氨酸-半醛脱氢酶、高丝氨酸激酶、苏氨酸合成酶、L-丝氨酸脱水酶或苏氨酸脱水酶分别由源自*ppc*、*pyc*、*aspC*、*thrA*、*asd*、*thrB*、*thrC*、*sdaAB*、*tdcB*基因，或其同源基因或其自然产生的突变基因的多聚核苷酸编码。

61.根据权利要求50所述的重组微生物，包含ATCC登记号为no._____(命名为CRS-BuOH23)的一种表型。

62.根据权利要求50所述的重组微生物，包含ATCC登记号为no._____(命名为CRS-BuOH23)的基因组。

63.一种重组微生物，与亲本微生物相比，含有以下酶的表达或活性的增加：

(a)α-异丙基苹果酸合成酶

(b)β-异丙基苹果酸脱氢酶

(c)异丙基苹果异构酶；

(d)苏氨酸脱水酶；

其中，此重组微生物产生正丙醇。

64.根据权利要求63所述的重组微生物，其中，α-异丙基苹果酸合成酶由一种源自*cimA*基因、其同源基因或其自然产生的突变基因的多聚核苷酸编码。

65.根据权利要求63所述的重组微生物，其中，β-异丙基苹果酸脱氢酶由一种源自*leuB*基因、其同源基因或自然产生的突变基因的多聚核苷酸编码。

66.根据权利要求63所述的重组微生物，其中，异丙基苹果酸异构酶由一种源自*leuCD*操纵子、其同源体或其自然产生的突变体的多聚核苷酸编码。

67.根据权利要求66所述的重组微生物，其中，*leuCD*操纵子的*leuC*基因编码异丙基苹果酸异构酶大亚基多肽。

68.根据权利要求 66 所述的重组微生物，其中，*leuCD* 操纵子的 *leuD* 基因编码异丙基苹果酸异构酶小亚基多肽。

69.根据权利要求 64 所述的重组微生物，其中，*cimA* 基因是一种詹氏甲烷暖球菌 *cimA* 基因。

70.一种重组微生物，与亲本微生物相比，含有以下酶的表达或活性的增加：

(a)苏氨酸脱水酶

(b)乙酰羟酸合成酶；

(c)乙酰羟酸还原异构酶；

(d)二羟酸脱水酶；

(e)2-酮酸脱羧酶；

(f)醇脱氢酶；

其中，此重组微生物产生 2-甲基-1-丁醇。

71.根据权利要求 70 所述的重组微生物，其中，苏氨酸脱水酶由一种源自 *ilvA* 基因、其同源基因或其自然产生的突变基因的多聚核苷酸编码。

72.根据权利要求 70 所述的重组微生物，其中，乙酰羟酸合成酶由一种源自 *ilvIH* 操纵子、其同源体或其自然产生的突变体的多聚核苷酸编码。

73.根据权利要求 70 所述的重组微生物，其中，乙酰羟酸还原异构酶由一种源自 *ilvC* 基因、其同源基因或其自然产生的突变基因的多聚核苷酸编码。

74.根据权利要求 70 所述的重组微生物，其中，二羟酸脱水酶由一种源自 *ilvD* 基因、其同源基因或其自然产生的突变基因的多聚核苷酸编码。

75.根据权利要求 70 所述的重组微生物，其中，2-酮酸脱羧酶由源自 *kivd* 基因、其同源基因或其自然产生的突变基因；或 *PDC6* 基因、其同源基因或其自然产生的突变基因；或 *THI3* 基因、其同源基因或其自然产生的突变基因的多聚核苷酸编码。

76.根据权利要求 70 到 75 中任何一条所述的重组微生物，其中，醇脱氢酶由一种源自 *ADH2* 基因，或其同源基因的多聚核苷酸编码。

77.一种重组微生物，与亲本微生物相比，含有以下酶的表达或活性的增加：

(a)乙酰羟酸合成酶或乙酰乳酸合成酶；

(b)乙酰羟酸还原异构酶；

(c)二羟酸脱水酶；

(d)2-异丙基苹果酸合成酶；

(e)异丙基苹果酸异构酶

(f)β-异丙基苹果酸脱氢酶

(g)2-酮酸脱羧酶;

(h)醇脱氢酶;

其中，此重组微生物产生 3-甲基-1-丁醇。

78.根据权利要求 77 所述的重组微生物，其中，乙酰羟酸合成酶由一种源自 *ilvIH* 操纵子、其同源体或其自然产生的突变体的多聚核苷酸编码。

79.根据权利要求 77 所述的重组微生物，其中，乙酰乳酸合成酶由一种源自 *ilvMG* 操纵子、其同源基因或其自然产生的突变基因的多聚核苷酸编码。

80.根据权利要求 77 所述的重组微生物，其中，乙酰乳酸合成酶由一种源自 *ilvNB* 操纵子、其同源体或其自然产生的突变体的多聚核苷酸编码。

81.根据权利要求 77 所述的重组微生物，其中，乙酰羟酸还原异构酶由一种源自 *ilvC* 基因、其同源基因或其自然产生的突变基因的多聚核苷酸编码。

82.根据权利要求 77 所述的重组微生物，其中，二羟酸脱水酶由一种源自 *ilvD* 基因、其同源基因或其自然产生的突变基因的多聚核苷酸编码。

83.根据权利要求 77 所述的重组微生物，其中，2-异丙基苹果酸合成酶由一种源自 *leuA* 基因、其同源基因或自然发生的突变基因的多聚核苷酸编码。

84.根据权利要求 77 所述的重组微生物，其中，异丙基苹果酸异构酶由一种源自 *leuCD* 操纵子、其同源基因或其自然产生的突变基因的多聚核苷酸编码。

85.根据权利要求 77 所述的重组微生物，其中，β-异丙基苹果酸脱氢酶由一种源自 *leuB* 基因、其同源基因或自然产生的突变基因的多聚核苷酸编码。

86.根据权利要求 77 所述的重组微生物，其中，2-酮酸脱羧酶由源自 *kivd* 基因、其同源基因；或 *PDC6* 基因、其同源基因；或 *THI3* 基因、其同源基因或其自然产生的突变基因的多聚核苷酸编码。

87.根据权利要求 77 到 86 中任一项所述的重组微生物，其中，醇脱氢酶由一种源自 *ADH2* 基因、其同源基因或其自然产生的突变基因的多聚核苷酸编码。

88.一种重组微生物，与亲本微生物相比，含有以下酶的表达或活性的增加:

(a)分支酸变位酶 P/预苯酸脱水酶;

(b)分支酸变位酶 T/预苯酸脱氢酶;

(c)2-酮酸脱羧酶；和

(d)醇脱氢酶；

其中，此重组微生物产生苯基乙醇。

89.根据权利要求 88 所述的重组微生物，其中，分支酸变位酶 P/预苯酸脱水酶由一种源自 *pheA* 基因、其同源基因或其自然产生的突变基因的多聚核苷酸编码。

90.根据权利要求 88 所述的重组微生物，其中，分支酸变位酶 T/预苯酸脱氢酶由一种源自 tyrA 基因、其同源基因或其自然产生的突变基因的多聚核苷酸编码。

91.根据权利要求 88 所述的重组微生物，其中，2-酮酸脱羧酶由源自 *kivd* 基因、其同源基因、或 *PDC6* 基因、其同源基因、或 *THI3* 基因、其同源基因或它们自然产生的突变基因的多聚核苷酸编码。

92.根据权利要求 88 到 91 中任一项所述的重组微生物，其中，醇脱氢酶由一种源自 *ADH2* 基因或其同源基因的多聚核苷酸编码。

93.一种生产一种能够将合适的底物或代谢中间物转化为正丁醇的重组微生物的方法，此方法包括用一种或多种编码包括以下活性的多肽的重组多聚核苷酸转化微生物：2-异丙基苹果酸合成酶活性、β-异丙基苹果酸脱氢酶活性、异丙基苹果酸异构酶活性和苏氨酸脱水酶活性。

94.根据权利要求 93 所述的方法，其中，按照此方法生产的重组微生物包括在存在葡萄糖时产生的醇的状况与 ATCC 登记号为 no._____（名称为 CRS-BuOH23）的微生物产生的醇的状况实质相似。

95.一种生产一种能够将合适的底物或代谢中间物转化为异丁醇的重组微生物的方法，此方法包括用一种或多种编码包括以下活性的多肽的重组多聚核苷酸转化一种微生物：乙酰羟酸合成酶活性、乙酰羟酸还原异构酶活性、二羟酸脱水酶活性、2-酮酸脱羧酶活性和醇脱氢酶活性。

96.根据权利要求 95 所述的方法，其中，按照此方法生产的重组微生物包括在存在葡萄糖时产生的醇状况与 ATCC 登记号为 no._____（名称为 SA237）的微生物产生的醇状况实质相似。

97.一种生产一种能够将合适的底物或代谢中间物转化为正丙醇的重组微生物的方法，此方法包含用一种或多种编码具有以下酶活性的多肽的重组多聚核

苷酸转化一种微生物：α-异丙基苹果酸合成酶活性、β-异丙基苹果酸脱氢酶活性、异丙基苹果酸异构酶活性和苏氨酸脱水酶活性。

98. 一种生产一种将合适的底物或代谢中间物转化为2-甲基-1-丁醇的重组微生物的方法，此方法包含用一种或多种编码具有以下活性的多肽的重组多聚核苷酸转化一种微生物：苏氨酸脱水酶活性、乙酰羟酸合成酶活性、乙酰羟酸还原异构酶活性、二羟酸脱水酶活性、2-酮酸脱羧酶活性和醇脱氢酶活性。

99. 一种生产一种能够将合适的底物或代谢中间物转化为3-甲基-1-丁醇的重组微生物的方法，此方法包含用一种或多种编码具有以下酶活性的多肽的重组多聚核苷酸转化一种微生物：乙酰羟酸合成酶活性或乙酰乳酸合成酶活性、乙酰羟酸还原异构酶活性、二羟酸脱水酶活性、2-异丙基苹果酸合成酶活性、异丙基苹果酸异构酶活性、β-异丙基苹果酸脱氢酶活性、2-酮酸脱羧酶活性和醇脱氢酶活性。

100. 一种生产能够将合适的底物或代谢中间物转化为苯基乙醇的一种重组微生物的方法，此方法包含用一种或多种编码具有以下酶活性的多肽的重组多聚核苷酸转化一种微生物：分支酸变位酶P/预苯酸脱水酶活性、分支酸变位酶T/预苯酸脱氢酶活性、2-酮酸脱羧酶活性和醇脱氢酶活性。

101. 一种命名为 SA237 且其 ATCC 登记号为 no._____的重组微生物。

102. 一种命名为 CRS-BuOH23 且具有 ATCC 登记号 no._____的重组微生物。

103. 一种生产醇的方法，此方法包含：

(a)提供权利要求 1、28、29、50、63、70、77、78、101 或 102 中任何一项所述的重组微生物；

(b)在 a) 存在一种合适的底物或代谢中间产物，并在将合适的底物转化为醇的条件下培养微生物；

(c)基本上纯化醇。

104. 根据权利要求 103 所述的方法，其中，醇选自正丙醇、异丁醇、正丁醇、2-甲基 1-丁醇、3-甲基-1-丁醇和 2-苯基乙醇。

105. 根据权利要求 103 所述的方法，其中，底物或代谢中间产物包含 2-酮酸。

106.根据权利要求 105 所述的方法，其中，2-酮酸选自 2-酮基丁酸、2-酮异戊酸、2-酮基戊酸、2-酮基-3-甲基戊酸、2-酮基-4-甲基-戊酸和苯丙酮酸。

107.根据权利要求 103 所述的方法，其中，底物包含葡萄糖，醇包含异丁醇。

108.根据权利要求 107 所述的方法，其中，产率包括约 0.12－0.41g 异丁醇/g 葡萄糖。

利用重组微生物生产生物燃料

对相关申请的交叉引用

[0001] 本申请要求 2007 年 2 月 9 日提交的第 60/900，477 号美国临时专利申请、2007 年 2 月 9 日提交的第 60/900，546 号美国临时专利申请、2007 年 4 月 4 日提交的第 60/921，927 号美国临时专利申请和 2007 年 8 月 17 日提交的第 60/956，634 号美国临时专利申请的优先权，这些申请的公开内容以引用的方式并入本申请。

技术领域

[0002] 本发明提供代谢修饰微生物以及生产这种微生物的方法。还提供通过利用合适的底物与代谢修饰微生物及其酶制剂接触生产生物燃料的方法。

背景技术

[0003] 由于经济和环境问题，用生物燃料代替石油的需求日益增长。常见的生物燃料如乙醇并不是一种理想的燃料，因为它的能量密度比汽油低，并且必须在限定的浓度范围内与汽油混合，才能作为运输用燃料。乙醇还具有吸湿性和腐蚀性，这就造成了贮存和配送系统方面的问题。

发明内容

[0004] 本发明提供一种能够产生选自下组的醇类的重组微生物：（a）正丙醇；（b）异丁醇，其产率为每克葡萄糖产生约 0.12 至 0.41g 异丁醇；（c）正丁醇；（d）2-甲基-1-丁醇；（e）3-甲基-1-丁醇；和（f）2-苯基乙醇，这种醇类从包含 2-酮酸的代谢物中产生。在一实施方案中，该微生物经过 112 小时左右的培养，产生的乙醇不到 240mg/L 。在另一实施方案中，在相似条件下培养时，该微生物的乙醇生产曲线与被命名为 SA237 或 CRS-BuOH23 的微生物的生产曲线基本相同。在另一实施方案中，经过大约 16 至 64 小时的培养，异丁醇的产率约为每克葡萄糖产生约 0.33 至

0.36g 异丁醇，或每克葡萄糖产生约 0.36 至 0.40 g 异丁醇。还有一实施方案中，每克葡萄糖的乙醇产率不足 0.0037g/g 左右。在一实施方案中，与亲本微生物相比，该微生物的乙醇生产能力较低。在另一实施方案中，该微生物将乙酰辅酶 A 转化为乙醇的能力降低或受到抑制。在另一实施方案中，重组微生物的乙醇脱氢酶含量减少，因此，其乙醇生产能力降低。在一实施方案中，该微生物能够表达将丙酮酸转化为 α-酮异戊酸的酶或增加此酶的表达。在另一实施方案中，此酶是 2-酮酸脱羧酶（例如，Pdc、Pdc1、Pdc5、Pdc6、Aro10、Thi3、Kivd、KdcA 和前述各种酶的同源体或变体，以及与前述任何一种酶有至少 60% 同一性且具有 2-酮酸脱羧酶活性的多肽）。在另一实施方案中，2-酮酸脱羧酶由一种多聚核苷酸编码，这种多聚核苷酸与选自下组的一种核酸有 60% 的同一性：*pdc*、*pdc1*、*pdc5*、*pdc6*、*aro10*、*thi3*、*kivd*、*kdcA* 和前述各种基因的同源基因或变体，或它们的片段，其中这种多聚核苷酸编码具有 2-酮酸脱羧酶活性的多肽。在一个具体实施方案中，2-酮酸脱羧酶由一个源自 *kivd* 基因或其同源基因的多聚核苷酸编码。在一实施方案中，与其亲本微生物相比，此微生物能够增加 2-酮酸脱羧酶和醇脱氢酶的表达或活性。在一实施方案中，醇脱氢酶选自下组：Adh1、Adh2、Adh3、Adh4、Adh5、Adh6、Sfa1 和前述各种酶的同源体或变体，以及与前述任何一种酶有至少 60% 同一性并具有醇脱氢酶活性的多肽。在另一实施方案中，醇脱氢酶由一种多聚核苷酸编码，这种多聚核苷酸与选自下组的一种核酸有至少 60% 同一性：*adh1*、*adh2*、*adh3*、*adh4*、*adh5*、*adh6*、*sfa1* 基因和前述各种基因的同源基因或变体，其中这种多聚核苷酸编码具有 2-醇脱氢酶活性的多肽。在一实施方案中，此重组微生物的一个基因发生一处或多处缺失或敲除，这个基因编码的酶能够催化乙酰辅酶 A 转化为乙醇；催化丙酮酸酯转化为乳酸酯；催化延胡索酸酯转化为琥珀酸酯；催化乙酰辅酶 A 和磷酸酯转化为辅酶 A 和乙酰磷酸；催化乙酰辅酶 A 和甲酸酯转化为辅酶 A 和丙酮酸酯，乙酰辅酶 A 的乙酰基和 3-甲基-2-氧代丁酸酯（2-氧代异戊酸酯）的缩合反应，2-异丙基苹果酸酯和 3-异丙基苹果酸酯的异构化反应；催化 α-酮酸转化为支链氨基酸，苯丙氨酸、酪氨酸、天冬氨酸或亮氨酸的合成；催化丙酮酸酯转化为乙酰辅酶 A；催化生成支链氨基酸；催化苏氨酸生成 α-酮丁酸酯的反应；催化蛋氨酸生物合成的第一步反应；以及催化苏氨酸的代谢。例如，这种微生物可能包

含选自下组的一种或多种基因缺失：*adhE*、*ldhA*、*frdBC*、*fnr*、*pta*、*pflB*、*leuA*、*leuB*、*leuC*、*leuD*、*ilvE*、*tyrB*、*poxB*、*ilvB*、*ilvI*、*ilvA*、*metA*、*tdh*、前述各种基因的同源基因及自然产生的突变。在一个具体实施方案中，此微生物的基因型选自下组： （a）选自下组基因的一种缺失或敲除，选自一组基因，这组基因包括： ΔadhE、ΔldhA、Δpta、ΔleuA、ΔleuB、ΔleuC、ΔleuD、ΔpoxB、ΔilvB、ΔilvI、ΔmetA、Δtdh 及其任何组合，并包含 kivd、ThrABC 和 adh2 的表达或增强表达，其中，此微生物产生正丙醇；（b）选自下组基因的一种缺失或敲除： ΔadhE、ΔldhA、ΔfrdB、ΔfrdC、Δfnr、Δpta、ΔpflB、ΔleuA、ΔilvE、ΔpoxB、ΔilvA 及其任何组合，并包含 kivd、ThrABC 和 adh2 的表达或增强表达，其中，此微生物产生异丁醇；（c）选自下组基因的一种缺失或敲除：ΔadhE、ΔldhA、Δpta、ΔpoxB、ΔilvB、ΔilvI、ΔmetA、Δtdh 及其任何组合，并包含 kivd、ThrABC 和 adh2 的表达或增强表达，其中，此微生物产生正丁醇；（d）选自下组基因的一种缺失或敲除：ΔilvB、ΔilvI、ΔmetA、Δtdh 及其任何组合，并包含 kivd、ThrABC 和 adh2 的表达或增强表达，其中，此微生物产生 2-甲基-1-丁醇；（e）选自下组基因的一种缺失或敲除：ΔadhE、ΔldhA、ΔfrdB、ΔfrdC、Δfnr、Δpta、ΔpflB、ΔilvE、ΔtyrB 及其任何组合，并包含 kivd、ThrABC 和 adh2 的表达或增强表达，其中，此微生物产生 3-甲基 1-丁醇；和（f）选自下组基因的一种缺失或敲除：ΔadhE、ΔldhA、ΔfrdB、ΔfrdC、Δfnr、Δpta、ΔpflB、ΔleuA、ΔilvE、ΔpoxB、ΔilvA 及其任何组合，并包含 kivd、ThrABC 和 adh2 的表达或增强表达，其中，此微生物产生 2-苯基乙醇。在另一实施方案中，ThrABC 包含抗反馈抑制的 ThrA*。在一实施方案中，此重组微生物含有被命名为 SA237 或 CRS-BuOH23 的微生物的表型。

[0005] 本文提供的是代谢修饰微生物，其包括生化重组途径，该途径利用合适的底物，通过代谢工程微生物生产生物燃料，包括异丁醇、2-甲基-1-丁醇、3-甲基-1-丁醇、2-苯基乙醇、正丙醇或正丁醇等。本发明还提供利用本文描述的微生物生产生物燃料的方法。

[0006] 在一实施方案中，提供了产生一种醇类的重组微生物。该醇类可以是正丙醇、异丁醇、正丁醇、2-甲基-1-丁醇、3-甲基-1-丁醇或 2-苯基乙醇。通常情况下，可利用含有 2-酮酸的代谢物并以发酵或非发酵方式（即在有氧或无氧条件下）生产醇类。在某些方面，2-酮酸包含 2-酮丁酸、2-

酮异戊酸、2-酮基戊酸、2-酮基-3-甲基戊酸、2-酮基-4-甲基戊酸或苯丙酮酸。在其他方面，与其亲本微生物相比，此重组微生物增加 2-酮酸脱羧酶和醇脱氢酶的表达或活性。2-酮酸脱羧酶可能是来自酿酒酵母(*Saccharomyces cerevisiae*) 的 Pdc6、或 Aro10、或 Thi3; 来自乳酸乳球菌(*Lactococcus lactis*) 的 Kivd; 来自丙酮丁醇梭菌 (*Clostridium acetobutylicum*) 的 Pdc，或它们的同源体。2-酮酸脱羧酶可以由源自以下一种基因的多聚核苷酸编码: 来自酿酒酵母的 *PDC6*、来自酿酒酵母的 *ARO10*、来自酿酒酵母的 *THI3*、来自乳酸乳球菌的 *kivd*、来自丙酮丁醇梭菌的 *pdc* ，或它们的同源基因。在某些方面，醇脱氢酶可能是来自酿酒酵母的 Adh2 或其同源基因，由源自酿酒酵母 *ADH2* 基因的多聚核苷酸编码。在另一实施方案中，提供了一种产生异丁醇的重组微生物。与其亲本微生物相比，此微生物包含增加乙酰羟酸合成酶、乙酰羟酸还原异构酶、二羟酸脱水酶、2-酮酸脱羧酶以及醇脱氢酶的表达或活性。在某些方面，此微生物能够包含增加乙酰乳酸合成酶、乙酰羟酸还原异构酶、二羟酸脱水酶、2-酮酸脱羧酶以及醇脱氢酶的表达。在某些方面，与其亲本微生物相比，此重组微生物还含有更高水平的丙酮酸酯。因此，此重组微生物还可以含有 *adhE*、*ldh*、*frd*、*fnr*、*pflB*、*ackA* 或 *pta* 基因或它们的任何组合的基因缺失或表达抑制。特别是，此重组微生物能够含有单独的 *adh*、*ldh*、*frd* 基因缺失或与 *fnr*、*fnr* 和 *pta*、*pta* 和 *pflb* 组合的基因缺失。在某些方面，此重组微生物还可以含有其基因组的部分缺失，例如缺失大肠杆菌 (*E. coli*) 基因组中约 1,397,551 个到约 1,439,877 个的核苷酸。一方面，乙酰羟酸合成酶可由源自以下基因的多聚核苷酸编码: *ilvIH* 操纵子、*ilvBN* 操纵子、大肠杆菌的 *ilvGM* 、来自枯草芽孢杆菌的 *alsS* 基因，或它们的同源基因。*ilvIH* 操纵子的 *ilvI* 基因编码乙酰羟酸合成酶的大亚基多肽，*ilvIH* 操纵子的 *ilvH* 基因编码乙酰羟酸合成酶的小亚基多肽。另一方面，乙酰羟酸还原异构酶可由源自大肠杆菌 *ilvC* 基因或其同源基因的多聚核苷酸编码。另外，二羟酸脱水酶可由源自 *ilvD* 基因或其同源基因的多聚核苷酸编码。另一方面，2-酮酸脱羧酶可由源自以下基因的多聚核苷酸编码: 乳酸乳球菌的 *kivd* 基因或其同源基因，或酿酒酵母的 *ARO10* 基因或其同源基因。另一方面，醇脱氢酶可由源自酿酒酵母的 *ADH2* 基因或其同源基因的多聚核苷酸编码。

[0007] 通常，大肠杆菌的 *ilvIH* 操纵子编码乙酰羟酸合成酶，这是异亮氨酸、缬氨酸和亮氨酸生物合成途径中的第一个酶。乙酰羟酸合成酶 III 同功酶由两个亚基组成，即 *ilvI* 基因产物（乙酰羟酸合成酶 III 的大亚基）和 *ilvH* 基因产物（乙酰羟酸合成酶的小亚基），能够催化大肠杆菌 K-12 中异亮氨酸、亮氨酸及缬氨酸生物合成的共同第一步反应。大肠杆菌的 *ilvC* 基因编码乙酰羟酸还原异构酶，这是平行进行的异亮氨酸-缬氨酸生物合成途径中的第二个酶。大肠杆菌的 *ilvD* 基因编码二羟酸脱水酶，这是异亮氨酸-缬氨酸生物合成途径中的第三个酶。在某些方面，此重组微生物包含乙酰乳酸合成酶的表达增加。乙酰乳酸合成酶可来自枯草芽孢杆菌 (*Bacillus subtilis*) 的 AlsS 。

[0008] 在一实施方案中，提供了一种产生正丁醇的重组微生物。与其亲本微生物相比，此微生物包含 2-异丙基苹果酸合成酶、β-异丙基苹果酸脱氢酶、异丙基苹果酸异构酶和苏氨酸脱水酶的表达或活性增加。在另一实施方案中，与其亲本微生物相比，此重组微生物还含有更高水平的 2-酮基戊酸。在另一实施方案中，与其亲本微生物相比，此重组微生物还含有较低水平的 2-酮异戊酸、2-酮基-3-甲基戊酸或 2-酮基-4-甲基戊酸或它们的任何组合。因此，与其亲本微生物相比，此微生物还可含有 *ilvD* 基因缺失或表达受抑制。一方面， 2-异丙基苹果酸合成酶可由源自 *leuA* 基因的多聚核苷酸或其同源体编码。另一方面， β-异丙基苹果酸脱氢酶可由源自 *leuB* 基因或其同源基因的多聚核苷酸编码。另一方面，异丙基苹果酸异构酶可由源自 *leuCD 操纵子或其同源操纵子*的多聚核苷酸编码。通常，*leuCD* 操纵子的 *leuC* 基因编码异丙基苹果酸异构酶的大亚基多肽，*leuCD* 操纵子的 *leuD* 基因编码异丙基苹果酸异构酶的小亚基多肽。另一方面，苏氨酸脱水酶可由源自 *ilvA* 基因或其同源基因的多聚核苷酸编码。另一方面，苏氨酸脱水酶可由源自 *tdcB* 基因或其同源基因的多聚核苷酸编码。在另一实施方案中，与其亲本微生物相比，此重组微生物还可含有磷酸烯醇式丙酮酸羧化酶、丙酮酸羧化酶、天冬氨酸转氨酶、高丝氨酸脱氢酶、天冬氨酸-半醛脱氢酶、高丝氨酸激酶、苏氨酸合成酶、L-丝氨酸脱水酶、苏氨酸脱水酶或它们的任何组合的表达或活性的增加。在某些方面，磷酸烯醇式丙酮酸羧化酶、丙酮酸羧化酶、天冬氨酸转氨酶、高丝氨酸脱氢酶、天冬氨酸-半醛脱氢酶、高丝氨酸激酶、苏氨酸合成酶、L-丝氨酸脱水酶或苏

氨酸脱水酶分别由源自以下基因或其同源基因的多聚核苷酸编码：*ppc*、*pyc*、*aspC*、*thrA*、*asd*、*thrB*、*thrC*、*sdaAB* 和 *tdcB* 基因。

[0009] 在一实施方案中，提供了一种产生正丙醇的重组微生物。与其亲本微生物相比，此微生物包含 α-异丙基苹果酸合成酶、问号钩端螺旋体(*Leptospira interrogans*) 的 LeuB、异丙基苹果酸异构酶和苏氨酸脱水酶的表达或活性的增加。一方面，α-异丙基苹果酸合成酶可由源自 *cimA* 基因或其同源基因的多聚核苷酸编码。*cimA* 基因可能是问号钩端螺旋体的 *cimA* 基因或詹氏甲烷嗳球菌 (*Methanocaldococcus jannaschii*) 的 *cimA* 基因。另一方面，β-异丙基苹果酸脱氢酶可由源自 *leuB* 基因或其同源基因的多聚核苷酸编码。另一方面，异丙基苹果酸异构酶可由源自 *leuCD 操纵子或其同源操纵子*的多聚核苷酸编码。在另一实施方案中，与其亲本微生物相比，此重组微生物还可含有以下酶的表达或活性的增加：磷酸烯醇式丙酮酸羧化酶、丙酮酸羧化酶、天冬氨酸转氨酶、高丝氨酸脱氢酶、天冬氨酸-半醛脱氢酶、高丝氨酸激酶、苏氨酸合成酶、L-丝氨酸脱水酶或苏氨酸脱水酶，或它们的任何组合。在某些方面，磷酸烯醇式丙酮酸羧化酶、丙酮酸羧化酶、天冬氨酸转氨酶、高丝氨酸脱氢酶、天冬氨酸-半醛脱氢酶、高丝氨酸激酶、苏氨酸合成酶、L-丝氨酸脱水酶或苏氨酸脱水酶分别由源自以下基因的多聚核苷酸编码：*ppc*、*pyc*、*aspC*、*thrA*、*asd*、*thrB*、*thrC*、*sdaAB* 和 *tdcB* 基因，或其同源基因。

[0010] 在另一实施方案中，提供了一种产生 2-甲基-1-丁醇的重组微生物。与其亲本微生物相比，此微生物包含苏氨酸脱水酶、乙酰羟酸合成酶、乙酰羟酸还原异构酶、二羟酸脱水酶、2-酮酸脱羧酶和醇脱氢酶的表达或活性的增加，并产生 2-甲基 1-丁醇。在某些方面，苏氨酸脱水酶可由源自 *ilvA* 基因或其同源基因的多聚核苷酸编码。另一方面，苏氨酸脱水酶可由源自 *tdcB* 基因或其同源基因的多聚核苷酸编码。在另一实施方案中，与其亲本微生物相比，此重组微生物还含有更高水平的 2-酮基-3-甲基戊酸。另一方面，2-酮酸脱羧酶可由源自以下基因的多聚核苷酸编码：*kivd* 基因或其同源基因；*PDC6* 基因或其同源基因；*THI3* 基因或其同源基因。

[0011] 在另一实施方案中，提供了一种产生 3-甲基-1-丁醇的重组微生物。与其亲本微生物相比，此微生物包含乙酰羟酸合成酶或乙酰乳酸合成酶、乙酰羟酸还原异构酶、二羟酸脱水酶，2-异丙基苹果酸合成酶、异丙

基苹果酸异构酶、β-异丙基苹果酸脱氢酶、2-酮酸脱羧酶以及醇脱氢酶的表达或活性的增加。在某些方面，乙酰羟酸合成酶可由源自 *ilvIH* 操纵子或其同源操纵子的多聚核苷酸编码。另一方面，乙酰乳酸合成酶可由源自 *alsS* 基因或其同源基因的多聚核苷酸编码。另一方面，乙酰乳酸合成酶可由源自 *ilvMG* 操纵子或其同源操纵子的多聚核苷酸编码。在另一实施方案中，与其亲本微生物相比，此重组微生物还含有更高水平的 2-酮异己酸。另一方面，乙酰乳酸合成酶可由源自 *ilvNB* 操纵子或其同源操纵子的多聚核苷酸编码。

[0012] 在另一个实施方案中，提供了一种生产苯基乙醇的重组微生物。与其亲本微生物相比，此微生物包含分支酸变位酶 P/预苯酸脱水酶、分支酸变位酶 T/预苯酸脱氢酶、2-酮酸脱羧酶和醇脱氢酶的表达或活性的增加。另一方面，分支酸变位酶 P/预苯酸脱水酶可由源自 *pheA* 基因或其同源基因的多聚核苷酸编码。另一方面，分支酸变位酶 T/预苯酸脱水酶可由源自 *tyrA* 基因或其同源基因的多聚核苷酸编码。在另一实施方案中，与其亲本微生物相比，此重组微生物还含有更高水平的苯丙酮酸。

[0013] 在一实施方案中，提供了一种生产重组微生物的方法，此微生物能够将合适的底物或代谢中间物转化为正丁醇。此方法包含用一种或多种编码包括以下活性的多肽的重组多聚核苷酸转化微生物的方法：2-异丙基苹果酸合成酶活性、β-异丙基苹果酸脱氢酶活性、异丙基苹果酸异构酶活性和苏氨酸脱水酶活性。

[0014] 在另一实施方案中，提供了一种生产重组微生物的方法，此微生物能够将合适的底物或代谢中间物转化为异丁醇。此方法包含用一种或多种编码包括以下活性的多肽的重组多聚核苷酸转化微生物：乙酰羟酸合成酶活性、乙酰羟酸还原异构酶活性、二羟酸脱水酶活性、2-酮酸脱羧酶活性和醇脱氢酶活性。

[0015] 在另一实施方案中，提供了一种生产重组微生物的方法，此微生物能够将合适的底物或代谢中间物转化为正丙醇。此方法包含用一种或多种编码包括以下活性的多肽的重组多聚核苷酸转化微生物：α-异丙基苹果酸合成酶活性、β-异丙基苹果酸脱氢酶活性、异丙基苹果酸异构酶活性和苏氨酸脱水酶活性。

[0016] 在一实施方案中，提供了一种生产重组微生物的方法，此微生物能够将合适的底物或代谢中间物转化为 2-甲基-1-丁醇。此方法包含用一种或多种编码包括以下活性的多肽的重组多聚核苷酸转化微生物：苏氨酸脱水酶活性、乙酰羟酸合成酶活性、乙酰羟酸还原异构酶活性、二羟酸脱水酶活性、2-酮酸脱羧酶活性和醇脱氢酶活性。

[0017] 在另一实施方案中，提供了一种生产重组微生物的方法，此微生物能够将合适的底物或代谢中间物转化为 3-甲基-1-丁醇。此方法包含用一种或多种编码包括以下活性的多肽的重组多聚核苷酸转化微生物：乙酰羟酸合成酶活性或乙酰乳酸合成酶活性、乙酰羟酸还原异构酶活性、二羟酸脱水酶活性、2-异丙基苹果酸合成酶活性、异丙基苹果酸异构酶活性、β-异丙基苹果酸脱氢酶活性、2-酮酸脱羧酶活性和醇脱氢酶活性。

[0018] 在另一实施方案中，提供了一种生产重组微生物的方法，此微生物能够将合适的底物或代谢中间物转化为苯基乙醇。此方法包含用一种或多种编码包括以下活性的多肽的重组多聚核苷酸转化微生物：分支酸变位酶 P/预苯酸脱水酶活性、分支酸变位酶 T/预苯酸脱氢酶活性、2-酮酸脱羧酶活性和醇脱氢酶活性。

[0019] 在另一实施方案中，提供了一种生产一种醇类的方法。此方法包含提供一种本文中提供的重组微生物；在适合于将底物转化为醇类的条件下，利用合适的底物或代谢中间物培养微生物；以及检测醇类的产量。在许多方面，该醇类可以是正丙醇、异丁醇、正丁醇、2-甲基 1-丁醇、3-甲基-1-丁醇和 2-苯基乙醇。另一方面，这些底物或代谢中间物包括一种 2-酮酸，例如 2-酮丁酸、2-酮异戊酸、2-酮基戊酸、2-酮基-3-甲基戊酸、2-酮基-4-甲基-戊酸或苯丙酮酸。

[0020] 本发明的一个或多个实施方案的细节通过附图和以下说明陈述。其他特点、目标和优点在说明、附图和权利要求中显而易见。

附图说明

[0021] 本附图并入本说明书中，并构成本说明书的一部分，用来描述本发明的一个或多个实施方案，并与具体实施方式一起解释本发明的原理与实现方式。

[0022] 图 1A-C 描述了有助于了解本发明的反应途径。（A）描述了利用 2-酮酸代谢生产醇类的一条代表性非发酵合成途径。（B）描述了基因工程大肠杆菌中醇类化合物的代表性生成途径。灰色箭头代表 2-酮酸降解途径。酶、LeuABCD、IlvA 和 IlvIHCD 代表氨基酸生物合成途径。双线代表氨基酸生物合成途径的副反应。（C）描述了微生物的一般反应途径，并标出了这个生物体内可被破坏或敲除的酶，以增加所需反应方向的通量（例如增加合成丙酮酸的通量，减少流向其他竞争性途径的丙酮酸通量）。

[0023] 图 2A-C 描述了经修饰的氨基酸生物合成途径，以改进异丁醇和正丁醇的生产。A 图显示了经改造的 *ilvIHCD* 途径存在或不存在的情况下异丁醇的产量。左边区域： 异丁醇产量；右边区域： 每克葡萄糖的异丁醇产率。异丁醇的最高产率的理论值为 0.41g/g。 B 图显示了经改造的 *ilvA-leuABCD* 途径存在或不存在的情况下由葡萄糖获得正丁醇的产量。左边区域： 正丁醇产量；右边区域： 相同菌株的正丙醇产量。图 C 显示了加入 L-苏氨酸时的正丁醇产量。左边区域： 正丁醇产量；右边区域： 相同菌株的正丙醇产量。

[0024] 图 3 描述了异丁醇对细胞生长的影响。上面显示含有或不含有 2% 异丁醇时，野生型菌株和高耐受性突变菌株的细胞生长时间过程。两种菌株在 LB 培养基中都生长到指数期。在 OD_{600} 为~3.5 时，向培养基加入 2% 异丁醇。三角形： 未加异丁醇的野生型菌株；菱形： 添加异丁醇的野生型菌株；方形： 未添加异丁醇的高耐受性突变菌株；圆形： 添加异丁醇的高耐受性突变菌株。

[0025] 图 4 描述了大肠杆菌中异丁醇的代表性生产途径。

[0026] 图 5 描述了异丁醇产量的质谱分析检测结果。

[0027] 图 6 描述了质谱分析数据。

[0028] 图 7 描述了由丙酮酸生成 2-酮基异戊酸的代表性途径。

[0029] 图 8 描述了亮氨酸的代表性生物合成途径。

[0030] 图 9 描述了异亮氨酸的代表性生物合成途径。

[0031] 图 10 描述了以 2-酮基丁酸为中间产物的丁醇代表性生物合成途径。

[0032] 图 11 描述了由丙酮酸生成丁醇的代表性生物合成途径。

[0033] 图 12 描述了以苏氨酸为中间产物的丁醇代表性生物合成途径。

[0034] **图 13** 描述了生产异丁醇类化合物（例如，2-甲基丙基醇）、3-甲基-1-丁醇、正丁醇、乙醇、2-甲基-1-丁醇和正丙醇的代表性生物合成途径。

[0035] **图 14** 描述了生产苯基乙醇、乙醇、3-甲基-1-丁醇和异丁醇类化合物（例如2-甲基丙基醇）的代表性生物合成途径。

[0036] **图 15** 描述了源自 *kivd* 基因的核酸序列（序列 ID 号：27），此基因编码具有2-酮酸脱羧酶活性的多肽。

[0037] **图 16** 描述了源自 *PDC6* 基因的核酸序列（序列 ID 号：29），此基因编码一段具有2-酮酸脱羧酶活性的多肽。

[0038] **图 17** 描述了源自 *AR010* 基因的核酸序列（序列 ID 号：31），此基因编码一段具有 2-酮酸脱羧酶活性的多肽。

[0039] **图 18** 描述了源自 *THI3* 基因的核酸序列（序列 ID 号：33），此基因编码一段具有 2-酮酸脱羧酶活性的多肽。

[0040] **图 19** 描述了源自 *pdc* 基因的核酸序列（序列 ID 号：35），此基因编码一段具有2-酮酸脱羧酶活性的多肽。

[0041] **图 20** 描述了源自 *ADH2* 基因的核酸序列（序列 ID 号：37），此基因编码一段具有醇脱氢酶活性的多肽。

[0042] **图 21** 描述了源自 *ilvI* 基因的核酸序列（序列 ID 号：39），此基因编码一段具有乙酰乳酸合成酶大亚基活性的多肽。

[0043] **图 22** 描述了源自 *ilvH* 基因的核酸序列（序列 ID 号：41），此基因编码一段具有乙酰乳酸合成酶小亚基活性的多肽。

[0044] **图 23** 描述了源自 *ilvC* 基因的核酸序列（序列 ID 号：43），此基因编码一段具有乙酰羟酸还原异构酶活性的多肽。

[0045] 图 24 描述了源自 *ilvD* 基因的核酸序列（序列 ID 号：45），此基因编码一段具有二羟酸脱水酶活性的多肽。

[0046] **图 25** 描述了源自 *ilvA* 基因的核酸序列（序列 ID 号：47），此基因编码一段具有苏氨酸脱水酶活性的多肽。

[0047] **图 26** 描述了源自 *leuA* 基因的核酸序列（序列 ID 号：49），此基因编码一段具有2-异丙基苹果酸合成酶活性的多肽。

[0048] **图 27** 描述了源自 *leuB* 基因的核酸序列（序列 ID 号：51），此基因编码一段具有β-异丙基苹果酸脱氢酶活性的多肽。

[0049] 图 28 描述了源自 *leuC* 基因的核酸序列（序列 ID 号：53），此基因编码一段具有异丙基苹果酸异构酶大亚基活性的多肽。

[0050] 图 29 描述了源自 *leuD* 基因的核酸序列（序列 ID 号：55），此基因编码一段具有异丙基苹果酸异构酶小亚基活性的多肽。

[0051] 图 30 描述了源自 *cimA* 基因的核酸序列（序列 ID 号：57），此基因编码一段具有 α-异丙基苹果酸合成酶活性的多肽。

[0052] 图 31 描述了源自 *ilvM* 基因的核酸序列（序列 ID 号：59），此基因编码一段具有乙酰乳酸合成酶大亚基活性的多肽。

[0053] 图 32 描述了源自 *ilvG* 基因的核酸序列（序列 ID 号：61），此基因编码一段具有乙酰乳酸合成酶小亚基活性的多肽。

[0054] 图 33 描述了源自 *ilvN* 基因的核酸序列（序列 ID 号：63），此基因编码一段具有乙酰乳酸合成酶大亚基活性的多肽。

[0055] 图 34 描述了源自 *ilvB* 基因的核酸序列（序列 ID 号：65），此基因编码一段具有乙酰乳酸合成酶小亚基活性的多肽。

[0056] 图 35 描述了源自 *adhE2* 基因的核酸序列（序列 ID 号：67），此基因编码一段具有醇脱氢酶活性的多肽。

[0057] 图 36 描述了源自 *Li-cimA* 基因的核酸序列（序列 ID 号：69），此基因编码一段具有 α-异丙基苹果酸合成酶活性的多肽。

[0058] 图 37 描述了源自 *Li-leuC* 基因的核酸序列（序列 ID 号：71），此基因编码一段具有异丙基苹果酸异构酶大亚基活性的多肽。

[0059] 图 38 描述了源自 *Li-leuD* 基因的核酸序列（序列 ID 号：73），此基因编码一段具有异丙基苹果酸异构酶小亚基活性的多肽。

[0060] 图 39 描述了源自 *Li-leuB* 基因的核酸序列（序列 ID 号：75），此基因编码一段具有 β-异丙基苹果酸脱氢酶活性的多肽。

[0061] 图 40 描述了源自 *pheA* 基因的核酸序列（序列 ID 号：77），此基因编码一段具有分支酸变位酶 P/预苯酸脱水酶活性的多肽。

[0062] 图 41 描述了源自 *TyrA* 基因的核酸序列（序列 ID 号：79），此基因编码一段具有分支酸变位酶 T/预苯酸脱水酶活性的多肽。

[0063] 图 42 描述了源自 *alsS* 基因的核酸序列（序列 ID 号：81），此基因编码一段具有乙酰乳酸合成酶活性的多肽。

[0064] 图 43 描述了不同微生物的敲除突变与异丙酸产量间的相关性。

[0065] **图 44** 描述了对普通摇瓶与带螺旋盖烧瓶内异丁醇产量的比较。

[0066] **图 45** 描述了超过 200 小时的分批补料过程中，添加不同的复合培养基组分对异丁醇产量的影响。

[0067] **图 46** 显示了从葡萄糖转化为 3-甲基-1-丁醇的代谢途径。除另有注明外，所有基因均来自大肠杆菌。BS = 枯草芽孢杆菌；LL = 乳酸乳球菌；SC = 酿酒酵母。

[0068] **图 47A-B** 显示了 3-甲基-1-丁醇的初期产量。方格柱形代表异丁醇；实心柱形代表 3-甲基-1-丁醇。（A）JCL16 中 3-甲基-1-丁醇的产量。对携带 ilvIH（EC）或 alsS（BS）基因并具有 leuABCD 基因染色体或质粒表达的菌株进行了醇产量测定。（B）JCL260 中 3-甲基-1-丁醇的产量。对携带 ilvIH(IAA92) 或 alsS（IAA85） 基因的菌株进行了醇产量测定。

[0069] **图 48A-B** 是描述 α-酮酸产量的图表。方格柱形代表 2-酮异戊酸；实心柱形代表 2-酮基异己酸。（A）JCL260 背景下 α-酮酸的产量。对菌株的 2-酮异戊酸（异丁醇）和 2-酮基异己酸（3-甲基-1-丁醇）的产量进行了测定。改变 RBS 仅用于 leuA 基因产物（IPMS）。（B）L-亮氨酸合成敲除背景中 α-酮酸的产量。‘Δ’ 表示缺失。

[0070] **图 49A-B** 显示了消除反馈抑制后 3-甲基-1-丁醇的产量。（A）对含有 WT IPMS（IAA88）和 IPMS（G462D）(IAA89) 的 JCL260 宿主的生长和醇产量进行了比较。（B）对 JCL260 ΔilvE ΔtyrB (IAA69) 背景中含有 WT IPMS（IAA90）菌株的生长和醇产量进行了定量检测。

[0071] **图 50** 是基因工程大肠杆菌内由苏氨酸生成丙醇和丁醇的生物合成途径示意图。图中描述了特定途径的中断；空心矩形框表示醇产物的前体。图中还描述了非正常的正缬氨酸途径，以及缬氨酸、亮氨酸、异亮氨酸、脯氨酸和蛋氨酸（这些氨基酸分别以缩写形式写在灰色框中）的生物合成途径。

[0072] **图 51** 显示了 thrA*BC 的过度表达对 BW WT 中醇和主要代谢产物的影响，并显示了 CRS-BuOH 12（实心菱形）和 CRS-BuOH 31（空心菱形）中丙醇、丁醇、主要副产物、微生物生长及葡萄糖消耗的时间过程。 CRS-BuOH 12 和 CRS-BuOH 31 均为 BW WT 菌株。CRS-BuOH 31 含有 pSA62 和 pSA55I，CRS-BuOH 12 还含有另外的质粒 pCS49，此质

粒携带的 thrA*BC 位于 PLlacO1 后。按照本发明所述的材料和方法培养细胞。

[0073] **图 52A-B** 对各种敲除菌株的醇产量进行了比较。A.菌株被从左到右依次编号为 CRS-BuOH 12、32、2、11、23。所有菌株均包含 pCS49、pSA62 和 pSA55I 质粒。如本文所述，将细胞培养 72 小时。图中所示为第 72 个小时时间点的数据。B. A 图中显示的各菌株生长时间过程。

[0074] **图 53** 对采用替代性抗反馈苏氨酸脱水酶和 2-异丙基苹果酸合成酶生产的丙醇和丁醇的产量进行了比较。比较时采用 BW ΔmetA、Δtdh、ΔilvB、ΔilvI 作为背景菌株。各菌株被编号为 CRS-BuOH 11、18、19、20，除图中所示的质粒外，所有菌株还含有 pSA62 和 pSA55I 质粒。质粒编号下方为表达特定的苏氨酸脱水酶和 2-异丙基苹果酶的基因名称。如本文所述，将细胞培养 72 小时。图中所示为第 72 个小时时间点的数据。

[0075] **图 54A-E** 显示了 CRS-BuOH 23 中丙醇、丁醇和代谢副产物的时间过程。A.正丙醇和正丁醇的产量。实心菱形代表丁醇，空心菱形代表丙醇。B.主要副产物醋酸的产量。C.次要副产物丙酮酸、乳酸和乙醇的产量。实心菱形代表丙酮酸，空心方块代表乳酸，交叉符号代表乙醇。D.葡萄糖的消耗过程。E.CRS-BuOH 23 在 100 小时内的生长过程。

[0076] 每张图里类似的参考符号代表类似的要素。

具体实施方式

[0077] 除非文中另有明确说明，否则，本文及随后所附的权利要求中使用的表示单数含义的“一种”、“和”及“此”等词语也包含复数指代。因此，例如提到“一种多聚核苷酸”时，也包含多种这样的多聚核苷酸；当提到“此微生物”时，也指一种或多种微生物等等。

[0078] 除非另有说明，本文中使用的所有技术术语与科学术语均与本发明所属技术领域的技术人员通常理解的意思相同。虽然与本文所述的方法和材料相似或相当的方法和材料也能用于本发明公开的方法和成分中，但是本文描述的是代表性方法、装置和材料。

[0079] 提供的上述及全文中的任何公开内容仅为在提交此申请之日公开的内容。本文不得被视为发明人无权凭借先有公开内容而先于这样的公开。

[0080] 丁醇是疏水性化合物，其挥发性比乙醇差。正丁醇的能量密度与汽油接近。浓度为 85% 的丁醇可用于汽车，并且不需对发动机做任何改动（与乙醇不同），它产生的动力比乙醇大，几乎与汽油相同。丁醇也被作为溶剂用于化学和纺织工艺、有机合成以及用作化学中间体。丁醇还被作为液压制动液中的一种成分使用，以及作为基础成分用于香水中。

[0081] 与乙醇相比，异丁醇具有优点且和正丁醇相同。此外，由于异丁醇带有分支碳链，所以其辛烷值比正丁醇高， 正丁醇作为发酵产物生产，并用作发动机燃料。虽然异丁醇已被用作发动机添加剂，但是还不能以高产率由可再生资源生产出异丁醇。异丁醇尚未被考虑作为汽油的替代品。

[0082] 能够生产正丁醇的天然菌种如丙酮丁醇梭菌，还能生产出作为发酵产物的诸如丙酮、乙醇及丁酸等副产物。但是，这些微生物较难操控，其基因操控工具不如大肠杆菌等易操控的宿主的基因操控工具有效，而且对它们的生理学和代谢调节的了解也较少，因此阻碍了其高效生产的步伐。此外，虽然已在微生物代谢副产物中发现少量的其他高级醇，例如异丁醇、2-甲基-1-丁醇、3-甲基-1-丁醇和 2-苯基乙醇，但是尚未发现可以由葡萄糖生成并能够达到工业利用量的高级醇的天然菌种。

[0083] 通过各种酵母菌 (包括酿酒酵母) 生产异丁醇和其他杂醇，这对酒精饮料制造商颇具吸引力。对酒精饮料而言，杂醇通常属于不期望的异常成分。用野生型酵母菌生产异丁醇已经在各种培养基上证实过，这些培养基包括制酒过程中的葡萄汁（Romano 等，Metabolic diversity of Saccharomyces cerevisiae strains from spontaneously fermented grape musts（来自自然发酵葡萄汁的各种酿酒酵母菌株的代谢多样性），19:311-315, 2003)到基本补充培养基(Oliviera 等，(2005)World Journal of Microbiology and Biotechnology 21:1569-1576)，前者可产生 12-219 mg/L 异丁醇，后者可产生 16-34 mg/L 异丁醇。Dickinson 等人的研究 （J Biol Chem. 272(43):26871-8, 1997 ）已确定支链氨基酸（例如，缬氨酸和亮氨酸）转化为异丁醇的反应途径中的酶反应步骤。

[0084] 本发明提供了代谢工程微生物，其包含利用合适底物生产高级醇（包括异丁醇、正丁醇、正丙醇、2-甲基-1-丁醇、3-甲基-1-丁醇和 2-苯基乙醇）的生物化学途径。本发明的一种代谢工程微生物包含该生物体基因组内的或该生物体基因组以外的一种或多种重组多聚核苷酸。此微生物能

够包含在野生型微生物内发现的基因的减少、破坏或敲除，和/或导入外源多聚核苷酸。

[0085] 本发明还包括利用生物体的天然氨基酸途径的代谢工程生物合成途径。利用天然氨基酸途径生产生物燃料有多种好处。它不仅避免了表达一大组的外源基因的困难，还将积累有毒中间产物的可能性降到最低。与在多种梭状芽孢杆菌 (*Clostridium*) 中发现的丁醇产生途径相反，用于生物燃料生产的经改造氨基酸生物合成路径避免了涉及氧气敏感酶和对辅酶 A 依赖的中间产物。本发明提供了一种宿主友好型生物燃料生产系统，其利用微生物氨基酸生物合成途径中的自然代谢产物生产生物燃料。

[0086] 一方面，本发明提供了一种重组微生物，与其亲本微生物相比，此重组微生物包含至少一种目标酶的表达增加，或编码在其亲本微生物中未发现的酶。另一方面，此微生物包含至少一种基因的减少、破坏或敲除，此基因编码的酶与生产期望的高级醇产物所需的代谢产物有竞争作用。重组微生物产生至少一种与异丁醇、正丁醇、正丙醇、2-甲基-1-丁醇、3-甲基-1-丁醇或 2-苯基乙醇的生物合成途径相关的代谢产物。一般情况下，重组微生物包含至少一种重组代谢途径，此途径包含一种目标酶，并且还可包含竞争性生物合成途径中一种酶的活性或表达降低。此途径的作用是修饰异丁醇、正丁醇、正丙醇、2-甲基-1-丁醇、3-甲基-1-丁醇或 2-苯基乙醇生产过程中的底物或代谢中间产物。目标酶通过源自合适的生物来源的多聚核苷酸进行编码和表达。在某些实施方案中，此多聚核苷酸包含源自一种细菌或酵母的一种基因，其经重组改造引入本发明的微生物中。

[0087] 正如本文中所使用的，“代谢工程”或“代谢改造”这一术语涉及生物合成基因、与操纵子有关的基因以及此类多聚核苷酸的调控元件的合理途径改造与组装，其目的是在一种微生物里产生期望的代谢产物，例如 2-酮酸或高级醇。“代谢工程”还包含利用基因工程和合适的培养基条件（包括与导向所需路径的中间产物相竞争的一种竞争性代谢途径的减少、破坏或敲除）对转录、翻译、蛋白质稳定性和蛋白质功能进行调节和优化，以优化代谢通量。一种生物合成基因可能与宿主微生物异源，可利用宿主异源的基因，或通过致突变作用、重组进行修饰，并且/或与存在于内源性宿主细胞的异源表达控制序列有关。一方面，如果此多聚核苷酸对宿主生物体是异源基因，则能对此多聚核苷酸进行密码子优化。

[0088] 术语“生物合成途径”，也称“代谢途径”，指将一种化合物转化（改变）为另一种化合物的一组合成代谢或分解代谢生化反应。如果基因产物平行或连续地作用于相同的底物，产生相同的产物或作用于或产生一种在该相同底物与代谢终产物之间的代谢中间物（例如代谢产物），则这些基因产物属于相同的“代谢途径”。

[0089] 例如，通过 L-亮氨酸固有的生物合成途径合成亮氨酸，此途径偏离 L-缬氨酸生物合成系统的中间产物（2-酮异戊酸）。在大肠杆菌中，L-缬氨酸的生物合成和 L-亮氨酸固有的生物合成是通过一组分别由 ilvGMEDA 操纵子和 leuABCD 操纵子编码的酶进行。

[0090] leuABCD 操纵子包括 leuA、leuB、leuC 和 leuD 基因。其中，leu A 编码 α-异丙基苹果酸合成酶，leuB 编码 β-异丙基苹果酸脱氢酶，leuC 和 leuD 编码 α-异丙基苹果酸异构酶。其中，α-异丙基苹果酸合成酶催化由 α-酮异戊酸生成 α-异丙基苹果酸的合成反应，α-异丙基苹果酸异构酶催化由 α-异丙基苹果酸生成 β-异丙基苹果酸的异构化反应，β-异丙基苹果酸脱氢酶催化由 β-异丙基苹果酸生成 α-酮基异己酸的脱氢反应，α-酮基异己酸是 L-亮氨酸生物合成的最终中间产物。大肠杆菌具有四种转氨酶，即由 aspC 基因编码的转氨酶 A（天冬氨酸-谷氨酸转氨酶）、由 ilvGMEDA 操纵子含有的 ilvE 基因编码的转氨酶 B（BCAA 转氨酶）、由 avtA 基因编码的转氨酶 C（丙氨酸-缬氨酸转氨酶）和由 tyrB 基因编码的转氨酶 D（酪氨酸转氨酶）。这些酶参与各种胺化反应。其中，转氨酶 B 和转氨酶 D 催化上述由 α-酮基异己酸生成 L-亮氨酸的胺化反应。转氨酶 C 和转氨酶 D 催化 L-缬氨酸生物合成途径的最后一步反应，这一步是 L-缬氨酸生物合成和 L-亮氨酸生物合成的共有反应。

[0091] 另外，leuABCD 操纵子的表达也受到 L-亮氨酸的抑制。编码乙酰羟酸合成酶 I 的 ilvBN 基因的表达受到 L-缬氨酸和 L-亮氨酸的协同抑制，编码乙酰羟酸合成酶 II 的 ilvGM 基因的表达受到 L-异亮氨酸、L-缬氨酸和 L-亮氨酸的协同抑制，编码乙酰羟酸合成酶 III 的 ilvIH 基因的表达受到 L-亮氨酸的抑制。

[0092] 术语“底物”或“合适的底物”指在酶的作用下转化为或要转化为另一种化合物的物质或化合物。此术语不仅包含一种化合物，也包含化合物组合，例如溶液、混合物或至少包含一种底物或其衍生物的其他物质。并

且，术语“底物”不仅包括适合作为起始物质的提供碳源的化合物，例如源自任何生物质的糖类，还包括本文所述的用于与代谢工程微生物相关途径的中间代谢产物和最终代谢产物，“源自生物质的糖类”包括但不限于例如葡萄糖、蔗糖、甘露糖、木糖和阿拉伯糖的分子。术语“源自生物质的糖类”包括通常被微生物利用的合适的碳底物，诸如 6 碳糖，包括但不限于葡萄糖、乳糖、山梨糖、果糖、艾杜糖、半乳糖和甘露糖，所有糖的构型均为 D 型或 L 型，或 6 碳糖组合如葡萄糖和果糖和/或 6 碳糖酸类，包括但不限于 2-酮基-L-古洛糖酸、艾杜糖酸（IA）、葡萄糖酸（GA）、6-磷酸葡萄糖酸、2-酮基-D-葡萄糖酸（2 KDG）、5-酮基-D 葡萄糖酸、2-酮基葡萄糖酸磷酸、2,5-二酮基-L-古洛糖酸、2,3-L-二酮基古洛糖酸、脱氢抗坏血酸、异抗坏血酸（EA）和 D-甘露糖酸。

[0093] 术语“醇类”包括正丙醇、异丁醇、正丁醇、2-甲基-1-丁醇、3-甲基-1-丁醇或 2-苯基乙醇。术语“1-丁醇”或“正丁醇”一般指末端碳原子带有醇官能团的直链异构体。内部碳原子带有醇官能团的直链异构体是仲丁醇或 2-丁醇。末端碳原子带有醇官能团的支链异构体是异丁醇，内部碳原子带有醇官能团的支链异构体是叔丁醇。

[0094] 本文提供的重组微生物能够表达多种目标酶，这些酶与由合适碳底物生成正丙醇、异丁醇、正丁醇、2-甲基-1-丁醇、3-甲基-1-丁醇或 2-苯基乙醇的合成途径相关。

[0095] 因此，通过将基因物质引入选择的某个宿主或亲本微生物，从而产生经代谢“改造”或“修饰”的微生物，以修改或改变此微生物的细胞生理学或生物化学特性。通过引入基因物质，亲本微生物获得了新性状，例如，能够产生新的或更多数量的胞内代谢产物。在一示例实施方案中，将基因物质引入亲本微生物中导致其获得新的能力或经改进的能力，能够生产一种醇，例如正丙醇、异丁醇、正丁醇、2-甲基-1-丁醇、3-甲基-1-丁醇或 2-苯基乙醇。引入亲本微生物的基因物质包括基因或部分基因，其编码一种或多种与醇类生产的生物合成途径相关的酶，并且还包含参与这些基因的表达和/或表达调控的其他要素，例如启动子序列。

[0096] 除了将基因物质引入宿主或亲本微生物外，在可选的方法中，一种经改造或修饰的微生物还能包括基因或多聚核苷酸的破坏、缺失或敲除，以改变微生物的细胞生理学和生物化学特性。通过基因或多聚核苷酸的减

少、破坏或敲除，微生物获得了新的或经改进的性状（例如能够产生一种新的或更多数量的胞内代谢产物，改进所需途径的代谢通量，并且/或降低不需要的副产物的产量）。

[0097] 本发明证实了在具有以下酶活性的多肽存在时编码具有酮酸脱羧酶和醇脱氢酶的多肽的一种或多种异源多聚核苷酸的表达或过度表达：α-异丙基苹果酸合成酶、β-异丙基苹果酸脱氢酶、α-异丙基苹果酸异构酶、苏氨酸脱水酶、高丝氨酸脱氢酶、高丝氨酸激酶活性以及苏氨酸合成酶。

[0098] 例如，本发明证实了通过异源 *kivd* 和 *adh2* ；大肠杆菌的 *ilvA*、*leuA*、*leuB*、*leuC*、*leuD*（或 *Leu* 操纵子，例如 *leuABCD*）；以及 *thrA*、*thrB*、*thrC* 或 *Thr* 操纵子（例如 *thrABC*，thrA 可能是抗反馈抑制多肽，例如 thrA*）的过表达可以生产出正丁醇和正丙醇。利用 2-酮戊酸生产正丁醇涉及到中间产物 2-酮基丁酸和非自然的正缬氨酸生物合成途径。由于 Kivd 与两种 2-酮酸的亲和力相近，且 2-酮基丁酸是 LeuA 的第二底物，因此，正丙醇与正丁醇一同被生产出来，且数量相近。

[0099] 本文提供的微生物经修饰能够大量产生亲本微生物中不能产生的代谢产物。“代谢产物”指微生物产生的任何物质或某代谢过程所需的或参与此过程的物质。一种代谢产物可以是作为起始物质的一种有机物（例如，葡萄糖或丙酮酸）；代谢过程中的一种中间产物（例如，2-酮酸）；或一种最终产物（例如，正丙醇、异丁醇、正丁醇、2-甲基-1-丁醇、3-甲基-1-丁醇或 2-苯基乙醇）。代谢产物可用来构建更复杂的分子，或被分解成更简单的分子。中间代谢产物可由其他代谢产物合成，可用来构建更复杂的物质，或被分解成更简单的化合物，通常伴随着化学能的释放。

[00100] 代表性代谢产物包括葡萄糖、丙酮酸、正丙醇、异丁醇、正丁醇、2-甲基 1-丁醇、3-甲基 1-丁醇或 2-苯基乙醇，以及 2-酮酸。如图 1A 所示，代表性的 2-酮酸中间产物包括 2-酮基丁酸、2-酮异戊酸、2-酮基戊酸、2-酮基 3-甲基戊酸、2-酮基 4-甲基-戊酸、和苯丙酮酸。图 1A 所示的代表性 2-酮酸可被用作正丙醇、异丁醇、正丁醇、2-甲基-1-丁醇、3-甲基-1-丁醇或 2-苯基乙醇生产过程中的代谢中间产物。例如，如图 1B 所示，一种重组微生物经代谢改造，可以提供表达水平增加的如由 Leu 操纵子（例如 LeuABCD）编码的 2-异丙基苹果酸合成酶、β-异丙基苹果酸脱氢酶和异丙基苹果酸异构酶，其催化 2-酮基丁酸生成 2-酮基戊酸。2-酮戊酸代谢物可

用于生产正丁醇，此过程由代谢修饰微生物产生的其他酶产生。此外，可通过重组微生物由 2-酮丁酸和 2-酮基-3-甲基戊酸生产 1-丙醇和 2-甲基-1-丁醇，此微生物经过代谢改造，可表达或过度表达乙酰羟酸合成酶、α-酮酸脱羧酶和醇脱氢酶，这些酶由 例如 *ilvIHDC*、*kdc* 和 *adh* 基因编码。此外，代谢产物 2-酮酸异戊酸可由经过代谢改造的重组微生物生成，此微生物能够表达或过度表达由例如 *ilvIHCD* 基因编码的乙酰羟酸合成酶。此代谢产物可继续用于异丁醇或 3-甲基-1-丁醇的生产。代谢产物丙酮酸和苯丙酮酸可用于由经过代谢改造的重组微生物生产 2-苯基乙醇，此微生物能够表达或过度表达 α-酮酸脱羧酶和醇脱氢酶，这些酶由 例如 *kdc* 和 *adh* 基因编码。其他代谢产物和基因见图 1B 。

[00101] 此外，本文提供的是生产异丁醇的重组微生物，在某些方面，可能包含目标酶的增加表达，例如 乙酰羟酸合成酶（例如 *ilvIH* 操纵子）、乙酰羟酸还原异构酶（例如 *ilvC*）、二羟酸脱水酶（例如 *ilvD*）、2-酮酸脱羧酶(例如 *PDC6*、*ARO10*、*THI3*、*kivd* 或 *pdc*)及醇脱氢酶(例如 *ADH2*)。此微生物还可含有乙醇脱氢酶（例如，*adhE*）、*ldh*（例如 *ldhA*）、 *frd*（例如 *frdB*、*frdC* 或 *frdBC*）、*fnr*、*leuA*、*ilvE*、*poxB*、*ilvA*、 *pflB* 、 *pta* 基因或它们的任何组合的表达缺失或受抑制，以增加丙酮酸的可利用性或减少会争夺所需的生物合成途径里的代谢物的酶。在某些方面，此重组微生物可以含有乙酰乳酸合成酶（例如 *alsS*）、乙酰羟酸还原异构酶（例如 *ilvC*）、二羟酸脱水酶（例如 *ilvD*）、2-酮酸脱羧酶（例如 *PDC6*、 *ARO10*、*TH13*、*kivd* 或 *pdc*），以及醇脱氢酶（例如 *ADH2*）的增加表达。关于醇脱氢酶方面，虽然乙醇脱氢酶属于醇脱氢酶，但是在此生物合成途径中，乙醇的合成是不需要的副反应。因此，当提到微生物里醇脱氢酶活性或表达增强时，明确不包括乙醇脱氢酶活性。

[00102] 本发明还提供产生正丁醇的重组微生物，可包括目标酶的增加表达，例如 2-异丙基苹果酸合成酶（例如 *leuA*）、β-异丙基苹果酸脱氢酶（例如 *leuB*）、异丙基苹果酸异构酶（例如 *leuC*、*leuD* 或 *leuCD* 操纵子）、丝氨酸脱水酶（例如 *ilvA*）。与亲本微生物相比，此微生物还可以含有 2-酮异戊酸、2-酮基-3-甲基-戊酸、 2-酮基-4-甲基戊酸或他们的任何组合的含量减少。此外，与亲本微生物相比，此微生物可含有二羟酸脱水酶（例如 *ilvD* 基因）表达破坏、缺失或敲除。产生正丁醇的重组微生物还可以含

有磷酸烯醇式丙酮酸羧化酶、丙酮酸羧化酶、天冬氨酸转氨酶、高丝氨酸脱氢酶、天冬氨酸-半醛脱氢酶、高丝氨酸激酶、苏氨酸合成酶、L-丝氨酸脱水酶和/或苏氨酸脱水酶的表达或活性增加，这些酶分别由源自 *ppc*、*pyc*、*aspC*、*thrA*、*asd*、*thrB*、*thrC*、*sdaAB*、*tdcB* 基因或其同源基因的核酸序列编码。

[00103] 本发明还提供产生正丙醇的重组微生物，可包括目标酶的增加表达，例如 α-异丙基苹果酸合成酶（*例如 cimA*）、β-异丙基苹果酸脱氢酶（*例如 leuB*）、异丙基苹果酸异构酶（*例如 leuCD 操纵子*）和丝氨酸脱水酶。

[00104] 本发明还提供产生 2-甲基-1-丁醇的重组微生物，可包含目标酶的增加表达，例如苏氨酸脱水酶（*例如 ilvA 或 tdcB*）、乙酰羟酸合成酶（*例如 ilvIH 操纵子*）、乙酰羟酸还原异构酶（*例如 ilvC*）、二羟酸脱水酶（*例如 ilvD*）,2-酮酸脱羧酶（*例如 PDC6*、*ARO10*、*THI3*、*kivd 和/或 pdc*）及醇脱氢酶（*例如 ADH2*）。

[00105] 本发明还提供产生 3-甲基-1-丁醇的重组微生物，可以含有目标酶的增加表达，例如乙酰乳酸合成酶（*例如 alsS*）、乙酰羟酸合成酶（*例如 ilvIH*）、乙酰乳酸合成酶（*例如 ilvMG*）或（*例如 ilvNB*）、乙酰羟酸还原异构酶（*例如 ilvC*）,二羟酸脱水酶（*例如 ilvD*）,2-异丙基苹果酸酶合成酶（*leuA*）、异丙基苹果酸异构酶（*例如 leuC*、*D* 或 *leuCD* 操纵子）、β-异丙基苹果酸脱氢酶（*例如 leuB*）、2-酮酸脱羧酶（*例如 kivd*、*PDC6* 或 *THI3*）及醇脱氢酶（*例如 ADH2*）。

[00106] 本发明还提供产生苯基乙醇的重组微生物，可以含有目标酶的增加表达，例如分支酸变位酶 P/预苯酸脱水酶（*例如 pheA*）、分支酸变位酶 T/预苯酸脱氢酶（*例如 tyrA*）、2-酮酸脱羧酶（*例如 kivd*、*PDC6*、或 *THI3*）和醇脱氢酶（*例如 ADH2*）。

[00107] 如前所述，本发明描述的目标酶一般都会生成代谢产物。例如，2-异丙基苹果酸合成酶（*leuA*）、β-异丙基苹果酸脱氢酶（*leuB*）和异丙基苹果酸异构酶（*leuC*、*leuD* 或 *leuCD* 操纵子）可由含有 2-酮基丁酸的底物生成 2-酮基戊酸。此外，本文所述的目标酶都由多聚核苷酸编码。例如，苏氨酸脱水酶由源自 *ilvA* 基因的多聚核苷酸编码。乙酰羟酸合成酶由源自 *ilvIH* 操纵子的多聚核苷酸编码。乙酰羟酸还原异构酶由源自 *ilvC* 基因的多聚核苷酸编码。二羟酸脱水酶由源自 *ilvD* 基因的多聚核苷酸编码。2-

酮酸脱羧酶由源自 *PDC6*、*ARO10*、*THI3*、*kivd* 和或 *pdc* 基因的多聚核苷酸编码。醇脱氢酶由源自 *ADH2* 基因的多聚核苷酸编码。其他酶和代表性基因详见本文。各种多肽与多聚核苷酸的同源体可以来自提供由合适多聚核苷酸编码的合适酶的任何生物来源。例如，可以通过参考各种数据库来确定同源体。

[00108] 本发明确定了对本发明的方法、组成成分和生物体有用的特定基因；但是，应当认识到，没必要完全识别这些基因。例如，为了改变活性，可以对含有编码某多肽或酶的序列的某种基因或多聚核苷酸进行改变或筛选。通常，此类改变包含保守突变和沉默突变。可以利用所属领域中的方法对这类修饰或突变多聚核苷酸和多肽进行功能酶活性表达筛选。

[00109] 由于此遗传密码固有的简并性，其他编码基本相同或功能等同的多肽的多聚核苷酸也可用于复制，并表达编码此类酶的多聚核苷酸。

[00110] 正如本发明所属领域的技术人员所理解的，修饰编码序列有利于增强在特定宿主中的表达。此遗传密码具有 64 个可能的密码子是冗余的，但是大多数生物体通常只利用其中的部分密码子。一个物种中最频繁利用的密码子称为最佳密码子，那些不经常被利用的称为稀有或低利用率密码子。密码子的取代反应了宿主对偏爱密码子的利用，此过程通常称为“密码子优化”或“物种密码子偏爱控制”。

[00111] 可以制备含有特定的原核宿主或真核宿主偏爱的密码子的优化编码序列（见 Murray 等， (1989) 《核酸研究》(Nucl. Acids Res.) 17:477-508），例如，以增加翻译率或生成具有所需的特性的重组 RNA 转录产物，例如与非优化序列产生的转录产物相比此转录产物具有更长的半衰期。也可根据宿主偏好修饰翻译终止密码子。例如，*酿酒酵母*和哺乳动物的典型终止密码子分别为 UAA 和 UGA。单子叶植物的典型终止密码子是 UGA，昆虫和大肠杆菌共同利用的终止密码子是 UAA (Dalphin 等， (1996) 《核酸研究》(Nucl. Acids Res.) 24: 216-218)。第 6,015,891 号美国专利以及本文引用的文献提供了用于植物中表达的优化核苷酸序列的方法。

[00112] 本发明所属领域的技术人员公认的，由于遗传密码的简并性，许多具有不同核苷酸序列的 DNA 化合物可用于编码本发明中的某种酶。本文对编码上述生物合成酶的天然 DNA 序列的引用仅为了说明本发明的一

个实施方案，公开的内容包括对本发明提供的方法中利用的酶和多肽的氨基酸序列进行编码的任何序列的 DNA 化合物。以类似的方式，通常可以允许多肽的氨基酸序列中发生一个或多个氨基酸的置换、缺失和插入，而不对其所需的活性造成损害或严重损害。本发明包含与本文所述的特定蛋白质具有不同氨基酸序列的多肽，这些经过修饰的多肽变体具有参照多肽的酶的合成代谢或催化代谢活性。此外，本文所示的 DNA 序列编码的氨基酸序列仅用于说明本发明的实施方案。

[00113] 此外，本文提供的方法和微生物包含可用于产生代谢物（例如，酮硫解酶、乙酰 CoA、乙酰基转移酶、羟丁酰 CoA 脱氢酶、巴豆酸酶、巴豆酰-CoA 还原酶、丁酰-CoA-脱氢酶、醇脱氢酶 (ADH)）的酶的同源体。针对第一个家族或物种的原始的酶或基因而言的所用术语“同源体”指第二个家族或物种的不同酶或基因，这些不同酶或基因通过进行功能、结构或基因组分析确定，对应于第一个家族或物种的原始酶或基因。同源体往往具有相似的功能、结构或基因组。通过基因探针和 PCR 能够易于克隆酶或基因的同源体的技术是已知的。通过功能分析或基因的基因组作图可确认复制的同源体序列的身份。

[00114] 如果编码一个蛋白质的核酸序列与编码第二个蛋白质的核酸序列相似，则该蛋白质与第二个蛋白质具有“同源性”或是“同源的”。或者，如果这两种蛋白质具有“相似”的氨基酸序列，则这两种蛋白质具有同源性（因此，术语“同源蛋白质”是指具有相似氨基酸序列的两种蛋白质）。

[00115] 如本文所使用的，当两种蛋白质（或蛋白质区域）的氨基酸序列具有至少 30%、40%、50%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98% 或 99%的同一性时，它们就基本同源。为了确定两种氨基酸序列或两种核酸序列的百分比同一性，将这两种序列进行比对以达到最优的比较目的（例如，可在第一和第二氨基酸或核酸之一或两者中插入空位以便于最优比对，并且通过比对能够忽略非同源序列）。在一实施方案中，进行比较目的时比对用的参考序列长度至少为该序列长度的 30%，通常为 40% 以上，较典型的长度为 50% 以上，更加典型的长度为 60% 以上，最典型的长度至少为 70%、80%、90%、100%。然后，比较相应位置的氨基酸残基或核苷酸。如果第一个序列某位置的氨基酸残基或核苷酸与第二个序列相应位置的相同，则这两个分子在

此位置上相同（如本文所使用的，氨基酸或核酸“相同”等同于氨基酸或核酸“同源”）。两种序列间的百分比同一性是这两种序列的相同位置数目的函数，同时要考虑空位数以及每段空位的长度，空位是为了便于最优比对两种序列而插入的。例如，*kivd* 基因包含来自其他生物体的编码具有基本相似酶活性的酶的同源基因（例如 *pdc6*、*aro10*、*thI3*、*pdc*、*kdcA*、*pdc1*、*pdc5*），以及与该参考基因的同一性至少为 30、40、50、60、70、80、85、90、95、98、或 99%并且它们编码的酶与参考基因编码的酶具有基本相似酶活性的基因。例如，乳酸克鲁维酵母(*Kluyveromyces lactis*)中的丙酮酸脱羧酶在氨基酸水平上与 Kivd 同一性为 37%； kivd 和 thI3 在核酸水平上的同一性为 32%；粟酒裂殖酵母(*Schizosaccharomyces pombe*)中的醇脱氢酶与酿酒酵母中的 ADH2 在氨基酸序列水平上的同一性为 52%；酿酒酵母 *adh2* 与 乳酸乳球菌 *adh* 的同一性为 49%；KIVD（乳酸乳球菌）与 PDC6（酿酒酵母）间的同一性为 36%（相似残基= 322/562 (57%)，空位 =24/562 (4%)）；KIVD（乳酸乳球菌）和 THI3（*酿酒酵母*）间的同一性为 32%（相似残基= 307/571 (53%)，空位 = 35/571 (6%)）；kivd（乳酸乳球菌）和 ARO10（*酿酒酵母*）间的同一性为 30%（相似残基= 296/598 (49%)，空位 = 65/598 (10%)）；ARO10（*酿酒酵母*）与 PDC6（*酿酒酵母*）间的同一性为 34%（相似残基= 320/616 (51%)，空位 = 61/616 (9%)）；ARO10（*酿酒酵母*）与 THI3（酿酒酵母）间的同一性为 30%（相似残基= 304/599 (50%)，空位 = 48/599 (8%)）；ARO10（*酿酒酵母*）与丙酮酸脱羧酶（丙酮丁醇梭菌 ATCC 824）间的同一性为 30%（相似残基= 291/613 (47%)，空位 = 73/613 (11%)）；PDC6（*酿酒酵母*）与 THI3（*酿酒酵母*）间的同一性为 50%（相似残基= 402/561 (71%)，空位 = 17/561 (3%)）；PDC6（*酿酒酵母*）与丙酮酸脱羧酶（丙酮丁醇梭菌 ATCC 824）间的同一性为 38%（相似残基= 328/570 (57%)，空位 = 30/570 (5%)）；THI3（*酿酒酵母*）与丙酮酸脱羧酶（丙酮丁醇梭菌 ATCC 824）间的同一性为 35%（相似残基= 284/521 (54%)，空位 = 25/521 (4%)）。本文所列的每个基因和多肽/酶的序列可通过互联网上的数据库轻易获得（例如 http:(//)eecoli.kaist.ac.kr/main.html）。此外，可利用本发明所属领域的常用算法对氨基酸序列和核酸序列进行比对，以确定其同一性。

[00116] 当“同源”一词用于蛋白质和多肽时，即承认不相同的残基位置通常其保守氨基酸置换不同。“保守氨基酸置换”指一个氨基酸残基被另一个具有侧链（R 基团）的氨基酸残基置换，此另一个氨基酸残基具有相似的化学性质（例如电荷、疏水性）。一般情况下，保守氨基酸置换不会显著改变蛋白质的功能特性。如果两个或多个氨基酸序列由于保守置换互不相同，则可以要向上调整序列同一性或同源度，以修正置换的保守性。本发明所属领域的技术人员很了解这种调整方法（见 Pearson 等，1994，以引用的方式并入本文）。

[00117] 以下六组中，每组的氨基酸都可以相互进行保守置换。1）丝氨酸（S）,苏氨酸（T）；2）天冬氨酸（D）、谷氨酸（E）；3）天冬酰胺（N）、谷胺酰胺（Q）；4）精氨酸（R）、赖氨酸（K）；5）异亮氨酸（I）、亮氨酸（L）、蛋氨酸（M）、丙氨酸（A）、缬氨酸（V）；6）苯丙氨酸（F）、酪氨酸（Y）、色氨酸（W）。

[00118] 多肽的序列同源也称为序列百分比同一性，通常利用序列分析软件对其进行测量。例如威斯康辛大学生物技术中心遗传学计算机集团（Genetics Computer Group）(GCG) 的序列分析软件包（Sequence Analysis Software Package），该公司位于威斯康星州麦迪逊大学街 910 号，邮编 53705。蛋白质分析软件通过测定各种置换、缺失及其他修饰(包括保守氨基酸置换)的同源性配对相似序列。例如，GCG 包含诸如“Gap(空位)”和“Bestfit(最佳配合)”之类的程序，他们能够结合使用默认参数来确定紧密相关的多肽例如来自不同种类生物体的同源多肽间的序列同源度或序列同一性，或野生型蛋白质与其突变体之间的序列同源度或序列同一性。见 GCG 6.1 版。

[00119] 计算机程序 BLAST（Altschul，1990；Gish 1993； Madden，1996；Altschul 1997；Zhang，1997）特别是 blastp 或 tblastn（Altschul，1997）是一种典型算法，它可以将一个分子序列与包含大量源自不同生物体的序列的数据库进行比较。BLASTp 的典型参数是： 期望值： 10（默认值）；过滤器: seg （默认值）；开放空位罚分： 11（默认值）；扩展空位罚分： 1（默认值）；最大比对: 100（默认值）；字长: 11（默认值）；描述量： 100（默认值）；矩阵罚分： BLOWSUM62。

[00120]在搜索包含来自大量不同生物体序列的数据库时，通常要比对氨基酸序列。可以利用本发明所属领域内除 blastp 以外的算法分析氨基酸序列搜索数据库。例如，可以利用 FASTA，GCG 6.1 版中的一个程序，来比对多肽序列。FASTA 提供查询和搜索序列间最佳重叠区域的比对和序列百分比同一性（Pearson，1990，以引用的方式并入本文）。例如，可以利用 FASTA 及 GCG 6.1 版中提供的默认参数（字长为 2，以及 PAM250 得分矩阵），确定氨基酸序列间的序列百分比同一性，此参数以引用的方式特此并入本文。

[00121]下表及公开内容提供了基因以及每种基因的同源基因的非限定性例子，这些基因的多聚核苷酸和多肽序列是本发明所属领域的技术人员可以得到的。

表 1： 描述了生产各种高级醇的重组途径（“+”= 表达、表达或活性增强/“-”=表达或活性降低或敲除*）。

酶	代表性基因	异丁醇	正丁醇（由 L-苏氨酸生成）	正丁醇（由丙酮酸生成）	正丙醇（由 L-苏氨酸生成）	正丙醇（由丙酮酸生成）	3-甲基-1-丁醇（由丙酮酸生成）	2-甲基-1-丁醇（由 L-苏氨酸生成）
乙醇脱氢酶	adhE	-	-	-	-	-	-	-
乳酸脱氢酶	ldhA	-	-	-	-	-	-	-
延胡索酸还原酶	frdBC	-					-	-
	fnr	-					-	-
乙酸激酶	ackA	-	-	-	-	-	-	-
磷酸乙酰基转移酶	pta	-	-	-	-	-	-	-
甲酸乙酰基转移酶	pflB	-					-	-
α-异丙基苹果酸合成酶	leuA	-	+	+			+	
β-异丙基苹果酸脱氢酶	leuB		+	+	-	+	+	
α-异丙基苹果酸异构酶	leuC		+	+		+	+	
α-异丙基苹果酸异构酶	leuD		+	+		+		
BCAA 转氨酶	ilvE	-					-	
酪氨酸转氨酶	tyrB, tyrAT						-	
丙酮酸脱氢酶	poxB	-	-	-	-	-		
乙酰乳酸合成酶	ilvB		-	-	-	-		
乙酰乳酸合成酶	ilvI, alsS		-	-	-	-		
苏氨酸脱水酶	ilvA, tdcB	-	+	+	+	+		+
高丝氨酸琥珀酰基转移酶	metA		-	-	-	-		-

L-苏氨酸 3-脱氢酶	tdh		-	-	-	-		-
乙酰羟酸合成酶	ilvHI, ilvNB, ilvGM, alsS	+					+	+
乙酰羟酸还原异构酶	ilvC, ilv5	+					+	+
二羟酸脱水酶	ilvD, ilv3	+					+	+
2-酮酸脱羧酶	pdc6, aro10, thI3, kivd, pdc, kdcA, pdc1, pdc5	+	+	+	+	+	+	+
醇脱氢酶	adh1, adh2, adh3, adh4, adh5, adh6, sfa1	+	+	+	+	+	+	+
柠苹酸合成酶	cimA			+		+		

* 在产物的合成过程中，敲除或表达降低是可以选择的，这种敲除增加了各种底物中间体的数量，并提高了产量。

[00122]表 2-9 阐明了本发明各种重组微生物的反应途径，并列出了代表性基因和同源基因及其生物来源。

表 2

异丁醇产生途径（由丙酮酸生成）

第一步反应
丙酮酸 -> 2-乙酰乳酸
基因
ilvHI（大肠杆菌）、*ilvNB（大肠杆菌）*、*ilvGM（大肠杆菌）*、*alsS（枯草芽孢杆菌）*或其同源基因
第二步反应
2-乙酰乳酸 -> 2,3-二羟基-异戊酸
基因
*ilvC（大肠杆菌）*或其同源基因
第三步反应
2,3-二羟基-异戊酸 -> 2-酮基-异戊酸
基因
*ilvD（大肠杆菌）*或其同源基因
第四步反应
2-酮基-异戊酸 -> 异丁醛
基因
kivd（乳酸乳球菌）、*kdcA（乳酸乳球菌）*、*PDC1（酿酒酵母）*、*PDC5（酿酒酵母）*、*PDC6（酿酒酵母）*、*THI3（酿酒酵母）*、*ARO10（酿酒酵母）*或其同源基因
第五步反应
异丁醛 -> 异丁醇
基因
ADH1（酿酒酵母）、*ADH2（酿酒酵母）*、*ADH3（酿酒酵母）*、*ADH4（酿酒酵母）*，*ADH5（酿酒酵母）*、*ADH6（酿酒酵母）*、*SFA1（酿酒酵母）*或其同源基因

表 3

由 L-苏氨酸生成正丁醇的反应途径

第一步反应
L-苏氨酸 -> 2-酮基-丁酸
基因
ilvA（大肠杆菌）、 *tdcB（大肠杆菌）*或其同源基因
第二步反应
2-酮基-丁酸 -> 2-苹果酸乙酯
基因
*leuA（大肠杆菌）*或其同源基因
第三步反应
2-苹果酸乙酯 -> 3-苹果酸乙酯
基因
*leuCD（大肠杆菌）*或其同源基因
第四步反应
3-苹果酸乙酯 -> 2-乙基-3-氧代琥珀酸
基因
*leuB（大肠杆菌）*或其同源基因
第五步反应
2-乙基-3-氧代琥珀酸 -> 2-酮基-戊酸
基因
无（自发反应）
第六步反应
2-酮基-戊酸 -> 丁醛
基因
kivd（乳酸乳球菌）、*kdcA（乳酸乳球菌）*、*PDC1（酿酒酵母）*、*PDC5（酿酒酵母）*、*PDC6（酿酒酵母）*、*THI3（酿酒酵母）*、*ARO10（酿酒酵母）*或其同源基因
第七步反应
丁醛 -> 正丁醇
基因
ADH1（酿酒酵母）、*ADH2（酿酒酵母）*、*ADH3（酿酒酵母）*、*ADH4（酿酒酵母）*，*ADH5（酿酒酵母）*、*ADH6（酿酒酵母）*、*SFA1（酿酒酵母）*或其同源基因

表 4

由丙酮酸生成正丁醇的反应途径

第一步反应
丙酮酸 + 乙酰 CoA -> (R)-柠苹酸
基因
cimA（詹氏甲烷暖球菌）、*cimA（问号钩端螺旋体）*或其同源基因
第二步反应
(R)-柠苹酸 -> 柠康酸
基因
leuCD（问号钩端螺旋体）、*leuCD（大肠杆菌）*或其同源基因
第三步反应
柠康酸 ->β-甲基-D-苹果酸
基因
leuCD（问号钩端螺旋体）、*leuCD（大肠杆菌）*或其同源基因
第四步反应
β-甲基-D-苹果酸 -> 2-酮基-丁酸
基因
leuB（问号钩端螺旋体）、*leuB（大肠杆菌）*或其同源基因
第五步反应
2-酮基-丁酸 -> 2-苹果酸乙酯
基因
*leuA（大肠杆菌）*或其同源基因
第三步反应
2-苹果酸乙酯 -> 3-苹果酸乙酯
基因
*leuCD（大肠杆菌）*或其同源基因
第四步反应
3-苹果酸乙酯 -> 2-乙基-3-氧代琥珀酸
基因
*leuB（大肠杆菌）*或其同源基因
第五步反应
2-乙基-3-氧代琥珀酸 -> 2-酮基-戊酸
基因
（自发反应）

表 4 （续）

第六步反应
2-酮基-戊酸 -> 丁醛
基因
kivd（乳酸乳球菌）、*kdcA（乳酸乳球菌）*、*PDC1（酿酒酵母）*、*PDC5（酿酒酵母）*，*PDC6（酿酒酵母）*、*THI3（酿酒酵母）*、*ARO10（酿酒酵母）*或其同源基因
第 7 步反应
丁醛 -> 正丁醇
基因
ADH1（酿酒酵母）、*ADH2（酿酒酵母）*、*ADH3（酿酒酵母）*、*ADH4（酿酒酵母）*、*ADH5（酿酒酵母）*、*ADH6（酿酒酵母）*、*SFA1（酿酒酵母）*或其同源基因

表 5

由 L-苏氨酸生成正丙醇的反应途径

第一步反应
L-苏氨酸 -> 2-酮基-丁酸
基因
ilvA（大肠杆菌）、*tdcB（大肠杆菌）*或其同源基因
第二步反应
2-酮基-丁酸 -> 丙醛
基因
kivd（乳酸乳球菌）、*kdcA（乳酸乳球菌）*、*PDC1（酿酒酵母）*、*PDC5（酿酒酵母）*、*PDC6（酿酒酵母）*、*THI3（酿酒酵母）*、*ARO10（酿酒酵母）*或其同源基因
第三步反应
丙醛 -> 正丙醇
基因
ADH1（酿酒酵母）、*ADH2（酿酒酵母）*、*ADH3（酿酒酵母）*、*ADH4（酿酒酵母）*、*ADH5（酿酒酵母）*、*ADH6（酿酒酵母）*、*SFA1（酿酒酵母）*或其同源基因

表 6

由丙酮酸生成正丙醇的反应途径

第一步反应
丙酮酸 + 乙酰 CoA -> (R)-柠苹酸
基因
cimA（詹氏甲烷嗳球菌）、*cimA（问号钩端螺旋体）*或其同源基因
第二步反应
(R)-柠苹酸 -> 柠康酸
基因
leuCD（问号钩端螺旋体）、*leuCD（大肠杆菌）*或其同源基因
第三步反应
柠康酸 -> β-甲基-D-苹果酸
基因
leuCD（问号钩端螺旋体）、*leuCD（大肠杆菌）*或其同源基因
第四步反应
β-甲基-D-苹果酸 -> 2-酮基-丁酸
基因
leuB（问号钩端螺旋体）、*leuB（大肠杆菌）*或其同源基因
第五步反应
2-酮基丁酸 -> 丁醛
基因
kivd（乳酸乳球菌）、*kdcA（乳酸乳球菌）*、*PDC1（酿酒酵母）*、*PDC5（酿酒酵母）*、*PDC6（酿酒酵母）*、*THI3（酿酒酵母）*、*ARO10（酿酒酵母）*或其同源基因
第七步反应
丁醛 -> 正丁醇
基因
ADH1（酿酒酵母）、*ADH2（酿酒酵母）*、*ADH3（酿酒酵母）*、*ADH4（酿酒酵母）*、*ADH5（酿酒酵母）*、*ADH6（酿酒酵母）*、*SFA1（酿酒酵母）*或其同源基因

表 7

由 L-苏氨酸生成 2-甲基-1-丁醇的反应途径

第一步反应
L-苏氨酸 -> 2-酮基-丁酸
基因
ilvA（*大肠杆菌*）、*tdcB*（*大肠杆菌*）或其同源基因
第二步反应
2-酮基-丁酸 -> 2-乙酰基-2-羟基-丁酸
基因
ilvHI（*大肠杆菌*）、*ilvNB*（*大肠杆菌*）、*ilvGM*（*大肠杆菌*）、*alsS*（*枯草芽孢杆菌*）或其同源基因
第三步反应
2-乙酰基-2-羟基-丁酸 -> 2,3-二羟基-3-甲基戊酸
基因
ilvcC（*大肠杆菌*）或其同源基因
第四步反应
2,3-二羟基-3-甲基戊酸 -> 2-酮基-3-甲基戊酸
基因
ilvD（*大肠杆菌*）或其同源基因
第五步反应
2-酮基-3-甲基戊酸 -> 2-甲基丁醛
基因
kivd（*乳酸乳球菌*）、*kdcA*（*乳酸乳球菌*）、*PDC1*（*酿酒酵母*）、*PDC5*（*酿酒酵母*）、*PDC6*（*酿酒酵母*）、*THI3*（*酿酒酵母*）、*ARO10*（*酿酒酵母*）或其同源基因
第六步反应
2-甲基丁醛 -> 2-甲基-1-丁醇
基因
ADH1（*酿酒酵母*）、*ADH2*（*酿酒酵母*）、*ADH3*（*酿酒酵母*）、*ADH4*（*酿酒酵母*），*ADH5*（*酿酒酵母*）、*ADH6*（*酿酒酵母*）、*SFA1*（*酿酒酵母*）或其同源基因

表 8

由丙酮酸生成3-甲基-1-丁醇的反应途径

第一步反应
丙酮酸-> 2-乙酰乳酸
基因
ilvHI（大肠杆菌）、*ilvNB（大肠杆菌）*、*ilvGM（大肠杆菌）*、*alsS（枯草芽孢杆菌）*或其同源基因
第二步反应
2-乙酰乳酸 -> 2,3-二羟基-异戊酸
基因
*ilvC（大肠杆菌）*或其同源基因
第三步反应
2,3-二羟基-异戊酸 -> 2-酮基-异戊酸
基因
*ilvD（大肠杆菌）*或其同源基因
第四步反应
2-酮基-异戊酸 -> 2-异丙基苹果酸
基因
*leuA（大肠杆菌）*或其同源基因
第五步反应
2-异丙基苹果酸 -> 3-异丙基苹果酸
基因
*leuCD（大肠杆菌）*或其同源基因
第六步反应
3-异丙基苹果酸 -> 2-异丙基-3-氧代琥珀酸
基因
*leuB（大肠杆菌）*或其同源基因
第七步反应
2-异丙基-3-氧代琥珀酸 -> 2-酮基异己酸
基因
（自发反应）
第八步反应
2-酮基异己酸 -> 3-甲基丁醛
基因
kivd（乳酸乳球菌）、*kdcA（乳酸乳球菌）*、*PDC1（酿酒酵母）*、*PDC5（酿酒酵母）*、*PDC6（酿酒酵母）*、*THI3（酿酒酵母）*、*ARO10（酿酒酵母）*或其同源基因
第九步反应
3-甲基丁醛 -> 3-甲基-1-丁醇
基因
ADH1（酿酒酵母）、*ADH2（酿酒酵母）*、*ADH3（酿酒酵母）*、*ADH4（酿酒酵母）*、*ADH5（酿酒酵母）*、*ADH6（酿酒酵母）*、*SFA1（酿酒酵母）*或其同源基因

表 9

由分支酸生成苯基乙醇的反应途径

第一步反应
分支酸 -> 预苯酸
基因
tyrA（大肠杆菌）、*pheA（大肠杆菌）*或其同源基因
第二步反应
预苯酸 -> 苯丙酮酸
基因
*pheA（大肠杆菌）*或其同源基因
第三步反应
苯丙酮酸 -> 苯甲醛
基因
kivd（乳酸乳球菌）、*kdcA（乳酸乳球菌）*、*PDC1（酿酒酵母）*、*PDC5（酿酒酵母）*、*PDC6（酿酒酵母）*、*THI3（酿酒酵母）*、*ARO10（酿酒酵母）*或其同源基因
第四步反应
苯甲醛 -> 2-苯基乙醇
基因
ADH1（酿酒酵母）、*ADH2（酿酒酵母）*、*ADH3（酿酒酵母）*、*ADH4（酿酒酵母）*，*ADH5（酿酒酵母）*、*ADH6（酿酒酵母）*、*SFA1（酿酒酵母）*或其同源基因

[00123] 本发明提供了可用于本文所述的重组微生物构建的各种基因、同源体及突变体的登记号。应当了解，本文所述的同源体和突变体为代表性的而非限制性的。本发明所属领域的技术人员可利用各种数据库获得其他同源体、突变体和序列，这些数据库包括例如，可通过互联网访问的美国国家生物技术信息中心（NCBI，National Center for Biotechnology Information）。

[00124] 乙醇脱氢酶（又称为醛醇脱氢酶 ）由大肠杆菌中的 *adhE* 编码。adhE 含有三种酶的活性：醇脱氢酶（ADH）；乙醛/乙酰-CoA 脱氢酶（ACDH）；丙酮酸-甲酸-裂解酶失活酶（PFL 失活酶）；PFL 失活酶的活性能够催化铁、NAD 和 CoA 依赖性反应中丙酮酸-甲酸-裂解酶催化剂的灭活。此项技术中的同源体为已知（见醛/醇脱氢酶（Polytomella 种 Pringsheim 198.80) gi|40644910|emb|CAD42653.2|(40644910)；醛/醇脱氢酶（A 型*肉毒梭菌* ATCC 3502） gi|148378348|ref|YP_001252889.1|(148378348)；醛/醇脱氢酶（*鼠疫耶尔*

森氏杆菌 CO92）gi|16122410|ref|NP_405723.1|(16122410)；醛/醇脱氢酶（*假结核耶尔森氏菌* IP 32953） gi|51596429|ref|YP_070620.1|(51596429)；醛/醇脱氢酶（*鼠疫耶尔森氏杆菌* CO92）gi|115347889|emb|CAL20810.1|(115347889)；醛/醇脱氢酶（*假结核耶尔森氏菌* IP 32953）
gi|51589711|emb|CAH21341.1|(51589711) 醛/醇脱氢酶（大肠杆菌 CFT073）gi|26107972|gb|AAN80172.1|AE016760_31(26107972)； 醛/醇脱氢酶（田鼠型鼠疫耶尔森氏杆菌生物变种 91001）gi|45441777|ref|NP_993316.1|(45441777)；醛/醇脱氢酶（田鼠型鼠疫耶尔森氏杆菌生物变种 91001）
gi|45436639|gb|AAS62193.1|(45436639)；醛/醇脱氢酶（*产气荚膜梭菌* ATCC 13124）
gi|110798574|ref|YP_697219.1|(110798574)（*希瓦氏菌* MR-1）
gi|24373696|ref|NP_717739.1|(24373696)；醛/醇脱氢酶（A 型*肉毒梭菌* ATCC 19397）gi|153932445|ref|YP_001382747.1|(153932445)；醛/醇脱氢酶（*古典型鼠疫耶尔森氏杆菌生物变种* E1979001) gi|165991833|gb|EDR44134.1|(165991833)；醛/醇脱氢酶（A 型*肉毒梭菌* Hall
株)gi|153937530|ref|YP_001386298.1|(153937530)；醛/醇脱氢酶(*肉毒梭菌* ATCC 13124）gi|110673221|gb|ABG82208.1|(110673221)；醛/醇脱氢酶（A 型*肉毒梭菌* Hall 株）gi|152933444|gb|ABS38943.1|(152933444)；醛/醇脱氢酶（*东方型鼠疫耶尔森氏杆菌生物变种* F1991016）gi|165920640|gb|EDR37888.1|(165920640)；醛/醇脱氢酶（*东方型鼠疫耶尔森氏杆菌生物变种* IP275）
gi|165913933|gb|EDR32551.1|(165913933)；醛/醇脱氢酶（*安哥拉鼠疫耶尔森氏杆菌*）gi|162419116|ref|YP_001606617.1|(162419116)；醛/醇脱氢酶（F 型*肉毒梭菌* Langeland）gi|153940830|ref|YP_001389712.1|(153940830)；醛/醇脱氢酶（*大肠杆菌* HS）gi|157160746|ref|YP_001458064.1|(157160746)；醛/醇脱氢酶（*大肠杆菌* E24377A）gi|157155679|ref|YP_001462491.1|(157155679)；醛/醇脱氢酶（*小肠结肠炎耶尔森氏杆菌小肠结肠炎亚种* 8081）
gi|123442494|ref|YP_001006472.1|(123442494)；醛/醇脱氢酶（聚球藻*种* JA-3-3Ab）gi|86605191|ref|YP_473954.1|(86605191)；醛/醇脱氢酶（*单核细胞增生李斯特氏菌* 4b F2365）gi|46907864|ref|YP_014253.1|(46907864)；醛/醇脱氢酶（*粪肠球菌* V583）gi|29375484|ref|NP_814638.1|(29375484)；醛/醇脱氢酶（*无乳链球菌* 2603V/R）gi|22536238|ref|NP_687089.1|(22536238)；醛/醇脱氢酶（A 型

肉毒梭菌 ATCC 19397）gi|152928489|gb|ABS33989.1|(152928489)；醛/醇脱氢酶（*大肠杆菌* E24377A）gi|157077709|gb|ABV17417.1|(157077709)；醛/醇脱氢酶（*大肠杆菌* HS）gi|157066426|gb|ABV05681.1|(157066426)；醛/醇脱氢酶（F型*肉毒梭菌* Langeland）gi|152936726|gb|ABS42224.1|(152936726))；醛/醇脱氢酶(*鼠疫耶尔森氏杆菌* CA88-4125）gi|149292312|gb|EDM42386.1|(149292312)；醛/醇脱氢酶（*小肠结肠炎耶尔森氏杆菌小肠结肠炎亚种* 8081）gi|122089455|emb|CAL12303.1|(122089455)；醛/醇脱氢酶（*菜茵衣藻*）gi|92084840|emb|CAF04128.1|(92084840)；醛/醇脱氢酶（聚球藻种 JA-3-3Ab）gi|86553733|gb|ABC98691.1|(86553733)；醛/醇脱氢酶（*希瓦氏菌* MR-1）gi|24348056|gb|AAN55183.1|AE015655_9(24348056)；醛/醇脱氢酶（*粪肠球菌* V583）gi|29342944|gb|AAO80708.1|(29342944)；醛/醇脱氢酶（*单核细胞增生李斯特氏菌* 4b F2365）gi|46881133|gb|AAT04430.1|(46881133)；醛/醇脱氢酶（*单核细胞增生李斯特氏菌* 1/2a F6854) gi|47097587|ref|ZP_00235115.1|(47097587)；醛/醇脱氢酶（*单核细胞增生李斯特氏菌* 4b H7858）gi|47094265|ref|ZP_00231973.1|(47094265)；醛/醇脱氢酶（*单核细胞增生李斯特氏菌* 4b H7858）gi|47017355|gb|EAL08180.1|(47017355)；醛/醇脱氢酶（*单核细胞增生李斯特氏菌* 1/2a F6854）gi|47014034|gb|EAL05039.1|(47014034)；醛/醇脱氢酶（*无乳链球菌* 2603V/R）gi|22533058|gb|AAM98961.1|AE014194_6(22533058)p；醛/醇脱氢酶（古典型*鼠疫耶尔森氏杆菌生物变种* E1979001）gi|166009278|ref|ZP_02230176.1|(166009278)；醛/醇脱氢酶（*东方型鼠疫耶尔森氏杆菌生物变种* IP275）gi|165938272|ref|ZP_02226831.1|(165938272)；醛/醇脱氢酶（*东方型鼠疫耶尔森氏杆菌生物变种* F1991016）gi|165927374|ref|ZP_02223206.1|(165927374)；醛/醇脱氢酶（*安哥拉鼠疫耶尔森氏杆菌*）gi|162351931|gb|ABX85879.1|(162351931)；醛/醇脱氢酶（*假结核耶尔森氏菌* IP 31758）gi|153949366|ref|YP_001400938.1|(153949366)；醛/醇脱氢酶（*假结核耶尔森氏菌* IP 31758）gi|152960861|gb|ABS48322.1|(152960861)；醛/醇脱氢酶（*鼠疫耶尔森氏杆菌* CA88-4125）gi|149365899|ref|ZP_01887934.1|(149365899)；乙醛脱氢酶（乙酰化）（*大肠杆菌* CFT073）gi|26247570|ref|NP_753610.1|(26247570)；醛/醇脱氢酶（包括：醇脱氢酶；乙醛脱氢酶（乙酰化）（EC 1.2.1.10）（acdh）；丙酮酸-甲酸-裂解酶失

活酶（pfl 失活酶））（A 型*肉毒梭菌* ATCC 3502）gi|148287832|emb|CAL81898.1|(148287832)；醛/醇脱氢酶（包括：醇脱氢酶（ADH）；乙醛脱氢酶（乙酰化）（ACDH）；丙酮酸-甲酸-裂解酶失活酶（PFL 失活酶））gi|71152980|sp|P0A9Q7.2|ADHE_ECOLI(71152980)；醛/醇脱氢酶（包括：醇脱氢酶、乙醛脱氢酶和丙酮酸-甲酸-裂解酶失活酶（*胡萝卜软腐欧式杆菌胡萝卜软腐亚种* SCRI1043）gi|50121254|ref|YP_050421.1|(50121254)；醛/醇脱氢酶（包括：醇脱氢酶、乙醛脱氢酶和丙酮酸-甲酸-裂解酶失活酶（*胡萝卜软腐欧式杆菌胡萝卜软腐亚种* SCRI1043）gi|49611780|emb|CAG75229.1|(49611780)；醛/醇脱氢酶（包括：醇脱氢酶（ADH）；乙醛脱氢酶（乙酰化）（ACDH））gi|19858620|sp|P33744.3|ADHE_CLOAB(19858620)；醛/醇脱氢酶（包括：醇脱氢酶（ADH）；乙醛脱氢酶（乙酰化）（ACDH）；丙酮酸-甲酸-裂解酶失活酶（PFL 失活酶））gi|71152683|sp|P0A9Q8.2|ADHE_ECO57(71152683)；醛/醇脱氢酶（包括：醇脱氢酶；乙醛脱氢酶（乙酰化）；丙酮酸-甲酸-裂解酶失活酶（*艰难梭菌* 630）gi|126697906|ref|YP_001086803.1|(126697906)；醛/醇脱氢酶（包括：醇脱氢酶；乙醛脱氢酶（乙酰化）；丙酮酸-甲酸-裂解酶失活酶（*艰难梭菌* 630）gi|115249343|emb|CAJ67156.1|(115249343)；醛/醇脱氢酶（包括：醇脱氢酶（ADH）；乙醛脱氢酶（乙酰化）（ACDH）；丙酮酸-甲酸-裂解酶失活酶（PFL 失活酶））（*发光杆菌* laumondii 亚种 TTO1）gi|37526388|ref|NP_929732.1|(37526388)；醛/醇脱氢酶 2（包括：醇脱氢酶；乙醛脱氢酶）（*化脓链球菌* Manfredo）gi|134271169|emb|CAM29381.1|(134271169)；醛/醇脱氢酶（包括：醇脱氢酶（ADH）；乙醛脱氢酶（乙酰化）（ACDH）；丙酮酸-甲酸-裂解酶失活酶（PFL 失活酶））（*发光杆菌* laumondii 亚种 TTO1）gi|36785819|emb|CAE14870.1|(36785819)；醛/醇脱氢酶（包括：醇脱氢酶和丙酮酸-甲酸-裂解酶失活酶（*艰难梭菌* 630）gi|126700586|ref|YP_001089483.1|(126700586)；醛/醇脱氢酶（包括：醇脱氢酶和丙酮酸-甲酸-裂解酶失活酶（*艰难梭菌* 630）gi|115252023|emb|CAJ69859.1|(115252023)；醛/醇脱氢酶 2（*化脓链球菌* Manfredo）gi|139472923|ref|YP_001127638.1|(139472923)；醛/醇脱氢酶 E（*产气荚膜梭菌* 13) gi|18311513|ref|NP_563447.1|(18311513)；醛/醇脱氢酶 E（*产气荚*

膜梭菌 13）gi|18146197|dbj|BAB82237.1|(18146197)；醛/醇脱氢酶 ADHE1（丙酮丁醇梭菌 ATCC 824）gi|15004739|ref|NP_149199.1|(15004739)；醛/醇脱氢酶 ADHE1 （丙酮丁醇梭菌 ATCC 824）gi|14994351|gb|AAK76781.1|AE001438_34(14994351)；醛/醇脱氢酶 2（包括 醇脱氢酶（ADH）；乙醛/乙酰 CoA 脱氢酶（ACDH））gi|2492737|sp|Q24803.1|ADH2_ENTHI(2492737)；醇脱氢酶（*肠道沙门氏菌*肠道伤寒杆菌亚种 CT18）gi|16760134|ref|NP_455751.1|(16760134)；以及醇脱氢酶（*肠道沙门氏菌*肠道伤寒杆菌亚种）gi|16502428|emb|CAD08384.1|(16502428))，每个带有登记号的序列全部以引用的方式并入本文。

[00125]乳酸脱氢酶（又称为 D-乳酸脱氢酶和发酵脱氢酶）由大肠杆菌中的 ldhA 编码，它能催化丙酮酸到乳酸的 NADH 依赖型的转化。ldhA 的同源基因和突变基因已知。事实上，目前通过 NCBI 可以查到 1664 种细菌乳酸脱氢酶。例如，这些同源基因和突变基因包括 D-乳酸脱氢酶（D-LDH）（发酵乳酸脱氢酶）gi|1730102|sp|P52643.1|LDHD_ECOLI(1730102)；D-乳酸脱氢酶 gi|1049265|gb|AAB51772.1|(1049265)；D-乳酸脱氢酶（*大肠杆菌* APEC O1）gi|117623655|ref|YP_852568.1|(117623655)；D-乳酸脱氢酶（*大肠杆菌* CFT073）gi|26247689|ref|NP_753729.1|(26247689)；D-乳酸脱氢酶（*大肠杆菌* O157:H7 EDL933）gi|15801748|ref|NP_287766.1|(15801748)；D-乳酸脱氢酶（*大肠杆菌* APEC O1）gi|115512779|gb|ABJ00854.1|(115512779))；D-乳酸脱氢酶（*大肠杆菌* CFT073）gi|26108091|gb|AAN80291.1|AE016760_150(26108091)；NAD 依赖性发酵 D-乳酸脱氢酶（*大肠杆菌* K12） gi|16129341|ref|NP_415898.1|(16129341)；NAD 依赖性发酵 D-乳酸脱氢酶（*大肠杆菌* UTI89）gi|91210646|ref|YP_540632.1|(91210646)；NAD 依赖性发酵 D-乳酸脱氢酶（*大肠杆菌* K12）gi|1787645|gb|AAC74462.1|(1787645)；NAD 依赖性发酵 D-乳酸脱氢酶（*大肠杆菌* W3110）gi|89108227|ref|AP_002007.1|(89108227)；NAD 依赖性发酵 D-乳酸脱氢酶（*大肠杆菌* W3110） gi|1742259|dbj|BAA14990.1|(1742259)；NAD 依赖性发酵 D-乳酸脱氢酶（*大肠杆菌* UTI89）gi|91072220|gb|ABE07101.1|(91072220)；NAD 依赖性发酵 D-乳酸脱氢酶（*大肠杆菌* O157:H7 EDL933）gi|12515320|gb|AAG56380.1|AE005366_6(12515320)；NAD 依赖性发酵 D-乳酸脱氢酶（*大肠杆菌* O157:H7 str. Sakai）gi|13361468|dbj|BAB35425.1|(13361468)；COG1052： 乳酸脱氢酶及相关脱氢酶

（*大肠杆菌* 101-1）gi|83588593|ref|ZP_00927217.1|(83588593); COG1052: 乳酸脱氢酶及相关脱氢酶（*大肠杆菌* 53638）
gi|75515985|ref|ZP_00738103.1|(75515985); COG1052: 乳酸脱氢酶及相关脱氢酶（*大肠杆菌* E22）gi|75260157|ref|ZP_00731425.1|(75260157); COG1052: 乳酸脱氢酶及相关脱氢酶（*大肠杆菌* F11）
gi|75242656|ref|ZP_00726400.1|(75242656); COG1052: 乳酸脱氢酶及相关脱氢酶（*大肠杆菌* E110019）gi|75237491|ref|ZP_00721524.1|(75237491); COG1052: 乳酸脱氢酶及相关脱氢酶（*大肠杆菌* B7A）
gi|75231601|ref|ZP_00717959.1|(75231601); 和 COG1052: 乳酸脱氢酶及相关脱氢酶（*大肠杆菌* B171）
gi|75211308|ref|ZP_00711407.1|(75211308)，每个带有登记号的序列全部以引用的方式并入本文。

[00126] 两种含有 FAD 的膜结合酶负责催化延胡索酸和琥珀酸的相互转化；延胡索酸还原酶用于厌氧生长，琥珀酸脱氢酶用于好氧生长。延胡索酸还原酶包含多个亚基（例如*大肠杆菌*中的 frdA、B、C）。对任一亚基进行修饰都会得到本文所需的活性。例如，本发明的方法可使用 frdB、frdC 或 frdBC 的敲除。Frd 同源体和突变体已知。例如，同源体和突变体包括如延胡索酸还原酶亚基 D（延胡索酸还原酶 13 kDa 疏水蛋白）
gi|67463543|sp|P0A8Q3.1|FRDD_ECOLI(67463543); 延胡索酸还原酶亚基 C（延胡索酸还原酶 15 kDa 疏水蛋白）
gi|1346037|sp|P20923.2|FRDC_PROVU(1346037); 延胡索酸还原酶亚基 D （延胡索酸还原酶 13 kDa 疏水蛋白)gi|120499|sp|P20924.1|FRDD_PROVU(120499); 延胡索酸还原酶亚基 C （延胡索酸还原酶 15 kDa 疏水蛋白）
gi|67463538|sp|P0A8Q0.1|FRDC_ECOLI(67463538); 延胡索酸还原酶铁-硫亚基（*大肠杆菌*）gi|145264|gb|AAA23438.1|(145264); 延胡索酸还原酶黄素蛋白亚基（大肠杆菌）gi|145263|gb|AAA23437.1|(145263); 延胡索酸还原酶黄素蛋白亚基 gi|37538290|sp|P17412.3|FRDA_WOLSU(37538290); 延胡索酸还原酶黄素蛋白亚基 gi|120489|sp|P00363.3|FRDA_ECOLI(120489); 延胡索酸还原酶黄素蛋白亚基 gi|120490|sp|P20922.1|FRDA_PROVU(120490); 黄素细胞色素 c）（黄素细胞色素 c3）（Fcc3）gi|119370087|sp|Q07WU7.2|FRDA_SHEFN(119370087); 延胡索酸还原酶铁-硫亚基 gi|81175308|sp|P0AC47.2|FRDB_ECOLI(81175308)延胡索酸

还原酶黄素蛋白亚基（黄素细胞色素 c）（黄素细胞色素 c3）（Fcc3）；gi|119370088|sp|P0C278.1|FRDA_SHEFR(119370088)；Frd 操纵子非典型蛋白质 C gi|140663|sp|P20927.1|YFRC_PROVU(140663)；Frd 操纵子可能的铁-硫亚基 A gi|140661|sp|P20925.1|YFRA_PROVU(140661)；延胡索酸还原酶铁-硫亚基 gi|120493|sp|P20921.2|FRDB_PROVU(120493)；延胡索酸还原酶黄素蛋白亚基；gi|2494617|sp|O06913.2|FRDA_HELPY(2494617)；延胡索酸还原酶黄素蛋白亚基前体（三价铁离子诱发的黄素细胞色素 c3）（Ifc3）gi|13878499|sp|Q9Z4P0.1|FRD2_SHEFN(13878499)；延胡索酸还原酶黄素蛋白亚基 gi|54041009|sp|P64174.1|FRDA_MYCTU(54041009)；延胡索酸还原酶黄素蛋白亚基 gi|54037132|sp|P64175.1|FRDA_MYCBO(54037132)；延胡索酸还原酶黄素蛋白亚基 gi|12230114|sp|Q9ZMP0.1|FRDA_HELPJ(12230114)；延胡索酸还原酶黄素蛋白亚基；延胡索酸还原酶黄素蛋白亚基 gi|1169737|sp|P44894.1|FRDA_HAEIN(1169737)；延胡索酸还原酶黄素蛋白亚基（*产琥珀酸沃廉菌*）gi|13160058|emb|CAA04214.2|(13160058)；延胡索酸还原酶黄素蛋白亚基前体（黄素细胞色素 c）（FL cyt）gi|25452947|sp|P83223.2|FRDA_SHEON(25452947)；延胡索酸还原酶铁-硫亚基（*产琥珀酸沃廉菌*）gi|2282000|emb|CAA04215.1|(2282000)；以及延胡索酸还原酶细胞色素 b 亚基（*产琥珀酸沃廉菌*）gi|2281998|emb|CAA04213.1|(2281998)，每个带有登记号的序列全部以引用的方式并入本文。

[00127] 乙酸激酶由大肠杆菌中的 ackA 编码。AckA 参与到乙酰-coA 转化为乙酸的反应。具体讲，即 ackA 催化乙酰磷酸转化为乙酸。AckA 的同源体和突变体已知。NCBI 数据库列出了约 1450 种用作细菌乙酸激酶的多肽。例如，此类同源体和突变体包括乙酸激酶（*天蓝色链霉菌* A3(2)）gi|21223784|ref|NP_629563.1|(21223784)；乙酸激酶（*天蓝色链霉菌* A3(2)）gi|6808417|emb|CAB70654.1|(6808417)；乙酸激酶（*化脓链球菌* M1 GAS）gi|15674332|ref|NP_268506.1|(15674332))；乙酸激酶（*空肠弯曲杆菌*空肠弯曲亚种 NCTC 11168）gi|15792038|ref|NP_281861.1|(15792038)；乙酸激酶（*化脓链球菌* M1 GAS）gi|13621416|gb|AAK33227.1|(13621416)；乙酸激酶（*波罗的海红小梨形菌* SH 1）gi|32476009|ref|NP_869003.1|(32476009)；乙酸激酶（*波罗的海红小梨形菌* SH 1）gi|32472045|ref|NP_865039.1|(32472045)；乙酸激酶（*空肠弯曲杆菌*空肠弯曲亚种 NCTC 11168）

gi|112360034|emb|CAL34826.1|(112360034); 乙酸激酶（*波罗的海红小梨形菌* SH 1） gi|32446553|emb|CAD76388.1|(32446553)); 乙酸激酶（*波罗的海红小梨形菌* SH 1） gi|32397417|emb|CAD72723.1|(32397417); AckA（*克氏梭菌* DSM 555） gi|153954016|ref|YP_001394781.1|(153954016)); 乙酸激酶（*双歧杆菌* NCC2705） gi|23465540|ref|NP_696143.1|(23465540); AckA （*克氏梭菌* DSM 555） gi|146346897|gb|EDK33433.1|(146346897); 乙酸激酶（*白喉杆菌*） gi|38200875|emb|CAE50580.1|(38200875); 乙酸激酶（*双歧杆菌* NCC2705） gi|23326203|gb|AAN24779.1|(23326203); 乙酸激酶（乙酰激酶） gi|67462089|sp|P0A6A3.1|ACKA_ECOLI(67462089); 和 AckA（*地衣芽孢杆菌* DSM 13）gi|52349315|gb|AAU41949.1|(52349315)，每个带有登记号的序列全部以引用的方式并入本文。

[00128] 磷酸乙酰基转移酶由大肠杆菌中的 *pta* 编码。PTA 参与到乙酸转化为乙酰 CoA 的反应。具体讲，即 PTA 催化乙酰 coA 转化为乙酰磷酸。PTA 的同源体和突变体已知。通过 NCBI 可以查到约 1075 种细菌磷酸乙酰基转移酶。例如，此类同源体和突变体包括磷酸乙酰基转移酶 Pta（猫立克次氏体 URRWXCal2）gi|67004021|gb|AAY60947.1|(67004021); 磷酸乙酰基转移酶（*aphidicola 布赫纳氏菌* Cc （雪松长足大蚜））gi|116256910|gb|ABJ90592.1|(116256910); pta（*aphidicola 布赫纳氏菌* Cc（雪松长足大蚜））gi|116515056|ref|YP_802685.1|(116515056); pta（*brevipalpis* 舌蝇的魏格沃菌共生体） gi|25166135|dbj|BAC24326.1|(25166135); Pta （多杀性巴氏杆菌多杀性亚种 Pm70）gi|12720993|gb|AAK02789.1|(12720993); Pta（*深红红螺菌*）gi|25989720|gb|AAN75024.1|(25989720); pta（6b 型*威氏李斯特氏菌变种* SLCC5334）gi|116742418|emb|CAK21542.1|(116742418); Pta（*鸟分支杆菌*副结核亚种 K-10）gi|41398816|gb|AAS06435.1|(41398816); 磷酸乙酰基转移酶（pta）（*博氏疏螺旋体* B31）gi|15594934|ref|NP_212723.1|(15594934); 磷酸乙酰基转移酶（pta）（*博氏疏螺旋体* B31）gi|2688508|gb|AAB91518.1|(2688508); 磷酸乙酰基转移酶（pta）（*流感嗜血杆菌* Rd KW20） gi|1574131|gb|AAC22857.1|(1574131); 磷酸乙酰基转移酶 Pta（*bellii 立克次氏体* RML369-C）gi|91206026|ref|YP_538381.1|(91206026); 磷酸乙酰基转移酶 Pta（*bellii 立克次氏体* RML369-C）

gi|91206025|ref|YP_538380.1|(91206025); 磷酸乙酰基转移酶 pta（*结核分支杆菌* F11）gi|148720131|gb|ABR04756.1|(148720131); 磷酸乙酰基转移酶 pta（*结核分支杆菌* Haarlem 型）gi|134148886|gb|EBA40931.1|(134148886); 磷酸乙酰基转移酶 pta（*结核分支杆菌* C）gi|124599819|gb|EAY58829.1|(124599819); 磷酸乙酰基转移酶 Pta（*bellii 立克次氏体* RML369-C）gi|91069570|gb|ABE05292.1|(91069570); 磷酸乙酰基转移酶 Pta（*bellii 立克次氏体* RML369-C）gi|91069569|gb|ABE05291.1|(91069569); 磷酸乙酰基转移酶（pta）（梅毒螺旋体梅毒亚种，Nichols 型）gi|15639088|ref|NP_218534.1|(15639088); 和磷酸乙酰基转移酶（pta）（梅毒螺旋体梅毒亚种，Nichols 型）gi|3322356|gb|AAC65090.1|(3322356)，每个带有登记号的序列全部以引用的方式并入本文。

[00129] 丙酮酸-甲酸裂解酶（甲酸乙酰基转移酶）能够催化丙酮酸转化为乙酰 CoA 和甲酸。此酶通过在厌氧条件下由 plf-活化酶诱导，产生有机自由基，在磷酸限制条件下此酶显著减少。甲酸乙酰基转移酶由大肠杆菌中的 *pflB* 编码。PFLB 的同源体和突变体已知。例如，此同源体和突变体包含甲酸乙酰基转移酶 1（丙酸-甲酸-裂解酶）gi|129879|sp|P09373.2|PFLB_ECOLI(129879); 甲酸乙酰基转移酶 1（*鼠疫耶尔森氏杆菌* CO92）gi|16121663|ref|NP_404976.1|(16121663); 甲酸乙酰基转移酶 1（*假结核耶尔森氏菌* IP 32953）gi|51595748|ref|YP_069939.1|(51595748); 甲酸乙酰基转移酶 1（田鼠型鼠疫耶尔森氏杆菌生物变种 91001）gi|45441037|ref|NP_992576.1|(45441037); 甲酸乙酰基转移酶 1（*鼠疫耶尔森氏杆菌* CO92）gi|115347142|emb|CAL20035.1|(115347142); 甲酸乙酰基转移酶 1（田鼠型*鼠疫耶尔森氏杆菌*生物变种 91001）gi|45435896|gb|AAS61453.1|(45435896); 甲酸乙酰基转移酶 1（*假结核耶尔森氏菌* IP 32953）gi|51589030|emb|CAH20648.1|(51589030); 甲酸乙酰基转移酶 1（*肠道沙门氏菌肠道伤寒亚种* Typhi *变种* CT18）gi|16759843|ref|NP_455460.1|(16759843); 甲酸乙酰基转移酶 1（*肠道沙门氏菌*肠道亚种伤寒 A 变种 ATCC 9150）gi|56413977|ref|YP_151052.1|(56413977); 甲酸乙酰基转移酶 1（*肠道沙门氏菌*肠道亚种伤寒变种）gi|16502136|emb|CAD05373.1|(16502136); 甲酸乙酰基转移酶 1（*肠道沙门氏菌*肠道亚种伤寒 A 变种 ATCC 9150）gi|56128234|gb|AAV77740.1|(56128234); 甲

酸乙酰基转移酶 1（痢疾志贺氏菌 Sd197）
gi|82777577|ref|YP_403926.1|(82777577)；甲酸乙酰基转移酶 1（2a 型 福氏志贺氏菌 2457T）gi|30062438|ref|NP_836609.1|(30062438)；甲酸乙酰基转移酶 1（2a 型 福氏志贺氏菌 2457T）gi|30040684|gb|AAP16415.1|(30040684)；甲酸乙酰基转移酶 1（5 型 福氏志贺氏菌 8401）gi|110614459|gb|ABF03126.1|(110614459)；甲酸乙酰基转移酶 1（痢疾志贺氏菌 Sd197）
gi|81241725|gb|ABB62435.1|(81241725)；甲酸乙酰基转移酶 1（大肠杆菌 O157:H7 EDL933）gi|12514066|gb|AAG55388.1|AE005279_8(12514066)；甲酸乙酰基转移酶 1（鼠疫耶尔森氏杆菌 KIM）
gi|22126668|ref|NP_670091.1|(22126668)；甲酸乙酰基转移酶 1（*无乳链球菌* A909）gi|76787667|ref|YP_330335.1|(76787667)；甲酸乙酰基转移酶 1（鼠疫耶尔森氏杆菌 KIM）gi|21959683|gb|AAM86342.1|AE013882_3(21959683)；甲酸乙酰基转移酶 1（*无乳链球菌* A909）gi|76562724|gb|ABA45308.1|(76562724)；甲酸乙酰基转移酶 1（*小肠结肠炎耶尔森氏杆菌小肠结肠炎亚种* 8081）
gi|123441844|ref|YP_001005827.1|(123441844)；甲酸乙酰基转移酶 1（5 型 福氏志贺氏菌 8401) gi|110804911|ref|YP_688431.1|(110804911)；甲酸乙酰基转移酶 1（*大肠杆菌* UTI89）gi|91210004|ref|YP_539990.1|(91210004)；甲酸乙酰基转移酶 1（*鲍氏志贺氏菌* Sb227）gi|82544641|ref|YP_408588.1|(82544641))；甲酸乙酰基转移酶 1(*宋内氏志贺氏菌* Ss046)gi|74311459|ref|YP_309878.1|(74311459)；甲酸乙酰基转移酶 1（*肺炎克伯氏菌肺炎亚种* MGH 78578）
gi|152969488|ref|YP_001334597.1|(152969488)；甲酸乙酰基转移酶 1（*肠道沙门氏菌肠道亚种伤寒变种* Ty2）gi|29142384|ref|NP_805726.1|(29142384)；甲酸乙酰基转移酶 1（2a *型福氏志贺氏菌* 301）
gi|24112311|ref|NP_706821.1|(24112311)；甲酸乙酰基转移酶 1（*大肠杆菌* O157:H7 EDL933)gi|15800764|ref|NP_286778.1|(15800764)；甲酸乙酰基转移酶 1（*肺炎克伯氏菌肺炎亚种* MGH 78578）
gi|150954337|gb|ABR76367.1|(150954337)；甲酸乙酰基转移酶 1（鼠疫耶尔森氏杆菌 CA88-4125）gi|149366640|ref|ZP_01888674.1|(149366640)；甲酸乙酰基转移酶 1（鼠疫耶尔森氏杆菌 CA88-4125）
gi|149291014|gb|EDM41089.1|(149291014)；甲酸乙酰基转移酶 1（*小肠结肠炎耶尔森氏杆菌小肠结肠炎亚种* 8081）gi|122088805|emb|CAL11611.1|(122088805)；

甲酸乙酰基转移酶 1（*宋内氏志贺氏菌* Ss046）
gi|73854936|gb|AAZ87643.1|(73854936); 甲酸乙酰基转移酶 1(*大肠杆菌* UTI89）gi|91071578|gb|ABE06459.1|(91071578); 甲酸乙酰基转移酶 1（*肠道沙门氏菌肠道亚种伤寒变种* Ty2）gi|29138014|gb|AAO69575.1|(29138014); 甲酸乙酰基转移酶 1（*鲍氏志贺氏菌* Sb227）gi|81246052|gb|ABB66760.1|(81246052); 甲酸乙酰基转移酶 1（2a 型*福氏志贺氏菌* 301）gi|24051169|gb|AAN42528.1|(24051169); 甲酸乙酰基转移酶 1（*大肠杆菌* O157:H7 301）
gi|24051169|gb|AAN42528.1|(13360445); 甲酸乙酰基转移酶 1（*大肠杆菌* O157:H7 Sakai) gi|15830240|ref|NP_309013.1|(15830240); 甲酸乙酰基转移酶 I（丙酮酸甲酸裂解酶 1）（*发光杆菌* laumondii 亚种 TTO1）
gi|36784986|emb|CAE13906.1|(36784986); 甲酸乙酰基转移酶 I（丙酮酸甲酸裂解酶 1）（*发光杆菌* laumondii 亚种 TTO1）
gi|37525558|ref|NP_928902.1|(37525558); 甲酸乙酰基转移酶（*金黄色葡萄球菌*金黄色亚种 Mu50）gi|14245993|dbj|BAB56388.1|(14245993)甲酸乙酰基转移酶（*金黄色葡萄球菌*金黄色亚种 Mu50）gi|15923216|ref|NP_370750.1|(15923216); 甲酸乙酰基转移酶（丙酮酸甲酸裂解酶）
gi|81706366|sp|Q7A7X6.1|PFLB_STAAN(81706366); 甲酸乙酰基转移酶（丙酮酸甲酸裂解酶）gi|81782287|sp|Q99WZ7.1|PFLB_STAAM(81782287); 甲酸乙酰基转移酶（丙酮酸甲酸裂解酶）
gi|81704726|sp|Q7A1W9.1|PFLB_STAAW(81704726); 甲酸乙酰基转移酶（*金黄色葡萄球菌*金黄色亚种 Mu3）gi|156720691|dbj|BAF77108.1|(156720691); 甲酸乙酰基转移酶（胡萝卜软腐欧式杆菌胡萝卜软腐亚种 SCRI1043）
gi|50121521|ref|YP_050688.1|(50121521); 甲酸乙酰基转移酶（胡萝卜软腐欧式杆菌胡萝卜软腐亚种 SCRI1043）gi|49612047|emb|CAG75496.1|(49612047); 甲酸乙酰基转移酶（*金黄色葡萄球菌*金黄色亚种 Newman）
gi|150373174|dbj|BAF66434.1|(150373174); 甲酸乙酰基转移酶（*希瓦氏菌* MR-1)gi|24374439|ref|NP_718482.1|(24374439); 甲酸乙酰基转移酶（*希瓦氏菌* MR-1)gi|24349015|gb|AAN55926.1|AE015730_3(24349015); 甲酸乙酰基转移酶（*胸膜肺炎放线杆菌 3 型变种* JL03）
gi|165976461|ref|YP_001652054.1|(165976461); 甲酸乙酰基转移酶（*胸膜肺炎放线杆菌 3 型变种* JL03) gi|165876562|gb|ABY69610.1|(165876562); 甲酸乙酰基

转移酶（*金黄色葡萄球菌*金黄色亚种 MW2）
gi|21203365|dbj|BAB94066.1|(21203365)；甲酸乙酰基转移酶（*金黄色葡萄球菌*金黄色亚种 N315）gi|13700141|dbj|BAB41440.1|(13700141)；甲酸乙酰基转移酶（*金黄色葡萄球菌*金黄色亚种 Newman)
gi|151220374|ref|YP_001331197.1|(151220374)；甲酸乙酰基转移酶（*金黄色葡萄球菌*金黄色亚种 Mu3）gi|156978556|ref|YP_001440815.1|(156978556)；甲酸乙酰基转移酶（聚球藻种 JA-2-3B'a(2-13)）gi|86607744|ref|YP_476506.1|(86607744)；甲酸乙酰基转移酶（聚球藻种 JA-3-3Ab)
gi|86605195|ref|YP_473958.1|(86605195)；甲酸乙酰基转移酶（肺炎链球菌 D39）gi|116517188|ref|YP_815928.1|(116517188)；甲酸乙酰基转移酶（聚球藻种 JA-2-3B'a(2-13)）gi|86556286|gb|ABD01243.1|(86556286)；甲酸乙酰基转移酶（聚球藻种 JA-3-3Ab) gi|86553737|gb|ABC98695.1|(86553737)；甲酸乙酰基转移酶（诺氏梭菌 NT）gi|118134908|gb|ABK61952.1|(118134908)；甲酸乙酰基转移酶（*金黄色葡萄球菌*金黄色亚种 MRSA252）
gi|49482458|ref|YP_039682.1|(49482458)；和甲酸乙酰基转移酶（*金黄色葡萄球菌*金黄色亚种 MRSA252）gi|49240587|emb|CAG39244.1|(49240587)，每个带有登记号的序列全部以引用的方式并入本文。

[00130] α-异丙基苹果酸合成酶（EC 2.3.3.13，有时称为 2-异丙基苹果酸合成酶，α-IPM 合成酶）催化乙酰-CoA 的乙酰基与 3-甲基-2-氧代丁酸（2-氧代异戊酸 ）缩合成 3-羧基-3-羟基-4-甲基戊酸（2-异丙基苹果酸）。α-异丙基苹果酸合成酶由大肠杆菌中的 *leuA* 编码。LeuA 的同源体和突变体已知。例如，此类同源体和突变体，例如 2-异丙基苹果酸合成酶（谷氨酸棒状杆菌）gi|452382|emb|CAA50295.1|(452382)；2-异丙基苹果酸合成酶（大肠杆菌 K12）gi|16128068|ref|NP_414616.1|(16128068)；2-异丙基苹果酸合成酶（大肠杆菌 K12）gi|1786261|gb|AAC73185.1|(1786261)；2-异丙基苹果酸合成酶（拟南芥）gi|15237194|ref|NP_197692.1|(15237194)；2-异丙基苹果酸合成酶（拟南芥）gi|42562149|ref|NP_173285.2|(42562149)；2-异丙基苹果酸合成酶（拟南芥）gi|15221125|ref|NP_177544.1|(15221125)；2-异丙基苹果酸合成酶（*天蓝色链霉菌* A3(2)）gi|32141173|ref|NP_733575.1|(32141173)；2-异丙基苹果酸合成酶（*波罗的海红小梨形菌* SH 1)gi|32477692|ref|NP_870686.1|(32477692)；2-异丙基苹果酸合成酶（*波罗的海红小梨形菌* SH 1)gi|32448246|emb|CAD77763.1|(32448246)；

2-异丙基苹果酸合成酶（*Akkermansia muciniphila* ATCC BAA-835）gi|166241432|gb|EDR53404.1|(166241432)；2-异丙基苹果酸合成酶（*橙色滑柱菌* ATCC 23779）gi|159900959|ref|YP_001547206.1|(159900959)；2-异丙基苹果酸合成酶（*Dinoroseobacter shibae* DFL 12）gi|159043149|ref|YP_001531943.1|(159043149)；2-异丙基苹果酸合成酶（*砂嗜盐产孢菌* CNS-205）gi|159035933|ref|YP_001535186.1|(159035933)；2-异丙基苹果酸合成酶（番茄细菌性溃疡病菌 NCPPB 382)
gi|148272757|ref|YP_001222318.1|(148272757)；2-异丙基苹果酸合成酶（大肠杆菌 B）gi|124530643|ref|ZP_01701227.1|(124530643)；2-异丙基苹果酸合成酶（大肠杆菌 C ATCC 8739）gi|124499067|gb|EAY46563.1|(124499067)；2-异丙基苹果酸合成酶（*百日咳博德特氏菌* Tohama I）
gi|33591386|ref|NP_879030.1|(33591386)；2-异丙基苹果酸合成酶
（*Polynucleobacter necessarius* STIR1）
gi|164564063|ref|ZP_02209880.1|(164564063)；2-异丙基苹果酸合成酶
（*Polynucleobacter necessarius* STIR1）
gi|164506789|gb|EDQ94990.1|(164506789)；2-异丙基苹果酸合成酶（*韦氏芽孢杆菌* KBAB4）gi|163939313|ref|YP_001644197.1|(163939313)，任何带有登记号的序列全部以引用的方式并入本文。

[00131] BCCA 转氨酶催化支链氨基酸（BCAA）的形成。一些已知的这类转氨酶由大肠杆菌中的 *ilvE* 编码。典型同源体和突变体包含下列登记号所指的序列：ilvE（铜绿微囊藻 PCC 7806）gi|159026756|emb|CAO86637.1|(159026756)；IlvE（*大肠杆菌*）gi|87117962|gb|ABD20288.1|(87117962)；IlvE（*大肠杆菌*）gi|87117960|gb|ABD20287.1|(87117960)；IlvE（*大肠杆菌*）gi|87117958|gb|ABD20286.1|(87117958)；IlvE（*福氏志贺氏菌*）gi|87117956|gb|ABD20285.1|(87117956)；IlvE（*福氏志贺氏菌*）gi|87117954|gb|ABD20284.1|(87117954)；IlvE（*福氏志贺氏菌*）gi|87117952|gb|ABD20283.1|(87117952)；IlvE（*福氏志贺氏菌*）gi|87117950|gb|ABD20282.1|(87117950)；IlvE（*福氏志贺氏菌*）gi|87117948|gb|ABD20281.1|(87117948)；IlvE（*福氏志贺氏菌*）gi|87117946|gb|ABD20280.1|(87117946)；IlvE（*福氏志贺氏菌*）

gi|87117944|gb|ABD20279.1|(87117944)；IlvE（*福氏志贺氏菌*）gi|87117942|gb|ABD20278.1|(87117942)；IlvE（*福氏志贺氏菌*）gi|87117940|gb|ABD20277.1|(87117940)；IlvE（*福氏志贺氏菌*）gi|87117938|gb|ABD20276.1|(87117938)；IlvE（*痢疾志贺氏菌*）gi|87117936|gb|ABD20275.1|(87117936)；IlvE（*痢疾志贺氏菌*）gi|87117934|gb|ABD20274.1|(87117934)；IlvE（*痢疾志贺氏菌*）gi|87117932|gb|ABD20273.1|(87117932)；IlvE（*痢疾志贺氏菌*）gi|87117930|gb|ABD20272.1|(87117930)；和 IlvE（*痢疾志贺氏菌*）gi|87117928|gb|ABD20271.1|(87117928)，每个带有登记号的序列全部以引用的方式并入本文。

[00132] 酪氨酸转氨酶催化二羧酸和芳香族氨基酸底物的转氨基作用。大肠杆菌中的酪氨酸转氨酶由 tyrB 基因编码。TyrB 的同源体和突变体已知。例如，此类同源体和突变体包括 tyrB（*博德特氏菌*）gi|163857093|ref|YP_001631391.1|（163857093）；tyrB（*博德特氏菌*）gi|163260821|emb|CAP43123.1|(163260821)；转氨酶 gi|551844|gb|AAA24704.1|(551844)；转氨酶（慢生根瘤菌 BTAi1）gi|146404387|gb|ABQ32893.1|（146404387）；酪氨酸转氨酶 TyrB（肠道沙门氏菌）gi|4775574|emb|CAB40973.2|(4775574)；酪氨酸转氨酶（鼠伤寒沙门氏菌 LT2）gi|16422806|gb|AAL23072.1|(16422806)；和酪氨酸转氨酶 gi|148085|gb|AAA24703.1|(148085)，其所有序列均以引用的方式并入本文。

[00133] 丙酮酸氧化酶催化丙酮酸转化为乙酸和 CO_2 的反应。在大肠杆菌中，丙酮酸氧化酶由 *poxB* 编码。PoxB 及其同源体和突变体包括如丙酮酸氧化酶；PoxB（大肠杆菌）gi|685128|gb|AAB31180.1||bbm|348451|bbs|154716(685128)；PoxB（*荧光假单胞菌*）gi|32815820|gb|AAP88293.1|(32815820)；PoxB（大肠杆菌）gi|25269169|emb|CAD57486.1|(25269169)；丙酮酸脱氢酶（*肠道沙门氏菌*肠道亚种伤寒变异型）gi|16502101|emb|CAD05337.1|(16502101)；丙酮酸氧化酶（植物乳杆菌）gi|41691702|gb|AAS10156.1|(41691702)；丙酮酸脱氢酶（慢生型大豆根瘤菌）gi|20257167|gb|AAM12352.1|(20257167)；丙酮酸脱氢酶（鼠疫耶尔森氏杆菌 KIM）gi|22126698|ref|NP_670121.1|(22126698)；丙酮酸脱氢酶（细胞色素）（古典型鼠疫耶尔森氏杆菌生物变异体 B42003004）gi|166211240|ref|ZP_02237275.1|(166211240)；丙酮酸脱氢酶（细胞色素）（古典

型鼠疫耶尔森氏杆菌生物变异型 B42003004)
gi|166207011|gb|EDR51491.1|(166207011); 丙酮酸脱氢酶（番茄细菌性叶斑病菌 DC3000）gi|28869703|ref|NP_792322.1|(28869703); 丙酮酸脱氢酶（*鼠伤寒沙门氏菌* LT2)gi|16764297|ref|NP_459912.1|(16764297); 丙酮酸脱氢酶（肠道沙门氏菌肠道亚种伤寒变异型 CT18）gi|16759808|ref|NP_455425.1|(16759808); 丙酮酸脱氢酶（细胞色素）（*贝纳特氏立克次体* Dugway 5J108-111）
gi|154706110|ref|YP_001424132.1|(154706110); 丙酮酸脱氢酶（*番茄细菌性溃疡病菌* NCPPB 382）gi|148273312|ref|YP_001222873.1|(148273312); 丙酮酸脱氢酶（*嗜酸乳杆菌* NCFM）gi|58338213|ref|YP_194798.1|(58338213); 和丙酮酸脱氢酶（鼠疫耶尔森氏杆菌 CO92）gi|16121638|ref|NP_404951.1|(16121638)，带有登记号的序列以引用方式并入本文。

[00134] L-苏氨酸-3-脱氢酶（EC 1.1.1.103）催化 L-苏氨酸转化成 L-2-氨基-3-氧代丁酸的反应。tdh 基因编码 L-苏氨酸-3-脱氢酶。NCBI 中认可的来自细菌的 L-苏氨酸-3-脱氢酶约有 700 种。tdh 的各种同源体和突变体包括如 L-苏氨酸-3-脱氢酶 gi|135560|sp|P07913.1|TDH_ECOLI(135560); L-苏氨酸-3-脱氢酶 gi|166227854|sp|A4TSC6.1|TDH_YERPP(166227854)； L-苏氨酸-3-脱氢酶 gi|166227853|sp|A1JHX8.1|TDH_YERE8(166227853)； L-苏氨酸-3-脱氢酶 gi|166227852|sp|A6UBM6.1|TDH_SINMW(166227852)； L-苏氨酸-3-脱氢酶 gi|166227851|sp|A1RE07.1|TDH_SHESW(166227851)； L-苏氨酸-3-脱氢酶 gi|166227850|sp|A0L2Q3.1|TDH_SHESA(166227850)； L-苏氨酸-3-脱氢酶 gi|166227849|sp|A4YCC5.1|TDH_SHEPC(166227849)； L-苏氨酸-3-脱氢酶 gi|166227848|sp|A3QJC8.1|TDH_SHELP(166227848)； L-苏氨酸-3-脱氢酶 gi|166227847|sp|A6WUG6.1|TDH_SHEB8(166227847)； L-苏氨酸-3-脱氢酶 gi|166227846|sp|A3CYN0.1|TDH_SHEB5(166227846)； L-苏氨酸-3-脱氢酶 gi|166227845|sp|A1S1Q3.1|TDH_SHEAM(166227845)； L-苏氨酸-3-脱氢酶 gi|166227844|sp|A4FND4.1|TDH_SACEN(166227844)； L-苏氨酸-3-脱氢酶 gi|166227843|sp|A1SVW5.1|TDH_PSYIN(166227843)； L-苏氨酸-3-脱氢酶 gi|166227842|sp|A5IGK7.1|TDH_LEGPC(166227842)； L-苏氨酸-3-脱氢酶 gi|166227841|sp|A6TFL2.1|TDH_KLEP7(166227841)； L-苏氨酸-3-脱氢酶 gi|166227840|sp|A4IZ92.1|TDH_FRATW(166227840)； L-苏氨酸-3-脱氢酶 gi|166227839|sp|A0Q5K3.1|TDH_FRATN(166227839)； L-苏氨酸-3-脱氢酶

gi|166227838|sp|A7NDM9.1|TDH_FRATF(166227838)； L-苏氨酸-3-脱氢酶 gi|166227837|sp|A7MID0.1|TDH_ENTS8(166227837)；和 L-苏氨酸-3-脱氢酶 gi|166227836|sp|A1AHF3.1|TDH_ECOK1(166227836)，带有登记号的序列以引用方式并入本文。

[00135] 乙酰羟酸合成酶（例如 ilvH）和乙酰乳酸合成酶（例如 alsS、ilvB、ilvI）催化支链氨基酸（缬氨酸、亮氨酸和异亮氨酸）的合成反应。*大肠杆菌*中的 IlvH 编码乙酰羟酸合成酶（见乙酰羟酸合成酶 AHAS III（IlvH）（*大肠杆菌*）gi|40846|emb|CAA38855.1|(40846)，以引用的方式并入本文）。ilvH 的同源体、突变体以及构成 ilvH 的操纵子已知，其包含如 ilvH（铜绿微囊藻 PCC 7806）gi|159026908|emb|CAO89159.1|(159026908)；IlvH（*淀粉液化芽孢杆菌* FZB42）gi|154686966|ref|YP_001422127.1|(154686966)；IlvH（淀粉液化芽孢杆菌 FZB42）gi|154352817|gb|ABS74896.1|(154352817)；IlvH（嗜线虫致病杆菌）gi|131054140|gb|ABO32787.1|(131054140)；IlvH（*鼠伤寒沙门氏菌*）gi|7631124|gb|AAF65177.1|AF117227_2(7631124)；ilvN（无害李斯特氏菌）gi|16414606|emb|CAC97322.1|(16414606)；ilvN（*单核细胞增生李斯特氏菌*）gi|16411438|emb|CAD00063.1|(16411438)；乙酰羟酸合成酶（*新月柄杆菌*）gi|408939|gb|AAA23048.1|(408939)；乙酰羟酸合成酶 I，小亚基（*肠道沙门氏菌肠道亚种伤寒变异型*）gi|16504830|emb|CAD03199.1|(16504830)；乙酰羟酸合成酶，小亚基（*Tropheryma whipplei* TW08/27）gi|28572714|ref|NP_789494.1|(28572714)；乙酰羟酸合成酶，小亚基（*Tropheryma whipplei* TW08/27）gi|28410846|emb|CAD67232.1|(28410846)；乙酰羟酸合成酶 I，小亚基（*肠道沙门氏菌肠道亚种伤寒 A 变异型*） ATCC 9150）gi|56129933|gb|AAV79439.1|(56129933)；乙酰羟酸合成酶小亚基；乙酰羟酸合成酶，小亚基 gi|551779|gb|AAA62430.1|(551779)；乙酰羟酸合成酶 I，小亚基（*肠道沙门氏菌肠道亚种伤寒变异型* Ty2）gi|29139650|gb|AAO71216.1|(29139650)；乙酰羟酸合成酶小亚基（肉桂地链霉菌）gi|5733116|gb|AAD49432.1|AF175526_1(5733116)；乙酰羟酸合成酶大亚基；和乙酰羟酸合成酶，大亚基 gi|400334|gb|AAA62429.1|(400334)，带有登记号的序列以引用方式并入本文。乙酰乳酸合成酶基因包括 alsS 和 ilvI。ilvI 和 alsS 的同源体已知，例如，乙酰乳酸合成酶小亚基（长双歧杆菌 NCC2705）gi|23325489|gb|AAN24137.1|(23325489)；乙酰乳酸合成酶小亚基（*嗜热脂肪芽孢*

杆菌）gi|19918933|gb|AAL99357.1|(19918933); 乙酰乳酸合成酶（固氮弧菌 BH72）gi|119671178|emb|CAL95091.1|(119671178); 乙酰乳酸合成酶小亚基（*白喉杆菌*）gi|38199954|emb|CAE49622.1|(38199954); 乙酰乳酸合成酶（固氮弧菌 BH72）gi|119669739|emb|CAL93652.1|(119669739); 乙酰乳酸合成酶小亚基（*詹氏棒状杆菌* K411）gi|68263981|emb|CAI37469.1|(68263981); 乙酰乳酸合成酶小亚基（*枯草芽孢杆菌*）gi|1770067|emb|CAA99562.1|(1770067); 乙酰乳酸合成酶同工酶 1 小亚基（AHAS-I）（乙酰羟酸合成酶同工酶 1 小亚基）（ALS-I）gi|83309006|sp|P0ADF8.1|ILVN_ECOLI(83309006); 乙酰乳酸合成酶大亚基;（*嗜热脂肪芽孢杆菌*）gi|19918932|gb|AAL99356.1|(19918932); 和乙酰乳酸合成酶，小亚基（*腾冲嗜热杆菌* MB4）gi|20806556|ref|NP_621727.1|(20806556)，带有登记号的序列以引用方式并入本文。NCBI 中所列 ilvB 同源体和突变体约为 1120 种。

[00136] 乙酰羟酸还原异构酶是异亮氨酸和缬氨酸生物合成平行途径中的第二个酶。大肠杆菌中的 *IlvC* 编码乙酰羟酸还原异构酶。ilvC 的同源体和突变体已知，包括如乙酰羟酸还原异构酶（粟酒裂殖酵母 972h-）gi|162312317|ref|NP_001018845.2|(162312317); 乙酰羟酸还原异构酶（粟酒裂殖酵母）gi|3116142|emb|CAA18891.1|(3116142); 乙酰羟酸还原异构酶（*酿酒酵母* YJM789）gi|151940879|gb|EDN59261.1|(151940879); Ilv5p: 乙酰羟酸还原异构酶（*酿酒酵母*）gi|609403|gb|AAB67753.1|(609403); ACL198Wp （棉病囊菌 ATCC 10895）gi|45185490|ref|NP_983206.1|(45185490); ACL198Wp （棉病囊菌 ATCC 10895）gi|44981208|gb|AAS51030.1|(44981208); 乙酰羟酸还原异构酶; Ilv5x（酿酒酵母）gi|957238|gb|AAB33579.1||bbm|369068|bbs|165406(957238); 乙酰羟酸还原异构酶; Ilv5g（酿酒酵母）gi|957236|gb|AAB33578.1||bbm|369064|bbs|165405(957236); 和乙酰羟酸还原异构酶（粟酒裂殖酵母）gi|2696654|dbj|BAA24000.1|(2696654)，每个带有登记号的序列均以引用方式并入本文。

[00137] 二羟酸脱水酶催化异亮氨酸和缬氨酸生化合成途径中的第四步反应，即 2,3-二羟基-异戊酸脱水生成 α-酮异戊酸。*IlvD* 和 *ilv3* 编码二羟酸脱水酶。二羟酸脱水酶的同源体和突变体已知，包含如 IlvD（*麻风分枝杆菌*）gi|2104594|emb|CAB08798.1|(2104594); 二羟酸脱水酶（Tropheryma whipplei TW08/27）gi|28410848|emb|CAD67234.1|(28410848); 二羟酸脱水酶（*麻风分枝*

杆菌）gi|13093837|emb|CAC32140.1|(13093837)；二羟酸脱水酶（*波罗的海红小梨形菌* SH 1）gi|32447871|emb|CAD77389.1|(32447871)；和推定的二羟酸脱水酶（金黄色葡萄球菌金黄色亚种 MRSA252）gi|49242408|emb|CAG41121.1|(49242408)，每个带有登记号的序列均以引用方式并入本文。

[00138] 2-酮酸脱羧酶催化 2-酮酸转化为相应醛的反应。例如，2-酮异戊酸脱羧酶催化 2-酮异戊酸转化为异丁醛的反应。已知一些 2-酮酸脱羧酶,例如由 pdc、pdc1、pdc5、pdc6、aro10、thI3、kdcA 和 kivd 基因编码的 2-酮酸脱羧酶。对 2-酮酸转化为相应醛的反应起催化作用的典型同源体和突变体包含下列登记号所指的序列及已鉴定的酶活性：gi|44921617|gb|AAS49166.1| 支链 α-酮酸脱羧酶（乳酸乳球菌）；gi|15004729|ref|NP_149189.1| 丙酮酸脱羧酶（*丙酮丁醇梭菌* ATCC 824）；gi|82749898|ref|YP_415639.1| 可能的丙酮酸脱羧酶（金黄色葡萄糖球菌 RF122）；gi|77961217|ref|ZP_00825060.1| COG3961：丙酮酸脱羧酶及相关的硫胺素焦磷酸所需的酶(*莫氏耶尔森氏菌* ATCC 43969)；gi|71065418|ref|YP_264145.1| 推定的丙酮酸脱羧酶（北极嗜冷杆菌 273-4）；gi|16761331|ref|NP_456948.1| 推定的脱羧酶（肠道沙门氏菌肠道亚种伤寒变异型） CT18）；gi|93005792|ref|YP_580229.1| 丙酮酸脱羧酶（cryohalolentis 嗜冷杆菌 K5）gi|23129016|ref|ZP_00110850.1| COG3961：及相关的硫胺素焦磷酸所需的酶(点形念珠藻 PCC 73102)；gi|16417060|gb|AAL18557.1|AF354297_1 丙酮酸脱羧酶（*胃八叠球菌*）；gi|15607993|ref|NP_215368.1|可能的丙酮酸或吲哚-3-丙酮酸脱羧酶 PDC（*结核分枝杆菌* H37Rv）；gi|41406881|ref|NP_959717.1| Pdc（*鸟分枝杆菌副结核亚种* K-10）；gi|91779968|ref|YP_555176.1| 推定的丙酮酸脱羧酶（*xenovorans* 伯克氏菌 LB400）；gi|15828161|ref|NP_302424.1| 丙酮酸(或吲哚丙酮酸)脱羧酶(*麻风分枝杆菌* TN)；gi|118616174|ref|YP_904506.1| 丙酮酸或吲哚-3-丙酮酸脱羧酶 Pdc（溃疡分枝杆菌 Agy99）；gi|67989660|ref|NP_001018185.1| 推定蛋白 SPAC3H8.01（*粟酒裂殖酵母* 972h-）；gi|21666011|gb|AAM73540.1|AF282847_1 丙酮酸脱羧酶 PdcB（*米根酶*）；gi|69291130|ref|ZP_00619161.1| 丙酮酸脱羧酶：丙酮酸脱羧酶（耐辐射 *Kineococcus* SRS30216）；gi|66363022|ref|XP_628477.1| 丙酮酸脱羧酶（爱荷华*微小隐孢子虫* II）；gi|70981398|ref|XP_731481.1| 丙酮酸脱羧酶(*烟曲霉* Af293)；gi|121704274|ref|XP_001270401.1| 丙酮酸脱羧酶，推定（*棒曲霉* NRRL

1); gi|119467089|ref|XP_001257351.1| 丙酮酸脱羧酶,推定(*赞氏新萨托菌* 181); gi|26554143|ref|NP_758077.1| 丙酮酸脱羧酶(*穿通支原体* HF-2); gi|21666009|gb|AAM73539.1|AF282846_1 丙酮酸脱羧酶 PdcA (*米根酶*)。

[00139] 醇脱氢酶(adh)催化氨基酸分解代谢的最后一步反应,即醛转化为长链或复杂醇。在此项技术中,各种 *adh* 基因已知。如本文所指出的,adhq1 同源体和突变体包括如 adh2、adh3、adh4、adh5、adh 6 和 sfa1 (见 SFA (*酿酒酵母*) gi|288591|emb|CAA48161.1| (288591); 带有登记号的序列以引用方式并入本文中)。

[00140] 柠苹酸合成酶催化丙酮酸和乙酸的缩合反应。*CimA* 编码 柠苹酸合成酶。同源体和突变体已知,包括如 柠苹酸合成酶(*双曲钩端螺旋体* Patoc 变异型) gi|116664687|gb|ABK13757.1|(116664687); 柠苹酸合成酶(*双曲钩端螺旋体* Monteralerio 变异型) gi|116664685|gb|ABK13756.1|(116664685); 柠苹酸合成酶(*问号钩端螺旋体* Hebdomadis 变异型) gi|116664683|gb|ABK13755.1|(116664683); 柠苹酸合成酶(*问号钩端螺旋体* Pomona 变异型) gi|116664681|gb|ABK13754.1|(116664681); 柠苹酸合成酶(*问号钩端螺旋体* Australis 变异型) gi|116664679|gb|ABK13753.1|(116664679); 柠苹酸合成酶(*问号钩端螺旋体* Autumnalis 变异型) gi|116664677|gb|ABK13752.1|(116664677); 柠苹酸合成酶(*问号钩端螺旋体* Pyrogenes 变异型) gi|116664675|gb|ABK13751.1|(116664675); 柠苹酸合成酶(*问号钩端螺旋体* Canicola 变异型) gi|116664673|gb|ABK13750.1|(116664673); 柠苹酸合成酶(*问号钩端螺旋体* Lai 变异型) gi|116664671|gb|ABK13749.1|(116664671); CimA (*梅氏钩端螺旋体* Semaranga 变异型)gi|119720987|gb|ABL98031.1|(119720987); (R)-柠苹酸合成酶 gi|2492795|sp|Q58787.1|CIMA_METJA(2492795); (R)-柠苹酸合成酶 gi|22095547|sp|P58966.1|CIMA_METMA(22095547); (R)-柠苹酸合成酶 gi|22001554|sp|Q8TJJ1.1|CIMA_METAC(22001554); (R)-柠苹酸合成酶 gi|22001553|sp|O26819.1|CIMA_METTH(22001553); (R)-柠苹酸合成酶 gi|22001555|sp|Q8TYB1.1|CIMA_METKA(22001555); (R)-柠苹酸合成酶 (海沼甲烷球菌 S2) gi|45358581|ref|NP_988138.1|(45358581); (R)-柠苹酸合成酶(海沼甲烷球菌 S2) gi|44921339|emb|CAF30574.1|(44921339); 以及与(R)-柠苹酸合成酶相似的酶(*Candidatus Kuenenia* stuttgartiensis)

gi|91203541|emb|CAJ71194.1|(91203541)，每个带有上述登记号的序列以引用的方式并入本文。

[00141]在一实施方案中，本发明提供的微生物可表征为含有rrnBT14DlacZWJ16 hsdR514 DaraBADAH33 DrhaBADLD78（F' 转导自 XL-1 BLUE，以提供 lacIq）、ΔadhE、ΔldhA、ΔfrdBC、Δfnr、Δpta 和 ΔpflB，及 pSA55 和 pSA69 的大肠杆菌，其中 pSA55 是一种从 ColE1 原点构建的质粒，其含有受 PLlacO1 调控的 *kivd*（*乳酸乳球菌*）和 *adh2*（*酿酒酵母*）基因和一氨苄西林抗性基因，*pSA69* 是一种源自 p15A 起点的质粒，其含有受 PLlacO1 调控的 *alsS*（*枯草芽孢杆菌*）、*ilvC*（*大肠杆菌*）和 *ilvD*（*大肠杆菌*）基因和一卡那霉素抗性基因。

[00142]在另一实施方案中，本发明提供的微生物可表征为含有rrnBT14DlacZWJ16 hsdR514 DaraBADAH33 DrhaBADLD78（F' 转导自 XL-1 BLUE，以提供 lacIq）、ΔmetA、Δtdh、ΔilvB、ΔilvI 和 ΔadhE，及 pCS49、pSA62 和 pSA55I 的大肠杆菌，其中 pSA55I 为从 ColE1 原点构建的质粒，其含有受 PLlacO1 调控的 *kivd*（乳酸乳球菌）和 *adh2*（酿*酒酵母*）基因和一种氨苄西林抗性基因，lacI 位于氨苄西林抗性基因之后，pSA62 是一种源自 p15A 起点的质粒，其含有受 PLlacO1 调控的 *ilvA*（*大肠杆菌*）和 *leuABCD*（*大肠杆菌*）基因和一卡那霉素抗性基因，pCS49 是一种源自 pSC101*起点的质粒，其含有受 PLlacO1 调控的 *thrA(fbr)BC*（*大肠杆菌*）基因和一种壮观霉素抗性基因。

[00143]本发明还提供了被保存的微生物。被保存的微生物只是代表性的，根据本发明，本发明所属领域的技术人员能够修饰其他不同种或不同基因型的亲本微生物，以获得本发明提供的能够产生异丁醇和 正丁醇的微生物。

[00144]本发明提供了一种被命名为 SA237 的重组微生物，其 ATCC 登记号为 no._____，于 2008 年 2 月 7 日由 ATCC 保存。本发明包含微生物的培养物，此培养物含有 ATCC 登记号为 no._____的一种微生物群体，其包括混合培养物。本发明还提供源自 ATCC 登记号为 no._____ 的多聚核苷酸片段，其用于制备一种生产异丁醇的微生物，产率为每克葡萄糖产出 0.12 至 0.41 克异丁醇。例如，此多聚核苷酸片段的碱基对数量可以从 1000 个左右到数百万个。本发明还包含异丁醇或苯基乙醇生产过程中的生物反应器，此反应器含有一定数量的微生物，这些微生物的 ATCC 登记号为 _____。使用被保

存微生物时，本领域技术人员能够很容易确定被保存微生物或其任何基因片段以及本文所述的多聚核苷酸片段的序列，包括定位敲除或破坏的基因。

[00145] 本发明还提供了一种被命名为 CRS-BuOH23 的重组微生物，其 ATCC 登记号为 no._____,于 2008 年 2 月 7 日由 ATCC 保存。本发明包含微生物的培养物，此培养物含有一定数量的 ATCC 登记号为 no. 的微生物，包括混合培养物。本发明还提供源自 ATCC 登记号为 no._____的多聚核苷酸片段，它将用于制备一种生产正丁醇的微生物。例如，此多聚核苷酸片段的碱基对数量可以从 1000 个左右到数百万个。本发明还包含 正丁醇生产过程中的生物反应器，此反应器含有一定数量的微生物，这些微生物的 ATCC 登记号为 _____。使用被保存微生物时，本领域技术人员能够很容易确定被保存微生物或其任何基因片段以及本文所述的多聚核苷酸片段的序列，包括定位基因敲除或基因破坏。

[00146] 应当了解，可以对一系列微生物进行修饰，使它们含有适合于生产正丙醇、异丁醇、正丁醇、2-甲基-1-丁醇、3-甲基-1-丁醇或 2-苯基乙醇的重组代谢途径。应当了解，有多种微生物都能作为编码合适目标酶的遗传物质的“来源”，这些目标酶适用于本发明提供的重组微生物。术语“微生物”包括古细菌域、细菌域和真核生物域内的原核和真核微生物物种，后者包括酵母菌和丝状真菌、原生动物、藻类或高级原生生物。术语“微生物细胞”和“微生物(microbe)”可以与“微生物(microorganism)”互换使用。

[00147] 术语“原核生物”指本发明所属领域公认的并且指不含细胞核或其他细胞器的细胞。“原核细胞”通常被归为细菌域和古细菌域这两个域中的一个。古细菌域和细菌域的明显区别在于 16S 核糖体 RNA 的核苷酸序列间的根本不同。

[00148] 术语“古细菌域”指疵壁菌门中的一类微生物，它们通常存在于特殊环境中，并且在许多方面不同于其他原核生物，包括核糖体蛋白的数量以及细胞壁中胞壁酸的缺乏。通过 ssrRNA 分析，古细菌包含两种在系统进化上截然不同的种群： 泉古菌门和广古菌门。根据它们的生理学特征，古细菌可被分为三种类型： 产甲烷菌（能产生甲烷的原核生物）；极端嗜盐菌（能在含有极高浓度的盐（Nacl）环境中生存的原核生物）；和极端（超）嗜热菌（能在极高温度环境中生存的原核生物）。这些原核生物除了具有与细菌不同的统一的古细菌特性（例如细胞壁中无胞壁质、细胞膜脂之间由酯质相连，等）外，它们还展

现出独特的结构或生化特性，这使它们能够适应特殊的生境。泉古菌门主要包括超嗜热硫依赖原核生物，广古菌门则包含产甲烷菌和极端嗜盐菌。

[00149]“细菌”或“真细菌”指原核生物域。细菌至少包括以下 11 个不同种群：（1）革兰氏阳性（gram+）细菌，其中主要有两个分支：（1）高 G+C 组（放线菌、分枝杆菌、微球菌等）（2）低 G+C 组（杆菌、梭菌、乳杆菌、葡萄球菌、链球菌、支原体）；（2）变形菌门，例如，紫色光合菌+非光合革兰氏阴性细菌（包括大多数“常见”的革兰氏阴性细菌）；（3）蓝细菌，例如生氧光能菌；（4）螺旋体及相关菌种；（5）浮霉状菌；（6）类杆菌、黄杆菌；（7）衣原体；（8）绿色硫细菌；（9）绿色无硫细菌（也称为厌氧光能菌）；（10）耐辐射微球菌及相关菌种；（11）热袍菌和嗜热栖热腔菌属。

[00150]“革兰氏阴性细菌”包括球菌、非肠道杆菌和肠道杆菌。革兰氏阴性细菌属包括如奈瑟菌、螺旋菌、巴氏杆菌、布鲁氏菌、耶尔森氏菌、弗朗西斯氏菌、嗜血杆菌、博德特氏菌、大肠杆菌、沙门氏菌、志贺氏菌、克雷伯氏菌、变形杆菌、弧菌、假单胞菌、类杆菌、醋酸杆菌、产气杆菌、土壤杆菌、固氮菌、螺旋菌、沙雷氏菌、弧菌、根瘤菌、衣原体、立克次氏体、密螺旋体和梭杆菌。

[00151]“革兰氏阳性细菌”包括球菌、无孢杆菌和有孢杆菌。革兰氏阳性细菌属包括如放线菌、杆菌、梭菌、棒状杆菌、丹毒丝菌、乳杆菌、李斯特氏菌、分枝杆菌、粘球菌、诺卡氏菌、葡萄球菌、链球菌和链霉菌。

[00152]术语“重组微生物”和“重组宿主细胞”在本文中互换使用，指微生物经过基因修饰，用于表达或过度表达内源性多聚核苷酸，或表达非内源性序列如存在于载体的序列或减少表达内源性基因。此多聚核苷酸通常编码与参与产生上述所需的代谢产物的代谢途径的目标酶。因此，本文中描述的重组微生物已经过基因改造，能够表达或过度表达目标酶，此目标酶未被亲本微生物表达或过度表达过。应当了解，术语“重组微生物”和“重组宿主细胞”不仅指特定的重组微生物，还指此微生物的子代或潜在子代。

[00153]“亲本微生物”指用来制造重组微生物的细胞。术语“亲本微生物”描述了一种自然产生的细胞，即未经过基因修饰的“野生型”细胞。术语“亲本微生物”还描述了一种经过基因修饰的细胞，但是此细胞不能表达或过度表达目标酶，例如参与生物合成途径以能够产生所需的例如以下代谢产物的酶：正丙醇、异丁醇、正丁醇、 2-甲基-1-丁醇、 3-甲基-1-丁醇或 2-苯基乙醇。例如，一种野

生型微生物通过基因修饰，能够表达或过度表达硫解酶等第一目标酶。此微生物能作为亲本微生物制造修饰微生物，以表达或过度表达第二目标酶，例如羟丁酰 CoA 脱氢酶。反过来，表达或过度表达如硫解酶和羟丁酰 CoA 脱氢酶的修饰微生物可经过再修饰，以表达巴豆酸酶等第三目标酶。因此，亲本微生物可作为连续基因修饰的参考细胞。每次修饰通过向参考细胞里引入核酸分子来实现。引入核酸分子有利于表达或过度表达目标酶。应当了解，术语"有利于"包含通过对亲本微生物体内启动子序列的基因修饰来激活编码目标酶的内源性多聚核苷酸。还应了解，术语"有利于"也包含将编码目标酶的外源性多聚核苷酸引入亲本微生物。

[00154] 在另一实施方案中，提供了一种生产重组微生物的方法，此微生物能将合适的碳底物转化为正丙醇、异丁醇、正丁醇、2-甲基-1-丁醇、3-甲基-1-丁醇或 2-苯基乙醇。此方法包括用编码包括以下多肽的一种或多种重组多聚核苷酸转化一种微生物：例如乙酰羟酸合成酶（如 *ilvIH* 操纵子）、乙酰羟酸还原异构酶（如 *ilvC*）、二羟酸脱水酶（如 *ilvD* ）、2-酮酸脱羧酶（如 *PDC6*、*ARO10*、*THI3*、*kivd* 或 *pdc*）、2-异丙基苹果酸合成酶（如 *leuA*）、β-异丙基苹果酸脱氢酶（如 *leuB*）、异丙基苹果酸异构酶（如 *leuCD* 操纵子）、苏氨酸脱水酶（如 *ilvA*）、α-异丙基苹果酸合成酶（如 *cimA*）、β-异丙基苹果酸脱氢酶（如 *leuB*）、异丙基苹果酸异构酶（如 leuCD 操纵子）、苏氨酸脱水酶（如 *ilvA*）、乙酰乳酸合成酶（如 *ilvMG* 或 *ilvNB*）、乙酰羟酸还原异构酶（如 *ilvC*）、二羟酸脱水酶（如 *ilvD*）、β-异丙基苹果酸脱氢酶（如 leuB）、分支酸变位酶 P/预苯酸脱水酶（如 *pheA*）、分支酸变位酶 T/预苯酸脱氢酶（如 *tyrA*）、2-酮酸脱羧酶（如 *kivd*、*PDC6* 或 *THI3*）和醇脱氢酶活性。编码用于生成代谢物的酶的多聚核苷酸用于重组核酸分子，这些酶包括其同源体、突变体、片段、相关融合蛋白或具有相同功能的物质，所述重组核酸分子引导此类多肽在合适的宿主细胞如细菌或酵母菌细胞内表达。应当了解，如果增加序列不会改变多聚核苷酸的编码活性，例如加入非功能性或非编码序列，则构成对基本核酸的保守性改变。酶的"活性"是测定其催化生成代谢物反应的能力，即"起作用"，可表示为反应生成代谢物的速率。例如，酶活性可用单位时间或单位酶产生的代谢物的量（例如浓度或质量）表示，或用亲和性或电离常数表示。

[00155] "蛋白质"或"多肽"在本文中互换使用，它们包含一条或多条化学结构单元，即通过化学键即肽键联结在一起的氨基酸。"酶"指任何全部或大部分由蛋

白质构成的任何物质，它能在一定程度上特异性地催化或促进一个或多个化学或生化反应。术语“酶”也可指具有催化能力的多聚核苷酸（例如 RNA 或 DNA）。“天然”或“野生型”蛋白质、酶、多聚核苷酸、基因或细胞指自然界中存在的蛋白质、酶、多聚核苷酸、基因或细胞。

[00156] 应当了解，上述多聚核苷酸包括“基因”，上述核酸分子包括“载体”或“质粒”。例如，编码酮基硫解酶的多聚核苷酸可由 *atoB* 基因或其同源基因编码，或由 *fadA* 或其同源基因编码。因此，术语“基因”又称为“结构基因”，指编码特定氨基酸序列的多聚核苷酸，此氨基酸序列构成全部或部分蛋白质或酶，并且可能包含调控的（非转录的） DNA 序列，例如启动子序列，此序列决定基因的表达条件。基因的转录区域可能包含非翻译区域，包括内含子、5'-非翻译区域（UTR）、 3'-UTR 及编码序列。术语“核酸”或“重组核酸”指多聚核苷酸，例如脱氧核糖核酸（DNA）以及在适当的情况下指核糖核酸（RNA）。关于基因序列的“表达”这一术语是指基因的转录，以及在适当的情况下指产生的 mRNA 转录物翻译为蛋白质这一过程。因此，通过本文可以清楚了解蛋白质的翻译是开放阅读框序列的转录以及翻译的结果。

[00157] 术语“操纵子”指从一个共同启动子转录为一个转录单元的两个或多个基因。在某些实施方案中，这些构成操纵子的基因为邻接基因。应当了解，通过修饰共同的启动子可以对整个操纵子的转录进行修饰（如增加、减少或缺失）此外，可以对操纵子内的任何基因或基因组合进行修饰，以改变被编码多肽的功能或活性。修饰可以提高被编码多肽的活性。此外，修饰可以赋予被编码蛋白质新活性。代表性的新活性包括能够使用替代性底物和/或能够在替代性环境条件下起作用。

[00158] “载体”指任何能够在生物体、细胞或细胞组分间传播和/或转移核酸的工具。载体包括病毒、噬菌体、前病毒、质粒、噬菌粒、转座子和人工染色体，例如 YACs （酵母菌人工染色体）、BACs （细菌人工染色体）和 PLACs （植物人工染色体），及 “游离体”类，即能够自主复制或整合进宿主细胞的染色体中。载体也可以是裸体 RNA 多聚核苷酸、裸体 DNA 多聚核苷酸、在同一条链上包含 DNA 和 RNA 的多聚核苷酸、聚赖氨酸结合的 DNA 或 RNA、肽结合 DNA 或 RNA、脂质体结合 DNA，或自然界中不能游离存在的一类多聚核苷酸，或者可以是包含一种或多种上述多聚核苷酸构建体的生物体，例如土壤杆菌或细菌。

[00159]“转化”指将载体引入宿主细胞的过程。可通过包括电穿孔、显微注射、生物弹射（或粒子轰击介导的递送）或农杆菌介导转化的任何一种方式实现转化（或转导、转染）。

[00160]如本发明所属领域的技术人员所公认的，由于遗传密码的简并性，许多具有不同核苷酸序列的 DNA 化合物可用于编码本发明中的某个氨基酸序列。本文对编码上述生物合成酶的天然 DNA 序列的引用仅用于描述本发明的一个实施方案，公开的内容包括对本发明提供的方法中利用的酶和多肽的氨基酸序列进行编码的任何序列的 DNA 化合物。按照类似方式，通常可以允许多肽的氨基酸序列中一个或多个氨基酸的替代、缺失和插入，而不对其所需的活性造成损害或严重损害。本发明包括具有可置换氨基酸序列的多肽，并且本文所示的 DNA 序列编码的氨基酸序列仅用于描述本发明的实施方案。

[00161]本发明以重组 DNA 表达载体或质粒的形式提供核酸分子，如下文详述，此核酸分子编码一种或多种目标酶。一般情况下，此类载体能够在宿主微生物的细胞质内复制或整合进宿主微生物的染色体 DNA 中。在这两种情况下，此载体是稳定载体（即，即使只存在选择压力，经过多次细胞分裂，此载体仍能存在下去，）或短暂载体（即随着细胞分裂，此载体会逐渐在宿主微生物中丢失）。本发明提供经分离（即不纯，以自然界中没有的丰度和/或浓度存在于制品中）和纯化（即基本上不含污染物质，或基本上不含有自然界中发现的与相应 DNA 在一起的物质）的 DNA 分子。

[00162]本文提供一个或多个生物合成基因的异源表达方法以及此方法所用的重组 DNA 表达载体，这些基因涉及正丙醇、异丁醇、正丁醇、2-甲基-1-丁醇、3-甲基-1-丁醇和/或 2-苯基乙醇的生物合成。因此，包含此类核酸的重组表达载体包括在本发明的范围内。术语“表达载体”指一种能够被引入宿主微生物或无细胞转录和翻译系统的核酸。表达载体可永久或暂时存在于微生物中，不论是作为染色体的一部分或作为其他 DNA 存在于微生物中或任何细胞器中，例如细胞质中的复制载体。一个表达载体同样含有启动 RNA 表达的启动子，在微生物或细胞抽提物中 RNA 通常被翻译成多肽。为了有效地将 RNA 翻译成蛋白质，表达载体通常还要含有核糖体结合点序列，此序列位于待表达基因编码序列起始密码子的上游。其他元件如增强子、分泌信号序列、转录终止序列以及一个或多个标记基因也可能存在于表达载体中，通过标记基因可以识别和/或选

择含有载体的宿主微生物。当转化细胞在合适的选择培养基中生长时，利用可选标记即具有抗生素抗性或对抗生素敏感的基因赋予转化细胞可选表型。

[00163]表达载体的各种组分可以有很大差异，这取决于载体的用途和供载体进行复制或启动表达的宿主细胞。适用于大肠杆菌、酵母菌、链霉菌及其他常用细胞中基因表达和载体维持的表达载体组分已广为人知，且可以购买。例如，本发明所述的载体含有能够在真核或原核宿主微生物中起作用的合适启动子。启动子可以含有调控序列，来调控与宿主微生物生长有关的表达，或在某种化学或物理刺激下引起基因表达的开启和关闭。对于大肠杆菌和其他细菌宿主而言，可以利用源自生物合成酶、抗生素抗性酶及噬菌体蛋白基因的启动子，这些启动子包括半乳糖启动子、乳糖（lac）启动子、麦芽糖启动子、色氨酸（trp）启动子、β-内酰胺酶（bla）启动子、噬菌体 λPL 启动子 和 T5 启动子。此外，也可以利用合成启动子，例如 tac 启动子（第 4,551,433 号美国专利）。对于大肠杆菌表达载体，包含大肠杆菌的复制起点例如 pUC、p1P、p1 和 pBR 是有用的 。

[00164]因此，重组表达载体包含至少一个表达系统，反过来，表达系统由至少部分 PKS 和/或其他可通过操作连结到启动子的生物合成基因编码序列和可选的终止序列组成，启动子和终止序列能够作用，使编码序列在相容的宿主细胞里表达。通过利用本发明所述重组 DNA 表达载体进行转化，从而对宿主细胞进行修饰，使其含有表达系统序列，此序列可能位于染色体外，或者被整合进染色体中。

[00165]可根据标准 PCR 扩增技术以及下述实施例中的流程，利用 cDNA、mRNA 或基因组 DNA 作为模板，并采用合适的低聚核苷酸引物，对本发明所述的核酸进行扩增。被扩增的核酸可被克隆到合适的载体上，并通过 DNA 序列分析表征。此外，可通过标准技术制备核苷酸序列的相应低聚核苷酸，例如利用 DNA 自动合成器。

[00166]应当了解，编码与本文所述的酶同源的多肽的分离核苷酸分子，能够通过向编码特定多肽的核苷酸序列引入一个或多个核苷酸置换、增加或缺失制备，这样可向编码蛋白质引入一个或多个氨基酸置换、增加或缺失。可利用标准技术，例如定点诱变和 PCR 介导诱变，将突变引入多聚核苷酸中。与那些可能需要进行非保守氨基酸置换（见上文）的位点不同，一些位点更适宜进行保守氨基酸置换。“保守氨基酸置换”指用具有相似侧链的氨基酸残基替换氨基酸。

本发明所属领域对具有相似侧链的氨基酸残基家族进行了定义。这些氨基酸家族包括碱性侧链氨基酸（如赖氨酸、精氨酸、组氨酸）、酸性侧链氨基酸（如天门冬氨酸、谷氨酸）、不带电荷的极性侧链氨基酸（如甘氨酸、天门冬酰胺、谷氨酰胺、丝氨酸、苏氨酸、酪氨酸、半胱氨酸）、非极性侧链氨基酸（如丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、苯丙氨酸、蛋氨酸、色氨酸）、β-分支侧链氨基酸（如苏氨酸、缬氨酸、异亮氨酸）和芳烃侧链氨基酸（如酪氨酸、苯丙氨酸、色氨酸、组氨酸）。

[00167] 在另一实施方案中，提供了生产正丙醇、异丁醇、正丁醇、2-甲基-1-丁醇、3-甲基-1-丁醇或 2-苯基乙醇的方法。此方法包含存在合适底物的情况下，以及在适于将底物转化为正丙醇、异丁醇、正丁醇、2-甲基 1-丁醇、3-甲基 1-丁醇或 2-苯基乙醇的条件下，培养本发明提供的重组微生物，可采用技术人员熟知的方法检测本发明提供的微生物产生的醇。此类方法包括质谱法，详见以下说明及表 6 。适于本发明提供的重组微生物的生长和保存的培养条件见下述实施例。正如技术人员所公认的，此类条件可以根据每种微生物的需求进行修改。

[00168] 如前所述，描述本发明所用的包括使用载体、启动子的分子生物学技术以及其他许多相关专题的一般性文档，包括 Berger 与 Kimmel 的《分子克隆技术指南》，第 152 卷《酶学方法》（Academic Press, Inc.，圣地亚哥，加利福尼亚) ("Berger"); Sambrook 等的《分子克隆实验手册》第二版，第 1-3 卷，冷泉港实验室，冷泉港，纽约，1989 ("Sambrook") 和 《现代分子生物学实验技术》，F. M. Ausubel 等编著，《现代实验技术》，Greene Publishing Associates, Inc. 和 John Wiley & Sons, Inc. 联合出版，（1999 年增补）("Ausubel")。足以帮助技术人员掌握体外扩增方法(包括聚合酶链式反应（PCR）、连接酶链式反应（LCR）、Qβ 复制酶扩增和其他 RNA 聚合酶介导技术（例如 NASBA）) 以例如用于生产本发明所述的同源核酸的实验技术实例可在以下文献中找到: Mullis 等人(1987) U.S. Pat. No. 4,683,202; Innis *et al.*, eds. (1990) PCR Protocols: A Guide to Methods and Applications (Academic Press Inc. San Diego, Calif.) ("Innis"); Arnheim & Levinson (Oct. 1, 1990) C&EN 36-47; The Journal Of NIH Research (1991) 3: 81-94; Kwoh 等人(1989) Proc. Natl. Acad. Sci. USA 86: 1173; Guatelli 等人(1990) Proc. Nat'l. Acad. Sci. USA 87: 1874; Lomell 等人(1989) J. Clin. Chem 35: 1826; Landegren 等人 (1988) Science 241: 1077-1080; Van Brunt (1990)

Biotechnology 8: 291-294; Wu and Wallace (1989) Gene 4:560; Barringer 等人(1990) Gene 89:117; and Sooknanan and Malek (1995) Biotechnology 13: 563-564. Improved methods for cloning *in vitro* amplified nucleic acids are described in Wallace *et al.*, U.S. Pat. No. 5,426,039. Improved methods for amplifying large nucleic acids by PCR are summarized in Cheng 等人(1994) Nature 369: 684-685 及其参考文献，其中产生了高达 40kb 的 PCR 扩增子。技术人员会认识到，从本质上讲，任何 RNA 都能用逆转录酶和聚合酶转变为双链 DNA，以适用于限定消化、PCR 扩增及测序。详见上述 Ausubel、Sambrook 和 Berger 等的文献。

[00169] 合适的培养条件包括培养基的 pH、离子强度、营养含量等；温度；氧气/CO_2/氮气含量；湿度及其允许通过宿主微生物即通过微生物的代谢活动生产化合物的其他培养条件。能够作为宿主细胞的微生物的合适培养条件已熟知。

[00170] 本发明通过以下实施例进行阐述，此实施例是以说明的方式而不是限制性的方式提供。

[00171] 本发明的代表性微生物根据《布达佩斯条约》于 2008 年 2 月 7 日保存在美国模式培养物保藏所，邮箱：弗吉尼亚州马纳萨斯 1549 号，邮编：20108，ATCC 号码为_____（名称为 SA237 的微生物）和 ATCC 号码为（名称为 CRS-BuOH23 的微生物）。保存的微生物存放在经认可的存放处。在保存方最近一次收到发放样本的请求后五年内；或自保存之日起至少三十年内；或处于相关专利的实施寿命内 (以最长时间计)，如微生物发生突变、没有活力或遭到破坏，则可对其进行更换。专利一经授权，即会不可撤消地去掉专利申请中关于对公众获得这些细胞系的所有限制。

实施例

[00172] 用于 PCR 反应的 KOD DNA 聚合酶购自 EMD Chemicals (圣地亚哥，加利福尼亚)。所有限制性内切酶和热敏磷酸酶(antarctic phosphatase)均来自 New England Biolabs （伊普斯威奇，马萨诸塞州）。快速 DNA 连接试剂盒来自 Roche（曼海姆，德国）。低聚核苷酸购自 Operon（亨茨维尔，阿拉巴马州） 培养基中的所有抗生素和试剂均购自 Sigma Aldrich（圣路易斯，密苏里州）或 Fisher Scientifics（休斯顿，田纳西州）。表 10 中列出了所使用的低聚核苷酸。

[00173] 表 10:

表 10 采用的菌株、质粒和低聚核苷酸

菌株	基因型
BW25113	*rrnB*$_{T14}$ Δ*lacZ*$_{WJ16}$ *hsdR*514 Δ*araBAD*$_{AH33}$ Δ*rhaBAD*$_{LD78}$
XL-1 BLUE	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* (F' *proAB lacI*^{q}*ZΔM15 Tn10* (Tet^R))
BW25113 F'	BW25113 (*traD36, proAB*+, *lacI*^{q} *ZΔM15* (Tet^R))
CRS 21	BW25113F' Δ*metA*
CRS 22	BW25113F' Δ*metA, tdh*
CRS 23	BW25113F' Δ*metA, tdh, ilvB*
CRS 24	BW25113F' Δ*metA, tdh, ilvB, ilvI*
CRS 31	BW25113F' Δ*metA, tdh, ilvB, ilvI, adhE*
CRS-BuOH 2	BW25113F' Δ*metA, tdh* + pCS49 / pSA62 / pSA55I
CRS-BuOH 11	BW25113F' Δ*metA, tdh, ilvB, ilvI* + pCS49 / pSA62 / pSA55I
CRS-BuOH 12	BW25113F' + pCS49 / pSA62 / pSA55I
CRS-BuOH 18	BW25113F' Δ*metA, tdh, ilvB, ilvI* + pCS49 / pCS51 / pSA55I
CRS-BuOH 19	BW25113F' Δ*metA, tdh, ilvB, ilvI* + pCS49 / pCS20 / pSA55I
CRS-BuOH 20	BW25113F' Δ*metA, tdh, ilvB, ilvI* + pCS49 / pCS50 / pSA55I
CRS-BuOH 23	BW25113F' Δ*metA, tdh, ilvB, ilvI, adhE* + pCS49 / pSA62 / pSA55I
CRS-BuOH 31	BW25113F' + pSA62 / pSA55I
CRS-BuOH 32	BW25113F' Δ*metA* + pCS49 / pSA62 / pSA55I

质粒	基因型
pZA31-luc	$P_L$tetO$_1$:: *luc* (VF); p15A ori; Cm^R
pZS24-MCS1	$P_L$lac/ara-$_1$:: MCS1; pSC101 ori; Kan^R
pCS20	$P_L$lacO$_1$:: *tdcB*(EC)-*leuABCD*(EC); p15A ori; Kan^R
pCS27	$P_L$lacO$_1$:: MCS1; p15A ori; Kan^R
pCS49	$P_L$lacO$_1$:: *thrA*BC* (EC ATCC 21277); pSC101 ori; Spec^R
pCS50	$P_L$lacO$_1$:: *tdcB* (EC)-*leuA*BCD*(EC G462D mut); p15A ori; Kan^R
pCS51	$P_L$lacO$_1$:: *ilvA* (EC)-*leuA*BCD*(EC G462D mut); p15A ori; Kan^R

pSA55I	$P_L lacO_1$:: *kivd* (LL)-*adh2* (SC), *lacI*; ColE1 ori; Amp^R
pCS59	$P_L lacO_1$:: *thrABC* (EC); pSC101 ori; $Spec^R$
pSA62	$P_L lacO_1$:: *ilvA* (EC)-*leuABCD*(EC); p15A ori; Kan^R

引物名称	序列 5' → 3'
lacI SacI f	CTAGAGCTCGAAGGAGATATACCATGAAACCAGTAACGTTATACGATG（序列 ID 号：83）
lacI SacI r	CTAGAGCTCTCACTGCCCGCTTTCCAGTC （序列 ID 号：84）
tdcB f *Acc65*	CGAGCGGTACCATGCATATTACATACGATCTGCCGG（序列 ID 号：86）
tdcB r *SalI*	ACGCAGTCGACTTAAGCGTCAACGAAACCGGTGATT （序列 ID 号：86）
thrA f *Acc65*	TCAGGTACCATGCGAGTGTTGAAGTTCGGCGGTACAT （序列 ID 号：87）
thrC r *HindIII*	TCAAAGCTTTTACTGATGATTCATCATCAATTTACGCAA（序列 ID 号：88）

[00174] *菌株。大肠杆菌* BW25113 ($rrnB_{T14}$ $\Delta lacZ_{WJ16}$ *hsdR514* $\Delta araBAD_{AH33}$ $\Delta rhaBAD_{LD78}$) 在比较时被指定为野生型 (WT) (Datsenko and Wanner, Proc. Natl. Acad. Sci. USA 97, 6640–6645, 2000)。在一些异丁醇实验中，JCL16 ($rrnB_{T14}$ $\Delta lacZ_{WJ16}$ *hsdR514* $\Delta araBAD_{AH33}$ $\Delta rhaBAD_{LD78}$ / F' (*traD36, proAB+, lacIq* $Z\Delta M15$)) 用作野生型 (WT)。利用 Keio 保藏菌株作为供体，通过 P1 转导实现宿主 *metA*、*tdh*、*ilvB*、*ilvI*、*adhE*、*pta*、*ldhA* 和 *pflB* 基因的缺失（Baba 等， Mol. Systems Biol. 2, 2006)。通过位于各连续敲除基因间的 pCP20 （Datsenko 和 Wanner，如上所述），以去除被插入到目标基因区域中的 kan^R。然后，采用合适的引物进行 菌落 PCR，以验证基因片段的去除情况。XL-1 BLUE（Stratagene，拉荷亚，加利福尼亚州 CA）用于扩增所有的质粒。

[00175] *质粒的构建。*按照本文其他地方所述设计和构建 pSA40、pSA55 和 pSA62。用引物 *lacI* SacI f 和 *lacI* SacI r 从大肠杆菌 MG 1655 基因组 DNA 扩增 *lacI* 基因。然后， PCR 产物经过 SacI 的消化，并连接到用相同的酶切割过的 pSA55 开环载体中，切割点位于氨苄西林抗性基因的启动子后，由此生成 pSA55I。

[00176] 利用 *大肠杆菌 BW25113* WT 基因组 DNA 和引物 *tdcB* f Acc65 和 *tdcB* r SalIPCR 扩增 *tdcB* 基因。其 PCR 产物经过凝胶纯化，并用 Acc65 和

SalI 消化。然后，将经过消化的片段连接到用同一对酶切割过的 pSA40 开放载体中，由此生成 pCS14。

[00177] 为将 pCS14 的复制起点从 colE1 置换为 p15A，用 SacI 和 AvrII 消化 pZA31-luc。较短的片段经过凝胶纯化，被克隆到用相同的酶切割过的 pCS14 质粒中，由此生成 pCS16。

[00178] 使用 A106 和 A109 作为引物、使用*大肠杆菌* BW25113 基因组 DNA 作为模板，对操纵子 *leuABCD* 进行扩增。其 PCR 产物经过 SalI 和 BglII 的切割，连接到经 SalI 和 BamHI 消化过的 pCS16 中，由此生成 pCS20。

[00179] 为构建与 pSA40 相同但具有 p15A 复制起点的表达质粒，我们将用 SacI 和 AvrII 消化 pZA31-luc 获得的 p15A 片段克隆到用相同限制性内切酶切割过的 pSA40 开放载体中，由此生成 pCS27。

[00180] 采用重叠延伸剪接技术（SOE）以 G462Df 和 G462Dr 为引物，大肠杆菌 BW25113 WT 基因组 DNA 为模板获得 *leuA*BCD* 构建 *leuA** G462D 突变体。然后，消化 SOE 产物，并将其克隆到限制性位点 Acc65 和 XbaI，以此生成 PZE_leuABCD。接下来，利用获得的质粒为模板，A106 和 A109 为引物，进行 PCR 扩增获取 *leuA*BCD*。此产物经过 SalI 和 BglII 的切割，连接到经 SalI 和 BamHI 消化过的 pCS27 中，由此生成 pCS48。

[00181] 以 A110 和 A112 为引物，从*大肠杆菌* BW25113 WT 基因组 DNA 扩增出 *ilvA* 基因。然后，用 Acc65 和 XhoI 切割 *ilvA* 基因，并将其连接到用 Acc65 和 SalI 消化过的 pCS48 开放载体中，由此生成 pCS51。

[00182] 以 *tdcB* f Acc65 和 *tdcB* r SalI 为引物，从*大肠杆菌 BW25113* WT 基因组 DNA PCR 扩增 *tdcB* 基因。得到的 PCR 产物经过凝胶纯化，并用 Acc65 和 SalI 进行消化，然后，连接到用同一对酶切割过的 pCS48 开放载体中，由此生成 pCS50。

[00183] 以 thrA f Acc65 和 thrC r HindIII 为引物，对 WT *thrABC* 进行 PCR 扩增。获得的产物经过 Acc65 和 HindIII 的消化，连接到用同一对酶切割过的 pSA40 上，由此生成 pCS41。

[00184] 为将 pCS14 的复制起点从 colE1 置换到 pSC101，用 SacI 和 AvrII 消化 pZS24-MCS1。较短的片段经过凝胶纯化，被克隆到用相同的酶切割过的 pCS41 质粒中，由此生成 pCS59。

[00185]以 thrA f Acc65 和 thrC r HindIII 为引物，从大量生产苏氨酸的微生物 ATCC 21277 中分离的基因组 DNA 进行 PCR 扩增抗反馈抑制突变体 *thrA** 以及 *thrB* 和 *thrC*。最终产物经过 Acc65 和 HindIII 的消化，连接到用同一对酶切割过的 pSA40 上，由此生成 pCS43。

[00186]为将 pCS43 的复制起点从 colE1 置换到 pSC101，用 SacI 和 AvrII 消化 pZS24-MCS1。较短的片段经过凝胶纯化，被克隆到用相同的酶切割过的 pCS43 质粒中，由此生成 pCt49。

[00187]通过 P1 转导，分别从菌株 JWXXX 和 JWXXX (Baba 等)中缺失支链氨基酸转氨酶（由 *ilvE* 编码）和酪氨酸转氨酶（由 *tyrB* 编码）。

[00188]为了克隆 L-缬氨酸生物合成基因 i) *ilvIHCD* (EC) 和 ii) *als* (BS) 以及 *ilvCD* (EC)，利用 SacI 和 AvrII 消化，将 pZS24-MCS1 中的低拷贝复制起点 (ori)转移并连接到 i) pSA54 和 ii) pSA69 的相应位点，由此分别生成质粒 pIAA1 和 pIAA11。

[00189]为了从 乳酸乳球菌 中克隆 *kivd* 以及从酿酒酵母中克隆 *ADH2*，利用 SacI 和 AvrII 消化以去除 pSA55 的 ColE1 复制起点，并用经相同限制性内切酶消化过的 pSA54 的 p15A 复制起点置换，由此生成 pIAA13。为了更好地控制这些基因的表达，以 lacISacIf 和 lacISacIr 为引物，用 KOD 从大肠杆菌 MG1655 基因组 DNA 扩增 *lacI*，并连接到准备与氨苄西林抗性基因、bla 一同表达的 pCS22 的 SacI 位点上，由此生成 pIAA12 质粒。

[00190]为了在染色体的 BW25113/F' 上过度表达 *leuABCD* 操纵子，用 $P_{LlacO-1}$ 启动子置换天然启动子和前导序列。以 lacO1KanSOEf 和 lacO1LeuA1r 为引物，用 KOD 聚合酶以从 pZE12-luc 中扩增 $P_{LlacO-1}$ 启动子。以 KanLeuO1f 和 KanlacO1SOEr 为引物，从 pKD13 中扩增编码抗卡那霉素的基因 *aph*。加入 1μL 的各反应产物作为模板，并以 KanLeuO2f 和 lacO1LeuA2r 为引物，采用 SOE 技术并以 KOD 聚合酶进行扩增。以 leuKOv1 和 leuKOv2 为引物，从卡那霉素耐受克隆的基因组 DNA 扩增出所述新构建体，然后，对其进行序列分析以确定克隆的准确性。为了从质粒中过度表达 *leuABCD* 操纵子，利用 SacI 和 AvrII 去除来自 pSA54 的 p15A 复制起点，并连接到 pCS22 的相应点位上（ColE1, Cm^R, $P_{LlacO-1}$: *leuABCD*），由此生成 pIAA2 质粒。为了更严格的表达，按照前文所述的生成 pIAA12 的方式对 *lacI* 进行扩增，并将其连接到与氯霉素抗性基因 *cat* 一同表达的 pCS22 上，由此生成 pIAA15 质粒。通过

连接经 XhoI 和 NdeI 消化过的 pIAA15 的 5.5kb 片段以及连接经相同限制性内切酶切割过的 pZE12-leuABCD (ColE1, Amp^R, $P_{LlacO-1}$: *leuA(G1385A)BCD*) 2.3kb 片段，生成含有编码 IPMS(G462D)的 *leuA(G1385A)* 的 pIAA16 质粒。为了控制表达水平，置换 pIAA15 中的 RBS，以与 pIAA16 相匹配。为此，将经过 HindIII 和 NdeI 消化过的 pIAA16 的 5.6kb 片段与经相同酶消化过的 pIAA15 的 2.2kb 片段连接起来，由此生成 pIAA17。

[00191] *培养基与培养*。对于最初的生产试验，菌株在改良的 M9 培养基(每升水含 6g Na_2HPO4，3g KH_2PO_4，1g NH_4Cl，0.5g NaCl，1mM $MgSO_4$，1mM $CaCl_2$，10mg 维生素 B1)中生长，此培养基包含 10g/L 的葡萄糖、 5g/L 的酵母抽提物以及 1000X 微量金属混合物 A5（每升水含 2.86g H_3BO_3，1.81g $MnCl_2·4H_2O$ ， 0.222g $ZnSO_4·7H_2O$ ， 0.39g $Na_2MoO_4·2H_2O$ ， 0.079g $CuSO_4·5H_2O$，49.4mg $Co(NO_3)_2·6H_2O$）。- 从 3ml 过夜的 LB 培养基中取 1% 接种到置于 125mL 螺旋盖瓶中的 10mL 新鲜培养基中，并在 37°C 旋转摇床上培养 4 小时。然后，向培养基中加入 1ml IPTG 在 30°C 下生长 18 个小时进行诱导：根据需要加入抗生素（氨苄西林 100μg/mL，氯霉素 35μg/mL，卡那霉素 50μg/mL）。

[00192] 对于一些醇发酵实验，从 LB 平板上挑出单个菌落，并将其接种到含有适量抗生素（氨苄西林 100μg/mL、卡那霉素 50μg/mL 和壮观霉素 50μg/mL）的 3ml LB 培养基中。然后，将此 LB 培养基置于 37°C 旋转摇床（250 rpm）上过夜培养。之后，再将该过夜培养物接种 (1% vol/vol) 到 置于 250ml 锥形瓶中的 20ml 含有以下物质的 M9 培养基中(6g Na_2HPO4，3g KH_2PO_4，0.5g NaCl，1g NH_4Cl，1mM $MgSO_4$，10mg 维生素 B1 和 0.1mM $CaCl_2$，每升水)：30g/L 的葡萄糖、5g/L 的酵母抽提物、适量抗生素以及 1000X 微量金属混合物 A5（每升水含 2.86g H_3BO_3， 1.81g $MnCl_2·4H_2O$， 0.222g $ZnSO_4·7H_2O$， 0.39g $Na_2MoO_4·2H_2O$，0.079g $CuSO_4·5H_2O$，49.4mg $Co(NO_3)_2·6H_2O$）。将此培养基置于 37°C 的旋转摇床（250 rpm）上培养，直到其 OD_{600} 达到 0.4～0.6。然后，取 12mL 培养基转移到 250ml 螺旋盖锥形瓶中，并加入 1mM IPTG 进行诱导。将被诱导的培养物置于 30°C 旋转摇床（240rpm）上培养。在接下来的三到四天里，打开螺旋盖锥形瓶取样，并采用离心分离或过滤的方式处理培养液，以回收上清液。在所述的一些实验中，在诱导的同时直接加入 8g/L 苏氨酸到细胞培养物中。

[00193] 除非另有说明，所有 α-酮酸实验均在“富”氧条件下进行。对于富氧实验，将3mL 的 LB 过夜培养物取 1% 接种到装在 250mL 带挡板摇瓶的 10mL 培养液中。对于缺氧实验，则如前所述对置于 125mL 螺旋盖烧瓶中的 10mL 培养液接种。所有培养物都在 37°C 下培养 4 小时，并加入 1mM IPTG 诱导，再在 30°C 下培养 18 个小时后收获培养物。

[00194] 如前所述执行终产物实验，不同之处为使用置于 250mL 螺旋盖烧瓶中的 20mL 含有 5g/L 葡萄糖的改良 M9 培养基。

[00195] *代谢物检测。*用配备火焰离子化检测器的气相色谱仪（GC）对产生的醇类化合物进行定量测定。此系统由 5890A GC 型（惠普，埃文代尔，宾夕法尼亚州）和 7376A 型自动进样器、取样器和控制器（惠普）组成。将培养液的上清液（0.1ml）以分流进样模式（分流比为 1:15）进样，并使用甲醇作为内标物。

[00196] 用购自 Agilent Technologies（圣克拉拉，加利福尼亚州）的 DB-WAX 毛细管柱（30 m，0.32 mm-内径, 0.50 μm-膜厚）分离醇产物。初始 GC 炉温为 40°C，保持 5min，然后，以 15°C/min 的梯度逐渐升温，直到升至 120°C。接着，以 50°C/min 的梯度逐渐升温，直到升至 230°C，保持 4min。以氦气作为载气，进气压力为 9.3psi。进样器和检测器保持在 225°C。以分流进样模式注射 0.5ul 培养物上清液，分流比为 1:15。使用甲醇作为内标物。

[00197] 对于其他分泌的代谢产物，向配备自动取样器（Agilent 科技）和 BioRad（Biorad 实验室，Hercules，加利福尼亚州）Aminex HPX87 柱（5mM H_2SO_4, 0.6ml/min, 柱温 65°C）的 Agilent 1100 HPLC 仪载入过滤的上清液（20uL）。通过折光率检测器检测葡萄糖，用光电二极管阵列检测器检测有机酸，检测波长为 210nm 。对标准曲线应用外推法测定浓度。

[00198] 对于其他分泌的代谢产物，将过滤的上清液（0.02mL）加样到配备自动取样器（Agilent 科技）和 BioRad（Biorad 实验室，Hercules，加利福尼亚州）Aminex HPX87 柱（0.5mM H_2SO_4, 0.6ml/min, 柱温 65°C）的 Agilent 1100 HPLC 仪上。通过折光率检测器检测葡萄糖，用光电二极管阵列检测器检测有机酸，检测波长为 210nm。对标准曲线应用外推法测定浓度。

[00199] *L-缬氨酸和 L-亮氨酸生物合成途径基因的表达能够产生 3-甲基-1-丁醇。*为了在大肠杆菌中生成 3-甲基-1-丁醇，从丙酮酸到 3-甲基-1-丁醇的整条途径都要过度表达。ilvIHCD（*大肠杆菌*）、kivd（*乳酸乳球菌*） ADH2（*酿酒酵*

母）都在 $P_{LlacO-1}$ 启动子的控制下通过质粒（pSA54 和 pSA55）表达。通过用 JCL16 中的 $P_{LlacO-1}$ 启动子置换 leuA 的上游非编码区域，使 leuABCD 操纵子得以过度表达。经过 18 小时的 IPTG 的诱导，此菌株可生成 56mg/L 3-甲基-1-丁醇（图 47A）。为了提高 3-甲基-1-丁醇的产量，使用来自枯草芽孢杆菌的 alsS 置换 ilvIH，置换后显示 3-甲基-1-丁醇的产量得以增加（67mg/L）(图 47A)。为了增强亮氨酸生物合成途径的表达水平，leuABCD 也被克隆到 p15A 衍生的质粒中，并在 $P_{LlacO-1}$ 启动子的控制下表达。leuABCD 基于质粒的表达提高了含有 ilvIH （177mg/L） 或 alsS (124mg/L) (图 47B)菌株中的 3-甲基-1-丁醇产量，但 alsS 的过度表达也导致异丁醇产量的急剧增加。

[00200] 宿主中争夺碳和还原能力的途径被缺失。与野生型（WT）背景相比，*adhE*、*frdBC*、*ldhA*、*pta*、*fnr* 和 *pflB* 的缺失提高了大肠杆菌中异丁醇的产量。对于表达 *alsS* 的菌株，当 3-甲基-1-丁醇途径被转入此菌株时，3-甲基-1-丁醇的最终滴定浓度为 76mg/L（图 47B）。虽然 3-甲基-1-丁醇的累积量减少，但醇产量是由异丁醇决定，其最终浓度超过 1.3g/L。由于此过程异丁醇的滴定浓度是目标产物的 10 倍，因此，对此过程及代谢途径进行检查以解释结果。

[00201] *苏氨酸生物合成的去调节。*如图 50 所示， 2-酮丁酸和 2-酮戊酸分别是正丙醇和正丁醇的前体。2-酮丁酸是源自苏氨酸的常见中间产物，并且是异亮氨酸生物合成的一个前体，2-酮戊酸则是少见的代谢产物，可用于细胞合成非天然氨基酸 - 正缬氨酸。为了增加用于生产正丁醇的 2-酮戊酸累积量，利用质粒 pSA62 过度表达来自大肠杆菌的基因 *ilvA* 和 *leuABCD* 以 i)引导 2-酮丁酸方向的较高代谢通量和 ii)使正缬氨酸化学合成途径成为主要的 2-酮戊酸生产途径（图 50）。Kivd 和 Adh2 也通过 pSA55I 得以过度表达，使两种酮酸转化为对应的醇。

[00202] 除了 *kivd* 和 ADH2 以外，*ilvA* 和 *leuABCD* 的过度表达使 BW WT 中正丙醇和正丁醇的产量提高了 5 倍，从实际上无法检测的量分别提高到约 60mg/L 和 30mg/L（图 51）。然而，氨基酸生物合成在转录和酶水平上的天然反馈调控对苏氨酸的生成有持续影响，从而，导致 24 小时后正丙醇和正丁醇产量处于稳定状态，异丁醇和乙醇产量则稳步增长，以消耗过量的 NADH。

[00203] 为了确认苏氨酸限制是不是主要瓶颈，在诱导时向大肠杆菌培养基中加入 8g/L 苏氨酸。结果证实了此设想： 在 72 小时内，丙醇和丁醇的量逐渐累积并达到 2g/L，这两种醇的量大约都增长到了 10 倍。既然苏氨酸造成的

ThrA 转录衰减和别构反馈抑制是主要调控机制，在非天然启动子后的 ThrA 的抗反馈突变体有助于解除对苏氨酸合成的调控，因此能够提高下游醇的产量。抗反馈 ThrA（标记为 ThrA*，是一种具有超级生产苏氨酸能力的菌株，ATCC 21277）。此菌株的 *thrA*BC* 操纵子被克隆到 pCS49 质粒并在 $P_L$lacO1 启动子的控制下进行表达。为了进行比较，WT *thrABC* 操纵子被克隆到 pCS 59 并在 $P_L$lacO1 的控制下表达。加入 ThrA*BC 过度表达后，正丙醇和正丁醇的产量比无 ThrA*BC 的菌株提高三到四倍（图 51）。在相同的 BW WT 背景下及在 BW Δ*metA*、Δ*tdh*、Δ*ilvB*、Δ*ilvI*, Δ*adhE* 条件下，带有 WT ThrABC 过度表达的菌株比带有 ThrA*BC 过度表达的相同菌株的两种目标醇产量低 10～20%（图 51 和图 54）。这证实了虽然胞内苏氨酸累积量较少，但仍然能在较低水平上影响 WT ThrA 的活性。正如异丙醇产量降低所解释的（图 51），*thrA*BC* 的存在有助于引导代谢产物通量更多地朝向苏氨酸途径，从而，提高了正丙醇和正丁醇的总生产能力。

[00204] *竞争性途径的消除。*为了进一步增加丙醇和丁醇生产的滴定浓度，去除参与竞争的不良反应的酶，以避免对期望中间产物的消耗及降解。由于苏氨酸的产生是 2-酮丁酸合成中的重要检验点，应首先失活高丝氨酸 O-琥珀酰转移酶 *metA* 和苏氨酸脱氢酶 *tdh*，以降低期望的蛋氨酸合成前体的损耗，并阻碍苏氨酸生成 2-氨基-3-酮丁酸的分解代谢。随着 *metA* 和 *tdh* 的破坏，正丙醇和正丁醇的总产量增加到约 1.2g/L (如图 52 所示)，其中正丙醇占主要部分。在正丁醇的生产上可以见到这两个基因的缺失影响不太显著，其可能是由于 2-酮丁酸被分向参与异亮氨酸途径和/或提供乙酰-CoA 的原因。为了进一步保留乙酰-CoA 和 2-酮丁酸这两个生成正丁醇的主要前体，需破坏缬氨酸、亮氨酸和异亮氨酸生物合成中的第一步酶反应。去除乙酰羟酸合成酶同功酶（AHAS I）的大催化亚基（由 *ilvB* 编码）以及 AHAS III 的催化亚基（由 *ilvI* 编码）能够导致产生上述氨基酸的营养缺陷体。这两种额外的缺失导致正丁醇产量增加两倍（图 52），正丙醇产量保持不变。另外，由于消除了异丁醇和（2-，3-）甲基-丁醇的前体，从而，也几乎中止其生产。异丁醇累积量很少可能是由 IlvE 催化的缬氨酸合成途径中最后一步反应的逆反应造成，这一途径导致培养基（补充酵母抽提物）中的缬氨酸被消耗并被反转化为它的前体 2-酮基-异戊酸。

[00205] 为了降低乙醇产量，需缺失*大肠杆菌 adhE* 基因。虽然 *adhE* 的破坏未能显著增加 C3 和 C4 醇的总产量，但是由于它使醇的生成量从 0.25g/L 降

到 0.1g/L 左右，因此，实际上增加了特异性。随着这些基因从基因组中消除，最终菌株（Δ*metA*、Δ*tdh*、Δ*ilvB*、Δ*ilvI*、Δ*adhE*）的正丙醇和正丁醇产量比例接近 1:1，且乙醇的积累量很少，异丁醇和(2-,3-)甲基-丁醇的产量处于基础水平。

[00206] *对替代性抗反馈抑制酶的测定*。由于本发明所述的正丙醇和正丁醇的生产强烈依赖于宿主的氨基酸生物合成机器，因此，需要证实苏氨酸下游的必要醇前体不受细胞中存在的各种氨基酸调控机制的限制，特别是异亮氨酸和亮氨酸分别抑制 IlvA 和 LeuA 的酶活性。

[00207] TdcB 是大肠杆菌的异化苏氨酸脱水酶，其可替代 IlvA 催化苏氨酸生成 2-酮丁酸的脱氨基作用，同时它又天生对异亮氨酸反馈抑制不敏感。为了评估这种替代性酶对生产的益处，*tdcB* 通过 pCS20 中的 *leuABCD* 在 $P_L$lacO1 后过度表达。结果显示，TdcB 催化的两种目标醇产量比 IlvA 低 70%（图 53）。很有可能通过添加酵母抽提物产生的微量异亮氨酸对 IlvA 酶活性的抑制作用并不显著。因此，在给定的实验条件下，TdcB 对反馈抑制的不敏感与其自身的酶活相比显得不太重要。

[00208] 类似地，存在于酵母抽提物中的亮氨酸对 LeuA 的反馈抑制可用于构建并检测 *leuA** 反馈不敏感突变体 G462D。通过使用 SOE 进行定点导向诱变，在 *leuA**上引入点突变，在 pCS51 质粒中过度表达产生的 *leuA*BCD* 操纵子。如图 53 所示，反馈不敏感的 LeuA* 未能提高正丙醇和正丁醇的产量。这再次证实了存在于细胞中的亮氨酸的量很可能低于能够对 LeuA 酶活性起不良作用的抑制水平。

[00209] *从 CRS-BuOH 23 共同产生正丙醇和正丁醇及其副产物*。图 54 显示了具有 pCS49、pSA62 和 pSA55I 质粒(CRS-BuOH 23)的最终菌株 BW Δ*metA*、Δ*tdh*、Δ*ilvB*、Δ*ilvI*、Δ*adhE* 产生醇和代谢产物的时间过程。在 72 小时内，丙醇和丁醇的产量几乎呈直线上升，到第三天结束时似乎达到高峰。同样的模式也可见于乙醇生产中。Δ*adhE* 背景中生成的乙醇可能是由于 Kivd 对丙酮酸的亲和性较小所致。另一方面，醇生产阶段后，主要代谢产物乙酸和乳酸的胞外产量继续显著增长，这可能是乙酰-Coa 和 NADH 过量所致。如图 54 所示，对葡萄糖的消耗主要发生于醇生产阶段，并且似乎与细胞生长无关。头 24 小时过后，细胞停止生长，并在醇生产阶段以后的几天似乎仍然保持停止。

[00210] *发酵产物的移除及其对生物燃料生产的影响* 主要的混合酸发酵基因 *adhE*、*ldhA*、*pta*、*pflB* 以各种组合形式被缺失，以进一步表征本发明的 C3/C4

醇生产体系。如表 12 所示，乙醇的生产主要消耗过量的 NADH，而 *adhE* 的破坏则导致乳酸更加显著的积累。当 *ldhA* 被缺失后，乙醇的分泌量约为 1g/L，而乙酸的产量也有所增加。在 *pta* 和 *ldha* 敲除菌株中均观察到葡萄糖消耗量的减少。*adhE*、*ldhA*、*pta* 的缺失及它们的组合的缺失都会导致丙醇和丁醇产量的降低，*pflB* 对这一方面的影响不太显著。这说明乙酰-CoA 的大量积累是由丙酮酸脱氢酶复合体（PDHc）而非 PflB 复合体的活性造成。

表 11

发酵基因敲除对目标醇及次要副产物水平的影响

	BW25113F'背景下，除 ΔmetA、tdh、ilvB、ilvI 外的基因缺失					
产物水平（mg/L）	--	*ΔadhE*	*Δpta*	*ΔldhA*	*ΔpflB*	*ΔadhE, pta*
丙醇	1017.0	1218.5	576.8	649.0	788.5	710.0
丁醇	1006.5	1094.5	770.5	581.5	932.5	853.5
乙醇	278.0	127.4	358.5	923.0	285.5	122.0
丙酮酸	17.9	16.7	23.8	40.2	9.9	8.4
乙酸	2882.8	2726.7	2895.5	4230.0	3390.0	2346.1
乳酸	0.0	494.6	0.0	0.0	0.0	0.0
甲酸	0.0	0.0	0.0	0.0	0.0	0.0
消耗的葡萄糖（g/L）	25.3	26.1	23.7	18.7	23.4	25.4

采用本文所述的材料和方法对细胞进行 72 小时的培养。表中所示为第 72 小时时间点的数据。

[00211] 通过异源 *kivd* 和 *adh2* 以及大肠杆菌 *ilvA*、*leuABCD* 和 *thrA*BC* 的过度表达，本发明阐述了正丙醇和正丁醇的生产。利用 2-酮戊酸生产正丁醇不可避免地会涉及到中间产物 2-酮基丁酸和非自然的正缬氨酸生物合成途径。由于 Kivd 对 2-酮酸和 2-酮丁酸的亲合力相近，并且 2-酮丁酸是 LeuA 的第二底物，所以正丙醇与正丁醇被一同生产出来，且数量也接近。

[00212] 对苏氨酸生物合成的去调节以及对 *metA* 和 *tdh* 催化的分支途径的消除成功地保持了苏氨酸的积累，并且提高了正丙醇和正丁醇生产的滴定浓度。另一方面，赖氨酸生物合成虽然与苏氨酸途径岔开，但是并没有消除其参与细菌细胞壁合成的重要中间体二氨基庚二酸。然而，要考虑赖氨酸营养缺陷体的重要性。通过合理的设计，苏氨酸超生产也能体现在 *lysA* 缺失的有益影响中。通过利用钝化的 AHAS I 和 III 中断缬氨酸、亮氨酸和异亮氨酸的生物合成途径，可以进一步提高正丁醇生产能力，使正丁醇的产量增加两倍，但是对正丙

醇没什么效果。这种选择性增加要归因于：i)可利用的 2-酮丁酸和乙酰-CoA 增多； ii）亮氨酸途径中，其天然底物释放必要的酶 LeuABCD。

[00213] 如图 54 所示，在实验条件下，每种醇的产量约为 13～15mg/L/hr。菌株对正丙醇和正丁醇耐受性的增强可进一步提高生产能力。

[00214] 转录调控与减弱是氨基酸调控的两种主要机制，由于本发明中的所有必要基因被克隆到非天然启动子之后并被表达且不具有前导序列，因此这两种机制对本发明中的关键酶影响很小。另一方面，不能忽略由酶的氨基酸产物特别是 IlvA 和 LeuA 引起的酶别构反馈抑制。TdcB 是生物降解型苏氨酸脱水酶，它在过量氨基酸和稀少葡萄糖存在情况下为厌氧生长的细胞提供代谢能。它的表达是通过分解代谢物抑制进行转录控制，并通过降低其对苏氨酸的 K_m 值由其别构效应因子 AMP 进行激活。由于高浓度的丙酮酸和一些 2-酮酸（包括 2-酮丁酸）会抑制 TdcB 的酶活，因此当没有足够的 AMP 来抵消负面影响时，这些中间代谢物的积累对 TdcB 酶活力有有害影响， 并且，当细胞内缺乏足够的 AMP 水平时，由于 TdcB 对苏氨酸的 K_m 值比 IlvA 的高，脱氨基速率会减慢，因此造成醇的总生产能力的下降，见图 53 （缺乏异亮氨酸的情况下，提纯的大肠杆菌 *IlvA* 的 Km = 8 mM，缺乏 AMP 的情况下，大肠杆菌 TdcB 的 Km = 20mM）。对于 LeuA* 突变体，有可能是在亮氨酸合成途径中朝 2-酮基-异戊酸方向选择的 G462D 突变体导致了它对 2-酮丁酸亲合力的降低。

[00215] 减少半好氧培养环境中发酵副产物的努力已经引起了对现有生产系统中 NADH(P)H 平衡的注意。苏氨酸的生物合成需要消耗三分子的 NADPH 和两分子的 ATP，而起始于 2-酮基丁酸的正丁醇和正丙醇的生产所消耗的 NADH 分别为零净分子和一净分子。磷酸烯醇式丙酮酸羧化酶由基因 *ppc* 编码，其组成性作用能够补充苏氨酸途径中消耗的草酰乙酸（OAA）并维持 TCA 循环的进行。在半好氧条件下，TCA 循环的部分活性会导致产生过多的还原能力，因此，乙醇和乳酸等发酵产物可用来消耗过多的 NADH。如表 11 所示，乙醇是更好的 NADH 消耗者，其消耗的 NADH 比乳酸产生的 NADH 多一分子。*adhE* 破坏后，乳酸的分泌会更显著；另一方面，当 *ldhA* 缺失时，约 1g/L 乙醇累积在培养液中。随着氧气逐渐减少造成的 TCA 循环活力降低，以及涉及糖酵解的下游途径的失效，乙酸的生产似乎是产生 ATP 的主要来源及消耗过量乙酰-CoA 和/或丙酮酸的来源。通过比较图 51 和 54 中的 WT 菌株和 CRS-BuOH 23 菌株，便能解释丁醇生产与乙酸分泌间的负相关性。正如预期的

那样，*pta* 的消除降低了葡萄糖的消耗以及细胞在特定培养体系中的生长。然而，不能消除乙酸的生产，原因正在调查。从培养液中甲酸的缺乏及 *pflB* 敲除对细菌的丁醇生产能力影响不显著可以看出，PDHc 似乎是乙醇生产阶段的乙酰-CoA 的主要提供者。

[00216] 最初的假设是，正丁醇的产量要远高于 3-甲基-1-丁醇的产量，这是由基因产物 *kivd* 和 *leuA* 对底物 α-酮异戊酸的争夺造成。为了验证这一假设，对异丁醇和 3-甲基-1-丁醇的 α-酮酸前体即 *kivd* 底物进行了检查。为了达到这一目的，在 $P_{LlacO-1}$ 的控制下将 *leuABCD* 和 *alsS-ilvCD* 在衍生自 ColE1 的质粒（源自 pSC101 的质粒）中进行表达。在缺氧及富氧条件下，此异丁醇前体和 *leuA* 底物即 α-酮异戊酸（KIV）是主要产物（图 48A）。KIV 的最终浓度分别为 0.25 和 0.29g/L，而 3-甲基-1-丁醇前体即 α-酮基异己酸（KIC）未检测到（< 5mg/L）。为了增加 KIC 的累积量，我们将 RBS 的天然序列改变成更加共有的序列来提高 *leuA* 的表达水平。*leuA* 基因产物的的增加表达使 KIC 的浓度增加到 0.20g/L，但 KIV 仍是主要产物（0.37g/L）（图 48A）。该结果表明，3-甲基-1-丁醇产量的降低并不完全出于对 KIV 的争夺，而是由于 *leuA* 基因产物（异丙基苹果酸合成酶）活性低造成。

[00217] 异丙基苹果酸合成酶（IPMS）催化 KIV 与乙酰-CoA 的缩合反应。α-酮异戊酸的累积可能是由于 α-酮基异己酸合成的游离 L-苏氨酸引起 IPMS 反馈抑制造成。为了解除 *leuA* 基因产物的反馈抑制，可以采取两种策略：第一种，采用 IPMS（IPMS (G462D）的反馈不敏感突变体；第二种，通过缺失 *ilvE*（支链氨基酸转移酶）和 *tyrB*（酪氨酸转氨酶），使 L-亮氨酸合成途径的最后一步反应失活，这两种同工酶负责将 α-酮基异己酸转化为 L-亮氨酸。

[00218] 当 IPMS（G462D）被表达时，产物的分配明显流向 KIC，最终产物浓度为 1.61g/L，而 KIV 累积量减少到 0.17g/L（图 48A）。在表达 WT IPMS 的菌株中，*ilvE* 失活使 KIC 的产量增加到 0.53g/L，*ilvE* 和 *tyrB* 的缺失则使 KIC 的累积量进一步增加到 1.23g/L（图 48B）。Δ*ilvE* 和 Δ*ilvE* Δ*tyrB* 背景下的 KIV 产量与 $ilvE^{+}$ $tyrB^{+}$ 菌株中的 KIV 产量接近，其最终浓度分别为 0.40g/L 和 0.37g/L。通过将 Δ*ilvE* 及 Δ*ilvE* Δ*tyrB* 宿主菌株与 IPMS（G462D）的表达结合起来，KIC 的产量可分别增加到 2.31g/L 和 1.95g/L（图 48B）。

[00219] 由于 KIC 产量增加，所以用 WT IPMS 或 IPMS（G462D）改变丙酮酸生成 3-甲基-1-丁醇的整条反应途径。与酮酸生产相似，在 $ilvE^{+}$ $tyrB^{+}$ 背景下，

带有 WT IPMS 的菌株仍然产生大量异丁醇（169mg/L），但 3-甲基-1-丁醇是主要产物（308mg/L）（图 49A）。正如预期的那样，当 IPMS（*G462D*）在 *ilvE*$^+$ *tyrB*$^+$ 背景下表达时，3-甲基-1-丁醇是主要产物，其最终滴定浓度为 459 mg/L，异丁醇的累积量仅为 15 mg/L（图 49）。通过突变解除 IPMS 反馈抑制后，产物分配从使用 WT IPMS 的 1.8:1（3-甲基-1-丁醇:异丁醇）变为超过 30:1 。其他常见代谢副产物包括丙酮酸、延胡索酸和乙酸的积累量极少（表 12）。

[00220]表 12

3-甲基-1-丁醇生成菌株的代谢副产物

菌株		代谢产物浓度（g/L）			
		葡萄糖	丙酮酸	延胡索酸	乙酸
IAA88	起始值（0 小时）	4.7	ND	ND	ND
IPMS (WT)	最终值（16 小时）	ND[a]	ND	ND	0.03
***ilvE*$^+$ *tyrB*$^+$**	最大值（小时）	4.7 (0 小时)	0.53 (4 小时)	0.02 (6 小时)	0.03 (16 小时)
IAA90	起始值（0 小时）	4.7	ND	ND	ND
IPMS (WT)	最终值（16 小时）	ND	ND	ND	0.06
***ilvE*$^-$ *tyrB*$^-$**	最大值（小时）	4.7 (0 小时)	0.66 (4 小时)	0.02 (6 小时)	0.07 (6 小时)
IAA89	起始值（0 小时）	4.7	ND	ND	ND
IPMS (*G462D*)	最终值（16 小时）	ND	ND	ND	0.02
***ilvE*$^+$ *tyrB*$^+$**	最大值（小时）	4.7 (0 小时)	0.54 (4 小时)	0.02 (6 小时)	0.02 (16 小时)

[a]**ND =** 未检测到

其他代谢产物例如乳酸和琥珀酸均未检测到。

[00221]当测定表达 WT *leuA* 基因产物的 Δ*ilvE* Δ*tyrB* 菌株生产的 3-甲基-1-丁醇时，测定结果与包含变异 IPMS 的菌株相似。3-甲基-1-丁醇的最终累积浓度为 553mg/L，而异丁醇仅为 42mg/L（图 49）。这相当于大于 13:1 的 3-甲基-1-丁醇和异丁醇的产物分配比。

[00222]本发明提供了生产高级醇的合成方法，此醇可以用作下一代生物燃料。本文提供的实施例以大肠杆菌为宿主细胞，对其进行代谢修饰，使其包含重组生物合成途径。但是，应当了解，其他微生物如酿酒酵母也可提供合适的宿主细胞，使其包含重组生物合成途径。这些宿主微生物生长速度快且均为兼性厌氧生物，这就保证了用于大规模生产、灵活经济的工艺设计的可行性。但是，从其他微生物引入非自然途径这一做法具有一定的缺点：异源途径的表达可能导致代谢失衡，异源代谢产物的累积可能引起细胞毒性。例如，丙酮丁醇梭菌的正丁醇生产途径具有三种带有辅酶 A（CoA）的中间代谢产物。这种途径的过度表达会导致游离 CoA 库被耗尽，从而，扰乱大肠杆菌的代谢。为了实现目

标外源产物的高产，需要寻找与宿主相容的途径。本文提供的修饰微生物利用了大肠杆菌的现有代谢能力，以及降解 2-酮酸的埃利希途径的最后两步反应的一系列底物，而不是直接将 1-丁醇生产的共同途径直接转入非自然宿主大肠杆菌。

[00223] 2-酮酸是氨基酸生物合成的中间产物。这些代谢物可通过具有广泛底物范围的 2-酮酸脱羧酶（KDC）转化为乙醛，然后，通过醇脱氢酶（ADH）转化为醇。采用这种策略需要两个非天然步骤，即通过将氨基酸生物合成途径的中间代谢物转入醇生产途径来生产生物燃料（图 1A）。氨基酸生物合成途径产生各种 2-酮酸（图 1B）。在当今研究中，有六种不同的 2-酮酸用于醇生产。异亮氨酸生物合成途径产生 2-酮丁酸和 2-酮基-3-甲基-戊酸，它们可分别转化为正丙醇和 2-甲基-1-丁醇。缬氨酸生物合成途径产生 2-酮基-异戊酸，此化合物是异丁醇的前体。亮氨酸生物合成途径产生 2-酮基-4-甲基-戊酸，此化合物是 3-甲基-1-丁醇的底物。苯丙氨酸生物合成途径产生苯丙酮酸，此化合物可生成 2-苯基乙醇。正缬氨酸生物合成途径是亮氨酸生物合成的副反应，产生正丁醇的底物即 2-酮基戊酸。

[00224] 2-酮酸脱羧酶活性可由以下基因之一产生： 来自*酿酒酵母*的 *PDC6*; 来自*乳酸乳球菌*的 *kivd*; 来自*酿酒酵母*的 *THI3* (α-*酮异己酸脱羧酶*)以及来自*丙酮丁醇梭菌*的 pdc。醇脱氢酶（Adh）活性由来自*酿酒酵母*的 *ADH2* 提供。

[00225] 缬氨酸由两个丙酮酸分子合成，其反应途径有四步且由四个酶催化，包括 AHAS（*ilvBN* 基因产物）、还原异构酶（*ilvC* 基因产物）、二羟酸脱水酶（*ilvD* 基因产物）和转氨酶 B（*ilvE* 基因产物）。与其他微生物一样，这些酶也能催化从丙酮酸和 2-酮基丙酸生成 L-异亮氨酸的合成反应。后者是通过苏氨酸脱水酶（*ilvA* 基因产物）由 L-苏氨酸生成。AHAS 是支链氨基酸合成的关键酶。缬氨酸引起 AHAS I 和 AHAS III 的反馈抑制，前者由 *ilvBN* 编码，后者由乙酰羟酸合成酶操纵子 *ilvIH* 编码，研究显示其小调控亚基 IlvN 和 IlvH 是对缬氨酸的敏感性所必需的。

[00226] 此醇类生产策略中的一个酶是 KDC，这种酶在植物、酵母和真菌中常见，但是在细菌中比较罕见。产生的乙醛可由 Adh 转化为醇，Adh 存在于许多微生物中。一些 KDCs 具有广泛的底物范围，另外一些则比较特异。为了测试内源 2-酮酸作为大肠杆菌 KDC 底物的能力，用五种 KDC 及酿酒酵母的醇脱氢酶 2（Adh2）进行了过度生产，这五种 KDC 包括来自*酿酒酵母*的 Pdc6、

Aro10、Thi3；*来自乳酸乳球菌的* Kivd；*以及来自丙酮丁醇梭菌的* Pdc。*表达这些外源基因的大肠杆菌培养物在含有* 0.2 M *葡萄糖的基本培养基中生长。* GC-MS *分析（见表* 13）*显示，表达* *kivd* *或* *AR010* *的菌株产生了所有预期的醇。酿酒酵母* *PDC6* *和丙酮丁醇梭菌* *pdc* *产生的种类没有这么多，而酿酒酵母* *THI3* *未表现出预期的活性。*

[00227]表 13 显示了用大肠杆菌 KDC 和 ADH 产生的以下醇。

KDC	Kivd	ARO10	PDC6	THI3	Pdc(C.A.)
质粒	pSA55	pSA56	pSA49	pSA57	pSA59
产物（μM）					
正丙醇	520	290	125	未检测到	未检测到
异丁醇	5242	2094	260	未检测到	75
正丁醇	220	95	未检测到	未检测到	未检测到
2-甲基-1-丁醇	766	652	56	未检测到	未检测到
3-甲基-1-丁醇	1495	1099	92	未检测到	未检测到
2-苯基乙醇	324	469	未检测到	未检测到	175

[00228]结果检测到微量的醛，这说明 Adh2 具有足够的活性。这个结果证实了 Kivd 是一种具有活性且功能全面的脱羧酶，因此，适于本发明的目的。

[00229]向表达 *kivd* 的大肠杆菌培养物中加入各种 2-酮酸（见表 12），证实了相应醇类的特异性产量增加 2-23 倍。2-酮酸的补充也显著降低了其他醇类的产量。结果表明，增加 2-酮酸的通量有利于提高醇的生产能力以及醇生产的特异性。

[00230]表 13 显示在补充以下 2-酮酸时的醇产量

	2-酮丁酸	2-酮基-异戊酸	2-酮戊酸	2-酮基-3-甲基戊酸	2-酮基-4-甲基戊酸	苯丙酮酸
产物（μM）						
正丙醇	2138	未检测到	未检测到	未检测到	未检测到	8
异丁醇	98	10016	未检测到	未检测到	未检测到	64
正丁醇	492	未检测到	3926	未检测到	未检测到	23
2-甲基-1-丁醇	1315	未检测到	未检测到	5284	未检测到	未检测到
3-甲基-1-丁醇	未检测到	未检测到	52	未检测到	3756	105
2-苯基乙醇	26	109	66	未检测到	未检测到	7269

[00231]对现有的大肠杆菌代谢途径进行基因改造，以提高特定 2-酮酸的产量，从而生产所需的醇。为了生产异丁醇，对 *ilvIH*CD 基因进行扩增，以增强 2-酮异戊酸的生物合成（图 1B）。大肠杆菌的 *ilvIH* 操纵子编码乙酰羟酸合成酶，这是异亮氨酸、缬氨酸和亮氨酸生物合成途径中的第一个酶。乙酰羟酸合成酶 III 同工酶由两个亚基即 *ilvI* 基因产物（乙酰乳酸合成酶 III 的大亚基）和 *ilvH* 基因产物（乙酰乳酸合成酶的小亚基）组成，其能够催化大肠杆菌 K-12 中的异亮氨酸、亮氨酸及缬氨酸的生物合成的第一步共同反应。大肠杆菌的 *ilvIH* 操纵子编码乙酰羟酸合成酶，这是异亮氨酸、缬氨酸和亮氨酸生物合成途径中的第一个酶。大肠杆菌的 *ilvC* 基因编码乙酰羟酸还原异构酶，这是平行进行的异亮氨酸-缬氨酸生物合成途径中的第二个酶。大肠杆菌的 ilvD 基因编码二羟酸脱水酶，这是异亮氨酸-缬氨酸生物合成途径中的第三个酶。

[00232]在质粒中构建由 $P_L$lacO1 控制的编码 *ilvI*、*ilvH*、*ilvC* 和 *ilvD* 基因的操纵子。然后，将扩增的 Ilv 途径与合成醇的生产途径合并（Kivd 和 Adh2）以实现异丁醇生产。作为 *ilvIHCD* 途径的表达结果，此菌株产生了 23mM 异丁醇，此产量比不具备 *ilvIHCD* 途径过度表达的菌株的产量高 5 倍（见下表 14 和图 2A）。

[00233]表 14 显示通过 ilvIHCD 途径及其过度表达的醇生产，如下：

KDC	kivd	kivd	kivd	kivd	kivd	kivd
菌株	JCL16	JCL16	JCL16	JCL16	JCL88	JCL88
质粒	pSA55	pSA55	pSA54 pSA55	pSA54 pSA55	pSA54 pSA55	pSA54 pSA55
时间（小时）	16	24	16	24	16	24
产物（μM）						
乙醇	1277	1744	2047	2228	850	1430
正丙醇	321	420	612	808	210	493
异丁醇	**2069**	**2172**	**12979**	**20813**	**13794**	**30910**
正丁醇	67	112	146	198	89	127
2-甲基-1-丁醇	223	629	262	345	195	239
3-甲基-1-丁醇	1096	1696	1103	1577	727	1333
2-苯基乙醇	545	1220	282	425	454	820

[00234]结果显示此合成途径有效，并且能够为高效的异丁醇生产提供所需的2-酮异戊酸。为了进一步提高异丁醇的产量，缺失一个或多个有助于副产物生成的基因，其中包括 *adh*、*ldh*、*frd*、*fnr*、*pflB* 和/或 *pta* 。这些缺失会提高可供 *ilvIHCD* 途径使用的丙酮酸水平。事实上，此菌株产出了 30 mM 异丁醇，说明这些缺失有利于异丁醇的生产。此外，此菌株将葡萄糖转化为异丁醇，产率为在 16 到 24 小时内每克葡萄糖产生 0.21g 异丁醇（图 2A，右图）。图 42 所示的其他数据还说明，这些缺失也能增加可供乙酰乳酸合成酶途径使用的丙酮酸水平，从而影响异丁醇的生产，使 64 小时后异丁醇的产量从 JCL16 的 4.5 g/L（61 mM)增加到 JCL260 的 13.2 g/L（356 mM）。结果说明了此策略的潜力，因为产率已达到理论最大值的 50%，并且并未对反应途径和生产条件进行详细优化。这种高产率得益于合成途径与宿主细胞生理特性的完全相容。下表中提供了菌株 SA237 的其他产率数据:

时间（小时-小时）	产率（g/g）	时间（小时）	(g/L)
0-16	~0.32	0	0.00
16-40	~0.4	16	~7.172
40-64	~0.36	40	~12.01
64-88	~0.32	64	~17.11
88-112	~0.33	88	~19.41
0-40	~0.34	112	~21.89
0-64	~0.35		
0-88	~0.34		
0-112	~0.34		

30°C
M9+0.5% YE
0.1mM IPTG

0-112小时内乙醇的产量为 0.0037 g/g(240 mg/L)。

	异丁醇（mM）					
时间	#1	#2	#3	平均值	标准差	异丁醇(g/L)
0						
16	98.45	108.29	83.56	96.77	12.45	7.172345
40	158.21	173.42	154.3	161.98	10.10	12.00571
64	225.54	253.33	213.67	230.85	20.36	17.11035
88	256.23	281.21	248.11	261.85	17.25	19.40832
112	296.43	312.56	277.12	295.37	17.74	21.89282

时间	葡萄糖（g/L）				
0	36	36	36	36.00	
16	13.2	12.5	14.1	13.27	0.80
40	1.2	0.9	1.5	1.20	0.30
40.1	31.2	30.9	31.5	31.20	
64	17.4	15.6	18.9	17.30	1.65
88	10.3	7.8	12.5	10.20	2.35
112	2.8	1.4	4.1	2.77	1.35

起始的 36 g/L 葡萄糖是 T=0 时的值，在 40.1 小时时，再向培养基中添加 30g 葡萄糖。不同小时数的葡萄糖值代表相应时间点的葡萄糖浓度 (g/L)（例如，在 16 小时，约 22.8g 的葡萄糖被代谢掉）。

[00235] 类似的策略被用于制备生产正丁醇的微生物。正丁醇在其天然微生物中的生产途径被用于发酵培养，此途径中的许多酶都是氧气敏感型及 CoA 依赖型。数据显示，通过在不具有正丁醇发酵途径的大肠杆菌中过度表达 *kivd* 或 *ARO10*，此细胞能够在非发酵培养中由葡萄糖生产少量正丁醇（见表 13），从而说明此细胞中存在相应的 2-酮酸前体即 2-酮戊酸。但是，2-酮戊酸不是大肠杆菌的常见代谢物。为了增加 2-酮戊酸合成量，利用具有广泛底物特异性的 *leuABCD* 途径，此途径的天然底物是 2-酮异戊酸（图 1B）。通过利用较小的底物如 2-酮丁酸 - 此底物的甲基数比 2-酮异戊酸少一个（图 1A），可按照与亮氨酸生物合成中类似的转化方式合成 2-酮戊酸。2-酮丁酸可由 L-苏氨酸在苏氨酸脱水酶的催化下生成，此酶由 *ilvA* 基因编码，或采用在问号钩端螺旋体血清型和詹氏甲烷暖球菌中发现的另一种途径，以乙酸和丙酮酸为起始物进行合成。在后一种途径中，2-酮丁酸由柠苹酸 通过异丙基苹果酸异构酶（LeuCD）和 β-异丙基苹果酸脱氢酶（LeuB）生成。

[00236] 为生产正丁醇，我们构建了一种在 PLlacO1 控制下编码 *ilvA-leuABCD* 途径的操纵子。具有 *ilvA-leuABCD* 途径的菌株产生了 0.6 mM 正丁醇，这一数量比不具有 *ilvA-leuABCD* 途径过度表达的菌株增加了 3 倍（见下表 15 和图 2B）。

[00237] **表 15 显示通过苏氨酸途径过度表达生产醇**

KDC	kivd	kivd	kivd	kivd
质粒	pSA55	pSA55 pSA62	pSA55 pSA62	pSA54 pSA62
菌株	JCL16	JCL16	JCL16	SA203
0.8% L-苏氨酸	-	-	+	+
产物（μM）				
乙醇	2450	2343	3243	1493
正丙醇	520	1356	7592	9849
异丁醇	2242	4322	1061	未检测到
正丁醇	220	583	3157	9232
2-甲基-1-丁醇	766	1444	2002	未检测到
3-甲基-1-丁醇	1495	4074	1349	未检测到
2-苯基乙醇	324	358	269	524

[00238] 此外，当在培养基中补加 0.8% 的 L-苏氨酸时，观察到正丁醇的产量会明显增加到 3.2mM，说明可通过 IlvA 介导的反应（见图 2C）由 L-苏氨酸生成 2-酮戊酸。

[00239] 为了进一步提高正丁醇产量，将 *ilvD* 基因缺失。此基因编码二羟酸脱水酶 - 一种既能生成 2-酮异戊酸（亮氨酸和缬氨酸的一种前体）又能生成 2-酮基-3-甲基戊酸（异亮氨酸的一种前体）的酶。由于以下两个原因此基因的缺失是有益的：首先，*ilvD* 的缺失消除了 *leuABCD* 途径的天然底物 2-酮异戊酸，从而，减少了竞争性底物抑制；其次，*ilvD* 的缺失消除了 Kivd 的竞争性底物 2-酮基-3-甲基戊酸和 2-酮基-4-甲基-戊酸。正如预期的那样，*ilvD* 的缺失提高了正丁醇产量（图 2C）。

[00240] 由于*大肠杆菌*中 L-苏氨酸的高产已用于商业化生产，所以可以利用上述策略对生产苏氨酸的菌株进行修饰，以制备出能产生正丁醇的菌株。为了进一步改善，有必要增加 *leuABCD* 途径对非天然底物 2-酮丁酸的活性，并提高 Kivd 对 2-酮戊酸的特异性。由于 2-酮丁酸也是正丙醇的底物（图 1B），因此，

增加可获得的 2-酮丁酸量也能提高正丙醇的产量（图 2B 和图 2C，右图）。从而，增加 *leuABCD* 活性及 KDC 的特异性对正丁醇的高效生产至关重要。

[00241] 正丁醇和正丙醇生产菌 CRS-BuOH23 产生的醇状况和产量如下表所示：

时间段	丙醇	丁醇	合计	乙醇
0-20	0.037	0.015	0.053	0.004
20-28	0.054	0.046	0.1	0.004
28-52	0.033	0.042	0.075	0.003
52-76	0.061	0.066	0.127	0.009
76-100	-0.016	-0.014	-0.03	0.003
0-76	0.043	0.038	0.081	0.005
0-28	0.042	0.024	0.066	0.004
20-76	0.046	0.05	0.097	0.005
28-76	0.044	0.052	0.096	0.005

合计=正丙醇 + 正丁醇；最终乙醇： 116.2 mg/L

[00242] 3-甲基-1-丁醇生产菌产生的醇状况和产量如下表所示：

3-甲基-1-丁醇生产菌		
时间（小时-小时）	3MB 产率（g/g）	醇总产率（g/g）
0-4	0.000	0.000
4-6	0.040	0.040
6-8	0.180	0.180
8-12	0.160	0.170
12-20	0.120	0.160
20-28	-0.050	-0.010
0-28	0.100	0.120

[00243] 两种不同 2-甲基-1-丁醇菌产生的醇状况和产率如下表所示：

2-甲基-1-丁醇菌 1（AFC-2MB-01BW25113: pAFC46: kivd(乳酸乳球菌) AHD2(酿酒酵母) ilvA(谷氨酸棒杆菌) pAFC3: ilvGM(鼠伤寒沙门氏菌) ilvCD(大肠杆菌) PCS49: thrA^{FBR}BC (大肠杆菌))				
时间（小时-小时）	2MB产率（g/g）	醇总产率（g/g）	乙醇产率	乙醇 mg/L
0-9	0.026	0.081	0.063	230.6
9-15	0.050	0.171	0.019	156.1
15-18	0.043	0.164	0.016	196.0
18-21	0.038	0.163	0.016	226.2
21-24	0.037	0.157	0.017	271.6
24-27	0.034	0.144	0.017	288.7
27-33	0.046	0.146	0.022	376.4
33-39	0.031	0.123	0.018	329.5
2-甲基-1-丁醇菌 2 (AFC-2MB-02: CRS22: BW25113 ΔmetA Δtdh pAFC46: kivd(乳酸乳球菌) AHD2(酿酒酵母) ilvA(谷氨酸棒杆菌) pAFC3: ilvGM(鼠伤寒沙门氏菌) ilvCD(大肠杆菌) PCS49: thrAFBRBC (大肠杆菌))				
时间（小时-小时）	2MB 产率	醇总产率	乙醇产率	乙醇 mg/L
0-9	0.183	0.305	0.243	301.7
9-15	0.089	0.161	0.014	133.4
15-18	0.083	0.164	0.014	166.5
18-21	0.080	0.169	0.016	214.8
21-24	0.073	0.161	0.016	256.7
24-27	0.065	0.149	0.017	315.1
27-33	0.058	0.138	0.017	326.5
33-39	0.064	0.151	0.018	344.8
醇总产率=2-甲基-1-丁醇、3-甲基-1-丁醇、异丁醇、正丁醇和正丙醇				

[00244] 大肠杆菌等非天然宿主缺乏对高级醇的耐受性。对微生物而言，异丁醇和正丁醇一样有毒，天然的正丁醇生产菌可以耐受浓度为 2% 的正丁醇。为了显示耐受性改善的潜在可能，采用连续转移的体外进化法富集能够提高异丙醇耐受性的突变体菌株。数据显示，1.5%的异丁醇抑制野生型大肠杆菌(JCL16)。但是，经过 5 轮逐步提高异丁醇浓度的继代转移后，在 2%的异丁醇中发现了生长的突变体(图 3)。这种水平的溶剂耐受性与正丁醇的天然生产菌相似或更好，说明大肠杆菌能够适应高浓度的长链醇。可以利用其它策略如 gTME ，以进一步改善耐受性。

[00245] 本文所述的策略用于用大肠杆菌和其他微生物生产生物燃料。本策略利用了氨基酸生产技术并驱动氨基酸代谢中间物向 2-酮酸降解途径进行，以用于醇类的生产。本策略避免了 CoA 介导的化学反应，此反应存在于天然菌生产正丁醇的途径中，并使工业规模的其他高级复杂醇合成成为可能。生产其他醇的具体策略可根据本文提供的合成路径方便地设计。例如，可通过改变苯丙氨酸的合成途径实现 2-苯基乙醇的生产，在大肠杆菌中此途径已被有效扩展。这些策略也可轻松应用于酵母或其他工业微生物。对于异丁醇的合成，其完整的合成途径不依赖于 CoA，且仅需丙酮酸作为前体。将这些策略应用于酵母时，这种特性避免了乙酰-CoA 的线粒体区室化问题。这些用于异丁醇和正丁醇生产的策略提供的理论产率（0.41g/g 葡萄糖）与天然正丁醇产生菌的相同。

[00246] 利用本发明所属领域的技术人员熟知的方法来完成基因缺失。简而言之，即使用 BW25113 ($rrnB_{T14}$ $\Delta lacZ_{WJ16}$ *hsdR514* $\Delta araBAD_{AH33}$ $\Delta rhaBAD_{LD78}$) 作为亲本（例如，野生型）微生物。如文献中所述，缺失 *adhE*、*ldhA*、*frdBC*、*fnr*、*pflB* 和 *ilvD* 序列（Datsenko 和 Wanner, Proc Natl Acad Sci U.S.A 97:6640 (2000)）。以 JW2294（Baba 等人，Mol Syst Biol 2:E1–E11 (2006)）作为供体，通过 P1 转导缺失 *pta* 序列。从 XL-1 BLUE 转移 F′ 以提供 *lacIq*。本研究使用的菌种列于表 16 中。具体地说，为了缺失 *pta*，去除了第 2,412,772 - 2,414,893 个核苷酸； 为了缺失 *frdBC*，去除了 第 4,377,400 - 4,378,540 个核苷酸；为了缺失 *adhE*，去除了第 1,294,669 - 1,297,344 个核苷酸；为了缺失 *ldhA*，去除了第 1,439,878 - 1,440,867 个核苷酸；为了缺失 fnr，去除了第 1,396,798 - 1,397,550 个核苷酸；为了缺失 *pflB*，去除了第 950,508 - 952,784 个核苷酸。

[00247]表 16

菌株	相关基因型
BW25113	*rrnB*$_{T14}$ Δ*lacZ*WJ16 *hsdR514* Δ*araBAD*$_{AH33}$ Δ*rhaBAD*$_{LD78}$
JCL16	BW25113, F' (traD36, proAB+, lacI^q ZDM15)
JCL88	JCL16, Δ*adhE*, Δ*ldhA-fnr*, Δ*frdBC*, Δ*pta*
JCL93	JCL16, Δ*adhE*, Δ*ldhA*, Δ*frdBC*
SA203	JCL16 Δ*ilvD*
JCL260	JCL16, Δ*adhE*, Δ*ldhA-fnr*, Δ*frdBC*, Δ*pflB*, Δ*pta*
JCL167	JCL16, Δ*adhE*, Δ*ldhA-fnr*, Δ*frdBC*
JCL274	JCL16, Δ*adhE*, Δ*ldhA*, Δ*frdBC*, Δ*pta*
JCL168	JCL16, Δ*adhE*, Δ*ldhA-fnr*, Δ*frdBC*, Δ*pflB*
JCL171	JCL16, Δ*adhE*, Δ*ldhA*, Δ*frdBC*, Δ*pflB*, Δ*pta*

[00248]pSA46 包含 *PDC6* 序列。以 A65 和 A66 作为引物，（ATCC）的基因组 DNA 为 PCR 模板扩增 *PDC6* 序列（见表 17），用 Acc65I 和 SphI 消化 PCR 产物。然后，克隆到用同样的酶切割过的 pZE12-luc (3) 中。

[00249]pSA49 包含 *ADH2* 序列。以 A67 和 A68 为引物，以酿酒酵母（ATCC）的基因组 DNA 为 PCR 模板来克隆 *ADH2*（见表 17），用 SphI 和 XbaI 消化 PCR 产物。然后，克隆到用同样的酶切割过的 pSA46 中。

[00250]通过用 p15A 置换 pSA49 的复制起点来构建 pSA53，用 SacI 和 AvrII 消化 pZA31-luc。较短的片段经过纯化克隆到用同样的酶切割过的 pSA49 质粒中。

[00251]pSA55 包含 *kivd* 序列，此序列通过以 A96 和 A97 为引物、以乳酸乳球菌基因组 DNA 为模板进行 PCR 反应获取 (见表 17)。然后，用 Acc65I 和 SphI 消化 PCR 产物，再克隆到用同样的酶切割过的 pSA49 中。

[00252]pSA56 包含 ARO10 序列。以 A98 和 A99 为引物，以酿酒酵母（ATCC）的基因组 DNA 为 PCR 模板扩增 ARO10 序列（见表 17）。用 Acc65I 和 SphI 消化 PCR 产物，然后克隆到用同样的酶切割过的 pSA49 中。

[00253]pSA57 包含 *THI3* 序列。以 A100 和 A101 为引物，以酿酒酵母（ATCC）的基因组 DNA 为 PCR 模板扩增 *THI3* 序列（见表 17）。用 Acc65I 消化 PCR 产物。用 SphI 消化 pSA49，用 Klenow 片段平端化，然后，用 Acc65I 消化。这条链与 PCR 产物相连。

[00254]pSA58 包含从丙酮丁醇梭菌获得的 *pdc* 序列。以 A102 和 A103 为引物，以基因组 DNA 为 PCR 模板扩增获得 *pdc* 序列（见表 17），用 Acc65I 和 SphI 消化 PCR 产物。然后，克隆到用同样的酶切割过的 pSA49 中。

[00255]通过用 $P_L$lacO1 取代 pZE21-MCS1 的 $P_L$tetO1 序列，以构建 pSA40。然后，用 AatII 和 Acc65I 消化 pZE12-luc，较短片段经过纯化并克隆到经相同酶切割过的 pZE21-MCS1 质粒中。

[00256]pSA45 包含 *ilvC* 序列。以A71 和 A72 为引物，以大肠杆菌 MG1655 的基因组 DNA 为模板，通过 PCR 反应获得 *ilvC* 序列（见表 17）。然后，用 SalI 和 XmaI 消化 PCR 产物，再克隆到用同样的酶切割过的 pSA40 中。

[00257]pSA47 包含 *ilvD* 序列。以 A74 和 A84 为引物，以大肠杆菌 MG1655 的基因组 DNA 为模板，通过PCR 反应获得 *ilvD* 序列（见表 17）。然后，用 BspEI 和 MluI 消化 PCR 产物，再克隆到用 SalI 和 MluI 酶切割过的 pSA45 中。

[00258]pSA51 包含 *ilvI* 和 *ilvH* 序列。以 A70 和 A83 为引物，以大肠杆菌 MG1655 的基因组DNA 为 PCR 模板扩增获得该序列（见表 17）。然后，用 BsaI 和 SalI 消化 PCR 产物，再克隆到用 Acc65I 和 SalI 酶切割过的 pSA40 中。

[00259]pSA52 包含 *ilvH* 下游的 *ilvC* 和 *ilvD* 序列。用 SalI 和 MluI 消化 pSA47。较短的片段经过纯化，克隆到用同样的酶切割过的 pSA51 质粒中。

[00260] 通过将 用 SacI 和 AvrII 消化的 pZA31-luc 的 p15A 的复制起点转移到 pSA52 质粒中构建 pSA54。

[00261]pSA59 包含 *leuABCD* 序列。以 A106 和 A109 为引物，以大肠杆菌 MG1655 的基因组 DNA 为 PCR 模板扩增该序列（见表 17），用 SalI 和 BglII 消化 PCR 产物。然后，克隆到用 SalI 和 BamHI 切割过的 pSA40 中。

[00262]pSA60 包含 *ilvA* 序列。以 A104 和 A105 为引物，以大肠杆菌 MG1655 的基因组 DNA 为 PCR 模板扩增该序列（见表 17）。然后，用 Acc65I 和 XhoI 消化 PCR 产物，再克隆到用 Acc65I 和 SalI 切割过的 pSA59 中。

[00263]将 pZA31-luc 中的 p15A 复制起点克隆到 pSA60 质粒中，以构建 pSA62。对本文提供的代表性质粒的部分描述列于表 18 中。

[00264]pSA66 包含 *alsS* 序列的 3' 片段。以 A123 和 A124 为引物，以枯草芽孢杆菌的基因组 DNA 为模板，通过PCR 反应获得 alsS 序列（见表 17）。然后，用 Acc65I 和 SalI 消化 PCR 产物，再克隆到用同样的酶切割过的 pSA40 中。

[00265] pSA67 包含 *alsS* 序列。以 A125 和 126 为引物，以枯草芽孢杆菌的基因组 DNA 为模板，通过 PCR 反应获得 alsS 序列的 5' 片段（见表 17）。然后，用 BsrGI 和 XbaI 消化 PCR 产物，再克隆到用 Acc65I 和 XbaI 切割过的 pSA66 中。

[00266] pSA68 包含 alsS 下游的 *ilvC* 和 *ilvD* 序列。用 SalI 和 MluI 消化 pSA47。较短的片段经过纯化，克隆到用同样的酶切割过的 pSA67 质粒中。

[00267] 用 SacI 和 AvrII 消化 pZA31-luc 中的复制起点 p15A，然后，转移至 pSA68 质粒中，以构建 pSA69。

[00268] 表 18

质粒	相关基因型
pSA54	Origin of p15A; $P_L$lacO1::*ilvIHCD*; kan^r
pSA55	Origin of ColE1; $P_L$lacO1::*kivd-ADH2*; amp^r
pSA69	Origin of p15A; $P_L$lacO1::*alsS-ilvCD*; kan^r

[00269] 用于 PCR 和克隆过程的低聚核苷酸代表性清单见表 17。应当了解，可以根据用于 PCR 和/或克隆过程的特定序列对此代表性性低聚核苷酸进行修饰。

[00270] 表 17

名称	序列	（序列 ID 号：）
A65	CGAGCGGTACCATGTCTGAAATTACTCTTGGAAAAT	(1)
A66	GCCTGCGCATGCTTATTGTTTGGCATTTGTAGCGGCA	(2)
A67	GCCTGCGCATGCAGGAGATATACCATGTCTATTCCAGAA ACTCAAAAAG	(3)
A68	GCTCTAGATTATTTAGAAGTGTCAACAACGTAT	(4)
A70	ACGCAGTCGACTCAACGCATTATTTTATCGCCGCGC	(5)
A71	ACGCAGTCGACGAGGAATCACCATGGCTAACTACTT	(6)
A72	AATAACCCGGGTTAACCCGCAACAGCAATACGTTTC	(7)
A74	CGAGCACGCGTTTAACCCCCCAGTTTCGATTTATCG	(8)
A83	GCCACCGGTCTCCGTACCATGGAGATGTTGTCTGGAGCC GAGA	(9)
A84	GGCTCCGGAAGGAGATATACCATGCCTAAGTACCGTTCC GCCACCA	(10)
A96	CGAGCGGTACCATGTATACAGTAGGAGATTACCTAT	(11)
A97	GCCTGCGCATGCTTATGATTTATTTTGTTCAGCAAAT	(12)
A98	CGAGCGGTACCATGGCACCTGTTACAATTGAAAAGT	(13)
A99	GCCTGCGCATGCCTATTTTTTATTTCTTTTAAGTGCCGC	(14)
A100	CGAGCGGTACCATGAATTCTAGCTATACACAGAGAT	(15)
A101	GGTCAGTATCCAACTTGATTTTTTTTTAGAAG	(16)

A102	CGAGCGGTACCATGAAGAGTGAATACACAATTGGAAG	(17)
A103	GCCTGCGCATGCCTAATTATTTTGATTTGCAAAACGT	(18)
A104	CGAGCGGTACCATGGCTGACTCGCAACCCCTGTCCG	(19)
A105	CCGCTCGAGCTAACCCGCCAAAAAGAACCTGAAC	(20)
A106	ACGCAGTCGACAAGAGACAAGGACCCAAACCATGAGC CAG	(21)
A109	GGAAGATCTTTAATTCATAAACGCAGGTTGTTTTGC	(22)
A123	GCCACCCGTCTCCGTACCATGTTGACAAAAGCAACAAA AGAAC	(23)
A124	ACGCAGTCGACCTAGAGAGCTTTCGTTTTCATGAGT	(24)
A125	CGAGCTGTACAATGTTGACAAAAGCAACAAAAGAAC	(25)
A126	TCTCTAGAAAGGGTACCGGCAGCTTG	(26)

[00271] 在异丁醇生产的代表性的过程中，用 pSA69 和 pSA55 质粒对宿主菌株 JCL260 进行转化。转化涂板培养基为 LB + 0.5% 葡萄糖 + 50 mg/L 卡那霉素 + 200 mg/L 氨苄西林。将新鲜转化子接种到 2mL 培养基中，培养基的成分为：LB + 0.5% 葡萄糖 + 50 mg/L 卡那霉素 + 200 mg/L 氨苄西林。然后，在 37°C 和 290 rpm 转速下对培养物进行过夜培养。

[00272] 从过夜培养物中取 200 μL接种到置于250mL 螺旋盖烧瓶的 20mL培养基 M9 + 1x 微量金属混合物 A5 + 7.2% 葡萄糖+ 3% 胰蛋白胨 + 30 mg/L 卡那霉素 + 100 mg/L 氨苄西林中。然后，将此培养物置于 37°C /260 rpm 下培养。当 OD600 在 1-2(~ 5hr) 之间时，用 0.1 mM IPTG 诱导培养物，再在 30°C/260rpm 下培养 24 小时。

[00273] 应当通过过滤法（硫胺素）或高压蒸汽消毒灭菌法对上述溶液分别灭菌。培养基采用过滤法灭菌。

[00274] 采用气相色谱法测定异丁醇的产量，测定结果见表 42-44。在图 42 中，培养方法如下：制备含有以下物质的 M9 培养基：7.2% 葡萄糖，0.5% 酵母提取物，100μg/mL 氨苄西林，30μg/mL 卡那霉素和 1000× 微量金属混合物 A5。各种包含 pSA55 和 pSA69 的敲除菌株在含有 3mL LB 培养基的试管内于 37°C 下在旋转摇床（250rpm）上过夜预培养。然后，按 1:100 的比例将此过夜的培养物稀释到装在 250ml 螺旋盖锥形瓶的 20mL 新鲜培养基中。细胞在 37°C 下生长至 OD_{600} 2.0 左右，然后，添加 0.1 mM IPTG。添加 IPTG 后，置于旋转摇床（250 rpm）上并在 30°C 下培养 64 小时。在 16 小时、40 小时和 64 小时时打开螺旋盖取样，用 GC/FID 分析培养样本。在图 43 中，培养过程如下：制备含有以下物质的 M9 培养基：3.6 % 葡萄糖，100μg/mL 氨苄

西林，30μg/mL 卡那霉素和 1000× 微量金属混合物 A5。SA237（JCL260，包含 pSA55 和 pSA69）在含有 3ml LB 培养基的试管内在 37°C 下在旋转摇床（250 rpm）上过夜预培养。然后，以 1:100 的比例将此过夜的培养物稀释到装在 250mL 螺旋盖锥形瓶（红色）或 250ml 锥形瓶（蓝色）的 20mL 新鲜培养基中。细胞在 37°C 下生长至 OD_{600} 0.3 左右。然后，添加 0.1mM IPTG。添加 IPTG 后，置于旋转摇床（250 rpm）上并在 30°C 下培养 72hr。在 24 小时、48 小时和 72 小时时对培养样本的异丁醇浓度、OD_{600} 和 pH 进行测定。在图 45 中，培养过程如下：制备用于培养的 M9 培养基，其含有： 3.6 % 葡萄糖，100μg/mL 氨苄西林，30μg/mL 卡那霉素和 1000×微量金属混合物 A5。再向此培养基中添加 0.8 % 酪蛋白氨基酸（红色）、2 % 胰蛋白胨（绿色）或 0.5 % 酵母提取物（蓝色）。SA237（JCL260，包含 pSA55 和 pSA69）在含有 3mL LB 培养基的试管内 37°C 下在旋转摇床上（250 rpm）过夜预培养。然后，以 1:100 的比例将此过夜的培养物稀释到装在 250mL 螺旋盖锥形瓶的 20mL 新鲜培养基中。细胞在 37°C 下生长至 OD_{600} 0.8 左右。然后，添加 0.1mM IPTG。添加 IPTG 后，置于旋转摇床（250 rpm）上并在 30°C 下培养 198 小时。在每个时间点测定培养物的异丁醇浓度、葡萄糖浓度、OD_{600} 和 pH。40 小时时，向所有培养物中添加 1mL 的 36% 葡萄糖。在 112 小时时，向含有酪蛋白氨基酸的培养物（红色）和含有胰蛋白胨的培养物（绿色）中加入 5mL 新鲜培养基；在 112 小时时，向含有酵母提取物的培养物（蓝色）添加 5mL 新鲜培养基和 2mL 的 36% 葡萄糖。

[00275] 上述实施例的目的在于向本发明所属领域技术人员完整公开和描述如何制造和使用公开内容中的设备、系统和方法的实施例，不限定发明者所考虑的发明范围。对本发明所属领域的技术人员显而易见的对实施本发明的上述方法的修改涵盖在下列权利要求的范围内。本说明书提到的所有专利和出版物表明了本发明所属领域的技术人员的当前技能水平。本发明引用的所有参考文献以引用方式并入本文中，引用程度如同单独、完整引用各参考文献。

[00276] 本发明已对一些实施方案进行了描述。然而，应当了解，可在不偏离本发明精神和范围情况下对其作出各种修改。因此，其他实施方案也在下述权利要求的范围内。

<110> 加利福尼亚大学董事会
詹姆士.C.里奥
渥美正太
凯文.M.史密斯
罗.普.克莱尔.沈
安东尼.F.卡恩
迈克尔.R.康诺

<120> 利用重组微生物生产生物燃料

<130> 0069697-000027

<140> 仍然未知
<141> 2008-02-08

<150> US 60/900,477
<151> 2007-02-09

<150> US 60/900,546
<151> 2007-02-09

<150> US 60/921,927
<151> 2007-04-04

<150> US 60/956,634
<151> 2007-08-17

<160> 88

<170> PatentIn version 3.4

<210> 1
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 1
cgagcggtac catgtctgaa attactcttg gaaaat 36

<210> 2
<211> 37
<212> DNA
<213> 人工序列

<220>

<223> 寡核苷酸引物序列

<400> 2
gcctgcgcat gcttattgtt tggcatttgt agcggca 37

<210> 3
<211> 49
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 3
gcctgcgcat gcaggagata taccatgtct attccagaaa ctcaaaaag 49

<210> 4
<211> 33
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 4
gctctagatt atttagaagt gtcaacaacg tat 33

<210> 5
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 5
acgcagtcga ctcaacgcat tattttatcg ccgcgc 36

<210> 6
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 6

acgcagtcga cgaggaatca ccatggctaa ctactt 36

<210> 7
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 7
aataacccgg gttaacccgc aacagcaata cgtttc 36

<210> 8
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 8
cgagcacgcg tttaaccccc cagtttcgat ttatcg 36

<210> 9
<211> 43
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 9
gccaccggtc tccgtaccat ggagatgttg tctggagccg aga 43

<210> 10
<211> 46
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 10
ggctccggaa ggagatatac catgcctaag taccgttccg ccacca 46

<210> 11
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 11
cgagcggtac catgtataca gtaggagatt acctat 36

<210> 12
<211> 37
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 12
gcctgcgcat gcttatgatt tattttgttc agcaaat 37

<210> 13
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 13
cgagcggtac catggcacct gttacaattg aaaagt 36

<210> 14
<211> 39
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 14
gcctgcgcat gcctattttt tatttctttt aagtgccgc 39

<210> 15
<211> 36
<212> DNA

<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 15
cgagcggtac catgaattct agctatacac agagat 36

<210> 16
<211> 32
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 16
ggtcagtatc caacttgatt ttttttaga ag 32

<210> 17
<211> 37
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 17
cgagcggtac catgaagagt gaatacacaa ttggaag 37

<210> 18
<211> 37
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 18
gcctgcgcat gcctaattat tttgatttgc aaaacgt 37

<210> 19
<211> 36
<212> DNA
<213> 人工序列

<220>

<223> 寡核苷酸引物序列

<400> 19
cgagcggtac catggctgac tcgcaacccc tgtccg 36

<210> 20
<211> 34
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 20
ccgctcgagc taacccgcca aaaagaacct gaac 34

<210> 21
<211> 40
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 21
acgcagtcga caagagacaa ggacccaaac catgagccag 40

<210> 22
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 22
ggaagatctt taattcataa acgcaggttg ttttgc 36

<210> 23
<211> 43
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 23

gccacccgtc tccgtaccat gttgacaaaa gcaacaaaag aac 43

<210> 24
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 24
acgcagtcga cctagagagc tttcgttttc atgagt 36

<210> 25
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 25
cgagctgtac aatgttgaca aaagcaacaa aagaac 36

<210> 26
<211> 26
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 26
tctctagaaa gggtaccggc agcttg 26

<210> 27
<211> 1647
<212> DNA
<213> 乳酸乳球菌

<220>
<221> CDS
<222> (1)..(1647)

<400> 27
atg tat aca gta gga gat tac cta tta gac cga tta cac gag tta gga 48
Met Tyr Thr Val Gly Asp Tyr Leu Leu Asp Arg Leu His Glu Leu Gly
1 5 10 15

```
att gaa gaa att ttt gga gtc cct gga gac tat aac tta caa ttt tta      96
Ile Glu Glu Ile Phe Gly Val Pro Gly Asp Tyr Asn Leu Gln Phe Leu
            20                  25                  30

gat caa att att tcc cgc aag gat atg aaa tgg gtc gga aat gct aat     144
Asp Gln Ile Ile Ser Arg Lys Asp Met Lys Trp Val Gly Asn Ala Asn
        35                  40                  45

gaa tta aat gct tca tat atg gct gat ggc tat gct cgt act aaa aaa     192
Glu Leu Asn Ala Ser Tyr Met Ala Asp Gly Tyr Ala Arg Thr Lys Lys
    50                  55                  60

gct gcc gca ttt ctt aca acc ttt gga gta ggt gaa ttg agt gca gtt     240
Ala Ala Ala Phe Leu Thr Thr Phe Gly Val Gly Glu Leu Ser Ala Val
65                  70                  75                  80

aat gga tta gca gga agt tac gcc gaa aat tta cca gta gta gaa ata     288
Asn Gly Leu Ala Gly Ser Tyr Ala Glu Asn Leu Pro Val Val Glu Ile
                85                  90                  95

gtg gga tca cct aca tca aaa gtt caa aat gaa gga aaa ttt gtt cat     336
Val Gly Ser Pro Thr Ser Lys Val Gln Asn Glu Gly Lys Phe Val His
            100                 105                 110

cat acg ctg gct gac ggt gat ttt aaa cac ttt atg aaa atg cac gaa     384
His Thr Leu Ala Asp Gly Asp Phe Lys His Phe Met Lys Met His Glu
        115                 120                 125

cct gtt aca gca gct cga act tta ctg aca gca gaa aat gca acc gtt     432
Pro Val Thr Ala Ala Arg Thr Leu Leu Thr Ala Glu Asn Ala Thr Val
    130                 135                 140

gaa att gac cga gta ctt tct gca cta tta aaa gaa aga aaa cct gtc     480
Glu Ile Asp Arg Val Leu Ser Ala Leu Leu Lys Glu Arg Lys Pro Val
145                 150                 155                 160

tat atc aac tta cca gtt gat gtt gct gct gca aaa gca gag aaa ccc     528
Tyr Ile Asn Leu Pro Val Asp Val Ala Ala Ala Lys Ala Glu Lys Pro
                165                 170                 175

tca ctc cct ttg aaa aaa gaa aac tca act tca aat aca agt gac caa     576
Ser Leu Pro Leu Lys Lys Glu Asn Ser Thr Ser Asn Thr Ser Asp Gln
            180                 185                 190

gag atc ttg aac aaa att caa gaa agc ttg aaa aat gcc aaa aaa cca     624
Glu Ile Leu Asn Lys Ile Gln Glu Ser Leu Lys Asn Ala Lys Lys Pro
        195                 200                 205

atc gtg att aca gga cat gaa ata att agt ttt ggc tta gaa aaa aca     672
Ile Val Ile Thr Gly His Glu Ile Ile Ser Phe Gly Leu Glu Lys Thr
    210                 215                 220
```

```
gtc tct caa ttt att tca aag aca aaa cta cct att acg aca tta aac      720
Val Ser Gln Phe Ile Ser Lys Thr Lys Leu Pro Ile Thr Thr Leu Asn
225                 230                 235                 240

ttt gga aaa agt tca gtt gat gaa gct ctc cct tca ttt tta gga atc      768
Phe Gly Lys Ser Ser Val Asp Glu Ala Leu Pro Ser Phe Leu Gly Ile
                245                 250                 255

tat aat ggt aaa ctc tca gag cct aat ctt aaa gaa ttc gtg gaa tca      816
Tyr Asn Gly Lys Leu Ser Glu Pro Asn Leu Lys Glu Phe Val Glu Ser
            260                 265                 270

gcc gac ttc atc ctg atg ctt gga gtt aaa ctc aca gac tct tca aca      864
Ala Asp Phe Ile Leu Met Leu Gly Val Lys Leu Thr Asp Ser Ser Thr
        275                 280                 285

gga gcc ttc act cat cat tta aat gaa aat aaa atg att tca ctg aat      912
Gly Ala Phe Thr His His Leu Asn Glu Asn Lys Met Ile Ser Leu Asn
    290                 295                 300

ata gat gaa gga aaa ata ttt aac gaa agc atc caa aat ttt gat ttt      960
Ile Asp Glu Gly Lys Ile Phe Asn Glu Ser Ile Gln Asn Phe Asp Phe
305                 310                 315                 320

gaa tcc ctc atc tcc tct ctc tta gac cta agc gaa ata gaa tac aaa     1008
Glu Ser Leu Ile Ser Ser Leu Leu Asp Leu Ser Glu Ile Glu Tyr Lys
                325                 330                 335

gga aaa tat atc gat aaa aag caa gaa gac ttt gtt cca tca aat gcg     1056
Gly Lys Tyr Ile Asp Lys Lys Gln Glu Asp Phe Val Pro Ser Asn Ala
            340                 345                 350

ctt tta tca caa gac cgc cta tgg caa gca gtt gaa aac cta act caa     1104
Leu Leu Ser Gln Asp Arg Leu Trp Gln Ala Val Glu Asn Leu Thr Gln
        355                 360                 365

agc aat gaa aca atc gtt gct gaa caa ggg aca tca ttc ttt ggc gct     1152
Ser Asn Glu Thr Ile Val Ala Glu Gln Gly Thr Ser Phe Phe Gly Ala
    370                 375                 380

tca tca att ttc tta aaa cca aag agt cat ttt att ggt caa ccc tta     1200
Ser Ser Ile Phe Leu Lys Pro Lys Ser His Phe Ile Gly Gln Pro Leu
385                 390                 395                 400

tgg gga tca att gga tat aca ttc cca gca gca tta gga agc caa att     1248
Trp Gly Ser Ile Gly Tyr Thr Phe Pro Ala Ala Leu Gly Ser Gln Ile
                405                 410                 415

gca gat aaa gaa agc aga cac ctt tta ttt att ggt gat ggt tca ctt     1296
Ala Asp Lys Glu Ser Arg His Leu Leu Phe Ile Gly Asp Gly Ser Leu
            420                 425                 430
```

caa ctt acg gtg caa gaa tta gga tta gca atc aga gaa aaa att aat 1344
Gln Leu Thr Val Gln Glu Leu Gly Leu Ala Ile Arg Glu Lys Ile Asn
435 440 445

cca att tgc ttt att atc aat aat gat ggt tat aca gtc gaa aga gaa 1392
Pro Ile Cys Phe Ile Ile Asn Asn Asp Gly Tyr Thr Val Glu Arg Glu
450 455 460

att cat gga cca aat caa agc tac aat gat att cca atg tgg aat tac 1440
Ile His Gly Pro Asn Gln Ser Tyr Asn Asp Ile Pro Met Trp Asn Tyr
465 470 475 480

tca aaa tta cca gaa tca ttt gga gca aca gaa gaa cga gta gtc tcg 1488
Ser Lys Leu Pro Glu Ser Phe Gly Ala Thr Glu Glu Arg Val Val Ser
485 490 495

aaa atc gtt aga act gaa aat gaa ttt gtg tct gtc atg aaa gaa gct 1536
Lys Ile Val Arg Thr Glu Asn Glu Phe Val Ser Val Met Lys Glu Ala
500 505 510

caa gca gat cca aat aga atg tac tgg att gag tta att ttg gca aaa 1584
Gln Ala Asp Pro Asn Arg Met Tyr Trp Ile Glu Leu Ile Leu Ala Lys
515 520 525

gaa gat gca cca aaa gta ctg aaa aaa atg ggc aaa cta ttt gct gaa 1632
Glu Asp Ala Pro Lys Val Leu Lys Lys Met Gly Lys Leu Phe Ala Glu
530 535 540

caa aat aaa tca taa 1647
Gln Asn Lys Ser
545

<210> 28
<211> 548
<212> PRT
<213> 乳酸乳球菌

<400> 28

Met Tyr Thr Val Gly Asp Tyr Leu Leu Asp Arg Leu His Glu Leu Gly
1 5 10 15

Ile Glu Glu Ile Phe Gly Val Pro Gly Asp Tyr Asn Leu Gln Phe Leu
20 25 30

Asp Gln Ile Ile Ser Arg Lys Asp Met Lys Trp Val Gly Asn Ala Asn
35 40 45

```
Glu Leu Asn Ala Ser Tyr Met Ala Asp Gly Tyr Ala Arg Thr Lys Lys
     50                  55                  60

Ala Ala Ala Phe Leu Thr Thr Phe Gly Val Gly Glu Leu Ser Ala Val
65                  70                  75                  80

Asn Gly Leu Ala Gly Ser Tyr Ala Glu Asn Leu Pro Val Val Glu Ile
                85                  90                  95

Val Gly Ser Pro Thr Ser Lys Val Gln Asn Glu Gly Lys Phe Val His
            100                 105                 110

His Thr Leu Ala Asp Gly Asp Phe Lys His Phe Met Lys Met His Glu
        115                 120                 125

Pro Val Thr Ala Ala Arg Thr Leu Leu Thr Ala Glu Asn Ala Thr Val
    130                 135                 140

Glu Ile Asp Arg Val Leu Ser Ala Leu Leu Lys Glu Arg Lys Pro Val
145                 150                 155                 160

Tyr Ile Asn Leu Pro Val Asp Val Ala Ala Ala Lys Ala Glu Lys Pro
                165                 170                 175

Ser Leu Pro Leu Lys Lys Glu Asn Ser Thr Ser Asn Thr Ser Asp Gln
            180                 185                 190

Glu Ile Leu Asn Lys Ile Gln Glu Ser Leu Lys Asn Ala Lys Lys Pro
        195                 200                 205

Ile Val Ile Thr Gly His Glu Ile Ile Ser Phe Gly Leu Glu Lys Thr
    210                 215                 220

Val Ser Gln Phe Ile Ser Lys Thr Lys Leu Pro Ile Thr Thr Leu Asn
225                 230                 235                 240

Phe Gly Lys Ser Ser Val Asp Glu Ala Leu Pro Ser Phe Leu Gly Ile
                245                 250                 255
```

Tyr Asn Gly Lys Leu Ser Glu Pro Asn Leu Lys Glu Phe Val Glu Ser
260 265 270

Ala Asp Phe Ile Leu Met Leu Gly Val Lys Leu Thr Asp Ser Ser Thr
275 280 285

Gly Ala Phe Thr His His Leu Asn Glu Asn Lys Met Ile Ser Leu Asn
290 295 300

Ile Asp Glu Gly Lys Ile Phe Asn Glu Ser Ile Gln Asn Phe Asp Phe
305 310 315 320

Glu Ser Leu Ile Ser Ser Leu Leu Asp Leu Ser Glu Ile Glu Tyr Lys
325 330 335

Gly Lys Tyr Ile Asp Lys Lys Gln Glu Asp Phe Val Pro Ser Asn Ala
340 345 350

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

Ser Asn Glu Thr Ile Val Ala Glu Gln Gly Thr Ser Phe Phe Gly Ala
370 375 380

Ser Ser Ile Phe Leu Lys Pro Lys Ser His Phe Ile Gly Gln Pro Leu
385 390 395 400

Trp Gly Ser Ile Gly Tyr Thr Phe Pro Ala Ala Leu Gly Ser Gln Ile
405 410 415

Ala Asp Lys Glu Ser Arg His Leu Leu Phe Ile Gly Asp Gly Ser Leu
420 425 430

Gln Leu Thr Val Gln Glu Leu Gly Leu Ala Ile Arg Glu Lys Ile Asn
435 440 445

Pro Ile Cys Phe Ile Ile Asn Asn Asp Gly Tyr Thr Val Glu Arg Glu
450 455 460

```
Ile His Gly Pro Asn Gln Ser Tyr Asn Asp Ile Pro Met Trp Asn Tyr
465                 470                 475                 480

Ser Lys Leu Pro Glu Ser Phe Gly Ala Thr Glu Glu Arg Val Val Ser
                485                 490                 495

Lys Ile Val Arg Thr Glu Asn Glu Phe Val Ser Val Met Lys Glu Ala
            500                 505                 510

Gln Ala Asp Pro Asn Arg Met Tyr Trp Ile Glu Leu Ile Leu Ala Lys
        515                 520                 525

Glu Asp Ala Pro Lys Val Leu Lys Lys Met Gly Lys Leu Phe Ala Glu
    530                 535                 540

Gln Asn Lys Ser
545
```

<210> 29
<211> 1692
<212> DNA
<213> 酿酒酵母

<220>
<221> CDS
<222> (1)..(1692)

<400> 29

```
atg tct gaa att act ctt gga aaa tac tta ttt gaa aga ttg aag caa       48
Met Ser Glu Ile Thr Leu Gly Lys Tyr Leu Phe Glu Arg Leu Lys Gln
1               5                   10                  15

gtt aat gtt aac acc att ttt ggg cta cca ggc gac ttc aac ttg tcc       96
Val Asn Val Asn Thr Ile Phe Gly Leu Pro Gly Asp Phe Asn Leu Ser
            20                  25                  30

cta ttg gac aag att tac gag gta gat gga ttg aga tgg gct ggt aat      144
Leu Leu Asp Lys Ile Tyr Glu Val Asp Gly Leu Arg Trp Ala Gly Asn
        35                  40                  45

gca aat gag ctg aac gcc gcc tat gcc gcc gat ggt tac gca cgc atc      192
Ala Asn Glu Leu Asn Ala Ala Tyr Ala Ala Asp Gly Tyr Ala Arg Ile
    50                  55                  60
```

```
aag ggt tta tct gtg ctg gta act act ttt ggc gta ggt gaa tta tcc      240
Lys Gly Leu Ser Val Leu Val Thr Thr Phe Gly Val Gly Glu Leu Ser
65                  70                  75                  80

gcc ttg aat ggt att gca gga tcg tat gca gaa cac gtc ggt gta ctg      288
Ala Leu Asn Gly Ile Ala Gly Ser Tyr Ala Glu His Val Gly Val Leu
                85                  90                  95

cat gtt gtt ggt gtc ccc tct atc tcc gct cag gct aag caa ttg ttg      336
His Val Val Gly Val Pro Ser Ile Ser Ala Gln Ala Lys Gln Leu Leu
            100                 105                 110

ttg cat cat acc ttg ggt aac ggt gat ttt acc gtt ttt cac aga atg      384
Leu His His Thr Leu Gly Asn Gly Asp Phe Thr Val Phe His Arg Met
        115                 120                 125

tcc gcc aat atc tca gaa act aca tca atg att aca gac att gct aca      432
Ser Ala Asn Ile Ser Glu Thr Thr Ser Met Ile Thr Asp Ile Ala Thr
    130                 135                 140

gcc cct tca gaa atc gat agg ttg atc agg aca aca ttt ata aca caa      480
Ala Pro Ser Glu Ile Asp Arg Leu Ile Arg Thr Thr Phe Ile Thr Gln
145                 150                 155                 160

agg cct agc tac ttg ggg ttg cca gcg aat ttg gta gat cta aag gtt      528
Arg Pro Ser Tyr Leu Gly Leu Pro Ala Asn Leu Val Asp Leu Lys Val
                165                 170                 175

cct ggt tct ctt ttg gaa aaa ccg att gat cta tca tta aaa cct aac      576
Pro Gly Ser Leu Leu Glu Lys Pro Ile Asp Leu Ser Leu Lys Pro Asn
            180                 185                 190

gat ccc gaa gct gaa aag gaa gtt att gat acc gta cta gaa ttg atc      624
Asp Pro Glu Ala Glu Lys Glu Val Ile Asp Thr Val Leu Glu Leu Ile
        195                 200                 205

cag aat tcg aaa aac cct gtt ata cta tcg gat gcc tgt gct tct agg      672
Gln Asn Ser Lys Asn Pro Val Ile Leu Ser Asp Ala Cys Ala Ser Arg
    210                 215                 220

cac aac gtt aaa aaa gaa acc cag aag tta att gat ttg acg caa ttc      720
His Asn Val Lys Lys Glu Thr Gln Lys Leu Ile Asp Leu Thr Gln Phe
225                 230                 235                 240

cca gct ttt gtg aca cct cta ggt aaa ggg tca ata gat gaa cag cat      768
Pro Ala Phe Val Thr Pro Leu Gly Lys Gly Ser Ile Asp Glu Gln His
                245                 250                 255

ccc aga tat ggc ggt gtt tat gtg gga acg ctg tcc aaa caa gac gtg      816
Pro Arg Tyr Gly Gly Val Tyr Val Gly Thr Leu Ser Lys Gln Asp Val
            260                 265                 270
```

```
aaa cag gcc gtt gag tcg gct gat ttg atc ctt tcg gtc ggt gct ttg      864
Lys Gln Ala Val Glu Ser Ala Asp Leu Ile Leu Ser Val Gly Ala Leu
        275                 280                 285

ctc tct gat ttt aac aca ggt tcg ttt tcc tac tcc tac aag act aaa      912
Leu Ser Asp Phe Asn Thr Gly Ser Phe Ser Tyr Ser Tyr Lys Thr Lys
    290                 295                 300

aat gta gtg gag ttt cat tcc gat tac gta aag gtg aag aac gct acg      960
Asn Val Val Glu Phe His Ser Asp Tyr Val Lys Val Lys Asn Ala Thr
305                 310                 315                 320

ttc ctc ggt gta caa atg aaa ttt gca cta caa aac tta ctg aag gtt     1008
Phe Leu Gly Val Gln Met Lys Phe Ala Leu Gln Asn Leu Leu Lys Val
                325                 330                 335

att ccc gat gtt gtt aag ggc tac aag agc gtt ccc gta cca acc aaa     1056
Ile Pro Asp Val Val Lys Gly Tyr Lys Ser Val Pro Val Pro Thr Lys
            340                 345                 350

act ccc gca aac aaa ggt gta cct gct agc acg ccc ttg aaa caa gag     1104
Thr Pro Ala Asn Lys Gly Val Pro Ala Ser Thr Pro Leu Lys Gln Glu
        355                 360                 365

tgg ttg tgg aac gaa ttg tcc aaa ttc ttg caa gaa ggt gat gtt atc     1152
Trp Leu Trp Asn Glu Leu Ser Lys Phe Leu Gln Glu Gly Asp Val Ile
    370                 375                 380

att tcc gag acc ggc acg tct gcc ttc ggt atc aat caa act atc ttt     1200
Ile Ser Glu Thr Gly Thr Ser Ala Phe Gly Ile Asn Gln Thr Ile Phe
385                 390                 395                 400

cct aag gac gcc tac ggt atc tcg cag gtg ttg tgg ggg tcc atc ggt     1248
Pro Lys Asp Ala Tyr Gly Ile Ser Gln Val Leu Trp Gly Ser Ile Gly
                405                 410                 415

ttt aca aca gga gca act tta ggt gct gcc ttt gcc gct gag gag att     1296
Phe Thr Thr Gly Ala Thr Leu Gly Ala Ala Phe Ala Ala Glu Glu Ile
            420                 425                 430

gac ccc aac aag aga gtc atc tta ttc ata ggt gac ggg tct ttg cag     1344
Asp Pro Asn Lys Arg Val Ile Leu Phe Ile Gly Asp Gly Ser Leu Gln
        435                 440                 445

tta acc gtc caa gaa atc tcc acc atg atc aga tgg ggg tta aag ccg     1392
Leu Thr Val Gln Glu Ile Ser Thr Met Ile Arg Trp Gly Leu Lys Pro
    450                 455                 460

tat ctt ttt gtc ctt aac aac gac ggc tac act atc gaa aag ctg att     1440
Tyr Leu Phe Val Leu Asn Asn Asp Gly Tyr Thr Ile Glu Lys Leu Ile
465                 470                 475                 480
```

```
cat ggg cct cac gca gag tac aac gaa atc cag acc tgg gat cac ctc     1488
His Gly Pro His Ala Glu Tyr Asn Glu Ile Gln Thr Trp Asp His Leu
                485                 490                 495

gcc ctg ttg ccc gca ttt ggt gcg aaa aag tac gaa aat cac aag atc     1536
Ala Leu Leu Pro Ala Phe Gly Ala Lys Lys Tyr Glu Asn His Lys Ile
            500                 505                 510

gcc act acg ggt gag tgg gat gcc tta acc act gat tca gag ttc cag     1584
Ala Thr Thr Gly Glu Trp Asp Ala Leu Thr Thr Asp Ser Glu Phe Gln
        515                 520                 525

aaa aac tcg gtg atc aga cta att gaa ctg aaa ctg ccc gtc ttt gat     1632
Lys Asn Ser Val Ile Arg Leu Ile Glu Leu Lys Leu Pro Val Phe Asp
    530                 535                 540

gct ccg gaa agt ttg atc aaa caa gcg caa ttg act gcc gct aca aat     1680
Ala Pro Glu Ser Leu Ile Lys Gln Ala Gln Leu Thr Ala Ala Thr Asn
545                 550                 555                 560

gcc aaa caa taa                                                    1692
Ala Lys Gln
```

<210> 30
<211> 563
<212> PRT
<213> 酿酒酵母

<400> 30

```
Met Ser Glu Ile Thr Leu Gly Lys Tyr Leu Phe Glu Arg Leu Lys Gln
1               5                   10                  15

Val Asn Val Asn Thr Ile Phe Gly Leu Pro Gly Asp Phe Asn Leu Ser
            20                  25                  30

Leu Leu Asp Lys Ile Tyr Glu Val Asp Gly Leu Arg Trp Ala Gly Asn
        35                  40                  45

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

Lys Gly Leu Ser Val Leu Val Thr Thr Phe Gly Val Gly Glu Leu Ser
65                  70                  75                  80
```

Ala Leu Asn Gly Ile Ala Gly Ser Tyr Ala Glu His Val Gly Val Leu
85 90 95

His Val Val Gly Val Pro Ser Ile Ser Ala Gln Ala Lys Gln Leu Leu
100 105 110

Leu His His Thr Leu Gly Asn Gly Asp Phe Thr Val Phe His Arg Met
115 120 125

Ser Ala Asn Ile Ser Glu Thr Thr Ser Met Ile Thr Asp Ile Ala Thr
130 135 140

Ala Pro Ser Glu Ile Asp Arg Leu Ile Arg Thr Thr Phe Ile Thr Gln
145 150 155 160

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

Pro Gly Ser Leu Leu Glu Lys Pro Ile Asp Leu Ser Leu Lys Pro Asn
180 185 190

Asp Pro Glu Ala Glu Lys Glu Val Ile Asp Thr Val Leu Glu Leu Ile
195 200 205

Gln Asn Ser Lys Asn Pro Val Ile Leu Ser Asp Ala Cys Ala Ser Arg
210 215 220

His Asn Val Lys Lys Glu Thr Gln Lys Leu Ile Asp Leu Thr Gln Phe
225 230 235 240

Pro Ala Phe Val Thr Pro Leu Gly Lys Gly Ser Ile Asp Glu Gln His
245 250 255

Pro Arg Tyr Gly Gly Val Tyr Val Gly Thr Leu Ser Lys Gln Asp Val
260 265 270

Lys Gln Ala Val Glu Ser Ala Asp Leu Ile Leu Ser Val Gly Ala Leu
275 280 285

```
Leu Ser Asp Phe Asn Thr Gly Ser Phe Ser Tyr Ser Tyr Lys Thr Lys
    290                 295                 300

Asn Val Val Glu Phe His Ser Asp Tyr Val Lys Val Lys Asn Ala Thr
305                 310                 315                 320

Phe Leu Gly Val Gln Met Lys Phe Ala Leu Gln Asn Leu Leu Lys Val
                325                 330                 335

Ile Pro Asp Val Val Lys Gly Tyr Lys Ser Val Pro Val Pro Thr Lys
            340                 345                 350

Thr Pro Ala Asn Lys Gly Val Pro Ala Ser Thr Pro Leu Lys Gln Glu
        355                 360                 365

Trp Leu Trp Asn Glu Leu Ser Lys Phe Leu Gln Glu Gly Asp Val Ile
    370                 375                 380

Ile Ser Glu Thr Gly Thr Ser Ala Phe Gly Ile Asn Gln Thr Ile Phe
385                 390                 395                 400

Pro Lys Asp Ala Tyr Gly Ile Ser Gln Val Leu Trp Gly Ser Ile Gly
                405                 410                 415

Phe Thr Thr Gly Ala Thr Leu Gly Ala Ala Phe Ala Ala Glu Glu Ile
            420                 425                 430

Asp Pro Asn Lys Arg Val Ile Leu Phe Ile Gly Asp Gly Ser Leu Gln
        435                 440                 445

Leu Thr Val Gln Glu Ile Ser Thr Met Ile Arg Trp Gly Leu Lys Pro
    450                 455                 460

Tyr Leu Phe Val Leu Asn Asn Asp Gly Tyr Thr Ile Glu Lys Leu Ile
465                 470                 475                 480

His Gly Pro His Ala Glu Tyr Asn Glu Ile Gln Thr Trp Asp His Leu
                485                 490                 495
```

Ala Leu Leu Pro Ala Phe Gly Ala Lys Lys Tyr Glu Asn His Lys Ile
500 505 510

Ala Thr Thr Gly Glu Trp Asp Ala Leu Thr Thr Asp Ser Glu Phe Gln
515 520 525

Lys Asn Ser Val Ile Arg Leu Ile Glu Leu Lys Leu Pro Val Phe Asp
530 535 540

Ala Pro Glu Ser Leu Ile Lys Gln Ala Gln Leu Thr Ala Ala Thr Asn
545 550 555 560

Ala Lys Gln

<210> 31
<211> 1908
<212> DNA
<213> 酿酒酵母

<220>
<221> CDS
<222> (1)..(1908)

<400> 31
atg gca cct gtt aca att gaa aag ttc gta aat caa gaa gaa cga cac 48
Met Ala Pro Val Thr Ile Glu Lys Phe Val Asn Gln Glu Glu Arg His
1 5 10 15

ctt gtt tcc aac cga tca gca aca att ccg ttt ggt gaa tac ata ttt 96
Leu Val Ser Asn Arg Ser Ala Thr Ile Pro Phe Gly Glu Tyr Ile Phe
20 25 30

aaa aga ttg ttg tcc atc gat acg aaa tca gtt ttc ggt gtt cct ggt 144
Lys Arg Leu Leu Ser Ile Asp Thr Lys Ser Val Phe Gly Val Pro Gly
35 40 45

gac ttc aac tta tct cta tta gaa tat ctc tat tca cct agt gtt gaa 192
Asp Phe Asn Leu Ser Leu Leu Glu Tyr Leu Tyr Ser Pro Ser Val Glu
50 55 60

tca gct ggc cta aga tgg gtc ggc acg tgt aat gaa ctg aac gcc gct 240
Ser Ala Gly Leu Arg Trp Val Gly Thr Cys Asn Glu Leu Asn Ala Ala
65 70 75 80

```
tat gcg gcc gac gga tat tcc cgt tac tct aat aag att ggc tgt tta      288
Tyr Ala Ala Asp Gly Tyr Ser Arg Tyr Ser Asn Lys Ile Gly Cys Leu
                85                  90                  95

ata acc acg tat ggc gtt ggt gaa tta agc gcc ttg aac ggt ata gcc      336
Ile Thr Thr Tyr Gly Val Gly Glu Leu Ser Ala Leu Asn Gly Ile Ala
            100                 105                 110

ggt tcg ttc gct gaa aat gtc aaa gtt ttg cac att gtt ggt gtg gcc      384
Gly Ser Phe Ala Glu Asn Val Lys Val Leu His Ile Val Gly Val Ala
        115                 120                 125

aag tcc ata gat tcg cgt tca agt aac ttt agt gat cgg aac cta cat      432
Lys Ser Ile Asp Ser Arg Ser Ser Asn Phe Ser Asp Arg Asn Leu His
    130                 135                 140

cat ttg gtc cca cag cta cat gat tca aat ttt aaa ggg cca aat cat      480
His Leu Val Pro Gln Leu His Asp Ser Asn Phe Lys Gly Pro Asn His
145                 150                 155                 160

aaa gta tat cat gat atg gta aaa gat aga gtc gct tgc tcg gta gcc      528
Lys Val Tyr His Asp Met Val Lys Asp Arg Val Ala Cys Ser Val Ala
                165                 170                 175

tac ttg gag gat att gaa act gca tgt gac caa gtc gat aat gtt atc      576
Tyr Leu Glu Asp Ile Glu Thr Ala Cys Asp Gln Val Asp Asn Val Ile
            180                 185                 190

cgc gat att tac aag tat tct aaa cct ggt tat att ttt gtt cct gca      624
Arg Asp Ile Tyr Lys Tyr Ser Lys Pro Gly Tyr Ile Phe Val Pro Ala
        195                 200                 205

gat ttt gcg gat atg tct gtt aca tgt gat aat ttg gtt aat gtt cca      672
Asp Phe Ala Asp Met Ser Val Thr Cys Asp Asn Leu Val Asn Val Pro
    210                 215                 220

cgt ata tct caa caa gat tgt ata gta tac cct tct gaa aac caa ttg      720
Arg Ile Ser Gln Gln Asp Cys Ile Val Tyr Pro Ser Glu Asn Gln Leu
225                 230                 235                 240

tct gac ata atc aac aag att act agt tgg ata tat tcc agt aaa aca      768
Ser Asp Ile Ile Asn Lys Ile Thr Ser Trp Ile Tyr Ser Ser Lys Thr
                245                 250                 255

cct gcg atc ctt gga gac gta ctg act gat agg tat ggt gtg agt aac      816
Pro Ala Ile Leu Gly Asp Val Leu Thr Asp Arg Tyr Gly Val Ser Asn
            260                 265                 270

ttt ttg aac aag ctt atc tgc aaa act ggg att tgg aat ttt tcc act      864
Phe Leu Asn Lys Leu Ile Cys Lys Thr Gly Ile Trp Asn Phe Ser Thr
        275                 280                 285
```

```
gtt atg gga aaa tct gta att gat gag tca aac cca act tat atg ggt      912
Val Met Gly Lys Ser Val Ile Asp Glu Ser Asn Pro Thr Tyr Met Gly
    290                 295                 300

caa tat aat ggt aaa gaa ggt tta aaa caa gtc tat gaa cat ttt gaa      960
Gln Tyr Asn Gly Lys Glu Gly Leu Lys Gln Val Tyr Glu His Phe Glu
305                 310                 315                 320

ctg tgc gac ttg gtc ttg cat ttt gga gtc gac atc aat gaa att aat     1008
Leu Cys Asp Leu Val Leu His Phe Gly Val Asp Ile Asn Glu Ile Asn
                325                 330                 335

aat ggg cat tat act ttt act tat aaa cca aat gct aaa atc att caa     1056
Asn Gly His Tyr Thr Phe Thr Tyr Lys Pro Asn Ala Lys Ile Ile Gln
            340                 345                 350

ttt cat ccg aat tat att cgc ctt gtg gac act agg cag ggc aat gag     1104
Phe His Pro Asn Tyr Ile Arg Leu Val Asp Thr Arg Gln Gly Asn Glu
        355                 360                 365

caa atg ttc aaa gga atc aat ttt gcc cct att tta aaa gaa cta tac     1152
Gln Met Phe Lys Gly Ile Asn Phe Ala Pro Ile Leu Lys Glu Leu Tyr
    370                 375                 380

aag cgc att gac gtt tct aaa ctt tct ttg caa tat gat tca aat gta     1200
Lys Arg Ile Asp Val Ser Lys Leu Ser Leu Gln Tyr Asp Ser Asn Val
385                 390                 395                 400

act caa tat acg aac gaa aca atg cgg tta gaa gat cct acc aat gga     1248
Thr Gln Tyr Thr Asn Glu Thr Met Arg Leu Glu Asp Pro Thr Asn Gly
                405                 410                 415

caa tca agc att att aca caa gtt cac tta caa aag acg atg cct aaa     1296
Gln Ser Ser Ile Ile Thr Gln Val His Leu Gln Lys Thr Met Pro Lys
            420                 425                 430

ttt ttg aac cct ggt gat gtt gtc gtt tgt gaa aca ggc tct ttt caa     1344
Phe Leu Asn Pro Gly Asp Val Val Val Cys Glu Thr Gly Ser Phe Gln
        435                 440                 445

ttc tct gtt cgt gat ttc gcg ttt cct tcg caa tta aaa tat ata tcg     1392
Phe Ser Val Arg Asp Phe Ala Phe Pro Ser Gln Leu Lys Tyr Ile Ser
    450                 455                 460

caa gga ttt ttc ctt tcc att ggc atg gcc ctt cct gcc gcc cta ggt     1440
Gln Gly Phe Phe Leu Ser Ile Gly Met Ala Leu Pro Ala Ala Leu Gly
465                 470                 475                 480

gtt gga att gcc atg caa gac cac tca aac gct cac atc aat ggt ggc     1488
Val Gly Ile Ala Met Gln Asp His Ser Asn Ala His Ile Asn Gly Gly
                485                 490                 495
```

```
aac gta aaa gag gac tat aag cca aga tta att ttg ttt gaa ggt gac     1536
Asn Val Lys Glu Asp Tyr Lys Pro Arg Leu Ile Leu Phe Glu Gly Asp
            500                 505                 510

ggt gca gca cag atg aca atc caa gaa ctg agc acc att ctg aag tgc     1584
Gly Ala Ala Gln Met Thr Ile Gln Glu Leu Ser Thr Ile Leu Lys Cys
        515                 520                 525

aat att cca cta gaa gtt atc att tgg aac aat aac ggc tac act att     1632
Asn Ile Pro Leu Glu Val Ile Ile Trp Asn Asn Asn Gly Tyr Thr Ile
    530                 535                 540

gaa aga gcc atc atg ggc cct acc agg tcg tat aac gac gtt atg tct     1680
Glu Arg Ala Ile Met Gly Pro Thr Arg Ser Tyr Asn Asp Val Met Ser
545                 550                 555                 560

tgg aaa tgg acc aaa cta ttt gaa gca ttc gga gac ttc gac gga aag     1728
Trp Lys Trp Thr Lys Leu Phe Glu Ala Phe Gly Asp Phe Asp Gly Lys
                565                 570                 575

tat act aat agc act ctc att caa tgt ccc tct aaa tta gca ctg aaa     1776
Tyr Thr Asn Ser Thr Leu Ile Gln Cys Pro Ser Lys Leu Ala Leu Lys
            580                 585                 590

ttg gag gag ctt aag aat tca aac aaa aga agc ggg ata gaa ctt tta     1824
Leu Glu Glu Leu Lys Asn Ser Asn Lys Arg Ser Gly Ile Glu Leu Leu
        595                 600                 605

gaa gtc aaa tta ggc gaa ttg gat ttc ccc gaa cag cta aag tgc atg     1872
Glu Val Lys Leu Gly Glu Leu Asp Phe Pro Glu Gln Leu Lys Cys Met
    610                 615                 620

gtt gaa gca gcg gca ctt aaa aga aat aaa aaa tag                     1908
Val Glu Ala Ala Ala Leu Lys Arg Asn Lys Lys
625                 630                 635
```

<210> 32
<211> 635
<212> PRT
<213> 酿酒酵母

<400> 32

```
Met Ala Pro Val Thr Ile Glu Lys Phe Val Asn Gln Glu Glu Arg His
1               5                   10                  15

Leu Val Ser Asn Arg Ser Ala Thr Ile Pro Phe Gly Glu Tyr Ile Phe
            20                  25                  30
```

```
Lys Arg Leu Leu Ser Ile Asp Thr Lys Ser Val Phe Gly Val Pro Gly
        35                  40                  45

Asp Phe Asn Leu Ser Leu Leu Glu Tyr Leu Tyr Ser Pro Ser Val Glu
    50                  55                  60

Ser Ala Gly Leu Arg Trp Val Gly Thr Cys Asn Glu Leu Asn Ala Ala
65                  70                  75                  80

Tyr Ala Ala Asp Gly Tyr Ser Arg Tyr Ser Asn Lys Ile Gly Cys Leu
                85                  90                  95

Ile Thr Thr Tyr Gly Val Gly Glu Leu Ser Ala Leu Asn Gly Ile Ala
            100                 105                 110

Gly Ser Phe Ala Glu Asn Val Lys Val Leu His Ile Val Gly Val Ala
        115                 120                 125

Lys Ser Ile Asp Ser Arg Ser Ser Asn Phe Ser Asp Arg Asn Leu His
    130                 135                 140

His Leu Val Pro Gln Leu His Asp Ser Asn Phe Lys Gly Pro Asn His
145                 150                 155                 160

Lys Val Tyr His Asp Met Val Lys Asp Arg Val Ala Cys Ser Val Ala
                165                 170                 175

Tyr Leu Glu Asp Ile Glu Thr Ala Cys Asp Gln Val Asp Asn Val Ile
            180                 185                 190

Arg Asp Ile Tyr Lys Tyr Ser Lys Pro Gly Tyr Ile Phe Val Pro Ala
        195                 200                 205

Asp Phe Ala Asp Met Ser Val Thr Cys Asp Asn Leu Val Asn Val Pro
    210                 215                 220

Arg Ile Ser Gln Gln Asp Cys Ile Val Tyr Pro Ser Glu Asn Gln Leu
225                 230                 235                 240
```

```
Ser Asp Ile Ile Asn Lys Ile Thr Ser Trp Ile Tyr Ser Ser Lys Thr
                245                 250                 255

Pro Ala Ile Leu Gly Asp Val Leu Thr Asp Arg Tyr Gly Val Ser Asn
            260                 265                 270

Phe Leu Asn Lys Leu Ile Cys Lys Thr Gly Ile Trp Asn Phe Ser Thr
        275                 280                 285

Val Met Gly Lys Ser Val Ile Asp Glu Ser Asn Pro Thr Tyr Met Gly
    290                 295                 300

Gln Tyr Asn Gly Lys Glu Gly Leu Lys Gln Val Tyr Glu His Phe Glu
305                 310                 315                 320

Leu Cys Asp Leu Val Leu His Phe Gly Val Asp Ile Asn Glu Ile Asn
                325                 330                 335

Asn Gly His Tyr Thr Phe Thr Tyr Lys Pro Asn Ala Lys Ile Ile Gln
            340                 345                 350

Phe His Pro Asn Tyr Ile Arg Leu Val Asp Thr Arg Gln Gly Asn Glu
        355                 360                 365

Gln Met Phe Lys Gly Ile Asn Phe Ala Pro Ile Leu Lys Glu Leu Tyr
    370                 375                 380

Lys Arg Ile Asp Val Ser Lys Leu Ser Leu Gln Tyr Asp Ser Asn Val
385                 390                 395                 400

Thr Gln Tyr Thr Asn Glu Thr Met Arg Leu Glu Asp Pro Thr Asn Gly
                405                 410                 415

Gln Ser Ser Ile Ile Thr Gln Val His Leu Gln Lys Thr Met Pro Lys
            420                 425                 430

Phe Leu Asn Pro Gly Asp Val Val Val Cys Glu Thr Gly Ser Phe Gln
        435                 440                 445
```

Phe Ser Val Arg Asp Phe Ala Phe Pro Ser Gln Leu Lys Tyr Ile Ser
450 455 460

Gln Gly Phe Phe Leu Ser Ile Gly Met Ala Leu Pro Ala Ala Leu Gly
465 470 475 480

Val Gly Ile Ala Met Gln Asp His Ser Asn Ala His Ile Asn Gly Gly
485 490 495

Asn Val Lys Glu Asp Tyr Lys Pro Arg Leu Ile Leu Phe Glu Gly Asp
500 505 510

Gly Ala Ala Gln Met Thr Ile Gln Glu Leu Ser Thr Ile Leu Lys Cys
515 520 525

Asn Ile Pro Leu Glu Val Ile Ile Trp Asn Asn Asn Gly Tyr Thr Ile
530 535 540

Glu Arg Ala Ile Met Gly Pro Thr Arg Ser Tyr Asn Asp Val Met Ser
545 550 555 560

Trp Lys Trp Thr Lys Leu Phe Glu Ala Phe Gly Asp Phe Asp Gly Lys
565 570 575

Tyr Thr Asn Ser Thr Leu Ile Gln Cys Pro Ser Lys Leu Ala Leu Lys
580 585 590

Leu Glu Glu Leu Lys Asn Ser Asn Lys Arg Ser Gly Ile Glu Leu Leu
595 600 605

Glu Val Lys Leu Gly Glu Leu Asp Phe Pro Glu Gln Leu Lys Cys Met
610 615 620

Val Glu Ala Ala Ala Leu Lys Arg Asn Lys Lys
625 630 635

<210> 33
<211> 1830
<212> DNA

<213> 酿酒酵母

<220>
<221> CDS
<222> (1)..(1830)

<400> 33

```
atg aat tct agc tat aca cag aga tat gca ctg ccg aag tgt ata gca      48
Met Asn Ser Ser Tyr Thr Gln Arg Tyr Ala Leu Pro Lys Cys Ile Ala
1               5                  10                  15

ata tca gat tat ctt ttc cat cgg ctc aac cag ctg aac ata cat acc      96
Ile Ser Asp Tyr Leu Phe His Arg Leu Asn Gln Leu Asn Ile His Thr
            20                  25                  30

ata ttt gga ctc tcc gga gaa ttt agc atg ccg ttg ctg gat aaa cta     144
Ile Phe Gly Leu Ser Gly Glu Phe Ser Met Pro Leu Leu Asp Lys Leu
        35                  40                  45

tac aac att ccg aac tta cga tgg gcc ggt aat tct aat gag tta aat     192
Tyr Asn Ile Pro Asn Leu Arg Trp Ala Gly Asn Ser Asn Glu Leu Asn
    50                  55                  60

gct gcc tac gca gca gat gga tac tca cga cta aaa ggc ttg gga tgt     240
Ala Ala Tyr Ala Ala Asp Gly Tyr Ser Arg Leu Lys Gly Leu Gly Cys
65                  70                  75                  80

ctc ata aca acc ttt ggt gta ggc gaa tta tcg gca atc aat ggc gtg     288
Leu Ile Thr Thr Phe Gly Val Gly Glu Leu Ser Ala Ile Asn Gly Val
                85                  90                  95

gcc gga tct tac gct gaa cat gta gga ata ctt cac ata gtg ggt atg     336
Ala Gly Ser Tyr Ala Glu His Val Gly Ile Leu His Ile Val Gly Met
            100                 105                 110

ccg cca aca agt gca caa acg aaa caa cta cta ctg cat cat act ctg     384
Pro Pro Thr Ser Ala Gln Thr Lys Gln Leu Leu Leu His His Thr Leu
        115                 120                 125

ggc aat ggt gat ttc acg gta ttt cat aga ata gcc agt gat gta gca     432
Gly Asn Gly Asp Phe Thr Val Phe His Arg Ile Ala Ser Asp Val Ala
    130                 135                 140

tgc tat aca aca ttg att att gac tct gaa tta tgt gcc gac gaa gtc     480
Cys Tyr Thr Thr Leu Ile Ile Asp Ser Glu Leu Cys Ala Asp Glu Val
145                 150                 155                 160

gat aag tgc atc aaa aag gct tgg ata gaa cag agg cca gta tac atg     528
Asp Lys Cys Ile Lys Lys Ala Trp Ile Glu Gln Arg Pro Val Tyr Met
                165                 170                 175
```

```
ggc atg cct gtc aac cag gta aat ctc ccg att gaa tca gca agg ctt      576
Gly Met Pro Val Asn Gln Val Asn Leu Pro Ile Glu Ser Ala Arg Leu
            180                 185                 190

aat aca cct ctg gat tta caa ttg cat aaa aac gac cca gac gta gag      624
Asn Thr Pro Leu Asp Leu Gln Leu His Lys Asn Asp Pro Asp Val Glu
        195                 200                 205

aaa gaa gtt att tct cga ata ttg agt ttt ata tac aaa agc cag aat      672
Lys Glu Val Ile Ser Arg Ile Leu Ser Phe Ile Tyr Lys Ser Gln Asn
    210                 215                 220

ccg gca atc atc gta gat gca tgt act agt cga cag aat tta atc gag      720
Pro Ala Ile Ile Val Asp Ala Cys Thr Ser Arg Gln Asn Leu Ile Glu
225                 230                 235                 240

gag act aaa gag ctt tgt aat agg ctt aaa ttt cca gtt ttt gtt aca      768
Glu Thr Lys Glu Leu Cys Asn Arg Leu Lys Phe Pro Val Phe Val Thr
                245                 250                 255

cct atg ggt aag ggt aca gta aac gaa aca gac ccg caa ttt ggg ggc      816
Pro Met Gly Lys Gly Thr Val Asn Glu Thr Asp Pro Gln Phe Gly Gly
            260                 265                 270

gta ttc acg ggc tcg ata tca gcc cca gaa gta aga gaa gta gtt gat      864
Val Phe Thr Gly Ser Ile Ser Ala Pro Glu Val Arg Glu Val Val Asp
        275                 280                 285

ttt gcc gat ttt atc atc gtc att ggt tgc atg ctc tcc gaa ttc agc      912
Phe Ala Asp Phe Ile Ile Val Ile Gly Cys Met Leu Ser Glu Phe Ser
    290                 295                 300

acg tca act ttc cac ttc caa tat aaa act aag aat tgt gcg cta cta      960
Thr Ser Thr Phe His Phe Gln Tyr Lys Thr Lys Asn Cys Ala Leu Leu
305                 310                 315                 320

tat tct aca tct gtg aaa ttg aaa aat gcc aca tat cct gac ttg agc     1008
Tyr Ser Thr Ser Val Lys Leu Lys Asn Ala Thr Tyr Pro Asp Leu Ser
                325                 330                 335

att aaa tta cta cta cag aaa ata tta gca aat ctt gat gaa tct aaa     1056
Ile Lys Leu Leu Leu Gln Lys Ile Leu Ala Asn Leu Asp Glu Ser Lys
            340                 345                 350

ctg tct tac caa cca agc gaa caa ccc agt atg atg gtt cca aga cct     1104
Leu Ser Tyr Gln Pro Ser Glu Gln Pro Ser Met Met Val Pro Arg Pro
        355                 360                 365

tac cca gca gga aat gtc ctc ttg aga caa gaa tgg gtc tgg aat gaa     1152
Tyr Pro Ala Gly Asn Val Leu Leu Arg Gln Glu Trp Val Trp Asn Glu
    370                 375                 380
```

```
ata tcc cat tgg ttc caa cca ggt gac ata atc ata aca gaa act ggt     1200
Ile Ser His Trp Phe Gln Pro Gly Asp Ile Ile Ile Thr Glu Thr Gly
385                 390                 395                 400

gct tct gca ttt gga gtt aac cag acc aga ttt ccg gta aat aca cta     1248
Ala Ser Ala Phe Gly Val Asn Gln Thr Arg Phe Pro Val Asn Thr Leu
                405                 410                 415

ggt att tcg caa gct ctt tgg gga tct gtc gga tat aca atg ggg gcg     1296
Gly Ile Ser Gln Ala Leu Trp Gly Ser Val Gly Tyr Thr Met Gly Ala
            420                 425                 430

tgt ctt ggg gca gaa ttt gct gtt caa gag ata aac aag gat aaa ttc     1344
Cys Leu Gly Ala Glu Phe Ala Val Gln Glu Ile Asn Lys Asp Lys Phe
        435                 440                 445

ccc gca act aaa cat aga gtt att ctg ttt atg ggt gac ggt gct ttc     1392
Pro Ala Thr Lys His Arg Val Ile Leu Phe Met Gly Asp Gly Ala Phe
    450                 455                 460

caa ttg aca gtt caa gaa tta tcc aca att gtt aag tgg gga ttg aca     1440
Gln Leu Thr Val Gln Glu Leu Ser Thr Ile Val Lys Trp Gly Leu Thr
465                 470                 475                 480

cct tat att ttt gtg atg aat aac caa ggt tac tct gtg gac agg ttt     1488
Pro Tyr Ile Phe Val Met Asn Asn Gln Gly Tyr Ser Val Asp Arg Phe
                485                 490                 495

ttg cat cac agg tca gat gct agt tat tac gat atc caa cct tgg aac     1536
Leu His His Arg Ser Asp Ala Ser Tyr Tyr Asp Ile Gln Pro Trp Asn
            500                 505                 510

tac ttg gga tta ttg cga gta ttt ggt tgc acg aac tac gaa acg aaa     1584
Tyr Leu Gly Leu Leu Arg Val Phe Gly Cys Thr Asn Tyr Glu Thr Lys
        515                 520                 525

aaa att att act gtt gga gaa ttc aga tcc atg atc agt gac cca aac     1632
Lys Ile Ile Thr Val Gly Glu Phe Arg Ser Met Ile Ser Asp Pro Asn
    530                 535                 540

ttt gcg acc aat gac aaa att cgg atg ata gag att atg cta cca cca     1680
Phe Ala Thr Asn Asp Lys Ile Arg Met Ile Glu Ile Met Leu Pro Pro
545                 550                 555                 560

agg gat gtt cca cag gct ctg ctt gac agg tgg gtg gta gaa aaa gaa     1728
Arg Asp Val Pro Gln Ala Leu Leu Asp Arg Trp Val Val Glu Lys Glu
                565                 570                 575

cag agc aaa caa gtg caa gag gag aac gaa aat tct agc gca gta aat     1776
Gln Ser Lys Gln Val Gln Glu Glu Asn Glu Asn Ser Ser Ala Val Asn
            580                 585                 590
```

```
acg cca act cca gaa ttc caa cca ctt cta aaa aaa aat caa gtt gga     1824
Thr Pro Thr Pro Glu Phe Gln Pro Leu Leu Lys Lys Asn Gln Val Gly
        595                 600                 605

tac tga                                                           1830
Tyr
```

<210> 34
<211> 609
<212> PRT
<213> 酿酒酵母

<400> 34

```
Met Asn Ser Ser Tyr Thr Gln Arg Tyr Ala Leu Pro Lys Cys Ile Ala
1               5                   10                  15

Ile Ser Asp Tyr Leu Phe His Arg Leu Asn Gln Leu Asn Ile His Thr
            20                  25                  30

Ile Phe Gly Leu Ser Gly Glu Phe Ser Met Pro Leu Leu Asp Lys Leu
        35                  40                  45

Tyr Asn Ile Pro Asn Leu Arg Trp Ala Gly Asn Ser Asn Glu Leu Asn
    50                  55                  60

Ala Ala Tyr Ala Ala Asp Gly Tyr Ser Arg Leu Lys Gly Leu Gly Cys
65                  70                  75                  80

Leu Ile Thr Thr Phe Gly Val Gly Glu Leu Ser Ala Ile Asn Gly Val
                85                  90                  95

Ala Gly Ser Tyr Ala Glu His Val Gly Ile Leu His Ile Val Gly Met
            100                 105                 110

Pro Pro Thr Ser Ala Gln Thr Lys Gln Leu Leu Leu His His Thr Leu
        115                 120                 125

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

```
Cys Tyr Thr Thr Leu Ile Ile Asp Ser Glu Leu Cys Ala Asp Glu Val
145                 150                 155                 160

Asp Lys Cys Ile Lys Lys Ala Trp Ile Glu Gln Arg Pro Val Tyr Met
                165                 170                 175

Gly Met Pro Val Asn Gln Val Asn Leu Pro Ile Glu Ser Ala Arg Leu
            180                 185                 190

Asn Thr Pro Leu Asp Leu Gln Leu His Lys Asn Asp Pro Asp Val Glu
        195                 200                 205

Lys Glu Val Ile Ser Arg Ile Leu Ser Phe Ile Tyr Lys Ser Gln Asn
    210                 215                 220

Pro Ala Ile Ile Val Asp Ala Cys Thr Ser Arg Gln Asn Leu Ile Glu
225                 230                 235                 240

Glu Thr Lys Glu Leu Cys Asn Arg Leu Lys Phe Pro Val Phe Val Thr
                245                 250                 255

Pro Met Gly Lys Gly Thr Val Asn Glu Thr Asp Pro Gln Phe Gly Gly
            260                 265                 270

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

Phe Ala Asp Phe Ile Ile Val Ile Gly Cys Met Leu Ser Glu Phe Ser
    290                 295                 300

Thr Ser Thr Phe His Phe Gln Tyr Lys Thr Lys Asn Cys Ala Leu Leu
305                 310                 315                 320

Tyr Ser Thr Ser Val Lys Leu Lys Asn Ala Thr Tyr Pro Asp Leu Ser
                325                 330                 335

Ile Lys Leu Leu Leu Gln Lys Ile Leu Ala Asn Leu Asp Glu Ser Lys
            340                 345                 350
```

```
Leu Ser Tyr Gln Pro Ser Glu Gln Pro Ser Met Met Val Pro Arg Pro
        355                 360                 365

Tyr Pro Ala Gly Asn Val Leu Leu Arg Gln Glu Trp Val Trp Asn Glu
    370                 375                 380

Ile Ser His Trp Phe Gln Pro Gly Asp Ile Ile Ile Thr Glu Thr Gly
385                 390                 395                 400

Ala Ser Ala Phe Gly Val Asn Gln Thr Arg Phe Pro Val Asn Thr Leu
                405                 410                 415

Gly Ile Ser Gln Ala Leu Trp Gly Ser Val Gly Tyr Thr Met Gly Ala
            420                 425                 430

Cys Leu Gly Ala Glu Phe Ala Val Gln Glu Ile Asn Lys Asp Lys Phe
        435                 440                 445

Pro Ala Thr Lys His Arg Val Ile Leu Phe Met Gly Asp Gly Ala Phe
    450                 455                 460

Gln Leu Thr Val Gln Glu Leu Ser Thr Ile Val Lys Trp Gly Leu Thr
465                 470                 475                 480

Pro Tyr Ile Phe Val Met Asn Asn Gln Gly Tyr Ser Val Asp Arg Phe
                485                 490                 495

Leu His His Arg Ser Asp Ala Ser Tyr Tyr Asp Ile Gln Pro Trp Asn
            500                 505                 510

Tyr Leu Gly Leu Leu Arg Val Phe Gly Cys Thr Asn Tyr Glu Thr Lys
        515                 520                 525

Lys Ile Ile Thr Val Gly Glu Phe Arg Ser Met Ile Ser Asp Pro Asn
    530                 535                 540

Phe Ala Thr Asn Asp Lys Ile Arg Met Ile Glu Ile Met Leu Pro Pro
545                 550                 555                 560
```

```
Arg Asp Val Pro Gln Ala Leu Leu Asp Arg Trp Val Val Glu Lys Glu
                565                 570                 575

Gln Ser Lys Gln Val Gln Glu Glu Asn Glu Asn Ser Ser Ala Val Asn
            580                 585                 590

Thr Pro Thr Pro Glu Phe Gln Pro Leu Leu Lys Lys Asn Gln Val Gly
        595                 600                 605

Tyr
```

<210> 35
<211> 1665
<212> DNA
<213> 丙酮丁醇梭菌

<220>
<221> CDS
<222> (1)..(1665)

<400> 35

```
ttg aag agt gaa tac aca att gga aga tat ttg tta gac cgt tta tca       48
Leu Lys Ser Glu Tyr Thr Ile Gly Arg Tyr Leu Leu Asp Arg Leu Ser
1               5                   10                  15

gag ttg ggt att cgg cat atc ttt ggt gta cct gga gat tac aat cta       96
Glu Leu Gly Ile Arg His Ile Phe Gly Val Pro Gly Asp Tyr Asn Leu
            20                  25                  30

tcc ttt tta gac tat ata atg gag tac aaa ggg ata gat tgg gtt gga      144
Ser Phe Leu Asp Tyr Ile Met Glu Tyr Lys Gly Ile Asp Trp Val Gly
        35                  40                  45

aat tgc aat gaa ttg aat gct ggg tat gct gct gat gga tat gca aga      192
Asn Cys Asn Glu Leu Asn Ala Gly Tyr Ala Ala Asp Gly Tyr Ala Arg
    50                  55                  60

ata aat gga att gga gcc ata ctt aca aca ttt ggt gtt gga gaa tta      240
Ile Asn Gly Ile Gly Ala Ile Leu Thr Thr Phe Gly Val Gly Glu Leu
65                  70                  75                  80

agt gcc att aac gca att gct ggg gca tac gct gag caa gtt cca gtt      288
Ser Ala Ile Asn Ala Ile Ala Gly Ala Tyr Ala Glu Gln Val Pro Val
                85                  90                  95
```

```
gtt aaa att aca ggt atc ccc aca gca aaa gtt agg gac aat gga tta      336
Val Lys Ile Thr Gly Ile Pro Thr Ala Lys Val Arg Asp Asn Gly Leu
            100                 105                 110

tat gta cac cac aca tta ggt gac gga agg ttt gat cac ttt ttt gaa      384
Tyr Val His His Thr Leu Gly Asp Gly Arg Phe Asp His Phe Phe Glu
        115                 120                 125

atg ttt aga gaa gta aca gtt gct gag gca tta cta agc gaa gaa aat      432
Met Phe Arg Glu Val Thr Val Ala Glu Ala Leu Leu Ser Glu Glu Asn
    130                 135                 140

gca gca caa gaa att gat cgt gtt ctt att tca tgc tgg aga caa aaa      480
Ala Ala Gln Glu Ile Asp Arg Val Leu Ile Ser Cys Trp Arg Gln Lys
145                 150                 155                 160

cgt cct gtt ctt ata aat tta ccg att gat gta tat gat aaa cca att      528
Arg Pro Val Leu Ile Asn Leu Pro Ile Asp Val Tyr Asp Lys Pro Ile
                165                 170                 175

aac aaa cca tta aag cca tta ctc gat tat act att tca agt aac aaa      576
Asn Lys Pro Leu Lys Pro Leu Leu Asp Tyr Thr Ile Ser Ser Asn Lys
            180                 185                 190

gag gct gca tgt gaa ttt gtt aca gaa ata gta cct ata ata aat agg      624
Glu Ala Ala Cys Glu Phe Val Thr Glu Ile Val Pro Ile Ile Asn Arg
        195                 200                 205

gca aaa aag cct gtt att ctt gca gat tat gga gta tat cgt tac caa      672
Ala Lys Lys Pro Val Ile Leu Ala Asp Tyr Gly Val Tyr Arg Tyr Gln
    210                 215                 220

gtt caa cat gtg ctt aaa aac ttg gcc gaa aaa acc gga ttt cct gtg      720
Val Gln His Val Leu Lys Asn Leu Ala Glu Lys Thr Gly Phe Pro Val
225                 230                 235                 240

gct aca cta agt atg gga aaa ggt gtt ttc aat gaa gca cac cct caa      768
Ala Thr Leu Ser Met Gly Lys Gly Val Phe Asn Glu Ala His Pro Gln
                245                 250                 255

ttt att ggt gtt tat aat ggt gat gta agt tct cct tat tta agg cag      816
Phe Ile Gly Val Tyr Asn Gly Asp Val Ser Ser Pro Tyr Leu Arg Gln
            260                 265                 270

cga gtt gat gaa gca gac tgc att att agc gtt ggt gta aaa ttg acg      864
Arg Val Asp Glu Ala Asp Cys Ile Ile Ser Val Gly Val Lys Leu Thr
        275                 280                 285

gat tca acc aca ggg gga ttt tct cat gga ttt tct aaa agg aat gta      912
Asp Ser Thr Thr Gly Gly Phe Ser His Gly Phe Ser Lys Arg Asn Val
    290                 295                 300
```

```
att cac att gat cct ttt tca ata aag gca aaa ggt aaa aaa tat gca      960
Ile His Ile Asp Pro Phe Ser Ile Lys Ala Lys Gly Lys Lys Tyr Ala
305                 310                 315                 320

cct att acg atg aaa gat gct tta aca gaa tta aca agt aaa att gag     1008
Pro Ile Thr Met Lys Asp Ala Leu Thr Glu Leu Thr Ser Lys Ile Glu
                325                 330                 335

cat aga aac ttt gag gat tta gat ata aag cct tac aaa tca gat aat     1056
His Arg Asn Phe Glu Asp Leu Asp Ile Lys Pro Tyr Lys Ser Asp Asn
            340                 345                 350

caa aag tat ttt gca aaa gag aag cca att aca caa aaa cgt ttt ttt     1104
Gln Lys Tyr Phe Ala Lys Glu Lys Pro Ile Thr Gln Lys Arg Phe Phe
        355                 360                 365

gag cgt att gct cac ttt ata aaa gaa aaa gat gta tta tta gca gaa     1152
Glu Arg Ile Ala His Phe Ile Lys Glu Lys Asp Val Leu Leu Ala Glu
    370                 375                 380

cag ggt aca tgc ttt ttt ggt gcg tca acc ata caa cta ccc aaa gat     1200
Gln Gly Thr Cys Phe Phe Gly Ala Ser Thr Ile Gln Leu Pro Lys Asp
385                 390                 395                 400

gca act ttt att ggt caa cct tta tgg gga tct att gga tac aca ctt     1248
Ala Thr Phe Ile Gly Gln Pro Leu Trp Gly Ser Ile Gly Tyr Thr Leu
                405                 410                 415

cct gct tta tta ggt tca caa tta gct gat caa aaa agg cgt aat att     1296
Pro Ala Leu Leu Gly Ser Gln Leu Ala Asp Gln Lys Arg Arg Asn Ile
            420                 425                 430

ctt tta att ggg gat ggt gca ttt caa atg aca gca caa gaa att tca     1344
Leu Leu Ile Gly Asp Gly Ala Phe Gln Met Thr Ala Gln Glu Ile Ser
        435                 440                 445

aca atg ctt cgt tta caa atc aaa cct att att ttt tta att aat aac     1392
Thr Met Leu Arg Leu Gln Ile Lys Pro Ile Ile Phe Leu Ile Asn Asn
    450                 455                 460

gat ggt tat aca att gaa cgt gct att cat ggt aga gaa caa gta tat     1440
Asp Gly Tyr Thr Ile Glu Arg Ala Ile His Gly Arg Glu Gln Val Tyr
465                 470                 475                 480

aac aat att caa atg tgg cga tat cat aat gtt cca aag gtt tta ggt     1488
Asn Asn Ile Gln Met Trp Arg Tyr His Asn Val Pro Lys Val Leu Gly
                485                 490                 495

cct aaa gaa tgc agc tta acc ttt aaa gta caa agt gaa act gaa ctt     1536
Pro Lys Glu Cys Ser Leu Thr Phe Lys Val Gln Ser Glu Thr Glu Leu
            500                 505                 510
```

```
gaa aag gct ctt tta gtg gca gat aag gat tgt gaa cat ttg att ttt     1584
Glu Lys Ala Leu Leu Val Ala Asp Lys Asp Cys Glu His Leu Ile Phe
        515                 520                 525

ata gaa gtt gtt atg gat cgt tat gat aaa ccc gag cct tta gaa cgt     1632
Ile Glu Val Val Met Asp Arg Tyr Asp Lys Pro Glu Pro Leu Glu Arg
    530                 535                 540

ctt tcg aaa cgt ttt gca aat caa aat aat tag                        1665
Leu Ser Lys Arg Phe Ala Asn Gln Asn Asn
545                 550
```

<210> 36
<211> 554
<212> PRT
<213> 丙酮丁醇梭菌

<400> 36

```
Leu Lys Ser Glu Tyr Thr Ile Gly Arg Tyr Leu Leu Asp Arg Leu Ser
1               5                   10                  15

Glu Leu Gly Ile Arg His Ile Phe Gly Val Pro Gly Asp Tyr Asn Leu
            20                  25                  30

Ser Phe Leu Asp Tyr Ile Met Glu Tyr Lys Gly Ile Asp Trp Val Gly
        35                  40                  45

Asn Cys Asn Glu Leu Asn Ala Gly Tyr Ala Ala Asp Gly Tyr Ala Arg
    50                  55                  60

Ile Asn Gly Ile Gly Ala Ile Leu Thr Thr Phe Gly Val Gly Glu Leu
65                  70                  75                  80

Ser Ala Ile Asn Ala Ile Ala Gly Ala Tyr Ala Glu Gln Val Pro Val
                85                  90                  95

Val Lys Ile Thr Gly Ile Pro Thr Ala Lys Val Arg Asp Asn Gly Leu
            100                 105                 110

Tyr Val His His Thr Leu Gly Asp Gly Arg Phe Asp His Phe Phe Glu
        115                 120                 125
```

```
Met Phe Arg Glu Val Thr Val Ala Glu Ala Leu Leu Ser Glu Glu Asn
    130                 135                 140

Ala Ala Gln Glu Ile Asp Arg Val Leu Ile Ser Cys Trp Arg Gln Lys
145                 150                 155                 160

Arg Pro Val Leu Ile Asn Leu Pro Ile Asp Val Tyr Asp Lys Pro Ile
                165                 170                 175

Asn Lys Pro Leu Lys Pro Leu Leu Asp Tyr Thr Ile Ser Ser Asn Lys
            180                 185                 190

Glu Ala Ala Cys Glu Phe Val Thr Glu Ile Val Pro Ile Ile Asn Arg
        195                 200                 205

Ala Lys Lys Pro Val Ile Leu Ala Asp Tyr Gly Val Tyr Arg Tyr Gln
    210                 215                 220

Val Gln His Val Leu Lys Asn Leu Ala Glu Lys Thr Gly Phe Pro Val
225                 230                 235                 240

Ala Thr Leu Ser Met Gly Lys Gly Val Phe Asn Glu Ala His Pro Gln
                245                 250                 255

Phe Ile Gly Val Tyr Asn Gly Asp Val Ser Ser Pro Tyr Leu Arg Gln
            260                 265                 270

Arg Val Asp Glu Ala Asp Cys Ile Ile Ser Val Gly Val Lys Leu Thr
        275                 280                 285

Asp Ser Thr Thr Gly Gly Phe Ser His Gly Phe Ser Lys Arg Asn Val
    290                 295                 300

Ile His Ile Asp Pro Phe Ser Ile Lys Ala Lys Gly Lys Lys Tyr Ala
305                 310                 315                 320

Pro Ile Thr Met Lys Asp Ala Leu Thr Glu Leu Thr Ser Lys Ile Glu
                325                 330                 335
```

```
His Arg Asn Phe Glu Asp Leu Asp Ile Lys Pro Tyr Lys Ser Asp Asn
            340                 345                 350

Gln Lys Tyr Phe Ala Lys Glu Lys Pro Ile Thr Gln Lys Arg Phe Phe
        355                 360                 365

Glu Arg Ile Ala His Phe Ile Lys Glu Lys Asp Val Leu Leu Ala Glu
    370                 375                 380

Gln Gly Thr Cys Phe Phe Gly Ala Ser Thr Ile Gln Leu Pro Lys Asp
385                 390                 395                 400

Ala Thr Phe Ile Gly Gln Pro Leu Trp Gly Ser Ile Gly Tyr Thr Leu
                405                 410                 415

Pro Ala Leu Leu Gly Ser Gln Leu Ala Asp Gln Lys Arg Arg Asn Ile
            420                 425                 430

Leu Leu Ile Gly Asp Gly Ala Phe Gln Met Thr Ala Gln Glu Ile Ser
        435                 440                 445

Thr Met Leu Arg Leu Gln Ile Lys Pro Ile Ile Phe Leu Ile Asn Asn
    450                 455                 460

Asp Gly Tyr Thr Ile Glu Arg Ala Ile His Gly Arg Glu Gln Val Tyr
465                 470                 475                 480

Asn Asn Ile Gln Met Trp Arg Tyr His Asn Val Pro Lys Val Leu Gly
                485                 490                 495

Pro Lys Glu Cys Ser Leu Thr Phe Lys Val Gln Ser Glu Thr Glu Leu
            500                 505                 510

Glu Lys Ala Leu Leu Val Ala Asp Lys Asp Cys Glu His Leu Ile Phe
        515                 520                 525

Ile Glu Val Val Met Asp Arg Tyr Asp Lys Pro Glu Pro Leu Glu Arg
    530                 535                 540
```

```
Leu Ser Lys Arg Phe Ala Asn Gln Asn Asn
545                 550
```

<210> 37
<211> 1056
<212> DNA
<213> 酿酒酵母

<220>
<221> CDS
<222> (1)..(1056)

<400> 37

```
atg cct tcg caa gtc att cct gaa aaa caa aag gct att gtc ttt tat       48
Met Pro Ser Gln Val Ile Pro Glu Lys Gln Lys Ala Ile Val Phe Tyr
1               5                   10                  15

gag aca gat gga aaa ttg gaa tat aaa gac gtc aca gtt ccg gaa cct       96
Glu Thr Asp Gly Lys Leu Glu Tyr Lys Asp Val Thr Val Pro Glu Pro
            20                  25                  30

aag cct aac gaa att tta gtc cac gtt aaa tat tct ggt gtt tgt cat      144
Lys Pro Asn Glu Ile Leu Val His Val Lys Tyr Ser Gly Val Cys His
        35                  40                  45

agt gac ttg cac gcg tgg cac ggt gat tgg cca ttt caa ttg aaa ttt      192
Ser Asp Leu His Ala Trp His Gly Asp Trp Pro Phe Gln Leu Lys Phe
    50                  55                  60

cca tta atc ggt ggt cac gaa ggt gct ggt gtt gtt gtt aag ttg gga      240
Pro Leu Ile Gly Gly His Glu Gly Ala Gly Val Val Val Lys Leu Gly
65                  70                  75                  80

tct aac gtt aag ggc tgg aaa gtc ggt gat ttt gca ggt ata aaa tgg      288
Ser Asn Val Lys Gly Trp Lys Val Gly Asp Phe Ala Gly Ile Lys Trp
                85                  90                  95

ttg aat ggg act tgc atg tcc tgt gaa tat tgt gaa gta ggt aat gaa      336
Leu Asn Gly Thr Cys Met Ser Cys Glu Tyr Cys Glu Val Gly Asn Glu
            100                 105                 110

tct caa tgt cct tat ttg gat ggt act ggc ttc aca cat gat ggt act      384
Ser Gln Cys Pro Tyr Leu Asp Gly Thr Gly Phe Thr His Asp Gly Thr
        115                 120                 125

ttt caa gaa tac gca act gcc gat gcc gtt caa gct gcc cat att cca      432
Phe Gln Glu Tyr Ala Thr Ala Asp Ala Val Gln Ala Ala His Ile Pro
    130                 135                 140
```

```
cca aac gtc aat ctt gct gaa gtt gcc cca atc ttg tgt gca ggt atc      480
Pro Asn Val Asn Leu Ala Glu Val Ala Pro Ile Leu Cys Ala Gly Ile
145                 150                 155                 160

act gtt tat aag gcg ttg aaa aga gcc aat gtg ata cca ggc caa tgg      528
Thr Val Tyr Lys Ala Leu Lys Arg Ala Asn Val Ile Pro Gly Gln Trp
                165                 170                 175

gtc act ata tcc ggt gca tgc ggt ggc ttg ggt tct ctg gca atc caa      576
Val Thr Ile Ser Gly Ala Cys Gly Gly Leu Gly Ser Leu Ala Ile Gln
            180                 185                 190

tac gcc ctt gct atg ggt tac agg gtc att ggt atc gat ggt ggt aat      624
Tyr Ala Leu Ala Met Gly Tyr Arg Val Ile Gly Ile Asp Gly Gly Asn
        195                 200                 205

gcc aag cga aag tta ttt gaa caa tta ggc gga gaa ata ttc atc gat      672
Ala Lys Arg Lys Leu Phe Glu Gln Leu Gly Gly Glu Ile Phe Ile Asp
    210                 215                 220

ttc acg gaa gaa aaa gac att gtt ggt gct ata ata aag gcc act aat      720
Phe Thr Glu Glu Lys Asp Ile Val Gly Ala Ile Ile Lys Ala Thr Asn
225                 230                 235                 240

ggc ggt tct cat gga gtt att aat gtg tct gtt tct gaa gca gct atc      768
Gly Gly Ser His Gly Val Ile Asn Val Ser Val Ser Glu Ala Ala Ile
                245                 250                 255

gag gct tct acg agg tat tgt agg ccc aat ggt act gtc gtc ctg gtt      816
Glu Ala Ser Thr Arg Tyr Cys Arg Pro Asn Gly Thr Val Val Leu Val
            260                 265                 270

ggt atg cca gct cat gct tac tgc aat tcc gat gtt ttc aat caa gtt      864
Gly Met Pro Ala His Ala Tyr Cys Asn Ser Asp Val Phe Asn Gln Val
        275                 280                 285

gta aaa tca atc tcc atc gtt gga tct tgt gtt gga aat aga gct gat      912
Val Lys Ser Ile Ser Ile Val Gly Ser Cys Val Gly Asn Arg Ala Asp
    290                 295                 300

aca agg gag gct tta gat ttc ttc gcc aga ggt ttg atc aaa tct ccg      960
Thr Arg Glu Ala Leu Asp Phe Phe Ala Arg Gly Leu Ile Lys Ser Pro
305                 310                 315                 320

atc cac tta gct ggc cta tcg gat gtt cct gaa att ttt gca aag atg     1008
Ile His Leu Ala Gly Leu Ser Asp Val Pro Glu Ile Phe Ala Lys Met
                325                 330                 335

gag aag ggt gaa att gtt ggt aga tat gtt gtt gag act tct aaa tga     1056
Glu Lys Gly Glu Ile Val Gly Arg Tyr Val Val Glu Thr Ser Lys
            340                 345                 350
```

<210> 38
<211> 351
<212> PRT
<213> 酿酒酵母

<400> 38

Met Pro Ser Gln Val Ile Pro Glu Lys Gln Lys Ala Ile Val Phe Tyr
1 5 10 15

Glu Thr Asp Gly Lys Leu Glu Tyr Lys Asp Val Thr Val Pro Glu Pro
20 25 30

Lys Pro Asn Glu Ile Leu Val His Val Lys Tyr Ser Gly Val Cys His
35 40 45

Ser Asp Leu His Ala Trp His Gly Asp Trp Pro Phe Gln Leu Lys Phe
50 55 60

Pro Leu Ile Gly Gly His Glu Gly Ala Gly Val Val Val Lys Leu Gly
65 70 75 80

Ser Asn Val Lys Gly Trp Lys Val Gly Asp Phe Ala Gly Ile Lys Trp
85 90 95

Leu Asn Gly Thr Cys Met Ser Cys Glu Tyr Cys Glu Val Gly Asn Glu
100 105 110

Ser Gln Cys Pro Tyr Leu Asp Gly Thr Gly Phe Thr His Asp Gly Thr
115 120 125

Phe Gln Glu Tyr Ala Thr Ala Asp Ala Val Gln Ala Ala His Ile Pro
130 135 140

Pro Asn Val Asn Leu Ala Glu Val Ala Pro Ile Leu Cys Ala Gly Ile
145 150 155 160

Thr Val Tyr Lys Ala Leu Lys Arg Ala Asn Val Ile Pro Gly Gln Trp
165 170 175

Val Thr Ile Ser Gly Ala Cys Gly Gly Leu Gly Ser Leu Ala Ile Gln
180 185 190

Tyr Ala Leu Ala Met Gly Tyr Arg Val Ile Gly Ile Asp Gly Gly Asn
195 200 205

Ala Lys Arg Lys Leu Phe Glu Gln Leu Gly Gly Glu Ile Phe Ile Asp
210 215 220

Phe Thr Glu Glu Lys Asp Ile Val Gly Ala Ile Ile Lys Ala Thr Asn
225 230 235 240

Gly Gly Ser His Gly Val Ile Asn Val Ser Val Ser Glu Ala Ala Ile
245 250 255

Glu Ala Ser Thr Arg Tyr Cys Arg Pro Asn Gly Thr Val Val Leu Val
260 265 270

Gly Met Pro Ala His Ala Tyr Cys Asn Ser Asp Val Phe Asn Gln Val
275 280 285

Val Lys Ser Ile Ser Ile Val Gly Ser Cys Val Gly Asn Arg Ala Asp
290 295 300

Thr Arg Glu Ala Leu Asp Phe Phe Ala Arg Gly Leu Ile Lys Ser Pro
305 310 315 320

Ile His Leu Ala Gly Leu Ser Asp Val Pro Glu Ile Phe Ala Lys Met
325 330 335

Glu Lys Gly Glu Ile Val Gly Arg Tyr Val Val Glu Thr Ser Lys
340 345 350

<210> 39
<211> 1725
<212> DNA
<213> 大肠杆菌

<220>

<221> CDS
<222> (1)..(1725)

<400> 39

```
atg gag atg ttg tct gga gcc gag atg gtc gtc cga tcg ctt atc gat      48
Met Glu Met Leu Ser Gly Ala Glu Met Val Val Arg Ser Leu Ile Asp
1               5                  10                  15

cag ggc gtt aaa caa gta ttc ggt tat ccc gga ggc gca gtc ctt gat      96
Gln Gly Val Lys Gln Val Phe Gly Tyr Pro Gly Gly Ala Val Leu Asp
            20                  25                  30

att tat gat gca ttg cat acc gtg ggt ggt att gat cat gta tta gtt     144
Ile Tyr Asp Ala Leu His Thr Val Gly Gly Ile Asp His Val Leu Val
        35                  40                  45

cgt cat gag cag gcg gcg gtg cat atg gcc gat ggc ctg gcg cgc gcg     192
Arg His Glu Gln Ala Ala Val His Met Ala Asp Gly Leu Ala Arg Ala
    50                  55                  60

acc ggg gaa gtc ggc gtc gtg ctg gta acg tcg ggt cca ggg gcg acc     240
Thr Gly Glu Val Gly Val Val Leu Val Thr Ser Gly Pro Gly Ala Thr
65                  70                  75                  80

aat gcg att act ggc atc gcc acc gct tat atg gat tcc att cca tta     288
Asn Ala Ile Thr Gly Ile Ala Thr Ala Tyr Met Asp Ser Ile Pro Leu
                85                  90                  95

gtt gtc ctt tcc ggg cag gta gcg acc tcg ttg ata ggt tac gat gcc     336
Val Val Leu Ser Gly Gln Val Ala Thr Ser Leu Ile Gly Tyr Asp Ala
            100                 105                 110

ttt cag gag tgc gac atg gtg ggg att tcg cga ccg gtg gtt aaa cac     384
Phe Gln Glu Cys Asp Met Val Gly Ile Ser Arg Pro Val Val Lys His
        115                 120                 125

agt ttt ctg gtt aag caa acg gaa gac att ccg cag gtg ctg aaa aag     432
Ser Phe Leu Val Lys Gln Thr Glu Asp Ile Pro Gln Val Leu Lys Lys
    130                 135                 140

gct ttc tgg ctg gcg gca agt ggt cgc cca gga cca gta gtc gtt gat     480
Ala Phe Trp Leu Ala Ala Ser Gly Arg Pro Gly Pro Val Val Val Asp
145                 150                 155                 160

tta ccg aaa gat att ctt aat ccg gcg aac aaa tta ccc tat gtc tgg     528
Leu Pro Lys Asp Ile Leu Asn Pro Ala Asn Lys Leu Pro Tyr Val Trp
                165                 170                 175

ccg gag tcg gtc agt atg cgt tct tac aat ccc act act acc gga cat     576
Pro Glu Ser Val Ser Met Arg Ser Tyr Asn Pro Thr Thr Thr Gly His
            180                 185                 190
```

```
aaa ggg caa att aag cgt gct ctg caa acg ctg gta gcg gca aaa aaa      624
Lys Gly Gln Ile Lys Arg Ala Leu Gln Thr Leu Val Ala Ala Lys Lys
        195                 200                 205

ccg gtt gtc tac gta ggc ggt ggg gca atc acg gcg ggc tgc cat cag      672
Pro Val Val Tyr Val Gly Gly Gly Ala Ile Thr Ala Gly Cys His Gln
    210                 215                 220

cag ttg aaa gaa acg gtg gag gcg ttg aat ctg ccc gtt gtt tgc tca      720
Gln Leu Lys Glu Thr Val Glu Ala Leu Asn Leu Pro Val Val Cys Ser
225                 230                 235                 240

ttg atg ggg ctg ggg gcg ttt ccg gca acg cat cgt cag gca ctg ggc      768
Leu Met Gly Leu Gly Ala Phe Pro Ala Thr His Arg Gln Ala Leu Gly
                245                 250                 255

atg ctg gga atg cac ggt acc tac gaa gcc aat atg acg atg cat aac      816
Met Leu Gly Met His Gly Thr Tyr Glu Ala Asn Met Thr Met His Asn
            260                 265                 270

gcg gat gtg att ttc gcc gtc ggg gta cga ttt gat gac cga acg acg      864
Ala Asp Val Ile Phe Ala Val Gly Val Arg Phe Asp Asp Arg Thr Thr
        275                 280                 285

aac aat ctg gca aag tac tgc cca aat gcc act gtt ctg cat atc gat      912
Asn Asn Leu Ala Lys Tyr Cys Pro Asn Ala Thr Val Leu His Ile Asp
    290                 295                 300

att gat cct act tcc att tct aaa acc gtg act gcg gat atc ccg att      960
Ile Asp Pro Thr Ser Ile Ser Lys Thr Val Thr Ala Asp Ile Pro Ile
305                 310                 315                 320

gtg ggg gat gct cgc cag gtc ctc gaa caa atg ctt gaa ctc ttg tcg     1008
Val Gly Asp Ala Arg Gln Val Leu Glu Gln Met Leu Glu Leu Leu Ser
                325                 330                 335

caa gaa tcc gcc cat caa cca ctg gat gag atc cgc gac tgg tgg cag     1056
Gln Glu Ser Ala His Gln Pro Leu Asp Glu Ile Arg Asp Trp Trp Gln
            340                 345                 350

caa att gaa cag tgg cgc gct cgt cag tgc ctg aaa tat gac act cac     1104
Gln Ile Glu Gln Trp Arg Ala Arg Gln Cys Leu Lys Tyr Asp Thr His
        355                 360                 365

agt gaa aag att aaa ccg cag gcg gtg atc gag act ctt tgg cgg ttg     1152
Ser Glu Lys Ile Lys Pro Gln Ala Val Ile Glu Thr Leu Trp Arg Leu
    370                 375                 380

acg aag gga gac gct tac gtg acg tcc gat gtc ggg cag cac cag atg     1200
Thr Lys Gly Asp Ala Tyr Val Thr Ser Asp Val Gly Gln His Gln Met
385                 390                 395                 400
```

```
ttt gct gca ctt tat tat cca ttc gac aaa ccg cgt cgc tgg atc aat      1248
Phe Ala Ala Leu Tyr Tyr Pro Phe Asp Lys Pro Arg Arg Trp Ile Asn
                405                 410                 415

tcc ggt ggc ctc ggc acg atg ggt ttt ggt tta cct gcg gca ctg ggc      1296
Ser Gly Gly Leu Gly Thr Met Gly Phe Gly Leu Pro Ala Ala Leu Gly
            420                 425                 430

gtc aaa atg gcg ttg cca gaa gaa acc gtg gtt tgc gtc act ggc gac      1344
Val Lys Met Ala Leu Pro Glu Glu Thr Val Val Cys Val Thr Gly Asp
        435                 440                 445

ggc agt att cag atg aac atc cag gaa ctg tct acc gcg ttg caa tac      1392
Gly Ser Ile Gln Met Asn Ile Gln Glu Leu Ser Thr Ala Leu Gln Tyr
    450                 455                 460

gag ttg ccc gta ctg gtg gtg aat ctc aat aac cgc tat ctg ggg atg      1440
Glu Leu Pro Val Leu Val Val Asn Leu Asn Asn Arg Tyr Leu Gly Met
465                 470                 475                 480

gtg aag cag tgg cag gac atg atc tat tcc ggc cgt cat tca caa tct      1488
Val Lys Gln Trp Gln Asp Met Ile Tyr Ser Gly Arg His Ser Gln Ser
                485                 490                 495

tat atg caa tcg cta ccc gat ttc gtc cgt ctg gcg gaa gcc tat ggg      1536
Tyr Met Gln Ser Leu Pro Asp Phe Val Arg Leu Ala Glu Ala Tyr Gly
            500                 505                 510

cat gtc ggg atc cag att tct cat ccg cat gag ctg gaa agc aaa ctt      1584
His Val Gly Ile Gln Ile Ser His Pro His Glu Leu Glu Ser Lys Leu
        515                 520                 525

agc gag gcg ctg gaa cag gtg cgc aat aat cgc ctg gtg ttt gtt gat      1632
Ser Glu Ala Leu Glu Gln Val Arg Asn Asn Arg Leu Val Phe Val Asp
    530                 535                 540

gtt acc gtc gat ggc agc gag cac gtc tac ccg atg cag att cgc ggg      1680
Val Thr Val Asp Gly Ser Glu His Val Tyr Pro Met Gln Ile Arg Gly
545                 550                 555                 560

ggc gga atg gat gaa atg tgg tta agc aaa acg gag aga acc tga          1725
Gly Gly Met Asp Glu Met Trp Leu Ser Lys Thr Glu Arg Thr
                565                 570
```

<210> 40
<211> 574
<212> PRT
<213> 大肠杆菌

<400> 40

```
Met Glu Met Leu Ser Gly Ala Glu Met Val Val Arg Ser Leu Ile Asp
1               5                  10                  15

Gln Gly Val Lys Gln Val Phe Gly Tyr Pro Gly Gly Ala Val Leu Asp
            20                  25                  30

Ile Tyr Asp Ala Leu His Thr Val Gly Gly Ile Asp His Val Leu Val
        35                  40                  45

Arg His Glu Gln Ala Ala Val His Met Ala Asp Gly Leu Ala Arg Ala
    50                  55                  60

Thr Gly Glu Val Gly Val Val Leu Val Thr Ser Gly Pro Gly Ala Thr
65                  70                  75                  80

Asn Ala Ile Thr Gly Ile Ala Thr Ala Tyr Met Asp Ser Ile Pro Leu
                85                  90                  95

Val Val Leu Ser Gly Gln Val Ala Thr Ser Leu Ile Gly Tyr Asp Ala
            100                 105                 110

Phe Gln Glu Cys Asp Met Val Gly Ile Ser Arg Pro Val Val Lys His
        115                 120                 125

Ser Phe Leu Val Lys Gln Thr Glu Asp Ile Pro Gln Val Leu Lys Lys
    130                 135                 140

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

Leu Pro Lys Asp Ile Leu Asn Pro Ala Asn Lys Leu Pro Tyr Val Trp
                165                 170                 175

Pro Glu Ser Val Ser Met Arg Ser Tyr Asn Pro Thr Thr Thr Gly His
            180                 185                 190

Lys Gly Gln Ile Lys Arg Ala Leu Gln Thr Leu Val Ala Ala Lys Lys
        195                 200                 205
```

Pro Val Val Tyr Val Gly Gly Gly Ala Ile Thr Ala Gly Cys His Gln
210 215 220

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

Leu Met Gly Leu Gly Ala Phe Pro Ala Thr His Arg Gln Ala Leu Gly
245 250 255

Met Leu Gly Met His Gly Thr Tyr Glu Ala Asn Met Thr Met His Asn
260 265 270

Ala Asp Val Ile Phe Ala Val Gly Val Arg Phe Asp Asp Arg Thr Thr
275 280 285

Asn Asn Leu Ala Lys Tyr Cys Pro Asn Ala Thr Val Leu His Ile Asp
290 295 300

Ile Asp Pro Thr Ser Ile Ser Lys Thr Val Thr Ala Asp Ile Pro Ile
305 310 315 320

Val Gly Asp Ala Arg Gln Val Leu Glu Gln Met Leu Glu Leu Leu Ser
325 330 335

Gln Glu Ser Ala His Gln Pro Leu Asp Glu Ile Arg Asp Trp Trp Gln
340 345 350

Gln Ile Glu Gln Trp Arg Ala Arg Gln Cys Leu Lys Tyr Asp Thr His
355 360 365

Ser Glu Lys Ile Lys Pro Gln Ala Val Ile Glu Thr Leu Trp Arg Leu
370 375 380

Thr Lys Gly Asp Ala Tyr Val Thr Ser Asp Val Gly Gln His Gln Met
385 390 395 400

Phe Ala Ala Leu Tyr Tyr Pro Phe Asp Lys Pro Arg Arg Trp Ile Asn
405 410 415

Ser Gly Gly Leu Gly Thr Met Gly Phe Gly Leu Pro Ala Ala Leu Gly
420 425 430

Val Lys Met Ala Leu Pro Glu Glu Thr Val Val Cys Val Thr Gly Asp
435 440 445

Gly Ser Ile Gln Met Asn Ile Gln Glu Leu Ser Thr Ala Leu Gln Tyr
450 455 460

Glu Leu Pro Val Leu Val Val Asn Leu Asn Asn Arg Tyr Leu Gly Met
465 470 475 480

Val Lys Gln Trp Gln Asp Met Ile Tyr Ser Gly Arg His Ser Gln Ser
485 490 495

Tyr Met Gln Ser Leu Pro Asp Phe Val Arg Leu Ala Glu Ala Tyr Gly
500 505 510

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

Ser Glu Ala Leu Glu Gln Val Arg Asn Asn Arg Leu Val Phe Val Asp
530 535 540

Val Thr Val Asp Gly Ser Glu His Val Tyr Pro Met Gln Ile Arg Gly
545 550 555 560

Gly Gly Met Asp Glu Met Trp Leu Ser Lys Thr Glu Arg Thr
565 570

<210> 41
<211> 492
<212> DNA
<213> 大肠杆菌

<220>
<221> CDS
<222> (1)..(492)

<400> 41

```
atg cgc cgg ata tta tca gtc tta ctc gaa aat gaa tca ggc gcg tta      48
Met Arg Arg Ile Leu Ser Val Leu Leu Glu Asn Glu Ser Gly Ala Leu
1               5                  10                  15

tcc cgc gtg att ggc ctt ttt tcc cag cgt ggc tac aac att gaa agc      96
Ser Arg Val Ile Gly Leu Phe Ser Gln Arg Gly Tyr Asn Ile Glu Ser
            20                  25                  30

ctg acc gtt gcg cca acc gac gat ccg aca tta tcg cgt atg acc atc     144
Leu Thr Val Ala Pro Thr Asp Asp Pro Thr Leu Ser Arg Met Thr Ile
        35                  40                  45

cag acc gtg ggc gat gaa aaa gta ctt gag cag atc gaa aag caa tta     192
Gln Thr Val Gly Asp Glu Lys Val Leu Glu Gln Ile Glu Lys Gln Leu
    50                  55                  60

cac aaa ctg gtc gat gtc ttg cgc gtg agt gag ttg ggg cag ggc gcg     240
His Lys Leu Val Asp Val Leu Arg Val Ser Glu Leu Gly Gln Gly Ala
65                  70                  75                  80

cat gtt gag cgg gaa atc atg ctg gtg aaa att cag gcc agc ggt tac     288
His Val Glu Arg Glu Ile Met Leu Val Lys Ile Gln Ala Ser Gly Tyr
                85                  90                  95

ggg cgt gac gaa gtg aaa cgt aat acg gaa ata ttc cgt ggg caa att     336
Gly Arg Asp Glu Val Lys Arg Asn Thr Glu Ile Phe Arg Gly Gln Ile
            100                 105                 110

atc gat gtc aca ccc tcg ctt tat acc gtt caa tta gca ggc acc agc     384
Ile Asp Val Thr Pro Ser Leu Tyr Thr Val Gln Leu Ala Gly Thr Ser
        115                 120                 125

ggt aag ctt gat gca ttt tta gca tcg att cgc gat gtg gcg aaa att     432
Gly Lys Leu Asp Ala Phe Leu Ala Ser Ile Arg Asp Val Ala Lys Ile
    130                 135                 140

gtg gag gtt gct cgc tct ggt gtg gtc gga ctt tcg cgc ggc gat aaa     480
Val Glu Val Ala Arg Ser Gly Val Val Gly Leu Ser Arg Gly Asp Lys
145                 150                 155                 160

ata atg cgt tga                                                    492
Ile Met Arg
```

<210> 42
<211> 163
<212> PRT
<213> 大肠杆菌

<400> 42

```
Met Arg Arg Ile Leu Ser Val Leu Leu Glu Asn Glu Ser Gly Ala Leu
1               5                  10                  15

Ser Arg Val Ile Gly Leu Phe Ser Gln Arg Gly Tyr Asn Ile Glu Ser
            20                  25                  30

Leu Thr Val Ala Pro Thr Asp Asp Pro Thr Leu Ser Arg Met Thr Ile
        35                  40                  45

Gln Thr Val Gly Asp Glu Lys Val Leu Glu Gln Ile Glu Lys Gln Leu
    50                  55                  60

His Lys Leu Val Asp Val Leu Arg Val Ser Glu Leu Gly Gln Gly Ala
65                  70                  75                  80

His Val Glu Arg Glu Ile Met Leu Val Lys Ile Gln Ala Ser Gly Tyr
                85                  90                  95

Gly Arg Asp Glu Val Lys Arg Asn Thr Glu Ile Phe Arg Gly Gln Ile
            100                 105                 110

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

Gly Lys Leu Asp Ala Phe Leu Ala Ser Ile Arg Asp Val Ala Lys Ile
    130                 135                 140

Val Glu Val Ala Arg Ser Gly Val Val Gly Leu Ser Arg Gly Asp Lys
145                 150                 155                 160

Ile Met Arg
```

<210> 43
<211> 1476
<212> DNA
<213> 大肠杆菌

<220>
<221> CDS
<222> (1)..(1476)

<400> 43

```
atg gct aac tac ttc aat aca ctg aat ctg cgc cag cag ctg gca cag      48
Met Ala Asn Tyr Phe Asn Thr Leu Asn Leu Arg Gln Gln Leu Ala Gln
1               5                  10                  15

ctg ggc aaa tgt cgc ttt atg ggc cgc gat gaa ttc gcc gat ggc gcg      96
Leu Gly Lys Cys Arg Phe Met Gly Arg Asp Glu Phe Ala Asp Gly Ala
            20                  25                  30

agc tac ctt cag ggt aaa aaa gta gtc atc gtc ggc tgt ggc gca cag     144
Ser Tyr Leu Gln Gly Lys Lys Val Val Ile Val Gly Cys Gly Ala Gln
        35                  40                  45

ggt ctg aac cag ggc ctg aac atg cgt gat tct ggt ctc gat atc tcc     192
Gly Leu Asn Gln Gly Leu Asn Met Arg Asp Ser Gly Leu Asp Ile Ser
    50                  55                  60

tac gct ctg cgt aaa gaa gcg att gcc gag aag cgc gcg tcc tgg cgt     240
Tyr Ala Leu Arg Lys Glu Ala Ile Ala Glu Lys Arg Ala Ser Trp Arg
65                  70                  75                  80

aaa gcg acc gaa aat ggt ttt aaa gtg ggt act tac gaa gaa ctg atc     288
Lys Ala Thr Glu Asn Gly Phe Lys Val Gly Thr Tyr Glu Glu Leu Ile
                85                  90                  95

cca cag gcg gat ctg gtg att aac ctg acg ccg gac aag cag cac tct     336
Pro Gln Ala Asp Leu Val Ile Asn Leu Thr Pro Asp Lys Gln His Ser
            100                 105                 110

gat gta gtg cgc acc gta cag cca ctg atg aaa gac ggc gcg gcg ctg     384
Asp Val Val Arg Thr Val Gln Pro Leu Met Lys Asp Gly Ala Ala Leu
        115                 120                 125

ggc tac tcg cac ggt ttc aac atc gtc gaa gtg ggc gag cag atc cgt     432
Gly Tyr Ser His Gly Phe Asn Ile Val Glu Val Gly Glu Gln Ile Arg
    130                 135                 140

aaa gat atc acc gta gtg atg gtt gcg ccg aaa tgc cca ggc acc gaa     480
Lys Asp Ile Thr Val Val Met Val Ala Pro Lys Cys Pro Gly Thr Glu
145                 150                 155                 160

gtg cgt gaa gag tac aaa cgt ggg ttc ggc gta ccg acg ctg att gcc     528
Val Arg Glu Glu Tyr Lys Arg Gly Phe Gly Val Pro Thr Leu Ile Ala
                165                 170                 175
```

```
gtt cac ccg gaa aac gat ccg aaa ggc gaa ggc atg gcg att gcc aaa      576
Val His Pro Glu Asn Asp Pro Lys Gly Glu Gly Met Ala Ile Ala Lys
            180                 185                 190

gcc tgg gcg gct gca acc ggt ggt cac cgt gcg ggt gtg ctg gaa tcg      624
Ala Trp Ala Ala Ala Thr Gly Gly His Arg Ala Gly Val Leu Glu Ser
        195                 200                 205

tcc ttc gtt gcg gaa gtg aaa tct gac ctg atg ggc gag caa acc atc      672
Ser Phe Val Ala Glu Val Lys Ser Asp Leu Met Gly Glu Gln Thr Ile
    210                 215                 220

ctg tgc ggt atg ttg cag gct ggc tct ctg ctg tgc ttc gac aag ctg      720
Leu Cys Gly Met Leu Gln Ala Gly Ser Leu Leu Cys Phe Asp Lys Leu
225                 230                 235                 240

gtg gaa gaa ggt acc gat cca gca tac gca gaa aaa ctg att cag ttc      768
Val Glu Glu Gly Thr Asp Pro Ala Tyr Ala Glu Lys Leu Ile Gln Phe
                245                 250                 255

ggt tgg gaa acc atc acc gaa gca ctg aaa cag ggc ggc atc acc ctg      816
Gly Trp Glu Thr Ile Thr Glu Ala Leu Lys Gln Gly Gly Ile Thr Leu
            260                 265                 270

atg atg gac cgt ctc tct aac ccg gcg aaa ctg cgt gct tat gcg ctt      864
Met Met Asp Arg Leu Ser Asn Pro Ala Lys Leu Arg Ala Tyr Ala Leu
        275                 280                 285

tct gaa cag ctg aaa gag atc atg gca ccc ctg ttc cag aaa cat atg      912
Ser Glu Gln Leu Lys Glu Ile Met Ala Pro Leu Phe Gln Lys His Met
    290                 295                 300

gac gac atc atc tcc ggc gaa ttc tct tcc ggt atg atg gcg gac tgg      960
Asp Asp Ile Ile Ser Gly Glu Phe Ser Ser Gly Met Met Ala Asp Trp
305                 310                 315                 320

gcc aac gat gat aag aaa ctg ctg acc tgg cgt gaa gag acc ggc aaa     1008
Ala Asn Asp Asp Lys Lys Leu Leu Thr Trp Arg Glu Glu Thr Gly Lys
                325                 330                 335

acc gcg ttt gaa acc gcg ccg cag tat gaa ggc aaa atc ggc gag cag     1056
Thr Ala Phe Glu Thr Ala Pro Gln Tyr Glu Gly Lys Ile Gly Glu Gln
            340                 345                 350

gag tac ttc gat aaa ggc gta ctg atg att gcg atg gtg aaa gcg ggc     1104
Glu Tyr Phe Asp Lys Gly Val Leu Met Ile Ala Met Val Lys Ala Gly
        355                 360                 365

gtt gaa ctg gcg ttc gaa acc atg gtc gat tcc ggc atc att gaa gag     1152
Val Glu Leu Ala Phe Glu Thr Met Val Asp Ser Gly Ile Ile Glu Glu
    370                 375                 380
```

```
tct gca tat tat gaa tca ctg cac gag ctg ccg ctg att gcc aac acc      1200
Ser Ala Tyr Tyr Glu Ser Leu His Glu Leu Pro Leu Ile Ala Asn Thr
385                 390                 395                 400

atc gcc cgt aag cgt ctg tac gaa atg aac gtg gtt atc tct gat acc      1248
Ile Ala Arg Lys Arg Leu Tyr Glu Met Asn Val Val Ile Ser Asp Thr
                405                 410                 415

gct gag tac ggt aac tat ctg ttc tct tac gct tgt gtg ccg ttg ctg      1296
Ala Glu Tyr Gly Asn Tyr Leu Phe Ser Tyr Ala Cys Val Pro Leu Leu
            420                 425                 430

aaa ccg ttt atg gca gag ctg caa ccg ggc gac ctg ggt aaa gct att      1344
Lys Pro Phe Met Ala Glu Leu Gln Pro Gly Asp Leu Gly Lys Ala Ile
        435                 440                 445

ccg gaa ggc gcg gta gat aac ggg caa ctg cgt gat gtg aac gaa gcg      1392
Pro Glu Gly Ala Val Asp Asn Gly Gln Leu Arg Asp Val Asn Glu Ala
    450                 455                 460

att cgc agc cat gcg att gag cag gta ggt aag aaa ctg cgc ggc tat      1440
Ile Arg Ser His Ala Ile Glu Gln Val Gly Lys Lys Leu Arg Gly Tyr
465                 470                 475                 480

atg aca gat atg aaa cgt att gct gtt gcg ggt taa                      1476
Met Thr Asp Met Lys Arg Ile Ala Val Ala Gly
                485                 490
```

<210> 44
<211> 491
<212> PRT
<213> 大肠杆菌

<400> 44

```
Met Ala Asn Tyr Phe Asn Thr Leu Asn Leu Arg Gln Gln Leu Ala Gln
1               5                   10                  15

Leu Gly Lys Cys Arg Phe Met Gly Arg Asp Glu Phe Ala Asp Gly Ala
            20                  25                  30

Ser Tyr Leu Gln Gly Lys Lys Val Val Ile Val Gly Cys Gly Ala Gln
        35                  40                  45

Gly Leu Asn Gln Gly Leu Asn Met Arg Asp Ser Gly Leu Asp Ile Ser
    50                  55                  60
```

Tyr Ala Leu Arg Lys Glu Ala Ile Ala Glu Lys Arg Ala Ser Trp Arg
65 70 75 80

Lys Ala Thr Glu Asn Gly Phe Lys Val Gly Thr Tyr Glu Glu Leu Ile
85 90 95

Pro Gln Ala Asp Leu Val Ile Asn Leu Thr Pro Asp Lys Gln His Ser
100 105 110

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

Gly Tyr Ser His Gly Phe Asn Ile Val Glu Val Gly Glu Gln Ile Arg
130 135 140

Lys Asp Ile Thr Val Val Met Val Ala Pro Lys Cys Pro Gly Thr Glu
145 150 155 160

Val Arg Glu Glu Tyr Lys Arg Gly Phe Gly Val Pro Thr Leu Ile Ala
165 170 175

Val His Pro Glu Asn Asp Pro Lys Gly Glu Gly Met Ala Ile Ala Lys
180 185 190

Ala Trp Ala Ala Ala Thr Gly Gly His Arg Ala Gly Val Leu Glu Ser
195 200 205

Ser Phe Val Ala Glu Val Lys Ser Asp Leu Met Gly Glu Gln Thr Ile
210 215 220

Leu Cys Gly Met Leu Gln Ala Gly Ser Leu Leu Cys Phe Asp Lys Leu
225 230 235 240

Val Glu Glu Gly Thr Asp Pro Ala Tyr Ala Glu Lys Leu Ile Gln Phe
245 250 255

Gly Trp Glu Thr Ile Thr Glu Ala Leu Lys Gln Gly Gly Ile Thr Leu
260 265 270

Met Met Asp Arg Leu Ser Asn Pro Ala Lys Leu Arg Ala Tyr Ala Leu
275 280 285

Ser Glu Gln Leu Lys Glu Ile Met Ala Pro Leu Phe Gln Lys His Met
290 295 300

Asp Asp Ile Ile Ser Gly Glu Phe Ser Ser Gly Met Met Ala Asp Trp
305 310 315 320

Ala Asn Asp Asp Lys Lys Leu Leu Thr Trp Arg Glu Glu Thr Gly Lys
325 330 335

Thr Ala Phe Glu Thr Ala Pro Gln Tyr Glu Gly Lys Ile Gly Glu Gln
340 345 350

Glu Tyr Phe Asp Lys Gly Val Leu Met Ile Ala Met Val Lys Ala Gly
355 360 365

Val Glu Leu Ala Phe Glu Thr Met Val Asp Ser Gly Ile Ile Glu Glu
370 375 380

Ser Ala Tyr Tyr Glu Ser Leu His Glu Leu Pro Leu Ile Ala Asn Thr
385 390 395 400

Ile Ala Arg Lys Arg Leu Tyr Glu Met Asn Val Val Ile Ser Asp Thr
405 410 415

Ala Glu Tyr Gly Asn Tyr Leu Phe Ser Tyr Ala Cys Val Pro Leu Leu
420 425 430

Lys Pro Phe Met Ala Glu Leu Gln Pro Gly Asp Leu Gly Lys Ala Ile
435 440 445

Pro Glu Gly Ala Val Asp Asn Gly Gln Leu Arg Asp Val Asn Glu Ala
450 455 460

Ile Arg Ser His Ala Ile Glu Gln Val Gly Lys Lys Leu Arg Gly Tyr
465 470 475 480

Met Thr Asp Met Lys Arg Ile Ala Val Ala Gly
485 490

<210> 45
<211> 1851
<212> DNA
<213> 大肠杆菌

<220>
<221> CDS
<222> (1)..(1851)

<400> 45
atg cct aag tac cgt tcc gcc acc acc act cat ggt cgt aat atg gcg 48
Met Pro Lys Tyr Arg Ser Ala Thr Thr Thr His Gly Arg Asn Met Ala
1 5 10 15

ggt gct cgt gcg ctg tgg cgc gcc acc gga atg acc gac gcc gat ttc 96
Gly Ala Arg Ala Leu Trp Arg Ala Thr Gly Met Thr Asp Ala Asp Phe
20 25 30

ggt aag ccg att atc gcg gtt gtg aac tcg ttc acc caa ttt gta ccg 144
Gly Lys Pro Ile Ile Ala Val Val Asn Ser Phe Thr Gln Phe Val Pro
35 40 45

ggt cac gtc cat ctg cgc gat ctc ggt aaa ctg gtc gcc gaa caa att 192
Gly His Val His Leu Arg Asp Leu Gly Lys Leu Val Ala Glu Gln Ile
50 55 60

gaa gcg gct ggc ggc gtt gcc aaa gag ttc aac acc att gcg gtg gat 240
Glu Ala Ala Gly Gly Val Ala Lys Glu Phe Asn Thr Ile Ala Val Asp
65 70 75 80

gat ggg att gcc atg ggc cac ggg ggg atg ctt tat tca ctg cca tct 288
Asp Gly Ile Ala Met Gly His Gly Gly Met Leu Tyr Ser Leu Pro Ser
85 90 95

cgc gaa ctg atc gct gat tcc gtt gag tat atg gtc aac gcc cac tgc 336
Arg Glu Leu Ile Ala Asp Ser Val Glu Tyr Met Val Asn Ala His Cys
100 105 110

gcc gac gcc atg gtc tgc atc tct aac tgc gac aaa atc acc ccg ggg 384
Ala Asp Ala Met Val Cys Ile Ser Asn Cys Asp Lys Ile Thr Pro Gly
115 120 125

atg ctg atg gct tcc ctg cgc ctg aat att ccg gtg atc ttt gtt tcc 432
Met Leu Met Ala Ser Leu Arg Leu Asn Ile Pro Val Ile Phe Val Ser
130 135 140

```
ggc ggc ccg atg gag gcc ggg aaa acc aaa ctt tcc gat cag atc atc      480
Gly Gly Pro Met Glu Ala Gly Lys Thr Lys Leu Ser Asp Gln Ile Ile
145                 150                 155                 160

aag ctc gat ctg gtt gat gcg atg atc cag ggc gca gac ccg aaa gta      528
Lys Leu Asp Leu Val Asp Ala Met Ile Gln Gly Ala Asp Pro Lys Val
                165                 170                 175

tct gac tcc cag agc gat cag gtt gaa cgt tcc gcg tgt ccg acc tgc      576
Ser Asp Ser Gln Ser Asp Gln Val Glu Arg Ser Ala Cys Pro Thr Cys
            180                 185                 190

ggt tcc tgc tcc ggg atg ttt acc gct aac tca atg aac tgc ctg acc      624
Gly Ser Cys Ser Gly Met Phe Thr Ala Asn Ser Met Asn Cys Leu Thr
        195                 200                 205

gaa gcg ctg ggc ctg tcg cag ccg ggc aac ggc tcg ctg ctg gca acc      672
Glu Ala Leu Gly Leu Ser Gln Pro Gly Asn Gly Ser Leu Leu Ala Thr
    210                 215                 220

cac gcc gac cgt aag cag ctg ttc ctt aat gct ggt aaa cgc att gtt      720
His Ala Asp Arg Lys Gln Leu Phe Leu Asn Ala Gly Lys Arg Ile Val
225                 230                 235                 240

gaa ttg acc aaa cgt tat tac gag caa aac gac gaa agt gca ctg ccg      768
Glu Leu Thr Lys Arg Tyr Tyr Glu Gln Asn Asp Glu Ser Ala Leu Pro
                245                 250                 255

cgt aat atc gcc agt aag gcg gcg ttt gaa aac gcc atg acg ctg gat      816
Arg Asn Ile Ala Ser Lys Ala Ala Phe Glu Asn Ala Met Thr Leu Asp
            260                 265                 270

atc gcg atg ggt gga tcg act aac acc gta ctt cac ctg ctg gcg gcg      864
Ile Ala Met Gly Gly Ser Thr Asn Thr Val Leu His Leu Leu Ala Ala
        275                 280                 285

gcg cag gaa gcg gaa atc gac ttc acc atg agt gat atc gat aag ctt      912
Ala Gln Glu Ala Glu Ile Asp Phe Thr Met Ser Asp Ile Asp Lys Leu
    290                 295                 300

tcc cgc aag gtt cca cag ctg tgt aaa gtt gcg ccg agc acc cag aaa      960
Ser Arg Lys Val Pro Gln Leu Cys Lys Val Ala Pro Ser Thr Gln Lys
305                 310                 315                 320

tac cat atg gaa gat gtt cac cgt gct ggt ggt gtt atc ggt att ctc     1008
Tyr His Met Glu Asp Val His Arg Ala Gly Gly Val Ile Gly Ile Leu
                325                 330                 335

ggc gaa ctg gat cgc gcg ggg tta ctg aac cgt gat gtg aaa aac gta     1056
Gly Glu Leu Asp Arg Ala Gly Leu Leu Asn Arg Asp Val Lys Asn Val
            340                 345                 350
```

```
ctt ggc ctg acg ttg ccg caa acg ctg gaa caa tac gac gtt atg ctg      1104
Leu Gly Leu Thr Leu Pro Gln Thr Leu Glu Gln Tyr Asp Val Met Leu
        355                 360                 365

acc cag gat gac gcg gta aaa aat atg ttc cgc gca ggt cct gca ggc      1152
Thr Gln Asp Asp Ala Val Lys Asn Met Phe Arg Ala Gly Pro Ala Gly
    370                 375                 380

att cgt acc aca cag gca ttc tcg caa gat tgc cgt tgg gat acg ctg      1200
Ile Arg Thr Thr Gln Ala Phe Ser Gln Asp Cys Arg Trp Asp Thr Leu
385                 390                 395                 400

gac gac gat cgc gcc aat ggc tgt atc cgc tcg ctg gaa cac gcc tac      1248
Asp Asp Asp Arg Ala Asn Gly Cys Ile Arg Ser Leu Glu His Ala Tyr
                405                 410                 415

agc aaa gac ggc ggc ctg gcg gtg ctc tac ggt aac ttt gcg gaa aac      1296
Ser Lys Asp Gly Gly Leu Ala Val Leu Tyr Gly Asn Phe Ala Glu Asn
            420                 425                 430

ggc tgc atc gtg aaa acg gca ggc gtc gat gac agc atc ctc aaa ttc      1344
Gly Cys Ile Val Lys Thr Ala Gly Val Asp Asp Ser Ile Leu Lys Phe
        435                 440                 445

acc ggc ccg gcg aaa gtg tac gaa agc cag gac gat gcg gta gaa gcg      1392
Thr Gly Pro Ala Lys Val Tyr Glu Ser Gln Asp Asp Ala Val Glu Ala
    450                 455                 460

att ctc ggc ggt aaa gtt gtc gcc gga gat gtg gta gta att cgc tat      1440
Ile Leu Gly Gly Lys Val Val Ala Gly Asp Val Val Val Ile Arg Tyr
465                 470                 475                 480

gaa ggc ccg aaa ggc ggt ccg ggg atg cag gaa atg ctc tac cca acc      1488
Glu Gly Pro Lys Gly Gly Pro Gly Met Gln Glu Met Leu Tyr Pro Thr
                485                 490                 495

agc ttc ctg aaa tca atg ggt ctc ggc aaa gcc tgt gcg ctg atc acc      1536
Ser Phe Leu Lys Ser Met Gly Leu Gly Lys Ala Cys Ala Leu Ile Thr
            500                 505                 510

gac ggt cgt ttc tct ggt ggc acc tct ggt ctt tcc atc ggc cac gtc      1584
Asp Gly Arg Phe Ser Gly Gly Thr Ser Gly Leu Ser Ile Gly His Val
        515                 520                 525

tca ccg gaa gcg gca agc ggc ggc agc att ggc ctg att gaa gat ggt      1632
Ser Pro Glu Ala Ala Ser Gly Gly Ser Ile Gly Leu Ile Glu Asp Gly
    530                 535                 540

gac ctg atc gct atc gac atc ccg aac cgt ggc att cag tta cag gta      1680
Asp Leu Ile Ala Ile Asp Ile Pro Asn Arg Gly Ile Gln Leu Gln Val
545                 550                 555                 560
```

```
agc gat gcc gaa ctg gcg gcg cgt cgt gaa gcg cag gac gct cga ggt      1728
Ser Asp Ala Glu Leu Ala Ala Arg Arg Glu Ala Gln Asp Ala Arg Gly
                565                 570                 575

gac aaa gcc tgg acg ccg aaa aat cgt gaa cgt cag gtc tcc ttt gcc      1776
Asp Lys Ala Trp Thr Pro Lys Asn Arg Glu Arg Gln Val Ser Phe Ala
            580                 585                 590

ctg cgt gct tat gcc agc ctg gca acc agc gcc gac aaa ggc gcg gtg      1824
Leu Arg Ala Tyr Ala Ser Leu Ala Thr Ser Ala Asp Lys Gly Ala Val
        595                 600                 605

cgc gat aaa tcg aaa ctg ggg ggt taa                                 1851
Arg Asp Lys Ser Lys Leu Gly Gly
    610                 615
```

<210> 46
<211> 616
<212> PRT
<213> 大肠杆菌

<400> 46

```
Met Pro Lys Tyr Arg Ser Ala Thr Thr Thr His Gly Arg Asn Met Ala
1               5                   10                  15

Gly Ala Arg Ala Leu Trp Arg Ala Thr Gly Met Thr Asp Ala Asp Phe
            20                  25                  30

Gly Lys Pro Ile Ile Ala Val Val Asn Ser Phe Thr Gln Phe Val Pro
        35                  40                  45

Gly His Val His Leu Arg Asp Leu Gly Lys Leu Val Ala Glu Gln Ile
    50                  55                  60

Glu Ala Ala Gly Gly Val Ala Lys Glu Phe Asn Thr Ile Ala Val Asp
65                  70                  75                  80

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

Arg Glu Leu Ile Ala Asp Ser Val Glu Tyr Met Val Asn Ala His Cys
            100                 105                 110
```

Ala Asp Ala Met Val Cys Ile Ser Asn Cys Asp Lys Ile Thr Pro Gly
115 120 125

Met Leu Met Ala Ser Leu Arg Leu Asn Ile Pro Val Ile Phe Val Ser
130 135 140

Gly Gly Pro Met Glu Ala Gly Lys Thr Lys Leu Ser Asp Gln Ile Ile
145 150 155 160

Lys Leu Asp Leu Val Asp Ala Met Ile Gln Gly Ala Asp Pro Lys Val
165 170 175

Ser Asp Ser Gln Ser Asp Gln Val Glu Arg Ser Ala Cys Pro Thr Cys
180 185 190

Gly Ser Cys Ser Gly Met Phe Thr Ala Asn Ser Met Asn Cys Leu Thr
195 200 205

Glu Ala Leu Gly Leu Ser Gln Pro Gly Asn Gly Ser Leu Leu Ala Thr
210 215 220

His Ala Asp Arg Lys Gln Leu Phe Leu Asn Ala Gly Lys Arg Ile Val
225 230 235 240

Glu Leu Thr Lys Arg Tyr Tyr Glu Gln Asn Asp Glu Ser Ala Leu Pro
245 250 255

Arg Asn Ile Ala Ser Lys Ala Ala Phe Glu Asn Ala Met Thr Leu Asp
260 265 270

Ile Ala Met Gly Gly Ser Thr Asn Thr Val Leu His Leu Leu Ala Ala
275 280 285

Ala Gln Glu Ala Glu Ile Asp Phe Thr Met Ser Asp Ile Asp Lys Leu
290 295 300

Ser Arg Lys Val Pro Gln Leu Cys Lys Val Ala Pro Ser Thr Gln Lys
305 310 315 320

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

Gly Glu Leu Asp Arg Ala Gly Leu Leu Asn Arg Asp Val Lys Asn Val
340 345 350

Leu Gly Leu Thr Leu Pro Gln Thr Leu Glu Gln Tyr Asp Val Met Leu
355 360 365

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

Ile Arg Thr Thr Gln Ala Phe Ser Gln Asp Cys Arg Trp Asp Thr Leu
385 390 395 400

Asp Asp Asp Arg Ala Asn Gly Cys Ile Arg Ser Leu Glu His Ala Tyr
405 410 415

Ser Lys Asp Gly Gly Leu Ala Val Leu Tyr Gly Asn Phe Ala Glu Asn
420 425 430

Gly Cys Ile Val Lys Thr Ala Gly Val Asp Asp Ser Ile Leu Lys Phe
435 440 445

Thr Gly Pro Ala Lys Val Tyr Glu Ser Gln Asp Asp Ala Val Glu Ala
450 455 460

Ile Leu Gly Gly Lys Val Val Ala Gly Asp Val Val Val Ile Arg Tyr
465 470 475 480

Glu Gly Pro Lys Gly Gly Pro Gly Met Gln Glu Met Leu Tyr Pro Thr
485 490 495

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

Asp Gly Arg Phe Ser Gly Gly Thr Ser Gly Leu Ser Ile Gly His Val
515 520 525

Ser Pro Glu Ala Ala Ser Gly Gly Ser Ile Gly Leu Ile Glu Asp Gly
530 535 540

Asp Leu Ile Ala Ile Asp Ile Pro Asn Arg Gly Ile Gln Leu Gln Val
545 550 555 560

Ser Asp Ala Glu Leu Ala Ala Arg Arg Glu Ala Gln Asp Ala Arg Gly
565 570 575

Asp Lys Ala Trp Thr Pro Lys Asn Arg Glu Arg Gln Val Ser Phe Ala
580 585 590

Leu Arg Ala Tyr Ala Ser Leu Ala Thr Ser Ala Asp Lys Gly Ala Val
595 600 605

Arg Asp Lys Ser Lys Leu Gly Gly
610 615

<210> 47
<211> 1545
<212> DNA
<213> 大肠杆菌

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

<400> 47
atg gct gac tcg caa ccc ctg tcc ggt gct ccg gaa ggt gcc gaa tat 48
Met Ala Asp Ser Gln Pro Leu Ser Gly Ala Pro Glu Gly Ala Glu Tyr
1 5 10 15

tta aga gca gtg ctg cgc gcg ccg gtt tac gag gcg gcg cag gtt acg 96
Leu Arg Ala Val Leu Arg Ala Pro Val Tyr Glu Ala Ala Gln Val Thr
20 25 30

ccg cta caa aaa atg gaa aaa ctg tcg tcg cgt ctt gat aac gtc att 144
Pro Leu Gln Lys Met Glu Lys Leu Ser Ser Arg Leu Asp Asn Val Ile
35 40 45

ctg gtg aag cgc gaa gat cgc cag cca gtg cac agc ttt aag ctg cgc 192
Leu Val Lys Arg Glu Asp Arg Gln Pro Val His Ser Phe Lys Leu Arg
50 55 60

```
ggc gca tac gcc atg atg gcg ggc ctg acg gaa gaa cag aaa gcg cac      240
Gly Ala Tyr Ala Met Met Ala Gly Leu Thr Glu Glu Gln Lys Ala His
65                  70                  75                  80

ggc gtg atc act gct tct gcg ggt aac cac gcg cag ggc gtc gcg ttt      288
Gly Val Ile Thr Ala Ser Ala Gly Asn His Ala Gln Gly Val Ala Phe
                85                  90                  95

tct tct gcg cgg tta ggc gtg aag gcc ctg atc gtt atg cca acc gcc      336
Ser Ser Ala Arg Leu Gly Val Lys Ala Leu Ile Val Met Pro Thr Ala
            100                 105                 110

acc gcc gac atc aaa gtc gac gcg gtg cgc ggc ttc ggc ggc gaa gtg      384
Thr Ala Asp Ile Lys Val Asp Ala Val Arg Gly Phe Gly Gly Glu Val
        115                 120                 125

ctg ctc cac ggc gcg aac ttt gat gaa gcg aaa gcc aaa gcg atc gaa      432
Leu Leu His Gly Ala Asn Phe Asp Glu Ala Lys Ala Lys Ala Ile Glu
    130                 135                 140

ctg tca cag cag cag ggg ttc acc tgg gtg ccg ccg ttc gac cat ccg      480
Leu Ser Gln Gln Gln Gly Phe Thr Trp Val Pro Pro Phe Asp His Pro
145                 150                 155                 160

atg gtg att gcc ggg caa ggc acg ctg gcg ctg gaa ctg ctc cag cag      528
Met Val Ile Ala Gly Gln Gly Thr Leu Ala Leu Glu Leu Leu Gln Gln
                165                 170                 175

gac gcc cat ctc gac cgc gta ttt gtg cca gtc ggc ggc ggc ggt ctg      576
Asp Ala His Leu Asp Arg Val Phe Val Pro Val Gly Gly Gly Gly Leu
            180                 185                 190

gct gct ggc gtg gcg gtg ctg atc aaa caa ctg atg ccg caa atc aaa      624
Ala Ala Gly Val Ala Val Leu Ile Lys Gln Leu Met Pro Gln Ile Lys
        195                 200                 205

gtg atc gcc gta gaa gcg gaa gac tcc gcc tgc ctg aaa gca gcg ctg      672
Val Ile Ala Val Glu Ala Glu Asp Ser Ala Cys Leu Lys Ala Ala Leu
    210                 215                 220

gat gcg ggt cat ccg gtt gat ctg ccg cgc gta ggg cta ttt gct gaa      720
Asp Ala Gly His Pro Val Asp Leu Pro Arg Val Gly Leu Phe Ala Glu
225                 230                 235                 240

ggc gta gcg gta aaa cgc atc ggt gac gaa acc ttc cgt tta tgc cag      768
Gly Val Ala Val Lys Arg Ile Gly Asp Glu Thr Phe Arg Leu Cys Gln
                245                 250                 255

gag tat ctc gac gac atc atc acc gtc gat agc gat gcg atc tgt gcg      816
Glu Tyr Leu Asp Asp Ile Ile Thr Val Asp Ser Asp Ala Ile Cys Ala
            260                 265                 270
```

```
gcg atg aag gat tta ttc gaa gat gtg cgc gcg gtg gcg gaa ccc tct      864
Ala Met Lys Asp Leu Phe Glu Asp Val Arg Ala Val Ala Glu Pro Ser
        275                 280                 285

ggc gcg ctg gcg ctg gcg gga atg aaa aaa tat atc gcc ctg cac aac      912
Gly Ala Leu Ala Leu Ala Gly Met Lys Lys Tyr Ile Ala Leu His Asn
    290                 295                 300

att cgc ggc gaa cgg ctg gcg cat att ctt tcc ggt gcc aac gtg aac      960
Ile Arg Gly Glu Arg Leu Ala His Ile Leu Ser Gly Ala Asn Val Asn
305                 310                 315                 320

ttc cac ggc ctg cgc tac gtc tca gaa cgc tgc gaa ctg ggc gaa cag     1008
Phe His Gly Leu Arg Tyr Val Ser Glu Arg Cys Glu Leu Gly Glu Gln
                325                 330                 335

cgt gaa gcg ttg ttg gcg gtg acc att ccg gaa gaa aaa ggc agc ttc     1056
Arg Glu Ala Leu Leu Ala Val Thr Ile Pro Glu Glu Lys Gly Ser Phe
            340                 345                 350

ctc aaa ttc tgc caa ctg ctt ggc ggg cgt tcg gtc acc gag ttc aac     1104
Leu Lys Phe Cys Gln Leu Leu Gly Gly Arg Ser Val Thr Glu Phe Asn
        355                 360                 365

tac cgt ttt gcc gat gcc aaa aac gcc tgc atc ttt gtc ggt gtg cgc     1152
Tyr Arg Phe Ala Asp Ala Lys Asn Ala Cys Ile Phe Val Gly Val Arg
    370                 375                 380

ctg agc cgc ggc ctc gaa gag cgc aaa gaa att ttg cag atg ctc aac     1200
Leu Ser Arg Gly Leu Glu Glu Arg Lys Glu Ile Leu Gln Met Leu Asn
385                 390                 395                 400

gac ggc ggc tac agc gtg gtt gat ctc tcc gac gac gaa atg gcg aag     1248
Asp Gly Gly Tyr Ser Val Val Asp Leu Ser Asp Asp Glu Met Ala Lys
                405                 410                 415

cta cac gtg cgc tat atg gtc ggc gga cgt cca tcg cat ccg ttg cag     1296
Leu His Val Arg Tyr Met Val Gly Gly Arg Pro Ser His Pro Leu Gln
            420                 425                 430

gaa cgc ctc tac agc ttc gaa ttc ccg gaa tca ccg ggc gcg ctg ctg     1344
Glu Arg Leu Tyr Ser Phe Glu Phe Pro Glu Ser Pro Gly Ala Leu Leu
        435                 440                 445

cgc ttc ctc aac acg ctg ggt acg tac tgg aac att tct ttg ttc cac     1392
Arg Phe Leu Asn Thr Leu Gly Thr Tyr Trp Asn Ile Ser Leu Phe His
    450                 455                 460

tat cgc agc cat ggc acc gac tac ggg cgc gta ctg gcg gcg ttc gaa     1440
Tyr Arg Ser His Gly Thr Asp Tyr Gly Arg Val Leu Ala Ala Phe Glu
465                 470                 475                 480
```

ctt ggc gac cat gaa ccg gat ttc gaa acc cgg ctg aat gag ctg ggc 1488
Leu Gly Asp His Glu Pro Asp Phe Glu Thr Arg Leu Asn Glu Leu Gly
485 490 495

tac gat tgc cac gac gaa acc aat aac ccg gcg ttc agg ttc ttt ttg 1536
Tyr Asp Cys His Asp Glu Thr Asn Asn Pro Ala Phe Arg Phe Phe Leu
500 505 510

gcg ggt tag 1545
Ala Gly

<210> 48
<211> 514
<212> PRT
<213> 大肠杆菌

<400> 48

Met Ala Asp Ser Gln Pro Leu Ser Gly Ala Pro Glu Gly Ala Glu Tyr
1 5 10 15

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

Pro Leu Gln Lys Met Glu Lys Leu Ser Ser Arg Leu Asp Asn Val Ile
35 40 45

Leu Val Lys Arg Glu Asp Arg Gln Pro Val His Ser Phe Lys Leu Arg
50 55 60

Gly Ala Tyr Ala Met Met Ala Gly Leu Thr Glu Glu Gln Lys Ala His
65 70 75 80

Gly Val Ile Thr Ala Ser Ala Gly Asn His Ala Gln Gly Val Ala Phe
85 90 95

Ser Ser Ala Arg Leu Gly Val Lys Ala Leu Ile Val Met Pro Thr Ala
100 105 110

Thr Ala Asp Ile Lys Val Asp Ala Val Arg Gly Phe Gly Gly Glu Val
115 120 125

Leu Leu His Gly Ala Asn Phe Asp Glu Ala Lys Ala Lys Ala Ile Glu
130 135 140

Leu Ser Gln Gln Gln Gly Phe Thr Trp Val Pro Pro Phe Asp His Pro
145 150 155 160

Met Val Ile Ala Gly Gln Gly Thr Leu Ala Leu Glu Leu Leu Gln Gln
165 170 175

Asp Ala His Leu Asp Arg Val Phe Val Pro Val Gly Gly Gly Gly Leu
180 185 190

Ala Ala Gly Val Ala Val Leu Ile Lys Gln Leu Met Pro Gln Ile Lys
195 200 205

Val Ile Ala Val Glu Ala Glu Asp Ser Ala Cys Leu Lys Ala Ala Leu
210 215 220

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

Gly Val Ala Val Lys Arg Ile Gly Asp Glu Thr Phe Arg Leu Cys Gln
245 250 255

Glu Tyr Leu Asp Asp Ile Ile Thr Val Asp Ser Asp Ala Ile Cys Ala
260 265 270

Ala Met Lys Asp Leu Phe Glu Asp Val Arg Ala Val Ala Glu Pro Ser
275 280 285

Gly Ala Leu Ala Leu Ala Gly Met Lys Lys Tyr Ile Ala Leu His Asn
290 295 300

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

Phe His Gly Leu Arg Tyr Val Ser Glu Arg Cys Glu Leu Gly Glu Gln
325 330 335

Arg Glu Ala Leu Leu Ala Val Thr Ile Pro Glu Glu Lys Gly Ser Phe
340 345 350

Leu Lys Phe Cys Gln Leu Leu Gly Gly Arg Ser Val Thr Glu Phe Asn
355 360 365

Tyr Arg Phe Ala Asp Ala Lys Asn Ala Cys Ile Phe Val Gly Val Arg
370 375 380

Leu Ser Arg Gly Leu Glu Glu Arg Lys Glu Ile Leu Gln Met Leu Asn
385 390 395 400

Asp Gly Gly Tyr Ser Val Val Asp Leu Ser Asp Asp Glu Met Ala Lys
405 410 415

Leu His Val Arg Tyr Met Val Gly Gly Arg Pro Ser His Pro Leu Gln
420 425 430

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

Arg Phe Leu Asn Thr Leu Gly Thr Tyr Trp Asn Ile Ser Leu Phe His
450 455 460

Tyr Arg Ser His Gly Thr Asp Tyr Gly Arg Val Leu Ala Ala Phe Glu
465 470 475 480

Leu Gly Asp His Glu Pro Asp Phe Glu Thr Arg Leu Asn Glu Leu Gly
485 490 495

Tyr Asp Cys His Asp Glu Thr Asn Asn Pro Ala Phe Arg Phe Phe Leu
500 505 510

Ala Gly

<210> 49
<211> 1572
<212> DNA

<213> 大肠杆菌

<220>
<221> CDS
<222> (1)..(1572)

<400> 49

```
atg agc cag caa gtc att att ttc gat acc aca ttg cgc gac ggt gaa      48
Met Ser Gln Gln Val Ile Ile Phe Asp Thr Thr Leu Arg Asp Gly Glu
1               5                  10                  15

cag gcg tta cag gca agc ttg agt gtg aaa gaa aaa ctg caa att gcg      96
Gln Ala Leu Gln Ala Ser Leu Ser Val Lys Glu Lys Leu Gln Ile Ala
            20                  25                  30

ctg gcc ctt gag cgt atg ggt gtt gac gtg atg gaa gtc ggt ttc ccc     144
Leu Ala Leu Glu Arg Met Gly Val Asp Val Met Glu Val Gly Phe Pro
        35                  40                  45

gtc tct tcg ccg ggc gat ttt gaa tcg gtg caa acc atc gcc cgc cag     192
Val Ser Ser Pro Gly Asp Phe Glu Ser Val Gln Thr Ile Ala Arg Gln
    50                  55                  60

gtt aaa aac agc cgc gta tgt gcg tta gct cgc tgc gtg gaa aaa gat     240
Val Lys Asn Ser Arg Val Cys Ala Leu Ala Arg Cys Val Glu Lys Asp
65                  70                  75                  80

atc gac gtg gcg gcc gaa tcc ctg aaa gtc gcc gaa gcc ttc cgt att     288
Ile Asp Val Ala Ala Glu Ser Leu Lys Val Ala Glu Ala Phe Arg Ile
                85                  90                  95

cat acc ttt att gcc act tcg cca atg cac atc gcc acc aag ctg cgc     336
His Thr Phe Ile Ala Thr Ser Pro Met His Ile Ala Thr Lys Leu Arg
            100                 105                 110

agc acg ctg gac gag gtg atc gaa cgc gct atc tat atg gtg aaa cgc     384
Ser Thr Leu Asp Glu Val Ile Glu Arg Ala Ile Tyr Met Val Lys Arg
        115                 120                 125

gcc cgt aat tac acc gat gat gtt gaa ttt tct tgc gaa gat gcc ggg     432
Ala Arg Asn Tyr Thr Asp Asp Val Glu Phe Ser Cys Glu Asp Ala Gly
    130                 135                 140

cgt aca ccc att gcc gat ctg gcg cga gtg gtc gaa gcg gcg att aat     480
Arg Thr Pro Ile Ala Asp Leu Ala Arg Val Val Glu Ala Ala Ile Asn
145                 150                 155                 160

gcc ggt gcc acc acc atc aac att ccg gac acc gtg ggc tac acc atg     528
Ala Gly Ala Thr Thr Ile Asn Ile Pro Asp Thr Val Gly Tyr Thr Met
                165                 170                 175
```

```
ccg ttt gag ttc gcc gga atc atc agc ggc ctg tat gaa cgc gtg cct      576
Pro Phe Glu Phe Ala Gly Ile Ile Ser Gly Leu Tyr Glu Arg Val Pro
            180                 185                 190

aac atc gac aaa gcc att atc tcc gta cat acc cac gac gat ttg ggc      624
Asn Ile Asp Lys Ala Ile Ile Ser Val His Thr His Asp Asp Leu Gly
        195                 200                 205

ctg gcg gtc gga aac tca ctg gcg gcg gta cat gcc ggt gca cgc cag      672
Leu Ala Val Gly Asn Ser Leu Ala Ala Val His Ala Gly Ala Arg Gln
    210                 215                 220

gtg gaa ggc gca atg aac ggg atc ggc gag cgt gcc gga aac tgt tcc      720
Val Glu Gly Ala Met Asn Gly Ile Gly Glu Arg Ala Gly Asn Cys Ser
225                 230                 235                 240

ctg gaa gaa gtc atc atg gcg atc aaa gtt cgt aag gat att ctc aac      768
Leu Glu Glu Val Ile Met Ala Ile Lys Val Arg Lys Asp Ile Leu Asn
                245                 250                 255

gtc cac acc gcc att aat cac cag gag ata tgg cgc acc agc cag tta      816
Val His Thr Ala Ile Asn His Gln Glu Ile Trp Arg Thr Ser Gln Leu
            260                 265                 270

gtt agc cag att tgt aat atg ccg atc ccg gca aac aaa gcc att gtt      864
Val Ser Gln Ile Cys Asn Met Pro Ile Pro Ala Asn Lys Ala Ile Val
        275                 280                 285

ggc agc ggc gca ttc gca cac tcc tcc ggt ata cac cag gat ggc gtg      912
Gly Ser Gly Ala Phe Ala His Ser Ser Gly Ile His Gln Asp Gly Val
    290                 295                 300

ctg aaa aac cgc gaa aac tac gaa atc atg aca cca gaa tct att ggt      960
Leu Lys Asn Arg Glu Asn Tyr Glu Ile Met Thr Pro Glu Ser Ile Gly
305                 310                 315                 320

ctg aac caa atc cag ctg aat ctg acc tct cgt tcg ggg cgt gcg gcg     1008
Leu Asn Gln Ile Gln Leu Asn Leu Thr Ser Arg Ser Gly Arg Ala Ala
                325                 330                 335

gtg aaa cat cgc atg gat gag atg ggg tat aaa gaa agt gaa tat aat     1056
Val Lys His Arg Met Asp Glu Met Gly Tyr Lys Glu Ser Glu Tyr Asn
            340                 345                 350

tta gac aat ttg tac gat gct ttc ctg aag ctg gcg gac aaa aaa ggt     1104
Leu Asp Asn Leu Tyr Asp Ala Phe Leu Lys Leu Ala Asp Lys Lys Gly
        355                 360                 365

cag gtg ttt gat tac gat ctg gag gcg ctg gcc ttc atc ggt aag cag     1152
Gln Val Phe Asp Tyr Asp Leu Glu Ala Leu Ala Phe Ile Gly Lys Gln
    370                 375                 380
```

caa gaa gag ccg gag cat ttc cgt ctg gat tac ttc agc gtg cag tct 1200
Gln Glu Glu Pro Glu His Phe Arg Leu Asp Tyr Phe Ser Val Gln Ser
385 390 395 400

ggc tct aac gat atc gcc acc gcc gcc gtc aaa ctg gcc tgt ggc gaa 1248
Gly Ser Asn Asp Ile Ala Thr Ala Ala Val Lys Leu Ala Cys Gly Glu
405 410 415

gaa gtc aaa gca gaa gcc gcc aac ggt aac ggt ccg gtc gat gcc gtc 1296
Glu Val Lys Ala Glu Ala Ala Asn Gly Asn Gly Pro Val Asp Ala Val
420 425 430

tat cag gca att aac cgc atc act gaa tat aac gtc gaa ctg gtg aaa 1344
Tyr Gln Ala Ile Asn Arg Ile Thr Glu Tyr Asn Val Glu Leu Val Lys
435 440 445

tac agc ctg acc gcc aaa ggc cac ggt aaa gat gcg ctg ggt cag gtg 1392
Tyr Ser Leu Thr Ala Lys Gly His Gly Lys Asp Ala Leu Gly Gln Val
450 455 460

gat atc gtc gct aac tac aac ggt cgc cgc ttc cac ggc gtc ggc ctg 1440
Asp Ile Val Ala Asn Tyr Asn Gly Arg Arg Phe His Gly Val Gly Leu
465 470 475 480

gct acc gat att gtc gag tca tct gcc aaa gcc atg gtg cac gtt ctg 1488
Ala Thr Asp Ile Val Glu Ser Ser Ala Lys Ala Met Val His Val Leu
485 490 495

aac aat atc tgg cgt gcc gca gaa gtc gaa aaa gag ttg caa cgc aaa 1536
Asn Asn Ile Trp Arg Ala Ala Glu Val Glu Lys Glu Leu Gln Arg Lys
500 505 510

gct caa cac aac gaa aac aac aag gaa acc gtg tga 1572
Ala Gln His Asn Glu Asn Asn Lys Glu Thr Val
515 520

<210> 50
<211> 523
<212> PRT
<213> 大肠杆菌

<400> 50

Met Ser Gln Gln Val Ile Ile Phe Asp Thr Thr Leu Arg Asp Gly Glu
1 5 10 15

Gln Ala Leu Gln Ala Ser Leu Ser Val Lys Glu Lys Leu Gln Ile Ala
20 25 30

Leu Ala Leu Glu Arg Met Gly Val Asp Val Met Glu Val Gly Phe Pro
35 40 45

Val Ser Ser Pro Gly Asp Phe Glu Ser Val Gln Thr Ile Ala Arg Gln
50 55 60

Val Lys Asn Ser Arg Val Cys Ala Leu Ala Arg Cys Val Glu Lys Asp
65 70 75 80

Ile Asp Val Ala Ala Glu Ser Leu Lys Val Ala Glu Ala Phe Arg Ile
85 90 95

His Thr Phe Ile Ala Thr Ser Pro Met His Ile Ala Thr Lys Leu Arg
100 105 110

Ser Thr Leu Asp Glu Val Ile Glu Arg Ala Ile Tyr Met Val Lys Arg
115 120 125

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

Arg Thr Pro Ile Ala Asp Leu Ala Arg Val Val Glu Ala Ala Ile Asn
145 150 155 160

Ala Gly Ala Thr Thr Ile Asn Ile Pro Asp Thr Val Gly Tyr Thr Met
165 170 175

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

Asn Ile Asp Lys Ala Ile Ile Ser Val His Thr His Asp Asp Leu Gly
195 200 205

Leu Ala Val Gly Asn Ser Leu Ala Ala Val His Ala Gly Ala Arg Gln
210 215 220

Val Glu Gly Ala Met Asn Gly Ile Gly Glu Arg Ala Gly Asn Cys Ser
225 230 235 240

Leu Glu Glu Val Ile Met Ala Ile Lys Val Arg Lys Asp Ile Leu Asn
245 250 255

Val His Thr Ala Ile Asn His Gln Glu Ile Trp Arg Thr Ser Gln Leu
260 265 270

Val Ser Gln Ile Cys Asn Met Pro Ile Pro Ala Asn Lys Ala Ile Val
275 280 285

Gly Ser Gly Ala Phe Ala His Ser Ser Gly Ile His Gln Asp Gly Val
290 295 300

Leu Lys Asn Arg Glu Asn Tyr Glu Ile Met Thr Pro Glu Ser Ile Gly
305 310 315 320

Leu Asn Gln Ile Gln Leu Asn Leu Thr Ser Arg Ser Gly Arg Ala Ala
325 330 335

Val Lys His Arg Met Asp Glu Met Gly Tyr Lys Glu Ser Glu Tyr Asn
340 345 350

Leu Asp Asn Leu Tyr Asp Ala Phe Leu Lys Leu Ala Asp Lys Lys Gly
355 360 365

Gln Val Phe Asp Tyr Asp Leu Glu Ala Leu Ala Phe Ile Gly Lys Gln
370 375 380

Gln Glu Glu Pro Glu His Phe Arg Leu Asp Tyr Phe Ser Val Gln Ser
385 390 395 400

Gly Ser Asn Asp Ile Ala Thr Ala Ala Val Lys Leu Ala Cys Gly Glu
405 410 415

Glu Val Lys Ala Glu Ala Ala Asn Gly Asn Gly Pro Val Asp Ala Val
420 425 430

Tyr Gln Ala Ile Asn Arg Ile Thr Glu Tyr Asn Val Glu Leu Val Lys
435 440 445

Tyr Ser Leu Thr Ala Lys Gly His Gly Lys Asp Ala Leu Gly Gln Val
450 455 460

Asp Ile Val Ala Asn Tyr Asn Gly Arg Arg Phe His Gly Val Gly Leu
465 470 475 480

Ala Thr Asp Ile Val Glu Ser Ser Ala Lys Ala Met Val His Val Leu
485 490 495

Asn Asn Ile Trp Arg Ala Ala Glu Val Glu Lys Glu Leu Gln Arg Lys
500 505 510

Ala Gln His Asn Glu Asn Asn Lys Glu Thr Val
515 520

<210> 51
<211> 1095
<212> DNA
<213> 大肠杆菌

<220>
<221> CDS
<222> (1)..(1095)

<400> 51
gtg atg tcg aag aat tac cat att gcc gta ttg ccg ggg gac ggt att 48
Val Met Ser Lys Asn Tyr His Ile Ala Val Leu Pro Gly Asp Gly Ile
1 5 10 15

ggt ccg gaa gtg atg acc cag gcg ctg aaa gtg ctg gat gcc gtg cgc 96
Gly Pro Glu Val Met Thr Gln Ala Leu Lys Val Leu Asp Ala Val Arg
20 25 30

aac cgc ttt gcg atg cgc atc acc acc agc cat tac gat gta ggc ggc 144
Asn Arg Phe Ala Met Arg Ile Thr Thr Ser His Tyr Asp Val Gly Gly
35 40 45

gca gcc att gat aac cac ggg caa cca ctg ccg cct gcg acg gtt gaa 192
Ala Ala Ile Asp Asn His Gly Gln Pro Leu Pro Pro Ala Thr Val Glu
50 55 60

ggt tgt gag caa gcc gat gcc gtg ctg ttt ggc tcg gta ggc ggc ccg 240
Gly Cys Glu Gln Ala Asp Ala Val Leu Phe Gly Ser Val Gly Gly Pro
65 70 75 80

```
aag tgg gaa cat tta cca cca gac cag caa cca gaa cgc ggc gcg ctg      288
Lys Trp Glu His Leu Pro Pro Asp Gln Gln Pro Glu Arg Gly Ala Leu
                85                  90                  95

ctg cct ctg cgt aag cac ttc aaa tta ttc agc aac ctg cgc ccg gca      336
Leu Pro Leu Arg Lys His Phe Lys Leu Phe Ser Asn Leu Arg Pro Ala
            100                 105                 110

aaa ctg tat cag ggg ctg gaa gca ttc tgt ccg ctg cgt gca gac att      384
Lys Leu Tyr Gln Gly Leu Glu Ala Phe Cys Pro Leu Arg Ala Asp Ile
        115                 120                 125

gcc gca aac ggc ttc gac atc ctg tgt gtg cgc gaa ctg acc ggc ggc      432
Ala Ala Asn Gly Phe Asp Ile Leu Cys Val Arg Glu Leu Thr Gly Gly
    130                 135                 140

atc tat ttc ggt cag cca aaa ggc cgc gaa ggt agc gga caa tat gaa      480
Ile Tyr Phe Gly Gln Pro Lys Gly Arg Glu Gly Ser Gly Gln Tyr Glu
145                 150                 155                 160

aaa gcc ttt gat acc gag gtg tat cac cgt ttt gag atc gaa cgt atc      528
Lys Ala Phe Asp Thr Glu Val Tyr His Arg Phe Glu Ile Glu Arg Ile
                165                 170                 175

gcc cgc atc gcg ttt gaa tct gct cgc aag cgt cgc cac aaa gtg acg      576
Ala Arg Ile Ala Phe Glu Ser Ala Arg Lys Arg Arg His Lys Val Thr
            180                 185                 190

tcg atc gat aaa gcc aac gtg ctg caa tcc tct att tta tgg cgg gag      624
Ser Ile Asp Lys Ala Asn Val Leu Gln Ser Ser Ile Leu Trp Arg Glu
        195                 200                 205

atc gtt aac gag atc gcc acg gaa tac ccg gat gtc gaa ctg gcg cat      672
Ile Val Asn Glu Ile Ala Thr Glu Tyr Pro Asp Val Glu Leu Ala His
    210                 215                 220

atg tac atc gac aac gcc acc atg cag ctg att aaa gat cca tca cag      720
Met Tyr Ile Asp Asn Ala Thr Met Gln Leu Ile Lys Asp Pro Ser Gln
225                 230                 235                 240

ttt gac gtt ctg ctg tgc tcc aac ctg ttt ggc gac att ctg tct gac      768
Phe Asp Val Leu Leu Cys Ser Asn Leu Phe Gly Asp Ile Leu Ser Asp
                245                 250                 255

gag tgc gca atg atc act ggc tcg atg ggg atg ttg cct tcc gcc agc      816
Glu Cys Ala Met Ile Thr Gly Ser Met Gly Met Leu Pro Ser Ala Ser
            260                 265                 270

ctg aac gag caa ggt ttt gga ctg tat gaa ccg gcg ggc ggc tcg gca      864
Leu Asn Glu Gln Gly Phe Gly Leu Tyr Glu Pro Ala Gly Gly Ser Ala
        275                 280                 285
```

```
cca gat atc gca ggc aaa aac atc gcc aac ccg att gca caa atc ctt      912
Pro Asp Ile Ala Gly Lys Asn Ile Ala Asn Pro Ile Ala Gln Ile Leu
    290                 295                 300

tcg ctg gca ctg ctg ctg cgt tac agc ctg gat gcc gat gat gcg gct      960
Ser Leu Ala Leu Leu Leu Arg Tyr Ser Leu Asp Ala Asp Asp Ala Ala
305                 310                 315                 320

tgc gcc att gaa cgc gcc att aac cgc gca tta gaa gaa ggc att cgc     1008
Cys Ala Ile Glu Arg Ala Ile Asn Arg Ala Leu Glu Glu Gly Ile Arg
                325                 330                 335

acc ggg gat tta gcc cgt ggc gct gcc gcc gtt agt acc gat gaa atg     1056
Thr Gly Asp Leu Ala Arg Gly Ala Ala Ala Val Ser Thr Asp Glu Met
            340                 345                 350

ggc gat atc att gcc cgc tat gta gca gaa ggg gtg taa                1095
Gly Asp Ile Ile Ala Arg Tyr Val Ala Glu Gly Val
        355                 360
```

<210> 52
<211> 364
<212> PRT
<213> 大肠杆菌

<400> 52

```
Val Met Ser Lys Asn Tyr His Ile Ala Val Leu Pro Gly Asp Gly Ile
1               5                   10                  15

Gly Pro Glu Val Met Thr Gln Ala Leu Lys Val Leu Asp Ala Val Arg
            20                  25                  30

Asn Arg Phe Ala Met Arg Ile Thr Thr Ser His Tyr Asp Val Gly Gly
        35                  40                  45

Ala Ala Ile Asp Asn His Gly Gln Pro Leu Pro Pro Ala Thr Val Glu
    50                  55                  60

Gly Cys Glu Gln Ala Asp Ala Val Leu Phe Gly Ser Val Gly Gly Pro
65                  70                  75                  80

Lys Trp Glu His Leu Pro Pro Asp Gln Gln Pro Glu Arg Gly Ala Leu
                85                  90                  95
```

Leu Pro Leu Arg Lys His Phe Lys Leu Phe Ser Asn Leu Arg Pro Ala
100 105 110

Lys Leu Tyr Gln Gly Leu Glu Ala Phe Cys Pro Leu Arg Ala Asp Ile
115 120 125

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

Ile Tyr Phe Gly Gln Pro Lys Gly Arg Glu Gly Ser Gly Gln Tyr Glu
145 150 155 160

Lys Ala Phe Asp Thr Glu Val Tyr His Arg Phe Glu Ile Glu Arg Ile
165 170 175

Ala Arg Ile Ala Phe Glu Ser Ala Arg Lys Arg Arg His Lys Val Thr
180 185 190

Ser Ile Asp Lys Ala Asn Val Leu Gln Ser Ser Ile Leu Trp Arg Glu
195 200 205

Ile Val Asn Glu Ile Ala Thr Glu Tyr Pro Asp Val Glu Leu Ala His
210 215 220

Met Tyr Ile Asp Asn Ala Thr Met Gln Leu Ile Lys Asp Pro Ser Gln
225 230 235 240

Phe Asp Val Leu Leu Cys Ser Asn Leu Phe Gly Asp Ile Leu Ser Asp
245 250 255

Glu Cys Ala Met Ile Thr Gly Ser Met Gly Met Leu Pro Ser Ala Ser
260 265 270

Leu Asn Glu Gln Gly Phe Gly Leu Tyr Glu Pro Ala Gly Gly Ser Ala
275 280 285

Pro Asp Ile Ala Gly Lys Asn Ile Ala Asn Pro Ile Ala Gln Ile Leu
290 295 300

Ser Leu Ala Leu Leu Leu Arg Tyr Ser Leu Asp Ala Asp Asp Ala Ala
305 310 315 320

Cys Ala Ile Glu Arg Ala Ile Asn Arg Ala Leu Glu Glu Gly Ile Arg
325 330 335

Thr Gly Asp Leu Ala Arg Gly Ala Ala Ala Val Ser Thr Asp Glu Met
340 345 350

Gly Asp Ile Ile Ala Arg Tyr Val Ala Glu Gly Val
355 360

<210> 53
<211> 1401
<212> DNA
<213> 大肠杆菌

<220>
<221> CDS
<222> (1)..(1401)

<400> 53
atg gct aag acg tta tac gaa aaa ttg ttc gac gct cac gtt gtg tac 48
Met Ala Lys Thr Leu Tyr Glu Lys Leu Phe Asp Ala His Val Val Tyr
1 5 10 15

gaa gcc gaa aac gaa acc cca ctg tta tat atc gac cgc cac ctg gtg 96
Glu Ala Glu Asn Glu Thr Pro Leu Leu Tyr Ile Asp Arg His Leu Val
20 25 30

cat gaa gtg acc tca ccg cag gcg ttc gat ggt ctg cgc gcc cac ggt 144
His Glu Val Thr Ser Pro Gln Ala Phe Asp Gly Leu Arg Ala His Gly
35 40 45

cgc ccg gta cgt cag ccg ggc aaa acc ttc gct acc atg gat cac aac 192
Arg Pro Val Arg Gln Pro Gly Lys Thr Phe Ala Thr Met Asp His Asn
50 55 60

gtc tct acc cag acc aaa gac att aat gcc tgc ggt gaa atg gcg cgt 240
Val Ser Thr Gln Thr Lys Asp Ile Asn Ala Cys Gly Glu Met Ala Arg
65 70 75 80

atc cag atg cag gaa ctg atc aaa aac tgc aaa gaa ttt ggc gtc gaa 288
Ile Gln Met Gln Glu Leu Ile Lys Asn Cys Lys Glu Phe Gly Val Glu
85 90 95

```
ctg tat gac ctg aat cac ccg tat cag ggg atc gtc cac gta atg ggg      336
Leu Tyr Asp Leu Asn His Pro Tyr Gln Gly Ile Val His Val Met Gly
            100                 105                 110

ccg gaa cag ggc gtc acc ttg ccg ggg atg acc att gtc tgc ggc gac      384
Pro Glu Gln Gly Val Thr Leu Pro Gly Met Thr Ile Val Cys Gly Asp
        115                 120                 125

tcg cat acc gcc acc cac ggc gcg ttt ggc gca ctg gcc ttt ggt atc      432
Ser His Thr Ala Thr His Gly Ala Phe Gly Ala Leu Ala Phe Gly Ile
    130                 135                 140

ggc act tcc gaa gtt gaa cac gta ctg gca acg caa acc ctg aaa cag      480
Gly Thr Ser Glu Val Glu His Val Leu Ala Thr Gln Thr Leu Lys Gln
145                 150                 155                 160

ggc cgc gca aaa acc atg aaa att gaa gtc cag ggc aaa gcc gcg ccg      528
Gly Arg Ala Lys Thr Met Lys Ile Glu Val Gln Gly Lys Ala Ala Pro
                165                 170                 175

ggc att acc gca aaa gat atc gtg ctg gca att atc ggt aaa acc ggt      576
Gly Ile Thr Ala Lys Asp Ile Val Leu Ala Ile Ile Gly Lys Thr Gly
            180                 185                 190

agc gca ggc ggc acc ggg cat gtg gtg gag ttt tgc ggc gaa gca atc      624
Ser Ala Gly Gly Thr Gly His Val Val Glu Phe Cys Gly Glu Ala Ile
        195                 200                 205

cgt gat tta agc atg gaa ggt cgt atg acc ctg tgc aat atg gca atc      672
Arg Asp Leu Ser Met Glu Gly Arg Met Thr Leu Cys Asn Met Ala Ile
    210                 215                 220

gaa atg ggc gca aaa gcc ggt ctg gtt gca ccg gac gaa acc acc ttt      720
Glu Met Gly Ala Lys Ala Gly Leu Val Ala Pro Asp Glu Thr Thr Phe
225                 230                 235                 240

aac tat gtc aaa ggc cgt ctg cat gcg ccg aaa ggc aaa gat ttc gac      768
Asn Tyr Val Lys Gly Arg Leu His Ala Pro Lys Gly Lys Asp Phe Asp
                245                 250                 255

gac gcc gtt gcc tac tgg aaa acc ctg caa acc gac gaa ggc gca act      816
Asp Ala Val Ala Tyr Trp Lys Thr Leu Gln Thr Asp Glu Gly Ala Thr
            260                 265                 270

ttc gat acc gtt gtc act ctg caa gca gaa gaa att tca ccg cag gtc      864
Phe Asp Thr Val Val Thr Leu Gln Ala Glu Glu Ile Ser Pro Gln Val
        275                 280                 285

acc tgg ggc acc aat ccc ggc cag gtg att tcc gtg aac gac aat att      912
Thr Trp Gly Thr Asn Pro Gly Gln Val Ile Ser Val Asn Asp Asn Ile
    290                 295                 300
```

```
ccc gat ccg gct tcg ttt gcc gat ccg gtt gaa cgc gcg tcg gca gaa      960
Pro Asp Pro Ala Ser Phe Ala Asp Pro Val Glu Arg Ala Ser Ala Glu
305                 310                 315                 320

aaa gcg ctg gcc tat atg ggg ctg aaa ccg ggt att ccg ctg acc gaa     1008
Lys Ala Leu Ala Tyr Met Gly Leu Lys Pro Gly Ile Pro Leu Thr Glu
                325                 330                 335

gtg gct atc gac aaa gtg ttt atc ggt tcc tgt acc aac tcg cgc att     1056
Val Ala Ile Asp Lys Val Phe Ile Gly Ser Cys Thr Asn Ser Arg Ile
            340                 345                 350

gaa gat tta cgc gcg gca gcg gag atc gcc aaa ggg cga aaa gtc gcg     1104
Glu Asp Leu Arg Ala Ala Ala Glu Ile Ala Lys Gly Arg Lys Val Ala
        355                 360                 365

cca ggc gtg cag gca ctg gtg gtt ccc ggc tct ggc ccg gta aaa gcc     1152
Pro Gly Val Gln Ala Leu Val Val Pro Gly Ser Gly Pro Val Lys Ala
    370                 375                 380

cag gcg gaa gcg gaa ggt ctg gat aaa atc ttt att gaa gcc ggt ttt     1200
Gln Ala Glu Ala Glu Gly Leu Asp Lys Ile Phe Ile Glu Ala Gly Phe
385                 390                 395                 400

gaa tgg cgc ttg cct ggc tgc tca atg tgt ctg gcg atg aac aac gac     1248
Glu Trp Arg Leu Pro Gly Cys Ser Met Cys Leu Ala Met Asn Asn Asp
                405                 410                 415

cgt ctg aat ccg ggc gaa cgt tgt gcc tcc acc agc aac cgt aac ttt     1296
Arg Leu Asn Pro Gly Glu Arg Cys Ala Ser Thr Ser Asn Arg Asn Phe
            420                 425                 430

gaa ggc cgc cag ggg cgc ggc ggg cgc acg cat ctg gtc agc ccg gca     1344
Glu Gly Arg Gln Gly Arg Gly Gly Arg Thr His Leu Val Ser Pro Ala
        435                 440                 445

atg gct gcc gct gct gct gtg acc gga cat ttc gcc gac att cgc aac     1392
Met Ala Ala Ala Ala Ala Val Thr Gly His Phe Ala Asp Ile Arg Asn
    450                 455                 460

att aaa taa                                                        1401
Ile Lys
465
```

<210> 54
<211> 466
<212> PRT
<213> 大肠杆菌

<400> 54

Met Ala Lys Thr Leu Tyr Glu Lys Leu Phe Asp Ala His Val Val Tyr
1 5 10 15

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

His Glu Val Thr Ser Pro Gln Ala Phe Asp Gly Leu Arg Ala His Gly
35 40 45

Arg Pro Val Arg Gln Pro Gly Lys Thr Phe Ala Thr Met Asp His Asn
50 55 60

Val Ser Thr Gln Thr Lys Asp Ile Asn Ala Cys Gly Glu Met Ala Arg
65 70 75 80

Ile Gln Met Gln Glu Leu Ile Lys Asn Cys Lys Glu Phe Gly Val Glu
85 90 95

Leu Tyr Asp Leu Asn His Pro Tyr Gln Gly Ile Val His Val Met Gly
100 105 110

Pro Glu Gln Gly Val Thr Leu Pro Gly Met Thr Ile Val Cys Gly Asp
115 120 125

Ser His Thr Ala Thr His Gly Ala Phe Gly Ala Leu Ala Phe Gly Ile
130 135 140

Gly Thr Ser Glu Val Glu His Val Leu Ala Thr Gln Thr Leu Lys Gln
145 150 155 160

Gly Arg Ala Lys Thr Met Lys Ile Glu Val Gln Gly Lys Ala Ala Pro
165 170 175

Gly Ile Thr Ala Lys Asp Ile Val Leu Ala Ile Ile Gly Lys Thr Gly
180 185 190

Ser Ala Gly Gly Thr Gly His Val Val Glu Phe Cys Gly Glu Ala Ile
195 200 205

Arg Asp Leu Ser Met Glu Gly Arg Met Thr Leu Cys Asn Met Ala Ile
210 215 220

Glu Met Gly Ala Lys Ala Gly Leu Val Ala Pro Asp Glu Thr Thr Phe
225 230 235 240

Asn Tyr Val Lys Gly Arg Leu His Ala Pro Lys Gly Lys Asp Phe Asp
245 250 255

Asp Ala Val Ala Tyr Trp Lys Thr Leu Gln Thr Asp Glu Gly Ala Thr
260 265 270

Phe Asp Thr Val Val Thr Leu Gln Ala Glu Glu Ile Ser Pro Gln Val
275 280 285

Thr Trp Gly Thr Asn Pro Gly Gln Val Ile Ser Val Asn Asp Asn Ile
290 295 300

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

Lys Ala Leu Ala Tyr Met Gly Leu Lys Pro Gly Ile Pro Leu Thr Glu
325 330 335

Val Ala Ile Asp Lys Val Phe Ile Gly Ser Cys Thr Asn Ser Arg Ile
340 345 350

Glu Asp Leu Arg Ala Ala Ala Glu Ile Ala Lys Gly Arg Lys Val Ala
355 360 365

Pro Gly Val Gln Ala Leu Val Val Pro Gly Ser Gly Pro Val Lys Ala
370 375 380

Gln Ala Glu Ala Glu Gly Leu Asp Lys Ile Phe Ile Glu Ala Gly Phe
385 390 395 400

Glu Trp Arg Leu Pro Gly Cys Ser Met Cys Leu Ala Met Asn Asn Asp
405 410 415

Arg Leu Asn Pro Gly Glu Arg Cys Ala Ser Thr Ser Asn Arg Asn Phe
420 425 430

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

Met Ala Ala Ala Ala Ala Val Thr Gly His Phe Ala Asp Ile Arg Asn
450 455 460

Ile Lys
465

<210> 55
<211> 606
<212> DNA
<213> 大肠杆菌

<220>
<221> CDS
<222> (1)..(606)

<400> 55
atg gca gag aaa ttt atc aaa cac aca ggc ctg gtg gtt ccg ctg gat 48
Met Ala Glu Lys Phe Ile Lys His Thr Gly Leu Val Val Pro Leu Asp
1 5 10 15

gcc gcc aat gtc gat acc gat gca atc atc ccg aaa cag ttt ttg cag 96
Ala Ala Asn Val Asp Thr Asp Ala Ile Ile Pro Lys Gln Phe Leu Gln
20 25 30

aaa gtg acc cgt acg ggt ttt ggc gcg cat ctg ttt aac gac tgg cgt 144
Lys Val Thr Arg Thr Gly Phe Gly Ala His Leu Phe Asn Asp Trp Arg
35 40 45

ttt ctg gat gaa aaa ggc caa cag cca aac ccg gac ttc gtg ctg aac 192
Phe Leu Asp Glu Lys Gly Gln Gln Pro Asn Pro Asp Phe Val Leu Asn
50 55 60

ttc ccg cag tat cag ggc gct tcc att ttg ctg gca cga gaa aac ttc 240
Phe Pro Gln Tyr Gln Gly Ala Ser Ile Leu Leu Ala Arg Glu Asn Phe
65 70 75 80

ggc tgt ggc tct tcg cgt gag cac gcg ccc tgg gca ttg acc gac tac 288
Gly Cys Gly Ser Ser Arg Glu His Ala Pro Trp Ala Leu Thr Asp Tyr
85 90 95

```
ggt ttt aaa gtg gtg att gcg ccg agt ttt gct gac atc ttc tac ggc      336
Gly Phe Lys Val Val Ile Ala Pro Ser Phe Ala Asp Ile Phe Tyr Gly
            100                 105                 110

aat agc ttt aac aac cag ctg ctg ccg gtg aaa tta agc gat gca gaa      384
Asn Ser Phe Asn Asn Gln Leu Leu Pro Val Lys Leu Ser Asp Ala Glu
        115                 120                 125

gtg gac gaa ctg ttt gcg ctg gtg aaa gct aat ccg ggg atc cat ttc      432
Val Asp Glu Leu Phe Ala Leu Val Lys Ala Asn Pro Gly Ile His Phe
    130                 135                 140

gac gtg gat ctg gaa gcg caa gag gtg aaa gcg gga gag aaa acc tat      480
Asp Val Asp Leu Glu Ala Gln Glu Val Lys Ala Gly Glu Lys Thr Tyr
145                 150                 155                 160

cgc ttt acc atc gat gcc ttc cgc cgc cac tgc atg atg aac ggt ctg      528
Arg Phe Thr Ile Asp Ala Phe Arg Arg His Cys Met Met Asn Gly Leu
                165                 170                 175

gac agt att ggg ctt acc ttg cag cac gac gac gcc att gcc gct tat      576
Asp Ser Ile Gly Leu Thr Leu Gln His Asp Asp Ala Ile Ala Ala Tyr
            180                 185                 190

gaa gca aaa caa cct gcg ttt atg aat taa                             606
Glu Ala Lys Gln Pro Ala Phe Met Asn
        195                 200
```

<210> 56
<211> 201
<212> PRT
<213> 大肠杆菌

<400> 56

```
Met Ala Glu Lys Phe Ile Lys His Thr Gly Leu Val Val Pro Leu Asp
1               5                   10                  15

Ala Ala Asn Val Asp Thr Asp Ala Ile Ile Pro Lys Gln Phe Leu Gln
            20                  25                  30

Lys Val Thr Arg Thr Gly Phe Gly Ala His Leu Phe Asn Asp Trp Arg
        35                  40                  45

Phe Leu Asp Glu Lys Gly Gln Gln Pro Asn Pro Asp Phe Val Leu Asn
    50                  55                  60
```

Phe Pro Gln Tyr Gln Gly Ala Ser Ile Leu Leu Ala Arg Glu Asn Phe
65 70 75 80

Gly Cys Gly Ser Ser Arg Glu His Ala Pro Trp Ala Leu Thr Asp Tyr
85 90 95

Gly Phe Lys Val Val Ile Ala Pro Ser Phe Ala Asp Ile Phe Tyr Gly
100 105 110

Asn Ser Phe Asn Asn Gln Leu Leu Pro Val Lys Leu Ser Asp Ala Glu
115 120 125

Val Asp Glu Leu Phe Ala Leu Val Lys Ala Asn Pro Gly Ile His Phe
130 135 140

Asp Val Asp Leu Glu Ala Gln Glu Val Lys Ala Gly Glu Lys Thr Tyr
145 150 155 160

Arg Phe Thr Ile Asp Ala Phe Arg Arg His Cys Met Met Asn Gly Leu
165 170 175

Asp Ser Ile Gly Leu Thr Leu Gln His Asp Asp Ala Ile Ala Ala Tyr
180 185 190

Glu Ala Lys Gln Pro Ala Phe Met Asn
195 200

<210> 57
<211> 1476
<212> DNA
<213> 詹氏甲烷球菌

<220>
<221> CDS
<222> (1)..(1476)

<400> 57
atg atg gta agg ata ttt gat aca aca ctt aga gat gga gag caa aca 48
Met Met Val Arg Ile Phe Asp Thr Thr Leu Arg Asp Gly Glu Gln Thr
1 5 10 15

```
cca gga gtt tct tta aca cca aat gat aag tta gag ata gca aaa aaa      96
Pro Gly Val Ser Leu Thr Pro Asn Asp Lys Leu Glu Ile Ala Lys Lys
            20                  25                  30

ttg gat gag ctt gga gtt gat gtt ata gag gca ggt tca gct ata act     144
Leu Asp Glu Leu Gly Val Asp Val Ile Glu Ala Gly Ser Ala Ile Thr
        35                  40                  45

tca aaa gga gag aga gaa gga ata aaa tta ata aca aaa gaa ggt tta     192
Ser Lys Gly Glu Arg Glu Gly Ile Lys Leu Ile Thr Lys Glu Gly Leu
    50                  55                  60

aat gca gaa atc tgc tca ttt gtt aga gct tta cct gta gat att gat     240
Asn Ala Glu Ile Cys Ser Phe Val Arg Ala Leu Pro Val Asp Ile Asp
65                  70                  75                  80

gct gcc tta gaa tgt gat gta gat agt gtc cat tta gta gtg cca aca     288
Ala Ala Leu Glu Cys Asp Val Asp Ser Val His Leu Val Val Pro Thr
                85                  90                  95

tct cca ata cac atg aaa tat aag ctt aga aaa aca gaa gat gag gtt     336
Ser Pro Ile His Met Lys Tyr Lys Leu Arg Lys Thr Glu Asp Glu Val
            100                 105                 110

tta gag aca gct tta aag gct gta gag tat gct aaa gaa cat gga ttg     384
Leu Glu Thr Ala Leu Lys Ala Val Glu Tyr Ala Lys Glu His Gly Leu
        115                 120                 125

att gtt gag tta tct gca gag gat gca aca aga agt gat gta aat ttc     432
Ile Val Glu Leu Ser Ala Glu Asp Ala Thr Arg Ser Asp Val Asn Phe
    130                 135                 140

tta ata aaa cta ttt aat gaa ggg gaa aag gtt gga gca gac aga gtt     480
Leu Ile Lys Leu Phe Asn Glu Gly Glu Lys Val Gly Ala Asp Arg Val
145                 150                 155                 160

tgt gtt tgt gac aca gta gga gtt tta act cca caa aag agt cag gaa     528
Cys Val Cys Asp Thr Val Gly Val Leu Thr Pro Gln Lys Ser Gln Glu
                165                 170                 175

tta ttt aaa aaa ata act gaa aat gtt aat tta ccg gtc tca gtt cat     576
Leu Phe Lys Lys Ile Thr Glu Asn Val Asn Leu Pro Val Ser Val His
            180                 185                 190

tgc cac aac gac ttt gga atg gct act gct aat act tgc tca gca gtt     624
Cys His Asn Asp Phe Gly Met Ala Thr Ala Asn Thr Cys Ser Ala Val
        195                 200                 205

tta ggt gga gct gtt cag tgc cac gta aca gtt aat ggt att gga gag     672
Leu Gly Gly Ala Val Gln Cys His Val Thr Val Asn Gly Ile Gly Glu
    210                 215                 220
```

```
aga gca gga aat gcc tca ttg gaa gag gtt gtt gct gct tta aaa ata      720
Arg Ala Gly Asn Ala Ser Leu Glu Glu Val Val Ala Ala Leu Lys Ile
225                 230                 235                 240

ctc tat ggc tat gat act aag ata aag atg gaa aag tta tat gag gtt      768
Leu Tyr Gly Tyr Asp Thr Lys Ile Lys Met Glu Lys Leu Tyr Glu Val
                245                 250                 255

tca aga att gtc tca aga ttg atg aaa ctt cct gtt cca cca aat aaa      816
Ser Arg Ile Val Ser Arg Leu Met Lys Leu Pro Val Pro Pro Asn Lys
            260                 265                 270

gca att gtt ggg gac aat gca ttt gct cat gaa gca gga ata cat gtt      864
Ala Ile Val Gly Asp Asn Ala Phe Ala His Glu Ala Gly Ile His Val
        275                 280                 285

gat gga tta ata aaa aat act gaa acc tat gag cca ata aaa cca gaa      912
Asp Gly Leu Ile Lys Asn Thr Glu Thr Tyr Glu Pro Ile Lys Pro Glu
    290                 295                 300

atg gtt ggg aat aga aga aga att att ttg ggt aag cat tct ggt aga      960
Met Val Gly Asn Arg Arg Arg Ile Ile Leu Gly Lys His Ser Gly Arg
305                 310                 315                 320

aaa gct tta aaa tac aaa ctt gat ttg atg ggc ata aac gtt agt gat     1008
Lys Ala Leu Lys Tyr Lys Leu Asp Leu Met Gly Ile Asn Val Ser Asp
                325                 330                 335

gag caa tta aat aaa ata tat gaa aga gtt aaa gaa ttt ggg gat ttg     1056
Glu Gln Leu Asn Lys Ile Tyr Glu Arg Val Lys Glu Phe Gly Asp Leu
            340                 345                 350

ggt aaa tac att tca gac gct gat ttg ttg gct ata gtt aga gaa gtt     1104
Gly Lys Tyr Ile Ser Asp Ala Asp Leu Leu Ala Ile Val Arg Glu Val
        355                 360                 365

act gga aaa ttg gta gaa gag aaa atc aaa tta gat gaa tta act gtt     1152
Thr Gly Lys Leu Val Glu Glu Lys Ile Lys Leu Asp Glu Leu Thr Val
    370                 375                 380

gta tct gga aat aaa ata aca cca att gca tct gtt aaa ctc cat tat     1200
Val Ser Gly Asn Lys Ile Thr Pro Ile Ala Ser Val Lys Leu His Tyr
385                 390                 395                 400

aaa gga gaa gat ata act tta ata gaa act gct tat ggt gtt gga ccg     1248
Lys Gly Glu Asp Ile Thr Leu Ile Glu Thr Ala Tyr Gly Val Gly Pro
                405                 410                 415

gta gat gca gca ata aat gct gtg aga aag gca ata agt gga gtt gca     1296
Val Asp Ala Ala Ile Asn Ala Val Arg Lys Ala Ile Ser Gly Val Ala
            420                 425                 430
```

```
gat att aag ttg gta gag tat aga gtt gaa gca att ggt gga gga act      1344
Asp Ile Lys Leu Val Glu Tyr Arg Val Glu Ala Ile Gly Gly Gly Thr
        435                 440                 445

gat gcg tta ata gag gtt gtt gtt aaa tta aga aaa gga act gaa att      1392
Asp Ala Leu Ile Glu Val Val Val Lys Leu Arg Lys Gly Thr Glu Ile
    450                 455                 460

gtt gaa gtt aga aaa tca gac gct gat ata ata agg gct tct gta gat      1440
Val Glu Val Arg Lys Ser Asp Ala Asp Ile Ile Arg Ala Ser Val Asp
465                 470                 475                 480

gct gta atg gaa gga atc aat atg tta ttg aat taa                      1476
Ala Val Met Glu Gly Ile Asn Met Leu Leu Asn
                485                 490
```

<210> 58
<211> 491
<212> PRT
<213> 詹氏甲烷球菌

<400> 58

```
Met Met Val Arg Ile Phe Asp Thr Thr Leu Arg Asp Gly Glu Gln Thr
1               5                   10                  15

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

Leu Asp Glu Leu Gly Val Asp Val Ile Glu Ala Gly Ser Ala Ile Thr
        35                  40                  45

Ser Lys Gly Glu Arg Glu Gly Ile Lys Leu Ile Thr Lys Glu Gly Leu
    50                  55                  60

Asn Ala Glu Ile Cys Ser Phe Val Arg Ala Leu Pro Val Asp Ile Asp
65                  70                  75                  80

Ala Ala Leu Glu Cys Asp Val Asp Ser Val His Leu Val Val Pro Thr
                85                  90                  95

Ser Pro Ile His Met Lys Tyr Lys Leu Arg Lys Thr Glu Asp Glu Val
            100                 105                 110
```

Leu Glu Thr Ala Leu Lys Ala Val Glu Tyr Ala Lys Glu His Gly Leu
115 120 125

Ile Val Glu Leu Ser Ala Glu Asp Ala Thr Arg Ser Asp Val Asn Phe
130 135 140

Leu Ile Lys Leu Phe Asn Glu Gly Glu Lys Val Gly Ala Asp Arg Val
145 150 155 160

Cys Val Cys Asp Thr Val Gly Val Leu Thr Pro Gln Lys Ser Gln Glu
165 170 175

Leu Phe Lys Lys Ile Thr Glu Asn Val Asn Leu Pro Val Ser Val His
180 185 190

Cys His Asn Asp Phe Gly Met Ala Thr Ala Asn Thr Cys Ser Ala Val
195 200 205

Leu Gly Gly Ala Val Gln Cys His Val Thr Val Asn Gly Ile Gly Glu
210 215 220

Arg Ala Gly Asn Ala Ser Leu Glu Glu Val Val Ala Ala Leu Lys Ile
225 230 235 240

Leu Tyr Gly Tyr Asp Thr Lys Ile Lys Met Glu Lys Leu Tyr Glu Val
245 250 255

Ser Arg Ile Val Ser Arg Leu Met Lys Leu Pro Val Pro Pro Asn Lys
260 265 270

Ala Ile Val Gly Asp Asn Ala Phe Ala His Glu Ala Gly Ile His Val
275 280 285

Asp Gly Leu Ile Lys Asn Thr Glu Thr Tyr Glu Pro Ile Lys Pro Glu
290 295 300

Met Val Gly Asn Arg Arg Arg Ile Ile Leu Gly Lys His Ser Gly Arg
305 310 315 320

Lys Ala Leu Lys Tyr Lys Leu Asp Leu Met Gly Ile Asn Val Ser Asp
325 330 335

Glu Gln Leu Asn Lys Ile Tyr Glu Arg Val Lys Glu Phe Gly Asp Leu
340 345 350

Gly Lys Tyr Ile Ser Asp Ala Asp Leu Leu Ala Ile Val Arg Glu Val
355 360 365

Thr Gly Lys Leu Val Glu Glu Lys Ile Lys Leu Asp Glu Leu Thr Val
370 375 380

Val Ser Gly Asn Lys Ile Thr Pro Ile Ala Ser Val Lys Leu His Tyr
385 390 395 400

Lys Gly Glu Asp Ile Thr Leu Ile Glu Thr Ala Tyr Gly Val Gly Pro
405 410 415

Val Asp Ala Ala Ile Asn Ala Val Arg Lys Ala Ile Ser Gly Val Ala
420 425 430

Asp Ile Lys Leu Val Glu Tyr Arg Val Glu Ala Ile Gly Gly Gly Thr
435 440 445

Asp Ala Leu Ile Glu Val Val Val Lys Leu Arg Lys Gly Thr Glu Ile
450 455 460

Val Glu Val Arg Lys Ser Asp Ala Asp Ile Ile Arg Ala Ser Val Asp
465 470 475 480

Ala Val Met Glu Gly Ile Asn Met Leu Leu Asn
485 490

<210> 59
<211> 264
<212> DNA
<213> 大肠杆菌

<220>

<221> CDS
<222> (1)..(264)

<400> 59
atg atg caa cat cag gtc aat gta tcg gct cgc ttc aat cca gaa acc 48
Met Met Gln His Gln Val Asn Val Ser Ala Arg Phe Asn Pro Glu Thr
1 5 10 15

tta gaa cgt gtt tta cgc gtg gtg cgt cat cgt ggt ttc cac gtc tgc 96
Leu Glu Arg Val Leu Arg Val Val Arg His Arg Gly Phe His Val Cys
20 25 30

tca atg aat atg gcc gcc gcc agc gat gca caa aat ata aat atc gaa 144
Ser Met Asn Met Ala Ala Ala Ser Asp Ala Gln Asn Ile Asn Ile Glu
35 40 45

ttg acc gtt gcc agc cca cgg tcg gtc gac tta ctg ttt agt cag tta 192
Leu Thr Val Ala Ser Pro Arg Ser Val Asp Leu Leu Phe Ser Gln Leu
50 55 60

aat aaa ctg gtg gac gtc gca cac gtt gcc atc tgc cag agc aca acc 240
Asn Lys Leu Val Asp Val Ala His Val Ala Ile Cys Gln Ser Thr Thr
65 70 75 80

aca tca caa caa atc cgc gcc tga 264
Thr Ser Gln Gln Ile Arg Ala
85

<210> 60
<211> 87
<212> PRT
<213> 大肠杆菌

<400> 60

Met Met Gln His Gln Val Asn Val Ser Ala Arg Phe Asn Pro Glu Thr
1 5 10 15

Leu Glu Arg Val Leu Arg Val Val Arg His Arg Gly Phe His Val Cys
20 25 30

Ser Met Asn Met Ala Ala Ala Ser Asp Ala Gln Asn Ile Asn Ile Glu
35 40 45

Leu Thr Val Ala Ser Pro Arg Ser Val Asp Leu Leu Phe Ser Gln Leu
50 55 60

Asn Lys Leu Val Asp Val Ala His Val Ala Ile Cys Gln Ser Thr Thr
65 70 75 80

Thr Ser Gln Gln Ile Arg Ala
85

<210> 61
<211> 582
<212> DNA
<213> 大肠杆菌

<220>
<221> CDS
<222> (1)..(582)

<400> 61
ttg ttg tta aaa caa ctg tcg gat cgt aaa cct gcg gat tgc gtc gtg 48
Leu Leu Leu Lys Gln Leu Ser Asp Arg Lys Pro Ala Asp Cys Val Val
1 5 10 15

acc aca gat gtg ggg cag cac cag atg tgg gct gcg cag cac atc gcc 96
Thr Thr Asp Val Gly Gln His Gln Met Trp Ala Ala Gln His Ile Ala
20 25 30

cac act cgc ccg gaa aat ttc atc acc tcc agc ggt tta ggt acc atg 144
His Thr Arg Pro Glu Asn Phe Ile Thr Ser Ser Gly Leu Gly Thr Met
35 40 45

ggt ttt ggt tta ccg gcg gcg gtt ggc gca caa gtc gcg cga ccg aac 192
Gly Phe Gly Leu Pro Ala Ala Val Gly Ala Gln Val Ala Arg Pro Asn
50 55 60

gat acc gtt gtc tgt atc tcc ggt gac ggc tct ttc atg atg aat gtg 240
Asp Thr Val Val Cys Ile Ser Gly Asp Gly Ser Phe Met Met Asn Val
65 70 75 80

caa gag ctg ggc acc gta aaa cgc aag cag tta ccg ttg aaa atc gtc 288
Gln Glu Leu Gly Thr Val Lys Arg Lys Gln Leu Pro Leu Lys Ile Val
85 90 95

tta ctc gat aac caa cgg tta ggg atg gtt cga caa tgg cag caa ctg 336
Leu Leu Asp Asn Gln Arg Leu Gly Met Val Arg Gln Trp Gln Gln Leu
100 105 110

ttt ttt cag gaa cga tac agc gaa acc acc ctt act gat aac ccc gat 384
Phe Phe Gln Glu Arg Tyr Ser Glu Thr Thr Leu Thr Asp Asn Pro Asp
115 120 125

```
ttc ctc atg tta gcc agc gcc ttc ggc atc cat ggc caa cac atc acc      432
Phe Leu Met Leu Ala Ser Ala Phe Gly Ile His Gly Gln His Ile Thr
    130                 135                 140

cgg aaa gac cag gtt gaa gcg gca ctc gac acc atg ctg aac agt gat      480
Arg Lys Asp Gln Val Glu Ala Ala Leu Asp Thr Met Leu Asn Ser Asp
145                 150                 155                 160

ggg cca tac ctg ctt cat gtc tca atc gac gaa ctt gag aac gtc tgg      528
Gly Pro Tyr Leu Leu His Val Ser Ile Asp Glu Leu Glu Asn Val Trp
                165                 170                 175

ccg ctg gtg ccg cct ggc gcc agt aat tca gaa atg ttg gag aaa tta      576
Pro Leu Val Pro Pro Gly Ala Ser Asn Ser Glu Met Leu Glu Lys Leu
            180                 185                 190

tca tga                                                             582
Ser
```

<210> 62
<211> 193
<212> PRT
<213> 大肠杆菌

<400> 62

```
Leu Leu Leu Lys Gln Leu Ser Asp Arg Lys Pro Ala Asp Cys Val Val
1               5                   10                  15

Thr Thr Asp Val Gly Gln His Gln Met Trp Ala Ala Gln His Ile Ala
            20                  25                  30

His Thr Arg Pro Glu Asn Phe Ile Thr Ser Ser Gly Leu Gly Thr Met
        35                  40                  45

Gly Phe Gly Leu Pro Ala Ala Val Gly Ala Gln Val Ala Arg Pro Asn
    50                  55                  60

Asp Thr Val Val Cys Ile Ser Gly Asp Gly Ser Phe Met Met Asn Val
65                  70                  75                  80

Gln Glu Leu Gly Thr Val Lys Arg Lys Gln Leu Pro Leu Lys Ile Val
                85                  90                  95
```

Leu Leu Asp Asn Gln Arg Leu Gly Met Val Arg Gln Trp Gln Gln Leu
100 105 110

Phe Phe Gln Glu Arg Tyr Ser Glu Thr Thr Leu Thr Asp Asn Pro Asp
115 120 125

Phe Leu Met Leu Ala Ser Ala Phe Gly Ile His Gly Gln His Ile Thr
130 135 140

Arg Lys Asp Gln Val Glu Ala Ala Leu Asp Thr Met Leu Asn Ser Asp
145 150 155 160

Gly Pro Tyr Leu Leu His Val Ser Ile Asp Glu Leu Glu Asn Val Trp
165 170 175

Pro Leu Val Pro Pro Gly Ala Ser Asn Ser Glu Met Leu Glu Lys Leu
180 185 190

Ser

<210> 63
<211> 291
<212> DNA
<213> 大肠杆菌

<220>
<221> CDS
<222> (1)..(291)

<400> 63
atg caa aac aca act cat gac aac gta att ctg gag ctc acc gtt cgc 48
Met Gln Asn Thr Thr His Asp Asn Val Ile Leu Glu Leu Thr Val Arg
1 5 10 15

aac cat ccg ggc gta atg acc cac gtt tgt ggc ctt ttt gcc cgc cgc 96
Asn His Pro Gly Val Met Thr His Val Cys Gly Leu Phe Ala Arg Arg
20 25 30

gct ttt aac gtt gaa ggc att ctt tgt ctg ccg att cag gac agc gac 144
Ala Phe Asn Val Glu Gly Ile Leu Cys Leu Pro Ile Gln Asp Ser Asp
35 40 45

aaa agc cat atc tgg cta ctg gtc aat gac gac cag cgt ctg gag cag 192
Lys Ser His Ile Trp Leu Leu Val Asn Asp Asp Gln Arg Leu Glu Gln
50 55 60

atg ata agc caa atc gat aag ctg gaa gat gtc gtg aaa gtg cag cgt 240
Met Ile Ser Gln Ile Asp Lys Leu Glu Asp Val Val Lys Val Gln Arg
65 70 75 80

aat cag tcc gat ccg acg atg ttt aac aag atc gcg gtg ttt ttt cag 288
Asn Gln Ser Asp Pro Thr Met Phe Asn Lys Ile Ala Val Phe Phe Gln
85 90 95

taa 291

<210> 64
<211> 96
<212> PRT
<213> 大肠杆菌

<400> 64

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

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

Ala Phe Asn Val Glu Gly Ile Leu Cys Leu Pro Ile Gln Asp Ser Asp
35 40 45

Lys Ser His Ile Trp Leu Leu Val Asn Asp Asp Gln Arg Leu Glu Gln
50 55 60

Met Ile Ser Gln Ile Asp Lys Leu Glu Asp Val Val Lys Val Gln Arg
65 70 75 80

Asn Gln Ser Asp Pro Thr Met Phe Asn Lys Ile Ala Val Phe Phe Gln
85 90 95

<210> 65
<211> 1689
<212> DNA
<213> 大肠杆菌

<220>
<221> CDS
<222> (1)..(1689)

<400> 65

```
atg gca agt tcg ggc aca aca tcg acg cgt aag cgc ttt acc ggc gca      48
Met Ala Ser Ser Gly Thr Thr Ser Thr Arg Lys Arg Phe Thr Gly Ala
1               5                  10                  15

gaa ttt atc gtt cat ttc ctg gaa cag cag ggc att aag att gtg aca      96
Glu Phe Ile Val His Phe Leu Glu Gln Gln Gly Ile Lys Ile Val Thr
            20                  25                  30

ggc att ccg ggc ggt tct atc ctg cct gtt tac gat gcc tta agc caa     144
Gly Ile Pro Gly Gly Ser Ile Leu Pro Val Tyr Asp Ala Leu Ser Gln
        35                  40                  45

agc acg caa atc cgc cat att ctg gcc cgt cat gaa cag ggc gcg ggc     192
Ser Thr Gln Ile Arg His Ile Leu Ala Arg His Glu Gln Gly Ala Gly
    50                  55                  60

ttt atc gct cag gga atg gcg cgc acc gac ggt aaa ccg gcg gtc tgt     240
Phe Ile Ala Gln Gly Met Ala Arg Thr Asp Gly Lys Pro Ala Val Cys
65                  70                  75                  80

atg gcc tgt agc gga ccg ggt gcg act aac ctg gtg acc gcc att gcc     288
Met Ala Cys Ser Gly Pro Gly Ala Thr Asn Leu Val Thr Ala Ile Ala
                85                  90                  95

gat gcg cgg ctg gac tcc atc ccg ctg att tgc atc act ggt cag gtt     336
Asp Ala Arg Leu Asp Ser Ile Pro Leu Ile Cys Ile Thr Gly Gln Val
            100                 105                 110

ccc gcc tcg atg atc ggc acc gac gcc ttc cag gaa gtg gac acc tac     384
Pro Ala Ser Met Ile Gly Thr Asp Ala Phe Gln Glu Val Asp Thr Tyr
        115                 120                 125

ggc atc tct atc ccc atc acc aaa cac aac tat ctg gtc aga cat atc     432
Gly Ile Ser Ile Pro Ile Thr Lys His Asn Tyr Leu Val Arg His Ile
    130                 135                 140

gaa gaa ctc ccg cag gtc atg agc gat gcc ttc cgc att gcg caa tca     480
Glu Glu Leu Pro Gln Val Met Ser Asp Ala Phe Arg Ile Ala Gln Ser
145                 150                 155                 160

ggc cgc cca ggc ccg gtg tgg ata gac att cct aag gat gtg caa acg     528
Gly Arg Pro Gly Pro Val Trp Ile Asp Ile Pro Lys Asp Val Gln Thr
                165                 170                 175
```

```
gca gtt ttt gag att gaa aca cag ccc gct atg gca gaa aaa gcc gcc      576
Ala Val Phe Glu Ile Glu Thr Gln Pro Ala Met Ala Glu Lys Ala Ala
            180                 185                 190

gcc ccc gcc ttt agc gaa gaa agc att cgt gac gca gcg gcg atg att      624
Ala Pro Ala Phe Ser Glu Glu Ser Ile Arg Asp Ala Ala Ala Met Ile
        195                 200                 205

aac gct gcc aaa cgc ccg gtg ctt tat ctg ggc ggc ggt gtg atc aat      672
Asn Ala Ala Lys Arg Pro Val Leu Tyr Leu Gly Gly Gly Val Ile Asn
    210                 215                 220

gcg ccc gca cgg gtg cgt gaa ctg gcg gag aaa gcg caa ctg cct acc      720
Ala Pro Ala Arg Val Arg Glu Leu Ala Glu Lys Ala Gln Leu Pro Thr
225                 230                 235                 240

acc atg act tta atg gcg ctg ggc atg ttg cca aaa gcg cat ccg ttg      768
Thr Met Thr Leu Met Ala Leu Gly Met Leu Pro Lys Ala His Pro Leu
                245                 250                 255

tcg ctg ggt atg ctg ggg atg cac ggc gtg cgc agc acc aac tat att      816
Ser Leu Gly Met Leu Gly Met His Gly Val Arg Ser Thr Asn Tyr Ile
            260                 265                 270

ttg cag gag gcg gat ttg ttg ata gtg ctc ggt gcg cgt ttt gat gac      864
Leu Gln Glu Ala Asp Leu Leu Ile Val Leu Gly Ala Arg Phe Asp Asp
        275                 280                 285

cgg gcg att ggc aaa acc gag cag ttc tgt ccg aat gcc aaa atc att      912
Arg Ala Ile Gly Lys Thr Glu Gln Phe Cys Pro Asn Ala Lys Ile Ile
    290                 295                 300

cat gtc gat atc gac cgt gca gag ctg ggt aaa atc aag cag ccg cac      960
His Val Asp Ile Asp Arg Ala Glu Leu Gly Lys Ile Lys Gln Pro His
305                 310                 315                 320

gtg gcg att cag gcg gat gtt gat gac gtg ctg gcg cag ttg atc ccg     1008
Val Ala Ile Gln Ala Asp Val Asp Asp Val Leu Ala Gln Leu Ile Pro
                325                 330                 335

ctg gtg gaa gcg caa ccg cgt gca gag tgg cac cag ttg gta gcg gat     1056
Leu Val Glu Ala Gln Pro Arg Ala Glu Trp His Gln Leu Val Ala Asp
            340                 345                 350

ttg cag cgt gag ttt ccg tgt cca atc ccg aaa gcg tgc gat ccg tta     1104
Leu Gln Arg Glu Phe Pro Cys Pro Ile Pro Lys Ala Cys Asp Pro Leu
        355                 360                 365

agc cat tac ggc ctg atc aac gcc gtt gcc gcc tgt gtc gat gac aat     1152
Ser His Tyr Gly Leu Ile Asn Ala Val Ala Ala Cys Val Asp Asp Asn
    370                 375                 380
```

```
gca att atc acc acc gac gtt ggt cag cat cag atg tgg acc gcg caa      1200
Ala Ile Ile Thr Thr Asp Val Gly Gln His Gln Met Trp Thr Ala Gln
385                 390                 395                 400

gct tat ccg ctc aat cgc cca cgc cag tgg ctg acc tcc ggt ggg ctg      1248
Ala Tyr Pro Leu Asn Arg Pro Arg Gln Trp Leu Thr Ser Gly Gly Leu
                405                 410                 415

ggc acg atg ggt ttt ggc ctg cct gcg gcg att ggc gct gcg ctg gcg      1296
Gly Thr Met Gly Phe Gly Leu Pro Ala Ala Ile Gly Ala Ala Leu Ala
            420                 425                 430

aac ccg gat cgc aaa gtg ttg tgt ttc tcc ggc gac ggc agc ctg atg      1344
Asn Pro Asp Arg Lys Val Leu Cys Phe Ser Gly Asp Gly Ser Leu Met
        435                 440                 445

atg aat att cag gag atg gcg acc gcc agt gaa aat cag ctg gat gtc      1392
Met Asn Ile Gln Glu Met Ala Thr Ala Ser Glu Asn Gln Leu Asp Val
    450                 455                 460

aaa atc att ctg atg aac aac gaa gcg ctg ggg ctg gtg cat cag caa      1440
Lys Ile Ile Leu Met Asn Asn Glu Ala Leu Gly Leu Val His Gln Gln
465                 470                 475                 480

cag agt ctg ttc tac gag caa ggc gtt ttt gcc gcc acc tat ccg ggc      1488
Gln Ser Leu Phe Tyr Glu Gln Gly Val Phe Ala Ala Thr Tyr Pro Gly
                485                 490                 495

aaa atc aac ttt atg cag att gcc gcc gga ttc ggc ctc gaa acc tgt      1536
Lys Ile Asn Phe Met Gln Ile Ala Ala Gly Phe Gly Leu Glu Thr Cys
            500                 505                 510

gat ttg aat aac gaa gcc gat ccg cag gct tca ttg cag gaa atc atc      1584
Asp Leu Asn Asn Glu Ala Asp Pro Gln Ala Ser Leu Gln Glu Ile Ile
        515                 520                 525

aat cgc cct ggc ccg gcg ctg atc cat gtg cgc att gat gcc gaa gaa      1632
Asn Arg Pro Gly Pro Ala Leu Ile His Val Arg Ile Asp Ala Glu Glu
    530                 535                 540

aaa gtt tac ccg atg gtg ccg cca ggt gcg gcg aat act gaa atg gtg      1680
Lys Val Tyr Pro Met Val Pro Pro Gly Ala Ala Asn Thr Glu Met Val
545                 550                 555                 560

ggg gaa taa                                                        1689
Gly Glu
```

<210> 66
<211> 562

<212> PRT
<213> 大肠杆菌

<400> 66

Met Ala Ser Ser Gly Thr Thr Ser Thr Arg Lys Arg Phe Thr Gly Ala
1 5 10 15

Glu Phe Ile Val His Phe Leu Glu Gln Gln Gly Ile Lys Ile Val Thr
20 25 30

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

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

Phe Ile Ala Gln Gly Met Ala Arg Thr Asp Gly Lys Pro Ala Val Cys
65 70 75 80

Met Ala Cys Ser Gly Pro Gly Ala Thr Asn Leu Val Thr Ala Ile Ala
85 90 95

Asp Ala Arg Leu Asp Ser Ile Pro Leu Ile Cys Ile Thr Gly Gln Val
100 105 110

Pro Ala Ser Met Ile Gly Thr Asp Ala Phe Gln Glu Val Asp Thr Tyr
115 120 125

Gly Ile Ser Ile Pro Ile Thr Lys His Asn Tyr Leu Val Arg His Ile
130 135 140

Glu Glu Leu Pro Gln Val Met Ser Asp Ala Phe Arg Ile Ala Gln Ser
145 150 155 160

Gly Arg Pro Gly Pro Val Trp Ile Asp Ile Pro Lys Asp Val Gln Thr
165 170 175

Ala Val Phe Glu Ile Glu Thr Gln Pro Ala Met Ala Glu Lys Ala Ala
180 185 190

Ala Pro Ala Phe Ser Glu Glu Ser Ile Arg Asp Ala Ala Ala Met Ile
195 200 205

Asn Ala Ala Lys Arg Pro Val Leu Tyr Leu Gly Gly Gly Val Ile Asn
210 215 220

Ala Pro Ala Arg Val Arg Glu Leu Ala Glu Lys Ala Gln Leu Pro Thr
225 230 235 240

Thr Met Thr Leu Met Ala Leu Gly Met Leu Pro Lys Ala His Pro Leu
245 250 255

Ser Leu Gly Met Leu Gly Met His Gly Val Arg Ser Thr Asn Tyr Ile
260 265 270

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

Arg Ala Ile Gly Lys Thr Glu Gln Phe Cys Pro Asn Ala Lys Ile Ile
290 295 300

His Val Asp Ile Asp Arg Ala Glu Leu Gly Lys Ile Lys Gln Pro His
305 310 315 320

Val Ala Ile Gln Ala Asp Val Asp Asp Val Leu Ala Gln Leu Ile Pro
325 330 335

Leu Val Glu Ala Gln Pro Arg Ala Glu Trp His Gln Leu Val Ala Asp
340 345 350

Leu Gln Arg Glu Phe Pro Cys Pro Ile Pro Lys Ala Cys Asp Pro Leu
355 360 365

Ser His Tyr Gly Leu Ile Asn Ala Val Ala Ala Cys Val Asp Asp Asn
370 375 380

Ala Ile Ile Thr Thr Asp Val Gly Gln His Gln Met Trp Thr Ala Gln
385 390 395 400

Ala Tyr Pro Leu Asn Arg Pro Arg Gln Trp Leu Thr Ser Gly Gly Leu
405 410 415

Gly Thr Met Gly Phe Gly Leu Pro Ala Ala Ile Gly Ala Ala Leu Ala
420 425 430

Asn Pro Asp Arg Lys Val Leu Cys Phe Ser Gly Asp Gly Ser Leu Met
435 440 445

Met Asn Ile Gln Glu Met Ala Thr Ala Ser Glu Asn Gln Leu Asp Val
450 455 460

Lys Ile Ile Leu Met Asn Asn Glu Ala Leu Gly Leu Val His Gln Gln
465 470 475 480

Gln Ser Leu Phe Tyr Glu Gln Gly Val Phe Ala Ala Thr Tyr Pro Gly
485 490 495

Lys Ile Asn Phe Met Gln Ile Ala Ala Gly Phe Gly Leu Glu Thr Cys
500 505 510

Asp Leu Asn Asn Glu Ala Asp Pro Gln Ala Ser Leu Gln Glu Ile Ile
515 520 525

Asn Arg Pro Gly Pro Ala Leu Ile His Val Arg Ile Asp Ala Glu Glu
530 535 540

Lys Val Tyr Pro Met Val Pro Pro Gly Ala Ala Asn Thr Glu Met Val
545 550 555 560

Gly Glu

<210> 67
<211> 2577
<212> DNA
<213> 丙酮丁醇梭菌

<220>

<221> CDS
<222> (1)..(2577)

<400> 67

```
atg aaa gtt aca aat caa aaa gaa cta aaa caa aag cta aat gaa ttg      48
Met Lys Val Thr Asn Gln Lys Glu Leu Lys Gln Lys Leu Asn Glu Leu
1               5                  10                  15

aga gaa gcg caa aag aag ttt gca acc tat act caa gag caa gtt gat      96
Arg Glu Ala Gln Lys Lys Phe Ala Thr Tyr Thr Gln Glu Gln Val Asp
            20                  25                  30

aaa att ttt aaa caa tgt gcc ata gcc gca gct aaa gaa aga ata aac     144
Lys Ile Phe Lys Gln Cys Ala Ile Ala Ala Ala Lys Glu Arg Ile Asn
        35                  40                  45

tta gct aaa tta gca gta gaa gaa aca gga ata ggt ctt gta gaa gat     192
Leu Ala Lys Leu Ala Val Glu Glu Thr Gly Ile Gly Leu Val Glu Asp
    50                  55                  60

aaa att ata aaa aat cat ttt gca gca gaa tat ata tac aat aaa tat     240
Lys Ile Ile Lys Asn His Phe Ala Ala Glu Tyr Ile Tyr Asn Lys Tyr
65                  70                  75                  80

aaa aat gaa aaa act tgt ggc ata ata gac cat gac gat tct tta ggc     288
Lys Asn Glu Lys Thr Cys Gly Ile Ile Asp His Asp Asp Ser Leu Gly
                85                  90                  95

ata aca aag gtt gct gaa cca att gga att gtt gca gcc ata gtt cct     336
Ile Thr Lys Val Ala Glu Pro Ile Gly Ile Val Ala Ala Ile Val Pro
            100                 105                 110

act act aat cca act tcc aca gca att ttc aaa tca tta att tct tta     384
Thr Thr Asn Pro Thr Ser Thr Ala Ile Phe Lys Ser Leu Ile Ser Leu
        115                 120                 125

aaa aca aga aac gca ata ttc ttt tca cca cat cca cgt gca aaa aaa     432
Lys Thr Arg Asn Ala Ile Phe Phe Ser Pro His Pro Arg Ala Lys Lys
    130                 135                 140

tct aca att gct gca gca aaa tta att tta gat gca gct gtt aaa gca     480
Ser Thr Ile Ala Ala Ala Lys Leu Ile Leu Asp Ala Ala Val Lys Ala
145                 150                 155                 160

gga gca cct aaa aat ata ata ggc tgg ata gat gag cca tca ata gaa     528
Gly Ala Pro Lys Asn Ile Ile Gly Trp Ile Asp Glu Pro Ser Ile Glu
                165                 170                 175

ctt tct caa gat ttg atg agt gaa gct gat ata ata tta gca aca gga     576
Leu Ser Gln Asp Leu Met Ser Glu Ala Asp Ile Ile Leu Ala Thr Gly
            180                 185                 190
```

```
ggt cct tca atg gtt aaa gcg gcc tat tca tct gga aaa cct gca att      624
Gly Pro Ser Met Val Lys Ala Ala Tyr Ser Ser Gly Lys Pro Ala Ile
        195                 200                 205

ggt gtt gga gca gga aat aca cca gca ata ata gat gag agt gca gat      672
Gly Val Gly Ala Gly Asn Thr Pro Ala Ile Ile Asp Glu Ser Ala Asp
    210                 215                 220

ata gat atg gca gta agc tcc ata att tta tca aag act tat gac aat      720
Ile Asp Met Ala Val Ser Ser Ile Ile Leu Ser Lys Thr Tyr Asp Asn
225                 230                 235                 240

gga gta ata tgc gct tct gaa caa tca ata tta gtt atg aat tca ata      768
Gly Val Ile Cys Ala Ser Glu Gln Ser Ile Leu Val Met Asn Ser Ile
                245                 250                 255

tac gaa aaa gtt aaa gag gaa ttt gta aaa cga gga tca tat ata ctc      816
Tyr Glu Lys Val Lys Glu Glu Phe Val Lys Arg Gly Ser Tyr Ile Leu
            260                 265                 270

aat caa aat gaa ata gct aaa ata aaa gaa act atg ttt aaa aat gga      864
Asn Gln Asn Glu Ile Ala Lys Ile Lys Glu Thr Met Phe Lys Asn Gly
        275                 280                 285

gct att aat gct gac ata gtt gga aaa tct gct tat ata att gct aaa      912
Ala Ile Asn Ala Asp Ile Val Gly Lys Ser Ala Tyr Ile Ile Ala Lys
    290                 295                 300

atg gca gga att gaa gtt cct caa act aca aag ata ctt ata ggc gaa      960
Met Ala Gly Ile Glu Val Pro Gln Thr Thr Lys Ile Leu Ile Gly Glu
305                 310                 315                 320

gta caa tct gtt gaa aaa agc gag ctg ttc tca cat gaa aaa cta tca     1008
Val Gln Ser Val Glu Lys Ser Glu Leu Phe Ser His Glu Lys Leu Ser
                325                 330                 335

cca gta ctt gca atg tat aaa gtt aag gat ttt gat gaa gct cta aaa     1056
Pro Val Leu Ala Met Tyr Lys Val Lys Asp Phe Asp Glu Ala Leu Lys
            340                 345                 350

aag gca caa agg cta ata gaa tta ggt gga agt gga cac acg tca tct     1104
Lys Ala Gln Arg Leu Ile Glu Leu Gly Gly Ser Gly His Thr Ser Ser
        355                 360                 365

tta tat ata gat tca caa aac aat aag gat aaa gtt aaa gaa ttt gga     1152
Leu Tyr Ile Asp Ser Gln Asn Asn Lys Asp Lys Val Lys Glu Phe Gly
    370                 375                 380

tta gca atg aaa act tca agg aca ttt att aac atg cct tct tca cag     1200
Leu Ala Met Lys Thr Ser Arg Thr Phe Ile Asn Met Pro Ser Ser Gln
385                 390                 395                 400
```

```
gga gca agc gga gat tta tac aat ttt gcg ata gca cca tca ttt act      1248
Gly Ala Ser Gly Asp Leu Tyr Asn Phe Ala Ile Ala Pro Ser Phe Thr
                405                 410                 415

ctt gga tgc ggc act tgg gga gga aac tct gta tcg caa aat gta gag      1296
Leu Gly Cys Gly Thr Trp Gly Gly Asn Ser Val Ser Gln Asn Val Glu
            420                 425                 430

cct aaa cat tta tta aat att aaa agt gtt gct gaa aga agg gaa aat      1344
Pro Lys His Leu Leu Asn Ile Lys Ser Val Ala Glu Arg Arg Glu Asn
        435                 440                 445

atg ctt tgg ttt aaa gtg cca caa aaa ata tat ttt aaa tat gga tgt      1392
Met Leu Trp Phe Lys Val Pro Gln Lys Ile Tyr Phe Lys Tyr Gly Cys
    450                 455                 460

ctt aga ttt gca tta aaa gaa tta aaa gat atg aat aag aaa aga gcc      1440
Leu Arg Phe Ala Leu Lys Glu Leu Lys Asp Met Asn Lys Lys Arg Ala
465                 470                 475                 480

ttt ata gta aca gat aaa gat ctt ttt aaa ctt gga tat gtt aat aaa      1488
Phe Ile Val Thr Asp Lys Asp Leu Phe Lys Leu Gly Tyr Val Asn Lys
                485                 490                 495

ata aca aag gta cta gat gag ata gat att aaa tac agt ata ttt aca      1536
Ile Thr Lys Val Leu Asp Glu Ile Asp Ile Lys Tyr Ser Ile Phe Thr
            500                 505                 510

gat att aaa tct gat cca act att gat tca gta aaa aaa ggt gct aaa      1584
Asp Ile Lys Ser Asp Pro Thr Ile Asp Ser Val Lys Lys Gly Ala Lys
        515                 520                 525

gaa atg ctt aac ttt gaa cct gat act ata atc tct att ggt ggt gga      1632
Glu Met Leu Asn Phe Glu Pro Asp Thr Ile Ile Ser Ile Gly Gly Gly
    530                 535                 540

tcg cca atg gat gca gca aag gtt atg cac ttg tta tat gaa tat cca      1680
Ser Pro Met Asp Ala Ala Lys Val Met His Leu Leu Tyr Glu Tyr Pro
545                 550                 555                 560

gaa gca gaa att gaa aat cta gct ata aac ttt atg gat ata aga aag      1728
Glu Ala Glu Ile Glu Asn Leu Ala Ile Asn Phe Met Asp Ile Arg Lys
                565                 570                 575

aga ata tgc aat ttc cct aaa tta ggt aca aag gcg att tca gta gct      1776
Arg Ile Cys Asn Phe Pro Lys Leu Gly Thr Lys Ala Ile Ser Val Ala
            580                 585                 590

att cct aca act gct ggt acc ggt tca gag gca aca cct ttt gca gtt      1824
Ile Pro Thr Thr Ala Gly Thr Gly Ser Glu Ala Thr Pro Phe Ala Val
        595                 600                 605
```

```
ata act aat gat gaa aca gga atg aaa tac cct tta act tct tat gaa      1872
Ile Thr Asn Asp Glu Thr Gly Met Lys Tyr Pro Leu Thr Ser Tyr Glu
    610                 615                 620

ttg acc cca aac atg gca ata ata gat act gaa tta atg tta aat atg      1920
Leu Thr Pro Asn Met Ala Ile Ile Asp Thr Glu Leu Met Leu Asn Met
625                 630                 635                 640

cct aga aaa tta aca gca gca act gga ata gat gca tta gtt cat gct      1968
Pro Arg Lys Leu Thr Ala Ala Thr Gly Ile Asp Ala Leu Val His Ala
                645                 650                 655

ata gaa gca tat gtt tcg gtt atg gct acg gat tat act gat gaa tta      2016
Ile Glu Ala Tyr Val Ser Val Met Ala Thr Asp Tyr Thr Asp Glu Leu
            660                 665                 670

gcc tta aga gca ata aaa atg ata ttt aaa tat ttg cct aga gcc tat      2064
Ala Leu Arg Ala Ile Lys Met Ile Phe Lys Tyr Leu Pro Arg Ala Tyr
        675                 680                 685

aaa aat ggg act aac gac att gaa gca aga gaa aaa atg gca cat gcc      2112
Lys Asn Gly Thr Asn Asp Ile Glu Ala Arg Glu Lys Met Ala His Ala
    690                 695                 700

tct aat att gcg ggg atg gca ttt gca aat gct ttc tta ggt gta tgc      2160
Ser Asn Ile Ala Gly Met Ala Phe Ala Asn Ala Phe Leu Gly Val Cys
705                 710                 715                 720

cat tca atg gct cat aaa ctt ggg gca atg cat cac gtt cca cat gga      2208
His Ser Met Ala His Lys Leu Gly Ala Met His His Val Pro His Gly
                725                 730                 735

att gct tgt gct gta tta ata gaa gaa gtt att aaa tat aac gct aca      2256
Ile Ala Cys Ala Val Leu Ile Glu Glu Val Ile Lys Tyr Asn Ala Thr
            740                 745                 750

gac tgt cca aca aag caa aca gca ttc cct caa tat aaa tct cct aat      2304
Asp Cys Pro Thr Lys Gln Thr Ala Phe Pro Gln Tyr Lys Ser Pro Asn
        755                 760                 765

gct aag aga aaa tat gct gaa att gca gag tat ttg aat tta aag ggt      2352
Ala Lys Arg Lys Tyr Ala Glu Ile Ala Glu Tyr Leu Asn Leu Lys Gly
    770                 775                 780

act agc gat acc gaa aag gta aca gcc tta ata gaa gct att tca aag      2400
Thr Ser Asp Thr Glu Lys Val Thr Ala Leu Ile Glu Ala Ile Ser Lys
785                 790                 795                 800

tta aag ata gat ttg agt att cca caa aat ata agt gcc gct gga ata      2448
Leu Lys Ile Asp Leu Ser Ile Pro Gln Asn Ile Ser Ala Ala Gly Ile
                805                 810                 815
```

```
aat aaa aaa gat ttt tat aat acg cta gat aaa atg tca gag ctt gct     2496
Asn Lys Lys Asp Phe Tyr Asn Thr Leu Asp Lys Met Ser Glu Leu Ala
            820                 825                 830

ttt gat gac caa tgt aca aca gct aat cct agg tat cca ctt ata agt     2544
Phe Asp Asp Gln Cys Thr Thr Ala Asn Pro Arg Tyr Pro Leu Ile Ser
        835                 840                 845

gaa ctt aag gat atc tat ata aaa tca ttt taa                        2577
Glu Leu Lys Asp Ile Tyr Ile Lys Ser Phe
    850                 855
```

<210> 68
<211> 858
<212> PRT
<213> 丙酮丁醇梭菌

<400> 68

```
Met Lys Val Thr Asn Gln Lys Glu Leu Lys Gln Lys Leu Asn Glu Leu
1               5                   10                  15

Arg Glu Ala Gln Lys Lys Phe Ala Thr Tyr Thr Gln Glu Gln Val Asp
            20                  25                  30

Lys Ile Phe Lys Gln Cys Ala Ile Ala Ala Ala Lys Glu Arg Ile Asn
        35                  40                  45

Leu Ala Lys Leu Ala Val Glu Glu Thr Gly Ile Gly Leu Val Glu Asp
    50                  55                  60

Lys Ile Ile Lys Asn His Phe Ala Ala Glu Tyr Ile Tyr Asn Lys Tyr
65                  70                  75                  80

Lys Asn Glu Lys Thr Cys Gly Ile Ile Asp His Asp Asp Ser Leu Gly
                85                  90                  95

Ile Thr Lys Val Ala Glu Pro Ile Gly Ile Val Ala Ala Ile Val Pro
            100                 105                 110

Thr Thr Asn Pro Thr Ser Thr Ala Ile Phe Lys Ser Leu Ile Ser Leu
        115                 120                 125
```

Lys Thr Arg Asn Ala Ile Phe Phe Ser Pro Ilis Pro Arg Ala Lys Lys
130 135 140

Ser Thr Ile Ala Ala Ala Lys Leu Ilc Leu Asp Ala Ala Val Lys Ala
145 150 155 160

Gly Ala Pro Lys Asn Ile Ile Gly Trp Ile Asp Glu Pro Ser Ile Glu
165 170 175

Leu Ser Gln Asp Leu Met Ser Glu Ala Asp Ile Ile Leu Ala Thr Gly
180 185 190

Gly Pro Ser Met Val Lys Ala Ala Tyr Ser Ser Gly Lys Pro Ala Ile
195 200 205

Gly Val Gly Ala Gly Asn Thr Pro Ala Ile Ile Asp Glu Ser Ala Asp
210 215 220

Ile Asp Met Ala Val Ser Ser Ile Ile Leu Ser Lys Thr Tyr Asp Asn
225 230 235 240

Gly Val Ile Cys Ala Ser Glu Gln Ser Ile Leu Val Met Asn Ser Ile
245 250 255

Tyr Glu Lys Val Lys Glu Glu Phe Val Lys Arg Gly Ser Tyr Ile Leu
260 265 270

Asn Gln Asn Glu Ile Ala Lys Ile Lys Glu Thr Met Phe Lys Asn Gly
275 280 285

Ala Ile Asn Ala Asp Ile Val Gly Lys Ser Ala Tyr Ile Ile Ala Lys
290 295 300

Met Ala Gly Ile Glu Val Pro Gln Thr Thr Lys Ile Leu Ile Gly Glu
305 310 315 320

Val Gln Ser Val Glu Lys Ser Glu Leu Phe Ser His Glu Lys Leu Ser
325 330 335

Pro Val Leu Ala Met Tyr Lys Val Lys Asp Phe Asp Glu Ala Leu Lys
340 345 350

Lys Ala Gln Arg Leu Ile Glu Leu Gly Gly Ser Gly His Thr Ser Ser
355 360 365

Leu Tyr Ile Asp Ser Gln Asn Asn Lys Asp Lys Val Lys Glu Phe Gly
370 375 380

Leu Ala Met Lys Thr Ser Arg Thr Phe Ile Asn Met Pro Ser Ser Gln
385 390 395 400

Gly Ala Ser Gly Asp Leu Tyr Asn Phe Ala Ile Ala Pro Ser Phe Thr
405 410 415

Leu Gly Cys Gly Thr Trp Gly Gly Asn Ser Val Ser Gln Asn Val Glu
420 425 430

Pro Lys His Leu Leu Asn Ile Lys Ser Val Ala Glu Arg Arg Glu Asn
435 440 445

Met Leu Trp Phe Lys Val Pro Gln Lys Ile Tyr Phe Lys Tyr Gly Cys
450 455 460

Leu Arg Phe Ala Leu Lys Glu Leu Lys Asp Met Asn Lys Lys Arg Ala
465 470 475 480

Phe Ile Val Thr Asp Lys Asp Leu Phe Lys Leu Gly Tyr Val Asn Lys
485 490 495

Ile Thr Lys Val Leu Asp Glu Ile Asp Ile Lys Tyr Ser Ile Phe Thr
500 505 510

Asp Ile Lys Ser Asp Pro Thr Ile Asp Ser Val Lys Lys Gly Ala Lys
515 520 525

Glu Met Leu Asn Phe Glu Pro Asp Thr Ile Ile Ser Ile Gly Gly Gly
530 535 540

Ser Pro Met Asp Ala Ala Lys Val Met His Leu Leu Tyr Glu Tyr Pro
545 550 555 560

Glu Ala Glu Ile Glu Asn Leu Ala Ile Asn Phe Met Asp Ile Arg Lys
565 570 575

Arg Ile Cys Asn Phe Pro Lys Leu Gly Thr Lys Ala Ile Ser Val Ala
580 585 590

Ile Pro Thr Thr Ala Gly Thr Gly Ser Glu Ala Thr Pro Phe Ala Val
595 600 605

Ile Thr Asn Asp Glu Thr Gly Met Lys Tyr Pro Leu Thr Ser Tyr Glu
610 615 620

Leu Thr Pro Asn Met Ala Ile Ile Asp Thr Glu Leu Met Leu Asn Met
625 630 635 640

Pro Arg Lys Leu Thr Ala Ala Thr Gly Ile Asp Ala Leu Val His Ala
645 650 655

Ile Glu Ala Tyr Val Ser Val Met Ala Thr Asp Tyr Thr Asp Glu Leu
660 665 670

Ala Leu Arg Ala Ile Lys Met Ile Phe Lys Tyr Leu Pro Arg Ala Tyr
675 680 685

Lys Asn Gly Thr Asn Asp Ile Glu Ala Arg Glu Lys Met Ala His Ala
690 695 700

Ser Asn Ile Ala Gly Met Ala Phe Ala Asn Ala Phe Leu Gly Val Cys
705 710 715 720

His Ser Met Ala His Lys Leu Gly Ala Met His His Val Pro His Gly
725 730 735

Ile Ala Cys Ala Val Leu Ile Glu Glu Val Ile Lys Tyr Asn Ala Thr
740 745 750

Asp Cys Pro Thr Lys Gln Thr Ala Phe Pro Gln Tyr Lys Ser Pro Asn
755 760 765

Ala Lys Arg Lys Tyr Ala Glu Ile Ala Glu Tyr Leu Asn Leu Lys Gly
770 775 780

Thr Ser Asp Thr Glu Lys Val Thr Ala Leu Ile Glu Ala Ile Ser Lys
785 790 795 800

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

Asn Lys Lys Asp Phe Tyr Asn Thr Leu Asp Lys Met Ser Glu Leu Ala
820 825 830

Phe Asp Asp Gln Cys Thr Thr Ala Asn Pro Arg Tyr Pro Leu Ile Ser
835 840 845

Glu Leu Lys Asp Ile Tyr Ile Lys Ser Phe
850 855

<210> 69
<211> 1551
<212> DNA
<213> 问号钩端螺旋体

<220>
<221> CDS
<222> (1)..(1551)

<400> 69
atg aca aaa gta gaa act cga ttg gaa att tta gac gta act ttg aga 48
Met Thr Lys Val Glu Thr Arg Leu Glu Ile Leu Asp Val Thr Leu Arg
1 5 10 15

gac ggg gag cag acc aga ggg gtc agt ttt tcc act tcc gaa aaa cta 96
Asp Gly Glu Gln Thr Arg Gly Val Ser Phe Ser Thr Ser Glu Lys Leu
20 25 30

aat atc gca aaa ttt cta tta caa aaa cta aat gta gat cgg gta gag 144
Asn Ile Ala Lys Phe Leu Leu Gln Lys Leu Asn Val Asp Arg Val Glu
35 40 45

```
att gcg tct gca aga gtt tct aaa ggg gaa ttg gaa acg gtc caa aaa      192
Ile Ala Ser Ala Arg Val Ser Lys Gly Glu Leu Glu Thr Val Gln Lys
    50                  55                  60

atc atg gaa tgg gct gca aca gaa cag ctt acg gaa aga atc gaa atc      240
Ile Met Glu Trp Ala Ala Thr Glu Gln Leu Thr Glu Arg Ile Glu Ile
65                  70                  75                  80

tta ggt ttt gta gac ggg aat aaa acc gta gat tgg atc aaa gat agt      288
Leu Gly Phe Val Asp Gly Asn Lys Thr Val Asp Trp Ile Lys Asp Ser
                85                  90                  95

ggg gct aag gtt tta aat ctt ttg act aag gga tcg ctt cat cat tta      336
Gly Ala Lys Val Leu Asn Leu Leu Thr Lys Gly Ser Leu His His Leu
            100                 105                 110

gaa aaa caa tta ggc aaa act ccg aaa gaa ttc ttt aca gac gtt tct      384
Glu Lys Gln Leu Gly Lys Thr Pro Lys Glu Phe Phe Thr Asp Val Ser
        115                 120                 125

ttt gta ata gaa tac gcg atc aaa agc gga ctt aaa ata aac gta tat      432
Phe Val Ile Glu Tyr Ala Ile Lys Ser Gly Leu Lys Ile Asn Val Tyr
    130                 135                 140

tta gaa gat tgg tcc aac ggt ttc aga aac agt cca gat tac gtc aaa      480
Leu Glu Asp Trp Ser Asn Gly Phe Arg Asn Ser Pro Asp Tyr Val Lys
145                 150                 155                 160

tcg ctc gta gaa cat cta agt aaa gaa cat ata gaa aga att ttt ctt      528
Ser Leu Val Glu His Leu Ser Lys Glu His Ile Glu Arg Ile Phe Leu
                165                 170                 175

cca gac acg tta ggc gtt ctt tcg cca gaa gag acg ttt caa gga gtg      576
Pro Asp Thr Leu Gly Val Leu Ser Pro Glu Glu Thr Phe Gln Gly Val
            180                 185                 190

gat tca ctc att caa aaa tac ccg gat att cat ttt gaa ttt cac gga      624
Asp Ser Leu Ile Gln Lys Tyr Pro Asp Ile His Phe Glu Phe His Gly
        195                 200                 205

cat aac gac tac gat ctt tcc gtg gca aat agt ctt caa gcg att cgt      672
His Asn Asp Tyr Asp Leu Ser Val Ala Asn Ser Leu Gln Ala Ile Arg
    210                 215                 220

gcc gga gtc aaa ggt ctt cac gct tct ata aat ggt ctc gga gaa aga      720
Ala Gly Val Lys Gly Leu His Ala Ser Ile Asn Gly Leu Gly Glu Arg
225                 230                 235                 240

gcc gga aat act ccg ttg gaa gca ctc gta acc acg att cat gat aag      768
Ala Gly Asn Thr Pro Leu Glu Ala Leu Val Thr Thr Ile His Asp Lys
                245                 250                 255
```

```
tct aac tct aaa acg aac ata aac gaa att gca att acg gaa gca agc      816
Ser Asn Ser Lys Thr Asn Ile Asn Glu Ile Ala Ile Thr Glu Ala Ser
            260                 265                 270

cgt ctt gta gaa gta ttc agc gga aaa aga att tct gca aat aga ccg      864
Arg Leu Val Glu Val Phe Ser Gly Lys Arg Ile Ser Ala Asn Arg Pro
        275                 280                 285

atc gta gga gaa gac gtg ttt act cag acc gcg gga gta cac gca gac      912
Ile Val Gly Glu Asp Val Phe Thr Gln Thr Ala Gly Val His Ala Asp
    290                 295                 300

gga gac aaa aaa gga aat tta tac gca aat cct att tta ccg gaa aga      960
Gly Asp Lys Lys Gly Asn Leu Tyr Ala Asn Pro Ile Leu Pro Glu Arg
305                 310                 315                 320

ttt ggt agg aaa aga agt tac gcg tta ggc aaa ctt gca ggt aag gcg     1008
Phe Gly Arg Lys Arg Ser Tyr Ala Leu Gly Lys Leu Ala Gly Lys Ala
                325                 330                 335

agt atc tcc gaa aat gta aaa caa ctc gga atg gtt tta agt gaa gtg     1056
Ser Ile Ser Glu Asn Val Lys Gln Leu Gly Met Val Leu Ser Glu Val
            340                 345                 350

gtt tta caa aag gtt tta gaa agg gtg atc gaa tta gga gat cag aat     1104
Val Leu Gln Lys Val Leu Glu Arg Val Ile Glu Leu Gly Asp Gln Asn
        355                 360                 365

aaa cta gtg aca cct gaa gat ctt cca ttt atc att gcg gac gtt tct     1152
Lys Leu Val Thr Pro Glu Asp Leu Pro Phe Ile Ile Ala Asp Val Ser
    370                 375                 380

gga aga acc gga gaa aag gta ctt aca atc aaa tct tgt aat att cat     1200
Gly Arg Thr Gly Glu Lys Val Leu Thr Ile Lys Ser Cys Asn Ile His
385                 390                 395                 400

tcc gga att gga att cgt cct cac gca caa att gaa ttg gaa tat cag     1248
Ser Gly Ile Gly Ile Arg Pro His Ala Gln Ile Glu Leu Glu Tyr Gln
                405                 410                 415

gga aag att cat aag gaa att tct gaa gga gac gga ggg tat gat gcg     1296
Gly Lys Ile His Lys Glu Ile Ser Glu Gly Asp Gly Gly Tyr Asp Ala
            420                 425                 430

ttt atg aat gca ctt act aaa att acg aat cgc ctc ggt att agt att     1344
Phe Met Asn Ala Leu Thr Lys Ile Thr Asn Arg Leu Gly Ile Ser Ile
        435                 440                 445

cct aaa ttg ata gat tac gaa gta agg att cct cct ggt gga aaa aca     1392
Pro Lys Leu Ile Asp Tyr Glu Val Arg Ile Pro Pro Gly Gly Lys Thr
    450                 455                 460
```

```
gat gca ctt gta gaa act agg atc acc tgg aac aag tcc tta gat tta      1440
Asp Ala Leu Val Glu Thr Arg Ile Thr Trp Asn Lys Ser Leu Asp Leu
465                 470                 475                 480

gaa gag gac cag act ttc aaa acg atg gga gtt cat ccg gat caa acg      1488
Glu Glu Asp Gln Thr Phe Lys Thr Met Gly Val His Pro Asp Gln Thr
                485                 490                 495

gtt gca gcg gtt cat gca act gaa aag atg ctc aat caa att cta caa      1536
Val Ala Ala Val His Ala Thr Glu Lys Met Leu Asn Gln Ile Leu Gln
            500                 505                 510

cca tgg caa atc taa                                                 1551
Pro Trp Gln Ile
        515
```

<210> 70
<211> 516
<212> PRT
<213> 问号钩端螺旋体

<400> 70

```
Met Thr Lys Val Glu Thr Arg Leu Glu Ile Leu Asp Val Thr Leu Arg
1               5                   10                  15

Asp Gly Glu Gln Thr Arg Gly Val Ser Phe Ser Thr Ser Glu Lys Leu
            20                  25                  30

Asn Ile Ala Lys Phe Leu Leu Gln Lys Leu Asn Val Asp Arg Val Glu
        35                  40                  45

Ile Ala Ser Ala Arg Val Ser Lys Gly Glu Leu Glu Thr Val Gln Lys
    50                  55                  60

Ile Met Glu Trp Ala Ala Thr Glu Gln Leu Thr Glu Arg Ile Glu Ile
65                  70                  75                  80

Leu Gly Phe Val Asp Gly Asn Lys Thr Val Asp Trp Ile Lys Asp Ser
                85                  90                  95

Gly Ala Lys Val Leu Asn Leu Leu Thr Lys Gly Ser Leu His His Leu
            100                 105                 110
```

Glu Lys Gln Leu Gly Lys Thr Pro Lys Glu Phe Phe Thr Asp Val Ser
115 120 125

Phe Val Ile Glu Tyr Ala Ile Lys Ser Gly Leu Lys Ile Asn Val Tyr
130 135 140

Leu Glu Asp Trp Ser Asn Gly Phe Arg Asn Ser Pro Asp Tyr Val Lys
145 150 155 160

Ser Leu Val Glu His Leu Ser Lys Glu His Ile Glu Arg Ile Phe Leu
165 170 175

Pro Asp Thr Leu Gly Val Leu Ser Pro Glu Glu Thr Phe Gln Gly Val
180 185 190

Asp Ser Leu Ile Gln Lys Tyr Pro Asp Ile His Phe Glu Phe His Gly
195 200 205

His Asn Asp Tyr Asp Leu Ser Val Ala Asn Ser Leu Gln Ala Ile Arg
210 215 220

Ala Gly Val Lys Gly Leu His Ala Ser Ile Asn Gly Leu Gly Glu Arg
225 230 235 240

Ala Gly Asn Thr Pro Leu Glu Ala Leu Val Thr Thr Ile His Asp Lys
245 250 255

Ser Asn Ser Lys Thr Asn Ile Asn Glu Ile Ala Ile Thr Glu Ala Ser
260 265 270

Arg Leu Val Glu Val Phe Ser Gly Lys Arg Ile Ser Ala Asn Arg Pro
275 280 285

Ile Val Gly Glu Asp Val Phe Thr Gln Thr Ala Gly Val His Ala Asp
290 295 300

Gly Asp Lys Lys Gly Asn Leu Tyr Ala Asn Pro Ile Leu Pro Glu Arg
305 310 315 320

Phe Gly Arg Lys Arg Ser Tyr Ala Leu Gly Lys Leu Ala Gly Lys Ala
325 330 335

Ser Ile Ser Glu Asn Val Lys Gln Leu Gly Met Val Leu Ser Glu Val
340 345 350

Val Leu Gln Lys Val Leu Glu Arg Val Ile Glu Leu Gly Asp Gln Asn
355 360 365

Lys Leu Val Thr Pro Glu Asp Leu Pro Phe Ile Ile Ala Asp Val Ser
370 375 380

Gly Arg Thr Gly Glu Lys Val Leu Thr Ile Lys Ser Cys Asn Ile His
385 390 395 400

Ser Gly Ile Gly Ile Arg Pro His Ala Gln Ile Glu Leu Glu Tyr Gln
405 410 415

Gly Lys Ile His Lys Glu Ile Ser Glu Gly Asp Gly Gly Tyr Asp Ala
420 425 430

Phe Met Asn Ala Leu Thr Lys Ile Thr Asn Arg Leu Gly Ile Ser Ile
435 440 445

Pro Lys Leu Ile Asp Tyr Glu Val Arg Ile Pro Pro Gly Gly Lys Thr
450 455 460

Asp Ala Leu Val Glu Thr Arg Ile Thr Trp Asn Lys Ser Leu Asp Leu
465 470 475 480

Glu Glu Asp Gln Thr Phe Lys Thr Met Gly Val His Pro Asp Gln Thr
485 490 495

Val Ala Ala Val His Ala Thr Glu Lys Met Leu Asn Gln Ile Leu Gln
500 505 510

Pro Trp Gln Ile
515

<210> 71
<211> 1398
<212> DNA
<213> 问号钩端螺旋体

<220>
<221> CDS
<222> (1)..(1398)

<400> 71

```
atg aag aca atg ttc gaa aaa att tgg gaa gat cat cta gtc gga gaa      48
Met Lys Thr Met Phe Glu Lys Ile Trp Glu Asp His Leu Val Gly Glu
1               5                  10                  15

cta gat gct gga tcc tat cta atc tat ata gat cgc cat ctc att cat      96
Leu Asp Ala Gly Ser Tyr Leu Ile Tyr Ile Asp Arg His Leu Ile His
            20                  25                  30

gaa gtt aca agt cct cag gcg ttt gaa gga ctt aaa ctt gca ggc aga     144
Glu Val Thr Ser Pro Gln Ala Phe Glu Gly Leu Lys Leu Ala Gly Arg
        35                  40                  45

aag gtt cgt cgt cct gaa gct act ttt gcc aca atg gat cat aac gtt     192
Lys Val Arg Arg Pro Glu Ala Thr Phe Ala Thr Met Asp His Asn Val
    50                  55                  60

tct act aga aca cgt gat tta agt ctg gcc gat cct gtt tcc gca att     240
Ser Thr Arg Thr Arg Asp Leu Ser Leu Ala Asp Pro Val Ser Ala Ile
65                  70                  75                  80

caa atg cag act tta aaa aag aac tgc gac gaa aac gga atc cgc gtt     288
Gln Met Gln Thr Leu Lys Lys Asn Cys Asp Glu Asn Gly Ile Arg Val
                85                  90                  95

tat gat ttt caa aac cct gac caa gga atc att cac gta atc gct cct     336
Tyr Asp Phe Gln Asn Pro Asp Gln Gly Ile Ile His Val Ile Ala Pro
            100                 105                 110

gaa atg gga ctg act cat cct gga atg aca atc gta tgc gga gat tct     384
Glu Met Gly Leu Thr His Pro Gly Met Thr Ile Val Cys Gly Asp Ser
        115                 120                 125

cat act tct aca cac ggt gcg ttt ggt gcg ctt gct ttc ggg atc gga     432
His Thr Ser Thr His Gly Ala Phe Gly Ala Leu Ala Phe Gly Ile Gly
    130                 135                 140

acc agc gaa gta gag cac gtt ctt gcg act caa acc tta gtt caa aaa     480
Thr Ser Glu Val Glu His Val Leu Ala Thr Gln Thr Leu Val Gln Lys
145                 150                 155                 160
```

```
aga gca aaa aca atg gag att aga gtc gat gga aaa ctt tcc gat aag      528
Arg Ala Lys Thr Met Glu Ile Arg Val Asp Gly Lys Leu Ser Asp Lys
                165                 170                 175

gtc aca gca aaa gac atc att ctt gcg atc att gga aaa att gga acc      576
Val Thr Ala Lys Asp Ile Ile Leu Ala Ile Ile Gly Lys Ile Gly Thr
            180                 185                 190

gca ggt gcg aca ggt tat gtg atc gaa tat aga ggt tct gca att caa      624
Ala Gly Ala Thr Gly Tyr Val Ile Glu Tyr Arg Gly Ser Ala Ile Gln
        195                 200                 205

gcc ctc agt atg gaa gct aga atg act att tgt aat atg tct atc gaa      672
Ala Leu Ser Met Glu Ala Arg Met Thr Ile Cys Asn Met Ser Ile Glu
    210                 215                 220

gcg gga gct aga gca ggt tta atc gca cca gat gaa act act ttt aat      720
Ala Gly Ala Arg Ala Gly Leu Ile Ala Pro Asp Glu Thr Thr Phe Asn
225                 230                 235                 240

tat att caa gga aag gac ttt tct cca aaa gga gtc gaa tgg gat ctt      768
Tyr Ile Gln Gly Lys Asp Phe Ser Pro Lys Gly Val Glu Trp Asp Leu
                245                 250                 255

gcg gtc aaa aaa tgg aaa cac tat gta acg gac gaa ggt gct aaa ttt      816
Ala Val Lys Lys Trp Lys His Tyr Val Thr Asp Glu Gly Ala Lys Phe
            260                 265                 270

gat aga acc gta att ctt cat gca gat gaa atc gct cct atg gta act      864
Asp Arg Thr Val Ile Leu His Ala Asp Glu Ile Ala Pro Met Val Thr
        275                 280                 285

tgg gga act tct ccc agt cag gtt gtt tcg ata aaa gga gtc gtt cca      912
Trp Gly Thr Ser Pro Ser Gln Val Val Ser Ile Lys Gly Val Val Pro
    290                 295                 300

gat cca aaa gat gca aat gat ccg gtg gaa aaa att gga att gag tct      960
Asp Pro Lys Asp Ala Asn Asp Pro Val Glu Lys Ile Gly Ile Glu Ser
305                 310                 315                 320

gcg ctt aaa tat atg gat ctc aaa tcg ggc cag aag ata gaa gac att     1008
Ala Leu Lys Tyr Met Asp Leu Lys Ser Gly Gln Lys Ile Glu Asp Ile
                325                 330                 335

tca att aat aaa gtg ttt atc ggt tcc tgt act aat tct aga atc gaa     1056
Ser Ile Asn Lys Val Phe Ile Gly Ser Cys Thr Asn Ser Arg Ile Glu
            340                 345                 350

gat tta aga gcg gcc gct gct acc gta aaa gga aaa aaa gtt tcc tct     1104
Asp Leu Arg Ala Ala Ala Ala Thr Val Lys Gly Lys Lys Val Ser Ser
        355                 360                 365
```

```
aag gtt cag gcg att gtg gtt ccc ggt tca ggc aga gtc aaa cgt cag      1152
Lys Val Gln Ala Ile Val Val Pro Gly Ser Gly Arg Val Lys Arg Gln
    370                 375                 380

gcg gaa caa gaa ggt ctg gat aaa att ttt acc gcg gcc ggt ttt gaa      1200
Ala Glu Gln Glu Gly Leu Asp Lys Ile Phe Thr Ala Ala Gly Phe Glu
385                 390                 395                 400

tgg aga aat cca ggc tgt tct atg tgt ctt gcg atg aac gac gac gta      1248
Trp Arg Asn Pro Gly Cys Ser Met Cys Leu Ala Met Asn Asp Asp Val
                405                 410                 415

tta gaa ccg gga gat cgt tgt gct tct act tct aac cga aac ttt gaa      1296
Leu Glu Pro Gly Asp Arg Cys Ala Ser Thr Ser Asn Arg Asn Phe Glu
            420                 425                 430

ggt cgt caa gga aaa ggt gga aga acc cat cta gta gga ccg gaa atg      1344
Gly Arg Gln Gly Lys Gly Gly Arg Thr His Leu Val Gly Pro Glu Met
        435                 440                 445

gcc gcc gcc gcg gct atc gaa ggc cat ttt gtg gat att cga aac tgg      1392
Ala Ala Ala Ala Ala Ile Glu Gly His Phe Val Asp Ile Arg Asn Trp
    450                 455                 460

aaa taa                                                            1398
Lys
465
```

<210> 72
<211> 465
<212> PRT
<213> 问号钩端螺旋体

<400> 72

```
Met Lys Thr Met Phe Glu Lys Ile Trp Glu Asp His Leu Val Gly Glu
1               5                   10                  15

Leu Asp Ala Gly Ser Tyr Leu Ile Tyr Ile Asp Arg His Leu Ile His
            20                  25                  30

Glu Val Thr Ser Pro Gln Ala Phe Glu Gly Leu Lys Leu Ala Gly Arg
        35                  40                  45

Lys Val Arg Arg Pro Glu Ala Thr Phe Ala Thr Met Asp His Asn Val
    50                  55                  60
```

Ser Thr Arg Thr Arg Asp Leu Ser Leu Ala Asp Pro Val Ser Ala Ile
65 70 75 80

Gln Met Gln Thr Leu Lys Lys Asn Cys Asp Glu Asn Gly Ile Arg Val
85 90 95

Tyr Asp Phe Gln Asn Pro Asp Gln Gly Ile Ile His Val Ile Ala Pro
100 105 110

Glu Met Gly Leu Thr His Pro Gly Met Thr Ile Val Cys Gly Asp Ser
115 120 125

His Thr Ser Thr His Gly Ala Phe Gly Ala Leu Ala Phe Gly Ile Gly
130 135 140

Thr Ser Glu Val Glu His Val Leu Ala Thr Gln Thr Leu Val Gln Lys
145 150 155 160

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

Val Thr Ala Lys Asp Ile Ile Leu Ala Ile Ile Gly Lys Ile Gly Thr
180 185 190

Ala Gly Ala Thr Gly Tyr Val Ile Glu Tyr Arg Gly Ser Ala Ile Gln
195 200 205

Ala Leu Ser Met Glu Ala Arg Met Thr Ile Cys Asn Met Ser Ile Glu
210 215 220

Ala Gly Ala Arg Ala Gly Leu Ile Ala Pro Asp Glu Thr Thr Phe Asn
225 230 235 240

Tyr Ile Gln Gly Lys Asp Phe Ser Pro Lys Gly Val Glu Trp Asp Leu
245 250 255

Ala Val Lys Lys Trp Lys His Tyr Val Thr Asp Glu Gly Ala Lys Phe
260 265 270

Asp Arg Thr Val Ile Leu His Ala Asp Glu Ile Ala Pro Met Val Thr
275 280 285

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

Asp Pro Lys Asp Ala Asn Asp Pro Val Glu Lys Ile Gly Ile Glu Ser
305 310 315 320

Ala Leu Lys Tyr Met Asp Leu Lys Ser Gly Gln Lys Ile Glu Asp Ile
325 330 335

Ser Ile Asn Lys Val Phe Ile Gly Ser Cys Thr Asn Ser Arg Ile Glu
340 345 350

Asp Leu Arg Ala Ala Ala Ala Ala Thr Val Lys Gly Lys Lys Val Ser Ser
355 360 365

Lys Val Gln Ala Ile Val Val Pro Gly Ser Gly Arg Val Lys Arg Gln
370 375 380

Ala Glu Gln Glu Gly Leu Asp Lys Ile Phe Thr Ala Ala Gly Phe Glu
385 390 395 400

Trp Arg Asn Pro Gly Cys Ser Met Cys Leu Ala Met Asn Asp Asp Val
405 410 415

Leu Glu Pro Gly Asp Arg Cys Ala Ser Thr Ser Asn Arg Asn Phe Glu
420 425 430

Gly Arg Gln Gly Lys Gly Gly Arg Thr His Leu Val Gly Pro Glu Met
435 440 445

Ala Ala Ala Ala Ala Ile Glu Gly His Phe Val Asp Ile Arg Asn Trp
450 455 460

Lys
465

<210> 73
<211> 621
<212> DNA
<213> 问号钩端螺旋体

<220>
<221> CDS
<222> (1)..(621)

<400> 73

```
atg aaa ccc ttt act ata tta aat gga att gcc gcc tta ctg gac aga      48
Met Lys Pro Phe Thr Ile Leu Asn Gly Ile Ala Ala Leu Leu Asp Arg
1               5                  10                  15

ccc aac gtg gat acg gat cag atc att cca aaa caa ttt tta cgg aag      96
Pro Asn Val Asp Thr Asp Gln Ile Ile Pro Lys Gln Phe Leu Arg Lys
            20                  25                  30

ata gaa cga acc ggt ttc gga gtt cat ctg ttt cac gat tgg aga tac     144
Ile Glu Arg Thr Gly Phe Gly Val His Leu Phe His Asp Trp Arg Tyr
        35                  40                  45

tta gac gac gcg ggt acc aaa ctc aat cct gat ttt tcc ctc aat caa     192
Leu Asp Asp Ala Gly Thr Lys Leu Asn Pro Asp Phe Ser Leu Asn Gln
    50                  55                  60

gaa cga tat aag gga gct tct atc ctt atc acc aga gat aac ttt ggt     240
Glu Arg Tyr Lys Gly Ala Ser Ile Leu Ile Thr Arg Asp Asn Phe Gly
65                  70                  75                  80

tgt gga tct tcc aga gaa cac gct cct tgg gct tta gaa gac tac ggg     288
Cys Gly Ser Ser Arg Glu His Ala Pro Trp Ala Leu Glu Asp Tyr Gly
                85                  90                  95

ttt agg gca atc att gct cct tct tac gcg gat att ttt ttc aac aac     336
Phe Arg Ala Ile Ile Ala Pro Ser Tyr Ala Asp Ile Phe Phe Asn Asn
            100                 105                 110

tgc ttt aaa aac gga atg ctt cca gtc att tta aaa tcg gaa gaa gta     384
Cys Phe Lys Asn Gly Met Leu Pro Val Ile Leu Lys Ser Glu Glu Val
        115                 120                 125

gaa gag ctg ttc cat ttg gtt tcg act aac gta gga gcg aaa gtc ata     432
Glu Glu Leu Phe His Leu Val Ser Thr Asn Val Gly Ala Lys Val Ile
    130                 135                 140

gtg gat ctg gac aaa caa act gta acc gga ccg act gga aaa ata tat     480
Val Asp Leu Asp Lys Gln Thr Val Thr Gly Pro Thr Gly Lys Ile Tyr
145                 150                 155                 160
```

```
tat ttt gaa gtg gat tct ttt cgt aaa tac tgt ctt tat aac gga ctt      528
Tyr Phe Glu Val Asp Ser Phe Arg Lys Tyr Cys Leu Tyr Asn Gly Leu
                165                 170                 175

gat gac ata ggt cta act cta aaa caa gaa agt aaa att gga gag ttt      576
Asp Asp Ile Gly Leu Thr Leu Lys Gln Glu Ser Lys Ile Gly Glu Phe
            180                 185                 190

gaa aaa aag cag aaa gaa gtt gaa cct tgg tta tac gcc ata taa          621
Glu Lys Lys Gln Lys Glu Val Glu Pro Trp Leu Tyr Ala Ile
        195                 200                 205
```

<210> 74
<211> 206
<212> PRT
<213> 问号钩端螺旋体

<400> 74

```
Met Lys Pro Phe Thr Ile Leu Asn Gly Ile Ala Ala Leu Leu Asp Arg
1               5                   10                  15

Pro Asn Val Asp Thr Asp Gln Ile Ile Pro Lys Gln Phe Leu Arg Lys
            20                  25                  30

Ile Glu Arg Thr Gly Phe Gly Val His Leu Phe His Asp Trp Arg Tyr
        35                  40                  45

Leu Asp Asp Ala Gly Thr Lys Leu Asn Pro Asp Phe Ser Leu Asn Gln
    50                  55                  60

Glu Arg Tyr Lys Gly Ala Ser Ile Leu Ile Thr Arg Asp Asn Phe Gly
65                  70                  75                  80

Cys Gly Ser Ser Arg Glu His Ala Pro Trp Ala Leu Glu Asp Tyr Gly
                85                  90                  95

Phe Arg Ala Ile Ile Ala Pro Ser Tyr Ala Asp Ile Phe Phe Asn Asn
            100                 105                 110

Cys Phe Lys Asn Gly Met Leu Pro Val Ile Leu Lys Ser Glu Glu Val
        115                 120                 125
```

Glu Glu Leu Phe His Leu Val Ser Thr Asn Val Gly Ala Lys Val Ile
130 135 140

Val Asp Leu Asp Lys Gln Thr Val Thr Gly Pro Thr Gly Lys Ile Tyr
145 150 155 160

Tyr Phe Glu Val Asp Ser Phe Arg Lys Tyr Cys Leu Tyr Asn Gly Leu
165 170 175

Asp Asp Ile Gly Leu Thr Leu Lys Gln Glu Ser Lys Ile Gly Glu Phe
180 185 190

Glu Lys Lys Gln Lys Glu Val Glu Pro Trp Leu Tyr Ala Ile
195 200 205

<210> 75
<211> 1077
<212> DNA
<213> 问号钩端螺旋体

<220>
<221> CDS
<222> (1)..(1077)

<400> 75
atg aag aat gta gca gta ctt tca gga gac gga atc gga ccg gaa gtc 48
Met Lys Asn Val Ala Val Leu Ser Gly Asp Gly Ile Gly Pro Glu Val
1 5 10 15

atg gag ata gcc atc tcc gtt ttg aaa aag gct ctc ggt gca aaa gtt 96
Met Glu Ile Ala Ile Ser Val Leu Lys Lys Ala Leu Gly Ala Lys Val
20 25 30

tcc gag ttt caa ttt aaa gaa gga ttt gta ggt gga atc gca atc gat 144
Ser Glu Phe Gln Phe Lys Glu Gly Phe Val Gly Gly Ile Ala Ile Asp
35 40 45

aaa act gga cac cca ctt cca ccg gaa act ctt aaa cta tgt gaa gaa 192
Lys Thr Gly His Pro Leu Pro Pro Glu Thr Leu Lys Leu Cys Glu Glu
50 55 60

tct tcc gca att ctt ttc gga agt gtg gga ggt cct aaa tgg gaa aca 240
Ser Ser Ala Ile Leu Phe Gly Ser Val Gly Gly Pro Lys Trp Glu Thr
65 70 75 80

```
ctc cct ccg gaa aaa caa ccg gaa cga ggg gca ctt cta cct ttg aga      288
Leu Pro Pro Glu Lys Gln Pro Glu Arg Gly Ala Leu Leu Pro Leu Arg
                85                  90                  95

aaa cat ttt gat cta ttt gca aac tta aga cct gcg atc att tat cca      336
Lys His Phe Asp Leu Phe Ala Asn Leu Arg Pro Ala Ile Ile Tyr Pro
            100                 105                 110

gag ttg aaa aat gct tct cca gtt cgt tct gat att att gga aac gga      384
Glu Leu Lys Asn Ala Ser Pro Val Arg Ser Asp Ile Ile Gly Asn Gly
        115                 120                 125

tta gat att ctc ata tta aga gag tta acc gga gga att tat ttt gga      432
Leu Asp Ile Leu Ile Leu Arg Glu Leu Thr Gly Gly Ile Tyr Phe Gly
    130                 135                 140

caa cca aaa gga aga gaa gga tca ggt cag gaa gaa ttt gca tac gac      480
Gln Pro Lys Gly Arg Glu Gly Ser Gly Gln Glu Glu Phe Ala Tyr Asp
145                 150                 155                 160

acg atg aag tat tcc aga aga gaa atc gaa agg att gct aaa gtc gca      528
Thr Met Lys Tyr Ser Arg Arg Glu Ile Glu Arg Ile Ala Lys Val Ala
                165                 170                 175

ttc cag gcg gcc aga aaa aga aat aat aaa gtg act agt atc gat aaa      576
Phe Gln Ala Ala Arg Lys Arg Asn Asn Lys Val Thr Ser Ile Asp Lys
            180                 185                 190

gca aac gtc ttg act act tcc gtt ttt tgg aag gaa gta gta atc gaa      624
Ala Asn Val Leu Thr Thr Ser Val Phe Trp Lys Glu Val Val Ile Glu
        195                 200                 205

ttg cat aag aaa gaa ttt tca gac gtc caa ttg aat cat ctt tat gtg      672
Leu His Lys Lys Glu Phe Ser Asp Val Gln Leu Asn His Leu Tyr Val
    210                 215                 220

gac aat gcg gcg atg cag tta atc gta aat ccg aaa caa ttc gac gtg      720
Asp Asn Ala Ala Met Gln Leu Ile Val Asn Pro Lys Gln Phe Asp Val
225                 230                 235                 240

gtt ctt tgt gag aat atg ttt ggt gat att ctt tcg gac gag gct tcc      768
Val Leu Cys Glu Asn Met Phe Gly Asp Ile Leu Ser Asp Glu Ala Ser
                245                 250                 255

atc att acg ggt tca atc gga atg ctt cct tct gcc tct ctt tcc gaa      816
Ile Ile Thr Gly Ser Ile Gly Met Leu Pro Ser Ala Ser Leu Ser Glu
            260                 265                 270

tct gga ttt gga ttg tat gaa cct tct ggt ggt tct gcg ccg gac ata      864
Ser Gly Phe Gly Leu Tyr Glu Pro Ser Gly Gly Ser Ala Pro Asp Ile
        275                 280                 285
```

gcc gga aaa gga gtg gca aat ccg att gct caa gta ttg agt gcg gcg 912
Ala Gly Lys Gly Val Ala Asn Pro Ile Ala Gln Val Leu Ser Ala Ala
290 295 300

ttg atg tta cgt tat tct ttt tct atg gaa gaa gaa gca aac aag ata 960
Leu Met Leu Arg Tyr Ser Phe Ser Met Glu Glu Glu Ala Asn Lys Ile
305 310 315 320

gaa acc gcc gtg cgt aaa acg att gcc tcc gga aaa aga acc aga gac 1008
Glu Thr Ala Val Arg Lys Thr Ile Ala Ser Gly Lys Arg Thr Arg Asp
325 330 335

ata gcg gaa gta gga tct acg atc gta gga act aaa gaa atc ggt caa 1056
Ile Ala Glu Val Gly Ser Thr Ile Val Gly Thr Lys Glu Ile Gly Gln
340 345 350

ttg atc gaa tcc ttt ctc taa 1077
Leu Ile Glu Ser Phe Leu
355

<210> 76
<211> 358
<212> PRT
<213> 问号钩端螺旋体

<400> 76

Met Lys Asn Val Ala Val Leu Ser Gly Asp Gly Ile Gly Pro Glu Val
1 5 10 15

Met Glu Ile Ala Ile Ser Val Leu Lys Lys Ala Leu Gly Ala Lys Val
20 25 30

Ser Glu Phe Gln Phe Lys Glu Gly Phe Val Gly Gly Ile Ala Ile Asp
35 40 45

Lys Thr Gly His Pro Leu Pro Pro Glu Thr Leu Lys Leu Cys Glu Glu
50 55 60

Ser Ser Ala Ile Leu Phe Gly Ser Val Gly Gly Pro Lys Trp Glu Thr
65 70 75 80

Leu Pro Pro Glu Lys Gln Pro Glu Arg Gly Ala Leu Leu Pro Leu Arg
85 90 95

Lys His Phe Asp Leu Phe Ala Asn Leu Arg Pro Ala Ile Ile Tyr Pro
100 105 110

Glu Leu Lys Asn Ala Ser Pro Val Arg Ser Asp Ile Ile Gly Asn Gly
115 120 125

Leu Asp Ile Leu Ile Leu Arg Glu Leu Thr Gly Gly Ile Tyr Phe Gly
130 135 140

Gln Pro Lys Gly Arg Glu Gly Ser Gly Gln Glu Glu Phe Ala Tyr Asp
145 150 155 160

Thr Met Lys Tyr Ser Arg Arg Glu Ile Glu Arg Ile Ala Lys Val Ala
165 170 175

Phe Gln Ala Ala Arg Lys Arg Asn Asn Lys Val Thr Ser Ile Asp Lys
180 185 190

Ala Asn Val Leu Thr Thr Ser Val Phe Trp Lys Glu Val Val Ile Glu
195 200 205

Leu His Lys Lys Glu Phe Ser Asp Val Gln Leu Asn His Leu Tyr Val
210 215 220

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

Val Leu Cys Glu Asn Met Phe Gly Asp Ile Leu Ser Asp Glu Ala Ser
245 250 255

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

Ser Gly Phe Gly Leu Tyr Glu Pro Ser Gly Gly Ser Ala Pro Asp Ile
275 280 285

Ala Gly Lys Gly Val Ala Asn Pro Ile Ala Gln Val Leu Ser Ala Ala
290 295 300

Leu Met Leu Arg Tyr Ser Phe Ser Met Glu Glu Glu Ala Asn Lys Ile
305 310 315 320

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

Ile Ala Glu Val Gly Ser Thr Ile Val Gly Thr Lys Glu Ile Gly Gln
340 345 350

Leu Ile Glu Ser Phe Leu
355

<210> 77
<211> 1161
<212> DNA
<213> 大肠杆菌

<220>
<221> CDS
<222> (1)..(1161)

<400> 77
atg aca tcg gaa aac ccg tta ctg gcg ctg cga gag aaa atc agc gcg 48
Met Thr Ser Glu Asn Pro Leu Leu Ala Leu Arg Glu Lys Ile Ser Ala
1 5 10 15

ctg gat gaa aaa tta tta gcg tta ctg gca gaa cgg cgc gaa ctg gcc 96
Leu Asp Glu Lys Leu Leu Ala Leu Leu Ala Glu Arg Arg Glu Leu Ala
20 25 30

gtc gag gtg gga aaa gcc aaa ctg ctc tcg cat cgc ccg gta cgt gat 144
Val Glu Val Gly Lys Ala Lys Leu Leu Ser His Arg Pro Val Arg Asp
35 40 45

att gat cgt gaa cgc gat ttg ctg gaa aga tta att acg ctc ggt aaa 192
Ile Asp Arg Glu Arg Asp Leu Leu Glu Arg Leu Ile Thr Leu Gly Lys
50 55 60

gcg cac cat ctg gac gcc cat tac att act cgc ctg ttc cag ctc atc 240
Ala His His Leu Asp Ala His Tyr Ile Thr Arg Leu Phe Gln Leu Ile
65 70 75 80

att gaa gat tcc gta tta act cag cag gct ttg ctc caa caa cat ctc 288
Ile Glu Asp Ser Val Leu Thr Gln Gln Ala Leu Leu Gln Gln His Leu
85 90 95

```
aat aaa att aat ccg cac tca gca cgc atc gct ttt ctc ggc ccc aaa      336
Asn Lys Ile Asn Pro His Ser Ala Arg Ile Ala Phe Leu Gly Pro Lys
            100                 105                 110

ggt tct tat tcc cat ctt gcg gcg cgc cag tat gct gcc cgt cac ttt      384
Gly Ser Tyr Ser His Leu Ala Ala Arg Gln Tyr Ala Ala Arg His Phe
        115                 120                 125

gag caa ttc att gaa agt ggc tgc gcc aaa ttt gcc gat att ttt aat      432
Glu Gln Phe Ile Glu Ser Gly Cys Ala Lys Phe Ala Asp Ile Phe Asn
    130                 135                 140

cag gtg gaa acc ggc cag gcc gac tat gcc gtc gta ccg att gaa aat      480
Gln Val Glu Thr Gly Gln Ala Asp Tyr Ala Val Val Pro Ile Glu Asn
145                 150                 155                 160

acc agc tcc ggt gcc ata aac gac gtt tac gat ctg ctg caa cat acc      528
Thr Ser Ser Gly Ala Ile Asn Asp Val Tyr Asp Leu Leu Gln His Thr
                165                 170                 175

agc ttg tcg att gtt ggc gag atg acg tta act atc gac cat tgt ttg      576
Ser Leu Ser Ile Val Gly Glu Met Thr Leu Thr Ile Asp His Cys Leu
            180                 185                 190

ttg gtc tcc ggc act act gat tta tcc acc atc aat acg gtc tac agc      624
Leu Val Ser Gly Thr Thr Asp Leu Ser Thr Ile Asn Thr Val Tyr Ser
        195                 200                 205

cat ccg cag cca ttc cag caa tgc agc aaa ttc ctt aat cgt tat ccg      672
His Pro Gln Pro Phe Gln Gln Cys Ser Lys Phe Leu Asn Arg Tyr Pro
    210                 215                 220

cac tgg aag att gaa tat acc gaa agt acg tct gcg gca atg gaa aag      720
His Trp Lys Ile Glu Tyr Thr Glu Ser Thr Ser Ala Ala Met Glu Lys
225                 230                 235                 240

gtt gca cag gca aaa tca ccg cat gtt gct gcg ttg gga agc gaa gct      768
Val Ala Gln Ala Lys Ser Pro His Val Ala Ala Leu Gly Ser Glu Ala
                245                 250                 255

ggc ggc act ttg tac ggt ttg cag gta ctg gag cgt att gaa gca aat      816
Gly Gly Thr Leu Tyr Gly Leu Gln Val Leu Glu Arg Ile Glu Ala Asn
            260                 265                 270

cag cga caa aac ttc acc cga ttt gtg gtg ttg gcg cgt aaa gcc att      864
Gln Arg Gln Asn Phe Thr Arg Phe Val Val Leu Ala Arg Lys Ala Ile
        275                 280                 285

aac gtg tct gat cag gtt ccg gcg aaa acc acg ttg tta atg gcg acc      912
Asn Val Ser Asp Gln Val Pro Ala Lys Thr Thr Leu Leu Met Ala Thr
    290                 295                 300
```

```
ggg caa caa gcc ggt gcg ctg gtt gaa gcg ttg ctg gta ctg cgc aac      960
Gly Gln Gln Ala Gly Ala Leu Val Glu Ala Leu Leu Val Leu Arg Asn
305                 310                 315                 320

cac aat ctg att atg acc cgt ctg gaa tca cgc ccg att cac ggt aat     1008
His Asn Leu Ile Met Thr Arg Leu Glu Ser Arg Pro Ile His Gly Asn
                325                 330                 335

cca tgg gaa gag atg ttc tat ctg gat att cag gcc aat ctt gaa tca     1056
Pro Trp Glu Glu Met Phe Tyr Leu Asp Ile Gln Ala Asn Leu Glu Ser
            340                 345                 350

gcg gaa atg caa aaa gca ttg aaa gag tta ggg gaa atc acc cgt tca     1104
Ala Glu Met Gln Lys Ala Leu Lys Glu Leu Gly Glu Ile Thr Arg Ser
        355                 360                 365

atg aag gta ttg ggc tgt tac cca agt gag aac gta gtg cct gtt gat     1152
Met Lys Val Leu Gly Cys Tyr Pro Ser Glu Asn Val Val Pro Val Asp
    370                 375                 380

cca acc tga                                                        1161
Pro Thr
385
```

<210> 78
<211> 386
<212> PRT
<213> 大肠杆菌

<400> 78

```
Met Thr Ser Glu Asn Pro Leu Leu Ala Leu Arg Glu Lys Ile Ser Ala
1               5                   10                  15

Leu Asp Glu Lys Leu Leu Ala Leu Leu Ala Glu Arg Arg Glu Leu Ala
            20                  25                  30

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

Ile Asp Arg Glu Arg Asp Leu Leu Glu Arg Leu Ile Thr Leu Gly Lys
    50                  55                  60

Ala His His Leu Asp Ala His Tyr Ile Thr Arg Leu Phe Gln Leu Ile
65                  70                  75                  80
```

Ile Glu Asp Ser Val Leu Thr Gln Gln Ala Leu Leu Gln Gln His Leu
85 90 95

Asn Lys Ile Asn Pro His Ser Ala Arg Ile Ala Phe Leu Gly Pro Lys
100 105 110

Gly Ser Tyr Ser His Leu Ala Ala Arg Gln Tyr Ala Ala Arg His Phe
115 120 125

Glu Gln Phe Ile Glu Ser Gly Cys Ala Lys Phe Ala Asp Ile Phe Asn
130 135 140

Gln Val Glu Thr Gly Gln Ala Asp Tyr Ala Val Val Pro Ile Glu Asn
145 150 155 160

Thr Ser Ser Gly Ala Ile Asn Asp Val Tyr Asp Leu Leu Gln His Thr
165 170 175

Ser Leu Ser Ile Val Gly Glu Met Thr Leu Thr Ile Asp His Cys Leu
180 185 190

Leu Val Ser Gly Thr Thr Asp Leu Ser Thr Ile Asn Thr Val Tyr Ser
195 200 205

His Pro Gln Pro Phe Gln Gln Cys Ser Lys Phe Leu Asn Arg Tyr Pro
210 215 220

His Trp Lys Ile Glu Tyr Thr Glu Ser Thr Ser Ala Ala Met Glu Lys
225 230 235 240

Val Ala Gln Ala Lys Ser Pro His Val Ala Ala Leu Gly Ser Glu Ala
245 250 255

Gly Gly Thr Leu Tyr Gly Leu Gln Val Leu Glu Arg Ile Glu Ala Asn
260 265 270

Gln Arg Gln Asn Phe Thr Arg Phe Val Val Leu Ala Arg Lys Ala Ile
275 280 285

Asn Val Ser Asp Gln Val Pro Ala Lys Thr Thr Leu Leu Met Ala Thr
290 295 300

Gly Gln Gln Ala Gly Ala Leu Val Glu Ala Leu Leu Val Leu Arg Asn
305 310 315 320

His Asn Leu Ile Met Thr Arg Leu Glu Ser Arg Pro Ile His Gly Asn
325 330 335

Pro Trp Glu Glu Met Phe Tyr Leu Asp Ile Gln Ala Asn Leu Glu Ser
340 345 350

Ala Glu Met Gln Lys Ala Leu Lys Glu Leu Gly Glu Ile Thr Arg Ser
355 360 365

Met Lys Val Leu Gly Cys Tyr Pro Ser Glu Asn Val Val Pro Val Asp
370 375 380

Pro Thr
385

<210> 79
<211> 1122
<212> DNA
<213> 大肠杆菌

<220>
<221> CDS
<222> (1)..(1122)

<400> 79
atg gtt gct gaa ttg acc gca tta cgc gat caa att gat gaa gtc gat 48
Met Val Ala Glu Leu Thr Ala Leu Arg Asp Gln Ile Asp Glu Val Asp
1 5 10 15

aaa gcg ctg ctg aat tta tta gcg aag cgt ctg gaa ctg gtt gct gaa 96
Lys Ala Leu Leu Asn Leu Leu Ala Lys Arg Leu Glu Leu Val Ala Glu
20 25 30

gtg ggc gag gtg aaa agc cgc ttt gga ctg cct att tat gtt ccg gag 144
Val Gly Glu Val Lys Ser Arg Phe Gly Leu Pro Ile Tyr Val Pro Glu
35 40 45

```
cgc gag gca tct atg ttg gcc tcg cgt cgt gca gag gcg gaa gct ctg      192
Arg Glu Ala Ser Met Leu Ala Ser Arg Arg Ala Glu Ala Glu Ala Leu
    50                  55                  60

ggt gta ccg cca gat ctg att gag gat gtt ttg cgt cgg gtg atg cgt      240
Gly Val Pro Pro Asp Leu Ile Glu Asp Val Leu Arg Arg Val Met Arg
65                  70                  75                  80

gaa tct tac tcc agt gaa aac gac aaa gga ttt aaa aca ctt tgt ccg      288
Glu Ser Tyr Ser Ser Glu Asn Asp Lys Gly Phe Lys Thr Leu Cys Pro
                85                  90                  95

tca ctg cgt ccg gtg gtt atc gtc ggc ggt ggc ggt cag atg gga cgc      336
Ser Leu Arg Pro Val Val Ile Val Gly Gly Gly Gly Gln Met Gly Arg
            100                 105                 110

ctg ttc gag aag atg ctg acc ctc tcg ggt tat cag gtg cgg att ctg      384
Leu Phe Glu Lys Met Leu Thr Leu Ser Gly Tyr Gln Val Arg Ile Leu
        115                 120                 125

gag caa cat gac tgg gat cga gcg gct gat att gtt gcc gat gcc gga      432
Glu Gln His Asp Trp Asp Arg Ala Ala Asp Ile Val Ala Asp Ala Gly
    130                 135                 140

atg gtg att gtt agt gtg cca atc cac gtt act gag caa gtt att ggc      480
Met Val Ile Val Ser Val Pro Ile His Val Thr Glu Gln Val Ile Gly
145                 150                 155                 160

aaa tta ccg cct tta ccg aaa gat tgt att ctg gtc gat ctg gca tca      528
Lys Leu Pro Pro Leu Pro Lys Asp Cys Ile Leu Val Asp Leu Ala Ser
                165                 170                 175

gtg aaa aat ggg cca tta cag gcc atg ctg gtg gcg cat gat ggt ccg      576
Val Lys Asn Gly Pro Leu Gln Ala Met Leu Val Ala His Asp Gly Pro
            180                 185                 190

gtg ctg ggg cta cac ccg atg ttc ggt ccg gac agc ggt agc ctg gca      624
Val Leu Gly Leu His Pro Met Phe Gly Pro Asp Ser Gly Ser Leu Ala
        195                 200                 205

aag caa gtt gtg gtc tgg tgt gat gga cgt aaa ccg gaa gca tac caa      672
Lys Gln Val Val Val Trp Cys Asp Gly Arg Lys Pro Glu Ala Tyr Gln
    210                 215                 220

tgg ttt ctg gag caa att cag gtc tgg ggc gct cgg ctg cat cgt att      720
Trp Phe Leu Glu Gln Ile Gln Val Trp Gly Ala Arg Leu His Arg Ile
225                 230                 235                 240

agc gcc gtc gag cac gat cag aat atg gcg ttt att cag gca ctg cgc      768
Ser Ala Val Glu His Asp Gln Asn Met Ala Phe Ile Gln Ala Leu Arg
                245                 250                 255
```

```
cac ttt gct act ttt gct tac ggg ctg cac ctg gca gaa gaa aat gtt      816
His Phe Ala Thr Phe Ala Tyr Gly Leu His Leu Ala Glu Glu Asn Val
            260                 265                 270

cag ctt gag caa ctt ctg gcg ctc tct tcg ccg att tac cgc ctt gag      864
Gln Leu Glu Gln Leu Leu Ala Leu Ser Ser Pro Ile Tyr Arg Leu Glu
        275                 280                 285

ctg gcg atg gtc ggg cga ctg ttt gct cag gat ccg cag ctt tat gcc      912
Leu Ala Met Val Gly Arg Leu Phe Ala Gln Asp Pro Gln Leu Tyr Ala
    290                 295                 300

gac atc att atg tcg tca gag cgt aat ctg gcg tta atc aaa cgt tac      960
Asp Ile Ile Met Ser Ser Glu Arg Asn Leu Ala Leu Ile Lys Arg Tyr
305                 310                 315                 320

tat aag cgt ttc ggc gag gcg att gag ttg ctg gag cag ggc gat aag     1008
Tyr Lys Arg Phe Gly Glu Ala Ile Glu Leu Leu Glu Gln Gly Asp Lys
                325                 330                 335

cag gcg ttt att gac agt ttc cgc aag gtg gag cac tgg ttc ggc gat     1056
Gln Ala Phe Ile Asp Ser Phe Arg Lys Val Glu His Trp Phe Gly Asp
            340                 345                 350

tac gca cag cgt ttt cag agt gaa agc cgc gtg tta ttg cgt cag gcg     1104
Tyr Ala Gln Arg Phe Gln Ser Glu Ser Arg Val Leu Leu Arg Gln Ala
        355                 360                 365

aat gac aat cgc cag taa                                            1122
Asn Asp Asn Arg Gln
    370
```

<210> 80
<211> 373
<212> PRT
<213> 大肠杆菌

<400> 80

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

Lys Ala Leu Leu Asn Leu Leu Ala Lys Arg Leu Glu Leu Val Ala Glu
            20                  25                  30

Val Gly Glu Val Lys Ser Arg Phe Gly Leu Pro Ile Tyr Val Pro Glu
        35                  40                  45
```

Arg Glu Ala Ser Met Leu Ala Ser Arg Arg Ala Glu Ala Glu Ala Leu
50 55 60

Gly Val Pro Pro Asp Leu Ile Glu Asp Val Leu Arg Arg Val Met Arg
65 70 75 80

Glu Ser Tyr Ser Ser Glu Asn Asp Lys Gly Phe Lys Thr Leu Cys Pro
85 90 95

Ser Leu Arg Pro Val Val Ile Val Gly Gly Gly Gly Gln Met Gly Arg
100 105 110

Leu Phe Glu Lys Met Leu Thr Leu Ser Gly Tyr Gln Val Arg Ile Leu
115 120 125

Glu Gln His Asp Trp Asp Arg Ala Ala Asp Ile Val Ala Asp Ala Gly
130 135 140

Met Val Ile Val Ser Val Pro Ile His Val Thr Glu Gln Val Ile Gly
145 150 155 160

Lys Leu Pro Pro Leu Pro Lys Asp Cys Ile Leu Val Asp Leu Ala Ser
165 170 175

Val Lys Asn Gly Pro Leu Gln Ala Met Leu Val Ala His Asp Gly Pro
180 185 190

Val Leu Gly Leu His Pro Met Phe Gly Pro Asp Ser Gly Ser Leu Ala
195 200 205

Lys Gln Val Val Val Trp Cys Asp Gly Arg Lys Pro Glu Ala Tyr Gln
210 215 220

Trp Phe Leu Glu Gln Ile Gln Val Trp Gly Ala Arg Leu His Arg Ile
225 230 235 240

Ser Ala Val Glu His Asp Gln Asn Met Ala Phe Ile Gln Ala Leu Arg
245 250 255

His Phe Ala Thr Phe Ala Tyr Gly Leu His Leu Ala Glu Glu Asn Val
260 265 270

Gln Leu Glu Gln Leu Leu Ala Leu Ser Ser Pro Ile Tyr Arg Leu Glu
275 280 285

Leu Ala Met Val Gly Arg Leu Phe Ala Gln Asp Pro Gln Leu Tyr Ala
290 295 300

Asp Ile Ile Met Ser Ser Glu Arg Asn Leu Ala Leu Ile Lys Arg Tyr
305 310 315 320

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

Gln Ala Phe Ile Asp Ser Phe Arg Lys Val Glu His Trp Phe Gly Asp
340 345 350

Tyr Ala Gln Arg Phe Gln Ser Glu Ser Arg Val Leu Leu Arg Gln Ala
355 360 365

Asn Asp Asn Arg Gln
370

<210> 81
<211> 1716
<212> DNA
<213> 枯草芽孢杆菌

<220>
<221> CDS
<222> (1)..(1716)

<400> 81
atg ttg aca aaa gca aca aaa gaa caa aaa tcc ctt gtg aaa aac aga 48
Met Leu Thr Lys Ala Thr Lys Glu Gln Lys Ser Leu Val Lys Asn Arg
1 5 10 15

ggg gcg gag ctt gtt gtt gat tgc tta gtg gag caa ggt gtc aca cat 96
Gly Ala Glu Leu Val Val Asp Cys Leu Val Glu Gln Gly Val Thr His
20 25 30

```
gta ttt ggc att cca ggt gca aaa att gat gcg gta ttt gac gct tta      144
Val Phe Gly Ile Pro Gly Ala Lys Ile Asp Ala Val Phe Asp Ala Leu
        35                  40                  45

caa gat aaa gga cct gaa att atc gtt gcc cgg cac gaa caa aac gca      192
Gln Asp Lys Gly Pro Glu Ile Ile Val Ala Arg His Glu Gln Asn Ala
    50                  55                  60

gca ttc atg gcc caa gca gtc ggc cgt tta act gga aaa ccg gga gtc      240
Ala Phe Met Ala Gln Ala Val Gly Arg Leu Thr Gly Lys Pro Gly Val
65                  70                  75                  80

gtg tta gtc aca tca gga ccg ggt gcc tct aac ttg gca aca ggc ctg      288
Val Leu Val Thr Ser Gly Pro Gly Ala Ser Asn Leu Ala Thr Gly Leu
                85                  90                  95

ctg aca gcg aac act gaa gga gac cct gtc gtt gcg ctt gct gga aac      336
Leu Thr Ala Asn Thr Glu Gly Asp Pro Val Val Ala Leu Ala Gly Asn
            100                 105                 110

gtg atc cgt gca gat cgt tta aaa cgg aca cat caa tct ttg gat aat      384
Val Ile Arg Ala Asp Arg Leu Lys Arg Thr His Gln Ser Leu Asp Asn
        115                 120                 125

gcg gcg cta ttc cag ccg att aca aaa tac agt gta gaa gtt caa gat      432
Ala Ala Leu Phe Gln Pro Ile Thr Lys Tyr Ser Val Glu Val Gln Asp
    130                 135                 140

gta aaa aat ata ccg gaa gct gtt aca aat gca ttt agg ata gcg tca      480
Val Lys Asn Ile Pro Glu Ala Val Thr Asn Ala Phe Arg Ile Ala Ser
145                 150                 155                 160

gca ggg cag gct ggg gcc gct ttt gtg agc ttt ccg caa gat gtt gtg      528
Ala Gly Gln Ala Gly Ala Ala Phe Val Ser Phe Pro Gln Asp Val Val
                165                 170                 175

aat gaa gtc aca aat acg aaa aac gtg cgt gct gtt gca gcg cca aaa      576
Asn Glu Val Thr Asn Thr Lys Asn Val Arg Ala Val Ala Ala Pro Lys
            180                 185                 190

ctc ggt cct gca gca gat gat gca atc agt gcg gcc ata gca aaa atc      624
Leu Gly Pro Ala Ala Asp Asp Ala Ile Ser Ala Ala Ile Ala Lys Ile
        195                 200                 205

caa aca gca aaa ctt cct gtc gtt ttg gtc ggc atg aaa ggc gga aga      672
Gln Thr Ala Lys Leu Pro Val Val Leu Val Gly Met Lys Gly Gly Arg
    210                 215                 220

ccg gaa gca att aaa gcg gtt cgc aag ctt ttg aaa aag gtt cag ctt      720
Pro Glu Ala Ile Lys Ala Val Arg Lys Leu Leu Lys Lys Val Gln Leu
225                 230                 235                 240
```

```
cca ttt gtt gaa aca tat caa gct gcc ggt acc ctt tct aga gat tta      768
Pro Phe Val Glu Thr Tyr Gln Ala Ala Gly Thr Leu Ser Arg Asp Leu
                245                 250                 255

gag gat caa tat ttt ggc cgt atc ggt ttg ttc cgc aac cag cct ggc      816
Glu Asp Gln Tyr Phe Gly Arg Ile Gly Leu Phe Arg Asn Gln Pro Gly
            260                 265                 270

gat tta ctg cta gag cag gca gat gtt gtt ctg acg atc ggc tat gac      864
Asp Leu Leu Leu Glu Gln Ala Asp Val Val Leu Thr Ile Gly Tyr Asp
        275                 280                 285

ccg att gaa tat gat ccg aaa ttc tgg aat atc aat gga gac cgg aca      912
Pro Ile Glu Tyr Asp Pro Lys Phe Trp Asn Ile Asn Gly Asp Arg Thr
    290                 295                 300

att atc cat tta gac gag att atc gct gac att gat cat gct tac cag      960
Ile Ile His Leu Asp Glu Ile Ile Ala Asp Ile Asp His Ala Tyr Gln
305                 310                 315                 320

cct gat ctt gaa ttg atc ggt gac att ccg tcc acg atc aat cat atc     1008
Pro Asp Leu Glu Leu Ile Gly Asp Ile Pro Ser Thr Ile Asn His Ile
                325                 330                 335

gaa cac gat gct gtg aaa gtg gaa ttt gca gag cgt gag cag aaa atc     1056
Glu His Asp Ala Val Lys Val Glu Phe Ala Glu Arg Glu Gln Lys Ile
            340                 345                 350

ctt tct gat tta aaa caa tat atg cat gaa ggt gag cag gtg cct gca     1104
Leu Ser Asp Leu Lys Gln Tyr Met His Glu Gly Glu Gln Val Pro Ala
        355                 360                 365

gat tgg aaa tca gac aga gcg cac cct ctt gaa atc gtt aaa gag ttg     1152
Asp Trp Lys Ser Asp Arg Ala His Pro Leu Glu Ile Val Lys Glu Leu
    370                 375                 380

cgt aat gca gtc gat gat cat gtt aca gta act tgc gat atc ggt tcg     1200
Arg Asn Ala Val Asp Asp His Val Thr Val Thr Cys Asp Ile Gly Ser
385                 390                 395                 400

cac gcc att tgg atg tca cgt tat ttc cgc agc tac gag ccg tta aca     1248
His Ala Ile Trp Met Ser Arg Tyr Phe Arg Ser Tyr Glu Pro Leu Thr
                405                 410                 415

tta atg atc agt aac ggt atg caa aca ctc ggc gtt gcg ctt cct tgg     1296
Leu Met Ile Ser Asn Gly Met Gln Thr Leu Gly Val Ala Leu Pro Trp
            420                 425                 430

gca atc ggc gct tca ttg gtg aaa ccg gga gaa aaa gtg gtt tct gtc     1344
Ala Ile Gly Ala Ser Leu Val Lys Pro Gly Glu Lys Val Val Ser Val
        435                 440                 445
```

tct ggt gac ggc ggt ttc tta ttc tca gca atg gaa tta gag aca gca 1392
Ser Gly Asp Gly Gly Phe Leu Phe Ser Ala Met Glu Leu Glu Thr Ala
450 455 460

gtt cga cta aaa gca cca att gta cac att gta tgg aac gac agc aca 1440
Val Arg Leu Lys Ala Pro Ile Val His Ile Val Trp Asn Asp Ser Thr
465 470 475 480

tat gac atg gtt gca ttc cag caa ttg aaa aaa tat aac cgt aca tct 1488
Tyr Asp Met Val Ala Phe Gln Gln Leu Lys Lys Tyr Asn Arg Thr Ser
485 490 495

gcg gtc gat ttc gga aat atc gat atc gtg aaa tat gcg gaa agc ttc 1536
Ala Val Asp Phe Gly Asn Ile Asp Ile Val Lys Tyr Ala Glu Ser Phe
500 505 510

gga gca act ggc ttg cgc gta gaa tca cca gac cag ctg gca gat gtt 1584
Gly Ala Thr Gly Leu Arg Val Glu Ser Pro Asp Gln Leu Ala Asp Val
515 520 525

ctg cgt caa ggc atg aac gct gaa ggt cct gtc atc atc gat gtc ccg 1632
Leu Arg Gln Gly Met Asn Ala Glu Gly Pro Val Ile Ile Asp Val Pro
530 535 540

gtt gac tac agt gat aac att aat tta gca agt gac aag ctt ccg aaa 1680
Val Asp Tyr Ser Asp Asn Ile Asn Leu Ala Ser Asp Lys Leu Pro Lys
545 550 555 560

gaa ttc ggg gaa ctc atg aaa acg aaa gct ctc tag 1716
Glu Phe Gly Glu Leu Met Lys Thr Lys Ala Leu
565 570

<210> 82
<211> 571
<212> PRT
<213> 枯草芽孢杆菌

<400> 82

Met Leu Thr Lys Ala Thr Lys Glu Gln Lys Ser Leu Val Lys Asn Arg
1 5 10 15

Gly Ala Glu Leu Val Val Asp Cys Leu Val Glu Gln Gly Val Thr His
20 25 30

Val Phe Gly Ile Pro Gly Ala Lys Ile Asp Ala Val Phe Asp Ala Leu
35 40 45

Gln Asp Lys Gly Pro Glu Ile Ile Val Ala Arg His Glu Gln Asn Ala
50 55 60

Ala Phe Met Ala Gln Ala Val Gly Arg Leu Thr Gly Lys Pro Gly Val
65 70 75 80

Val Leu Val Thr Ser Gly Pro Gly Ala Ser Asn Leu Ala Thr Gly Leu
85 90 95

Leu Thr Ala Asn Thr Glu Gly Asp Pro Val Val Ala Leu Ala Gly Asn
100 105 110

Val Ile Arg Ala Asp Arg Leu Lys Arg Thr His Gln Ser Leu Asp Asn
115 120 125

Ala Ala Leu Phe Gln Pro Ile Thr Lys Tyr Ser Val Glu Val Gln Asp
130 135 140

Val Lys Asn Ile Pro Glu Ala Val Thr Asn Ala Phe Arg Ile Ala Ser
145 150 155 160

Ala Gly Gln Ala Gly Ala Ala Phe Val Ser Phe Pro Gln Asp Val Val
165 170 175

Asn Glu Val Thr Asn Thr Lys Asn Val Arg Ala Val Ala Ala Pro Lys
180 185 190

Leu Gly Pro Ala Ala Asp Asp Ala Ile Ser Ala Ala Ile Ala Lys Ile
195 200 205

Gln Thr Ala Lys Leu Pro Val Val Leu Val Gly Met Lys Gly Gly Arg
210 215 220

Pro Glu Ala Ile Lys Ala Val Arg Lys Leu Leu Lys Lys Val Gln Leu
225 230 235 240

Pro Phe Val Glu Thr Tyr Gln Ala Ala Gly Thr Leu Ser Arg Asp Leu
245 250 255

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

Asp Leu Leu Leu Glu Gln Ala Asp Val Val Leu Thr Ile Gly Tyr Asp
275 280 285

Pro Ile Glu Tyr Asp Pro Lys Phe Trp Asn Ile Asn Gly Asp Arg Thr
290 295 300

Ile Ile His Leu Asp Glu Ile Ile Ala Asp Ile Asp His Ala Tyr Gln
305 310 315 320

Pro Asp Leu Glu Leu Ile Gly Asp Ile Pro Ser Thr Ile Asn His Ile
325 330 335

Glu His Asp Ala Val Lys Val Glu Phe Ala Glu Arg Glu Gln Lys Ile
340 345 350

Leu Ser Asp Leu Lys Gln Tyr Met His Glu Gly Glu Gln Val Pro Ala
355 360 365

Asp Trp Lys Ser Asp Arg Ala His Pro Leu Glu Ile Val Lys Glu Leu
370 375 380

Arg Asn Ala Val Asp Asp His Val Thr Val Thr Cys Asp Ile Gly Ser
385 390 395 400

His Ala Ile Trp Met Ser Arg Tyr Phe Arg Ser Tyr Glu Pro Leu Thr
405 410 415

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

Ala Ile Gly Ala Ser Leu Val Lys Pro Gly Glu Lys Val Val Ser Val
435 440 445

Ser Gly Asp Gly Gly Phe Leu Phe Ser Ala Met Glu Leu Glu Thr Ala
450 455 460

Val Arg Leu Lys Ala Pro Ile Val His Ile Val Trp Asn Asp Ser Thr
465 470 475 480

Tyr Asp Met Val Ala Phe Gln Gln Leu Lys Lys Tyr Asn Arg Thr Ser
485 490 495

Ala Val Asp Phe Gly Asn Ile Asp Ile Val Lys Tyr Ala Glu Ser Phe
500 505 510

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

Leu Arg Gln Gly Met Asn Ala Glu Gly Pro Val Ile Ile Asp Val Pro
530 535 540

Val Asp Tyr Ser Asp Asn Ile Asn Leu Ala Ser Asp Lys Leu Pro Lys
545 550 555 560

Glu Phe Gly Glu Leu Met Lys Thr Lys Ala Leu
565 570

<210> 83
<211> 48
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 83
ctagagctcg aaggagatat accatgaaac cagtaacgtt atacgatg 48

<210> 84
<211> 29
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 84
ctagagctct cactgcccgc tttccagtc 29

<210> 85
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 85
cgagcggtac catgcatatt acatacgatc tgccgg 36

<210> 86
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 86
acgcagtcga cttaagcgtc aacgaaaccg gtgatt 36

<210> 87
<211> 37
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 87
tcaggtacca tgcgagtgtt gaagttcggc ggtacat 37

<210> 88
<211> 39
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸引物序列

<400> 88
tcaaagcttt tactgatgat tcatcatcaa tttacgcaa 39

申请人或代理机构案号：0069697-000027	国际申请号：PCT/US08/53514

有关保藏的微生物或其他生物材料的指明

（PCT 条约第 13bis 条）

A. 以下是记录了被保藏的微生物或其他生物材料在说明书中指明的位置 页码 全文，行____	
B. 保藏证明	其他保藏物在附页中记录☒
保藏单位名称 美国典型菌种保藏中心	
保藏机构地址（*包含邮编及国家*） 大学路 10801 号 马纳萨斯,弗吉尼亚州 20110-2209 美国	
保藏日期 2008 年 2 月 7 日	保藏编号： PTA-8944
C. 其他指明（*如果不适用则无需填写*）	此信息在附页中继续☒
埃希氏杆菌（E. coli）保藏人提供的识别号为 CRS-BuOH 23	
D. 此指明适用下述指定国（如果指明不针对全部指定国做出）	
E. 随后公开的指明信息（*如果不适用则无需填写*）	
下列指明信息会随后提交至国际局（*说明指示物的分类，如“保藏编号”*）	

仅供受理局填写	仅供国际局填写
☐ 此页与国际申请一并收到	☐ 国际局于下述日期收到此页
经授权官员	经授权官员

PCT/RO/134 表（1998 年 7 月；2004 年 1 月再版）

申请人或代理机构案号：0069697-000027	国际申请号：PCT/US08/53514

有关保藏的微生物或其他生物材料的指明
（PCT 条约第 13bis 条）

A. 以下是记录了被保藏的微生物或其他生物材料在说明书中指明的位置 页数 全文，行____	
B. 保藏证明	其他保藏物在附页中记录☐
保藏单位名称 美国典型菌种保藏中心	
保藏机构地址（*包含邮编及国家*） 大学路 10801 号 马纳萨斯,弗吉尼亚州 20110-2209 美国	
保藏日期 2008 年 2 月 7 日	保藏编号： PTA-8945
C. 其他知名（*如果不适用则无需填写*）	此信息在附页中继续☒
埃希氏杆菌（E. coli）保藏人提供的识别号为 SA237	
D. 此指明适用下述指定国（如果指明不针对全部指定国做出）	
E. 随后公开的指明信息（*如果不适用则无需填写*）	
下列指明信息会随后提交至国际局（*说明指示物的分类，如“保藏编号”*）	

仅供受理局填写	仅供国际局填写
☐ 此页与国际申请一并收到	☐ 国际局于下述日期收到此页
经授权官员	经授权官员

PCT/RO/134 表（1998 年 7 月；2004 年 1 月再版）

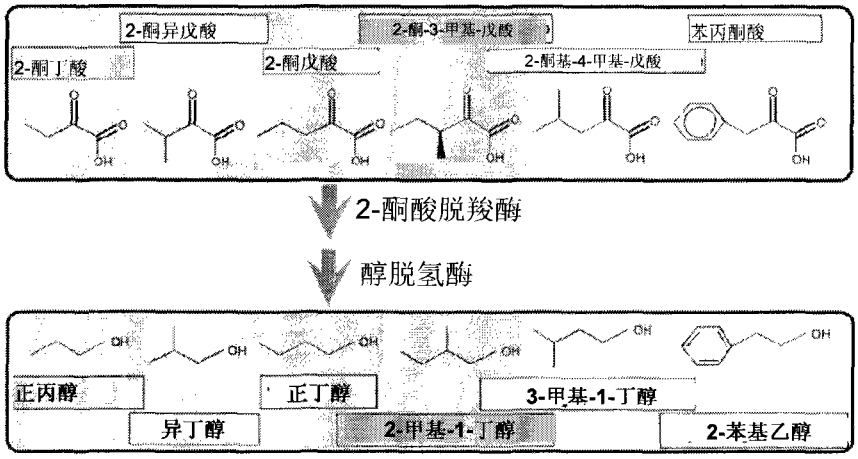

2-酮异戊酸
2-酮-3-甲基-戊酸
苯丙酮酸
2-酮丁酸
2-酮戊酸
2-酮基-4-甲基-戊酸
2-酮酸脱羧酶
醇脱氢酶
正丙醇
正丁醇
3-甲基-1-丁醇
异丁醇
2-甲基-1-丁醇
2-苯基乙醇

图 1A

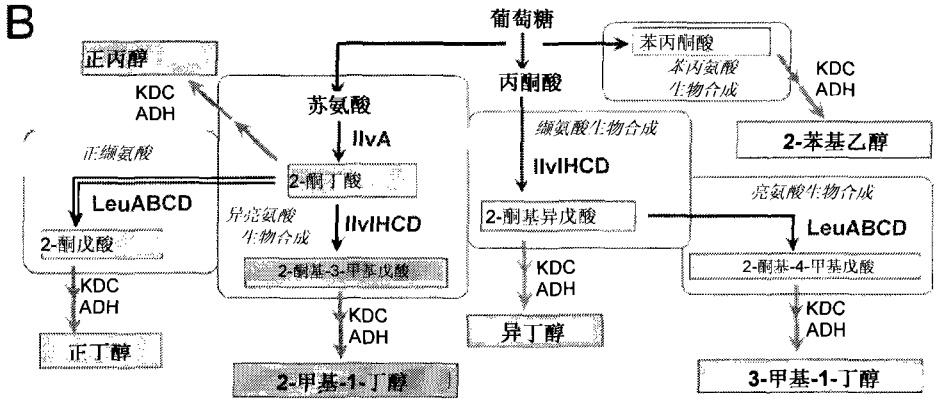

B
葡萄糖
正丙醇
KDC
ADH
苏氨酸
IlvA
正缬氨酸
2-酮丁酸
LeuABCD
异亮氨酸
生物合成
IlvIHCD
2-酮戊酸
2-酮基-3-甲基戊酸
KDC
ADH
正丁醇
2-甲基-1-丁醇
丙酮酸
缬氨酸生物合成
IlvIHCD
2-酮基异戊酸
KDC
ADH
异丁醇
苯丙酮酸
苯丙氨酸
生物合成
KDC
ADH
2-苯基乙醇
亮氨酸生物合成
LeuABCD
2-酮基-4-甲基戊酸
KDC
ADH
3-甲基-1-丁醇

图 1B

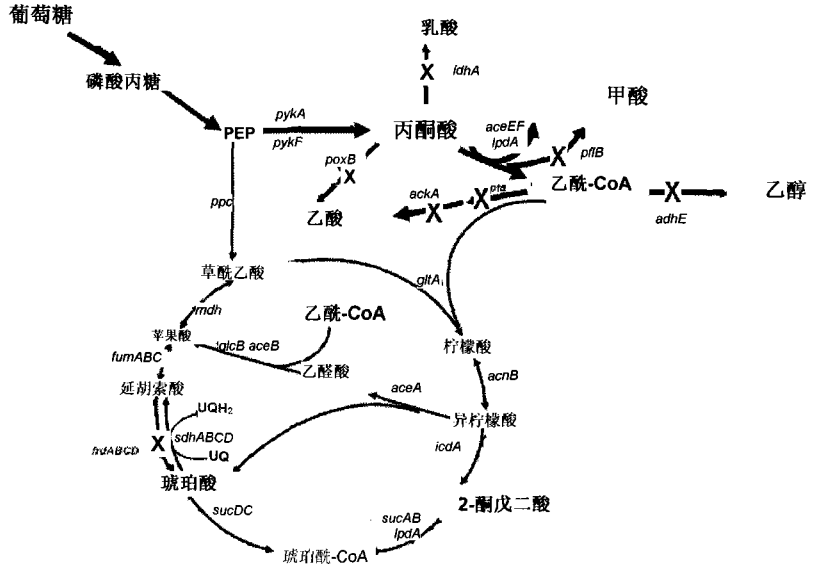

葡萄糖
磷酸丙糖
PEP
pykA
pykF
丙酮酸
乳酸
ldhA
甲酸
poxB
乙酸
aceEF
lpdA
pflB
乙酰-CoA
adhE
乙醇
ackA
pta
ppc
草酰乙酸
gltA
乙酰-CoA
mdh
苹果酸
fumABC
glcB aceB
乙醛酸
柠檬酸
acnB
aceA
异柠檬酸
icdA
延胡索酸
UQH2
sdhABCD
UQ
frdABCD
琥珀酸
2-酮戊二酸
sucDC
sucAB
lpdA
琥珀酰-CoA

图 1C

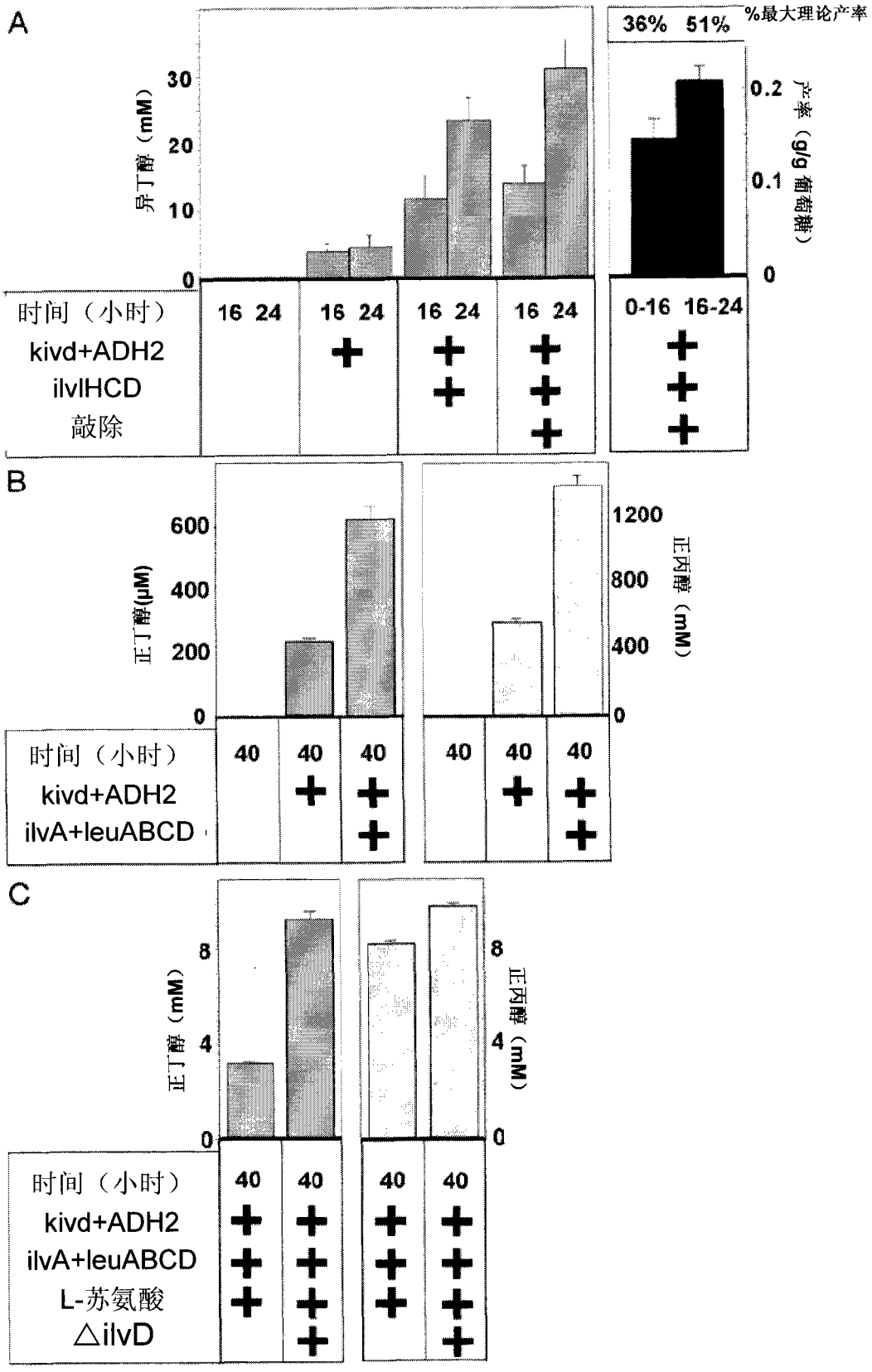

A
异丁醇（mM）
30
20
10
0
36% 51%
%最大理论产率
0.2
0.1
0
产率（g/g 葡萄糖）
时间（小时）
16 24
16 24
16 24
16 24
0-16 16-24
kivd+ADH2
ilvIHCD
敲除
B
正丁醇(μM)
600
400
200
0
1200
800
400
0
正丙醇（mM）
时间（小时）
40
kivd+ADH2
ilvA+leuABCD
C
正丁醇（mM）
8
4
0
正丙醇（mM）
时间（小时）
kivd+ADH2
ilvA+leuABCD
L-苏氨酸
△ilvD

图 2

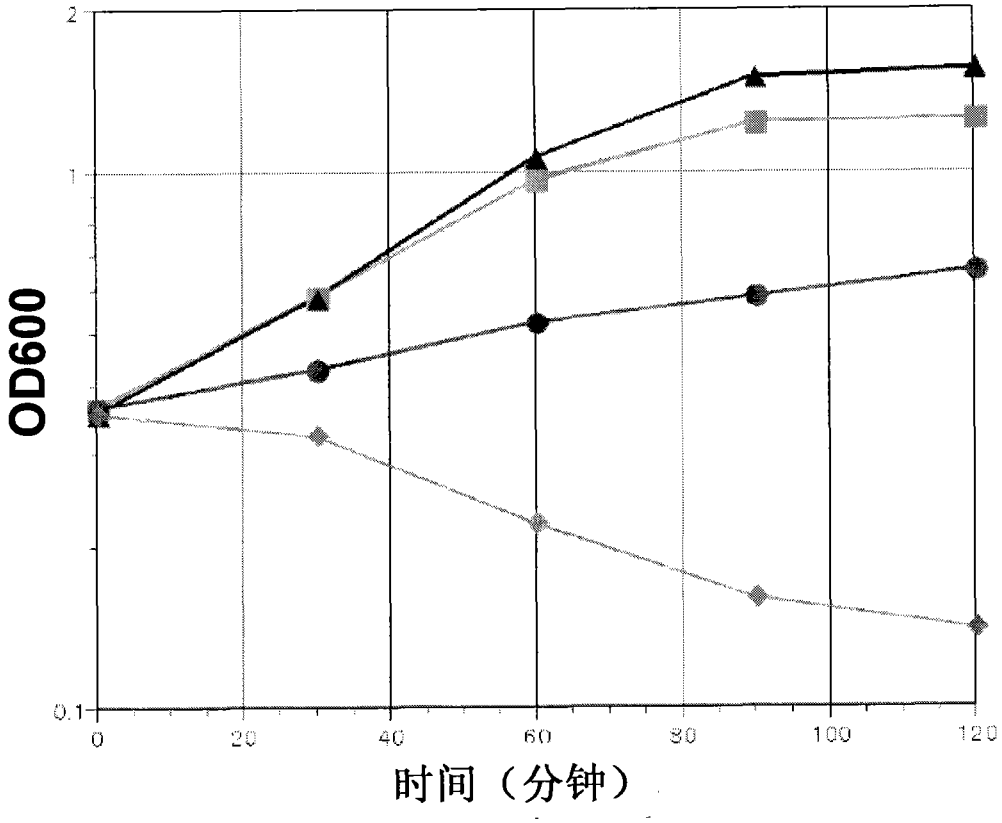


图 3

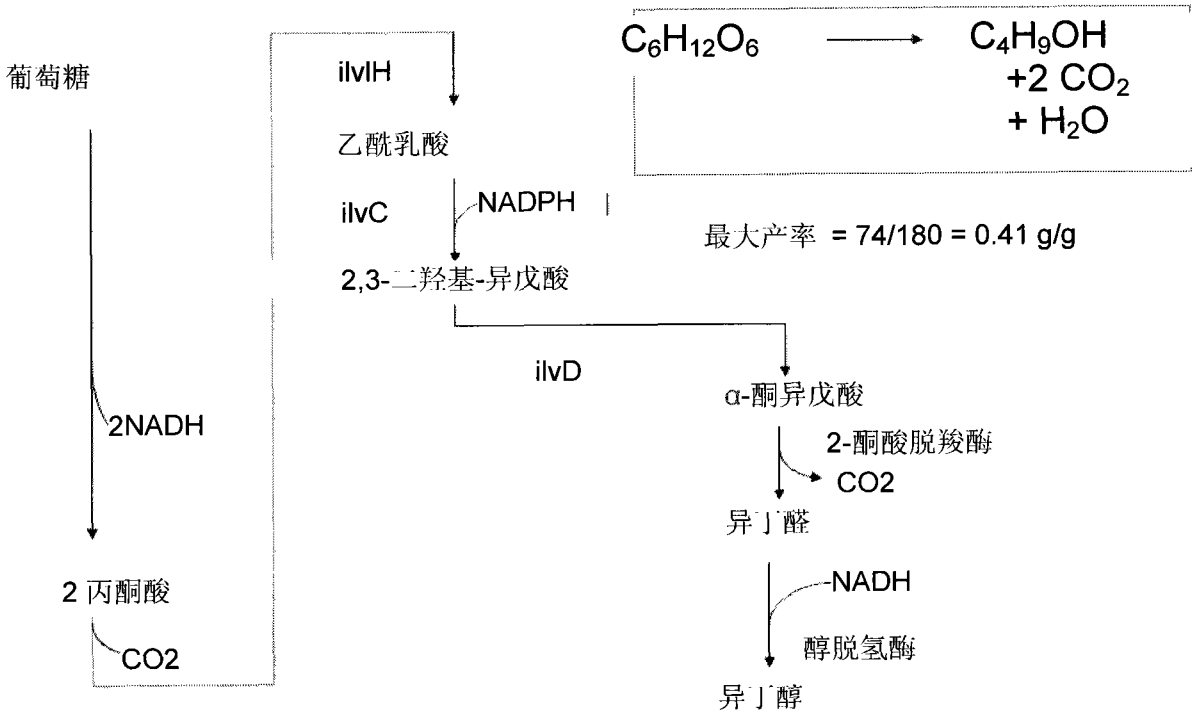

葡萄糖
ilvlH
乙酰乳酸
ilvC
NADPH
2,3-二羟基-异戊酸
ilvD
α-酮异戊酸
2-酮酸脱羧酶
CO2
异丁醛
NADH
醇脱氢酶
异丁醇
2NADH
2 丙酮酸
CO2
$C_6H_{12}O_6$ → C_4H_9OH +2 CO_2 + H_2O
最大产率 = 74/180 = 0.41 g/g

图 4

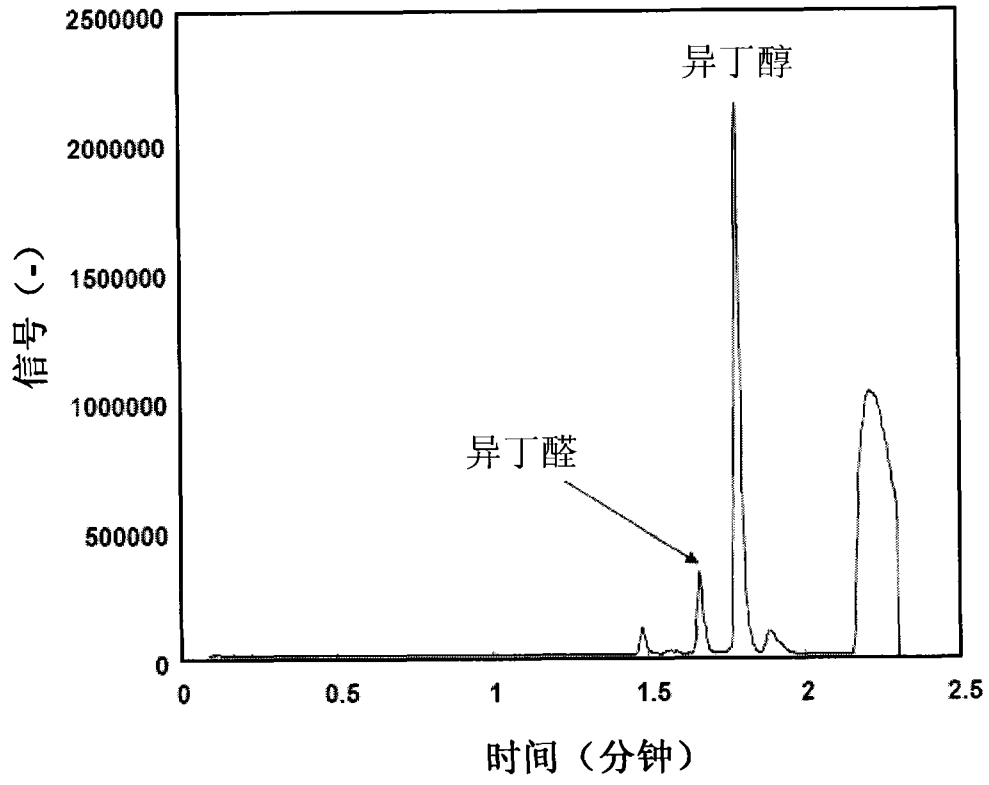

2500000
2000000
1500000
1000000
500000
0
信号（-）
异丁醇
异丁醛
0
0.5
1
1.5
2
2.5
时间（分钟）

图 5

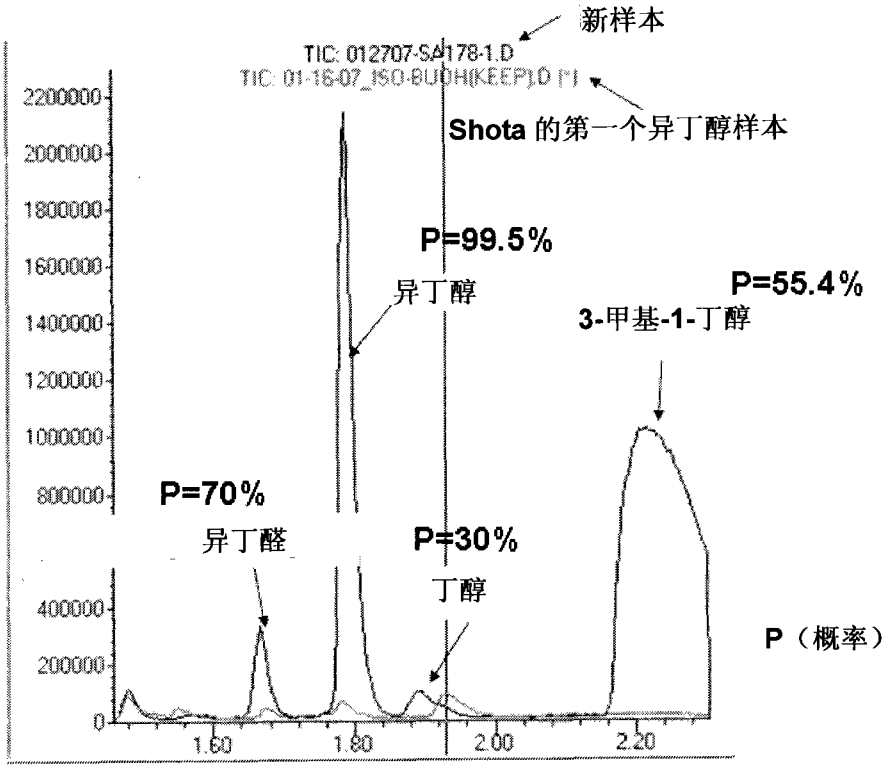

新样本
TIC: 012707-SA178-1.D
TIC: 01-16-07_ISO-BUOH(KEEP).D
Shota 的第一个异丁醇样本
P=99.5%
异丁醇
P=55.4%
3-甲基-1-丁醇
P=70%
异丁醛
P=30%
丁醇
P（概率）
2200000
2000000
1800000
1600000
1400000
1200000
1000000
800000
400000
200000
0
1.60
1.80
2.00
2.20

图 6

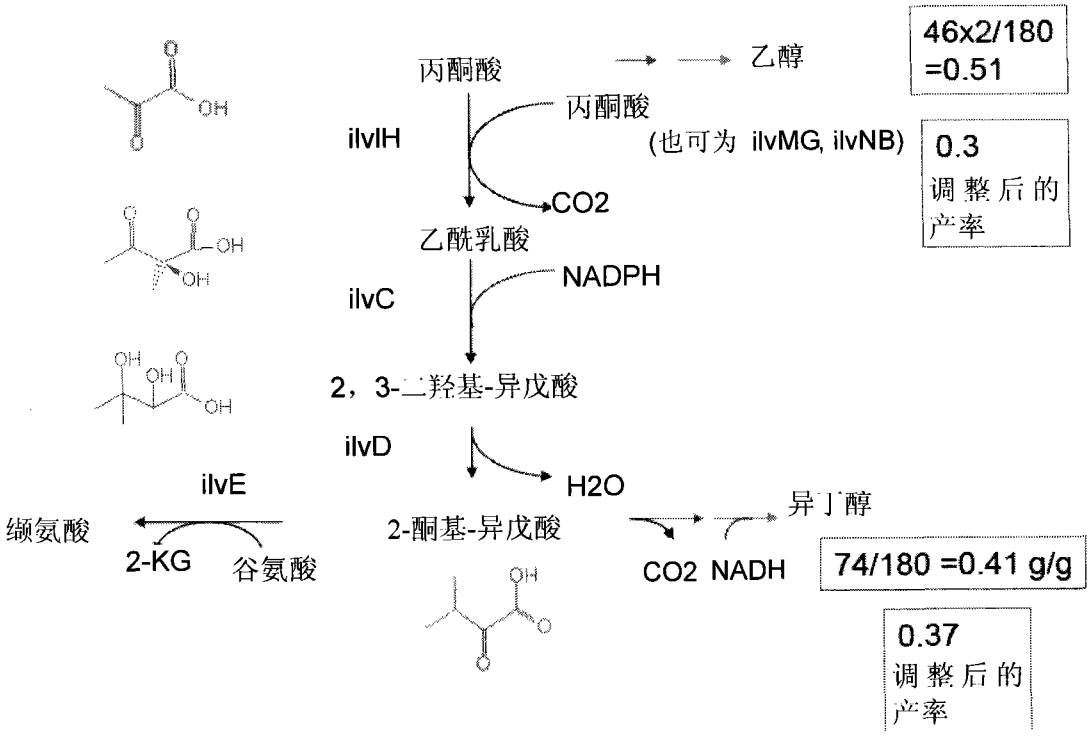

丙酮酸
乙醇
46x2/180
=0.51
ilvIH
丙酮酸
(也可为 ilvMG, ilvNB)
0.3
调整后的
产率
CO2
乙酰乳酸
ilvC
NADPH
2，3-二羟基-异戊酸
ilvD
H2O
ilvE
缬氨酸
2-KG
谷氨酸
2-酮基-异戊酸
异丁醇
CO2 NADH
74/180 =0.41 g/g
0.37
调整后的
产率

图 7

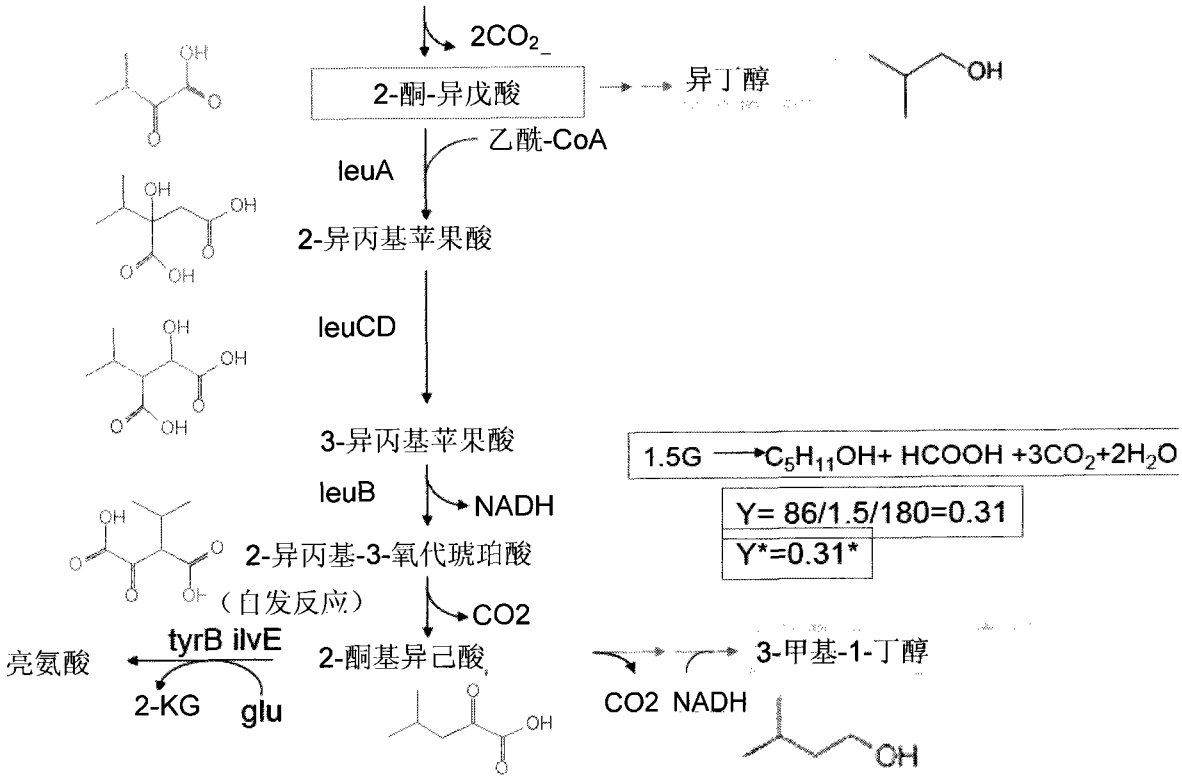

2CO2
2-酮-异戊酸
异丁醇
OH
乙酰-CoA
leuA
2-异丙基苹果酸
leuCD
3-异丙基苹果酸
1.5G → C5H11OH+ HCOOH +3CO2+2H2O
leuB
NADH
Y= 86/1.5/180=0.31
Y*=0.31*
2-异丙基-3-氧代琥珀酸
（白发反应）
CO2
tyrB ilvE
亮氨酸
2-KG
glu
2-酮基异己酸
3-甲基-1-丁醇
CO2 NADH
OH

图 8

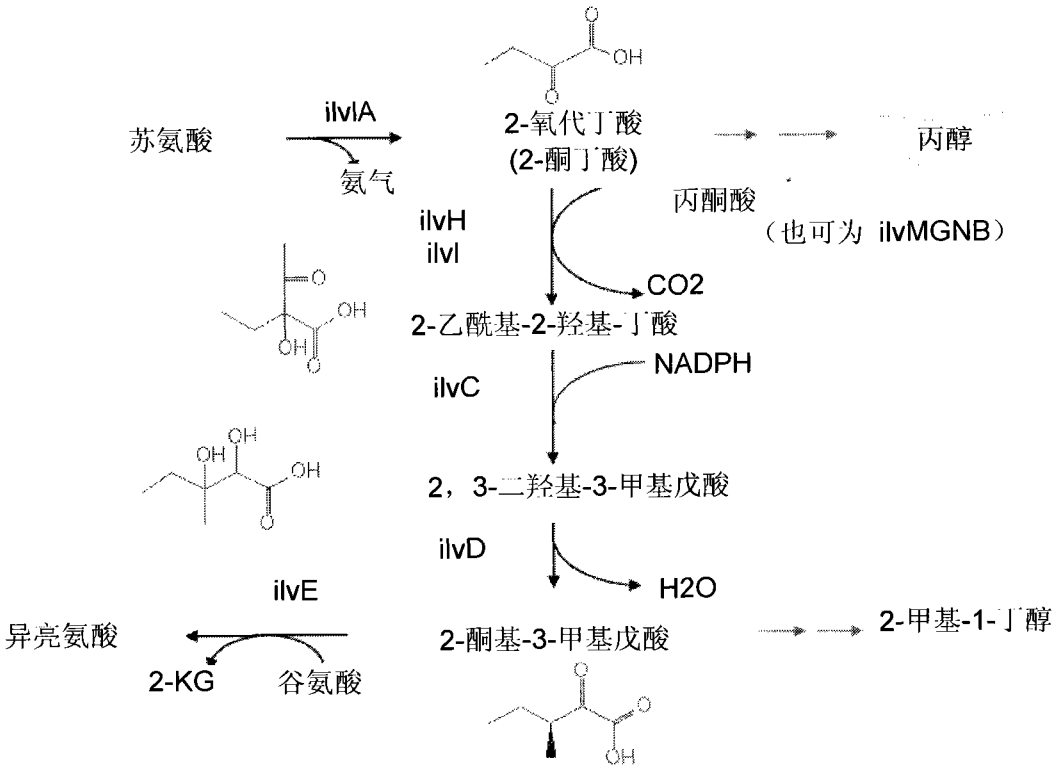

苏氨酸
ilvIA
氨气
2-氧代丁酸
(2-酮丁酸)
丙醇
丙酮酸
（也可为 ilvMGNB）
ilvH
ilvI
CO2
2-乙酰基-2-羟基-丁酸
ilvC
NADPH
2，3-二羟基-3-甲基戊酸
ilvD
H2O
ilvE
异亮氨酸
2-KG
谷氨酸
2-酮基-3-甲基戊酸
2-甲基-1-丁醇

图 9

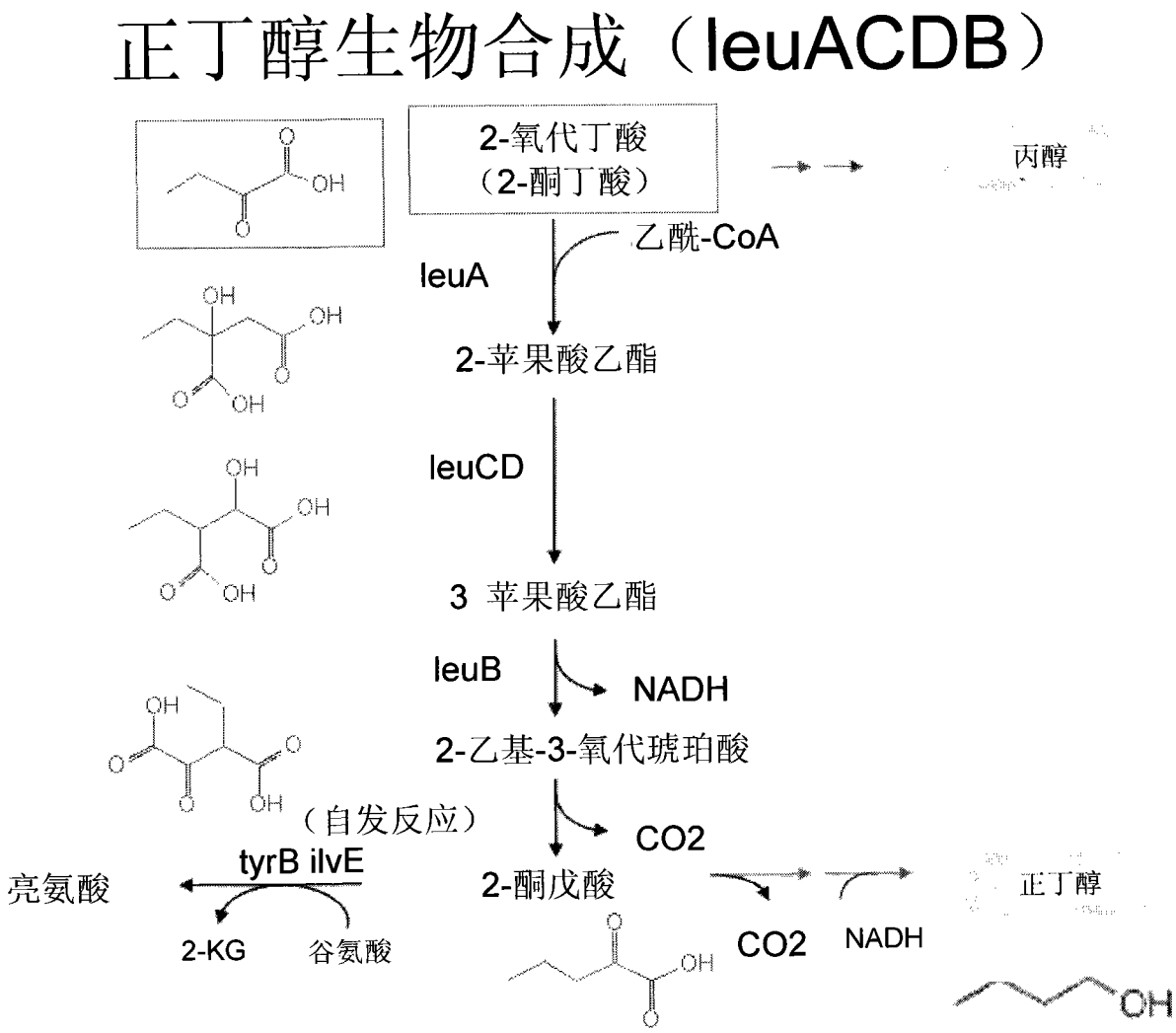
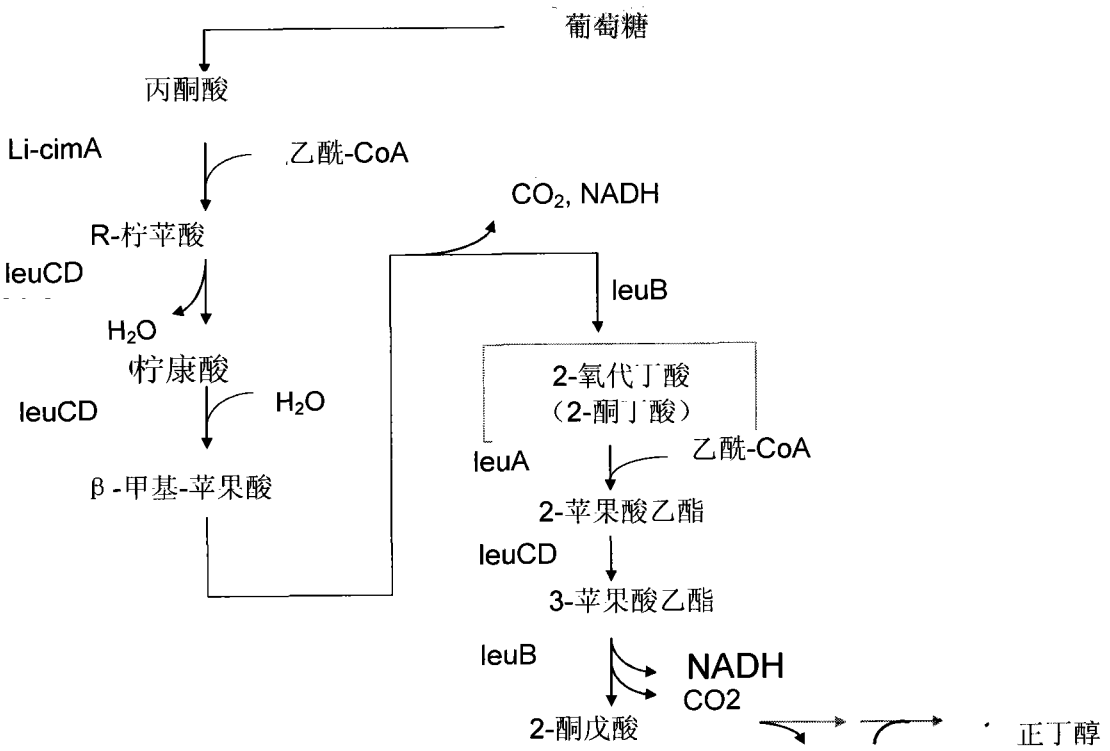

正丁醇生物合成（leuACDB）
2-氧代丁酸
（2-酮丁酸）
丙醇
乙酰-CoA
leuA
2-苹果酸乙酯
leuCD
3 苹果酸乙酯
leuB
NADH
2-乙基-3-氧代琥珀酸
（自发反应）
CO2
亮氨酸
tyrB ilvE
2-酮戊酸
正丁醇
2-KG
谷氨酸
CO2
NADH
OH

图 10



Li: 问号钩端螺旋体

图 11

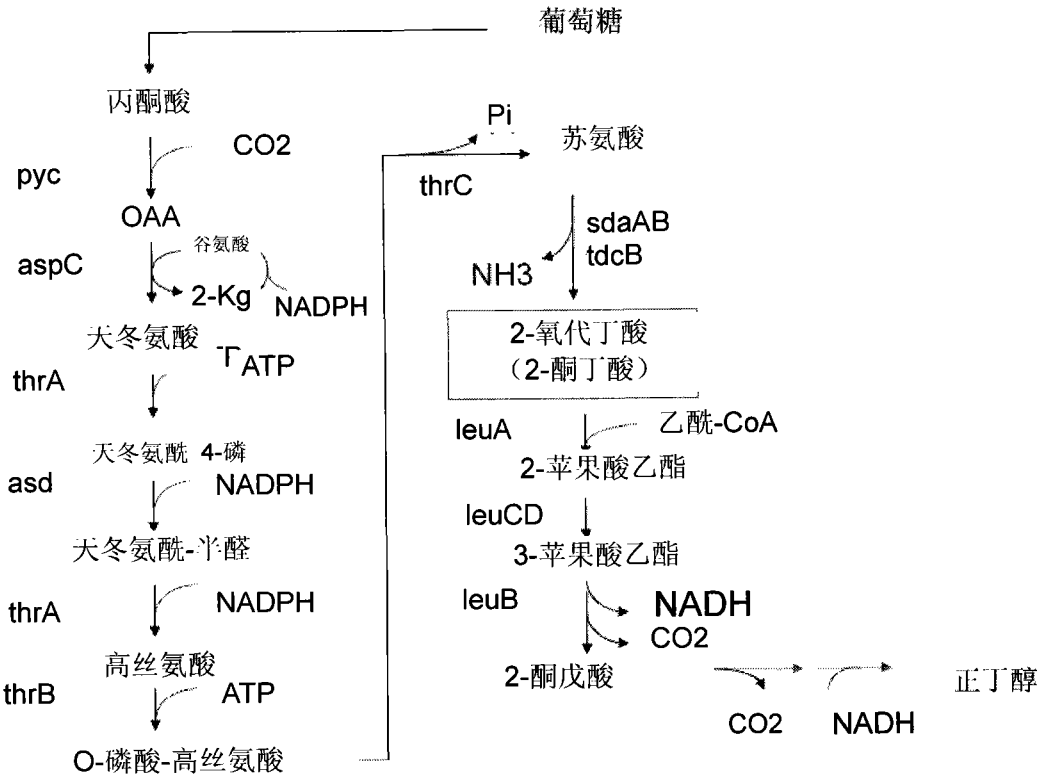


图 12

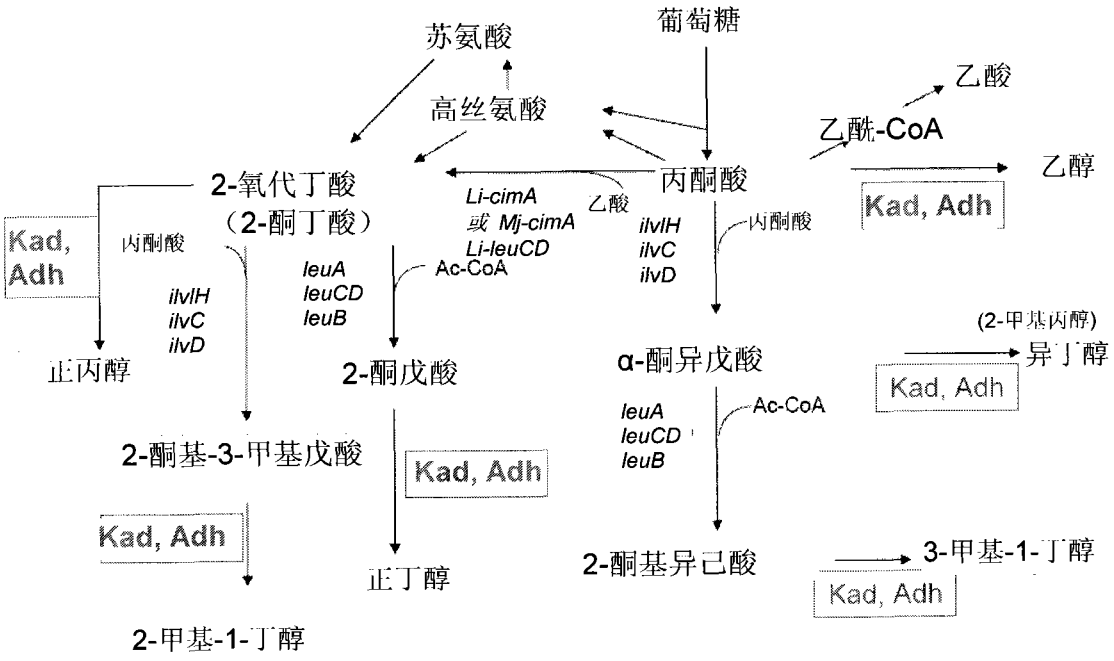

苏氨酸
葡萄糖
高丝氨酸
乙酸
乙酰-CoA
乙醇
2-氧代丁酸
（2-酮丁酸）
丙酮酸
Kad, Adh
Li-cimA
或 Mj-cimA
Li-leuCD
乙酸
ilvIH
ilvC
ilvD
丙酮酸
Kad, Adh
丙酮酸
leuA
leuCD
leuB
Ac-CoA
(2-甲基丙醇)
异丁醇
正丙醇
2-酮戊酸
α-酮异戊酸
Kad, Adh
leuA
leuCD
leuB
Ac-CoA
2-酮基-3-甲基戊酸
Kad, Adh
Kad, Adh
正丁醇
2-酮基异己酸
3-甲基-1-丁醇
Kad, Adh
2-甲基-1-丁醇

图 13

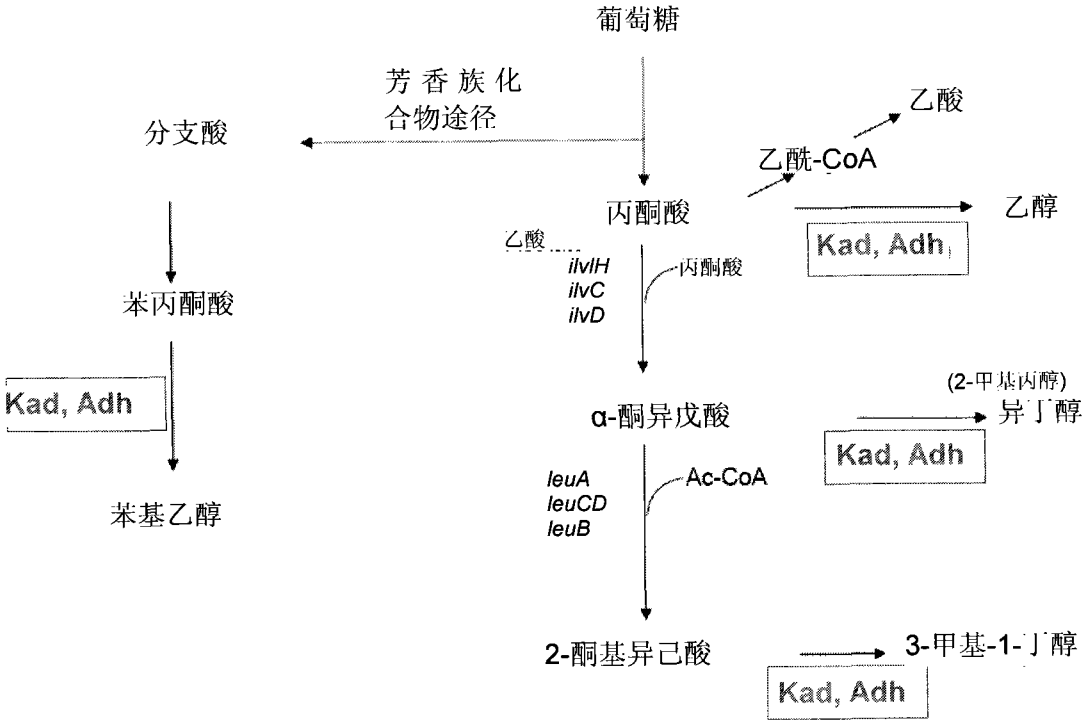

葡萄糖
芳香族化
合物途径
分支酸
乙酸
乙酰-CoA
丙酮酸
乙醇
Kad, Adh
乙酸
ilvIH
ilvC
ilvD
丙酮酸
苯丙酮酸
(2-甲基丙醇)
异丁醇
Kad, Adh
α-酮异戊酸
Kad, Adh
leuA
leuCD
leuB
Ac-CoA
苯基乙醇
2-酮基异己酸
3-甲基-1-丁醇
Kad, Adh

图 14

kivd: 内酮酸脱羧酶（*乳酸乳球菌*）（序列 ID 号：27）

atgtatacagtaggagattacctattagaccgattacacgagttaggaattgaagaaatttttggagtccctggagactataacttacaattttta
gatcaaattatttcccgcaaggatatgaaatgggtcggaaatgctaatgaattaaatgcttcatatatggctgatggctatgctcgtactaaaaaaa
gctgccgcatttcttacaacctttggagtaggtgaattgagtgcagttaatggattagcaggaagttacgccgaaaatttaccagtagtagaaata
gtgggatcacctacatcaaaagttcaaaatgaaggaaaatttgttcatcatacgctggctgacggtgattttaaacactttatgaaaatgcacgaa
cctgttacagcagctcgaactttactgacagcagaaaatgcaaccgttgaaattgaccgagtactttctgcactattaaaagaaagaaaacctgtc
tatatcaacttaccagttgatgttgctgctgcaaaagcagagaaaccctcactcctttgaaaaaagaaaactcaacttcaaatacaagtgaccaa
gagatcttgaacaaaattcaagaaagcttgaaaaatgccaaaaaaccaatcgtgattacaggacatgaaataattagttttggcttagaaaaaaca
gtctctcaatttatttcaaagacaaaactacctattacgacattaaactttggaaaaagttcagttgatgaagctctcccttcatttttaggaatc
tataatggtaaactctcagagcctaatcttaaagaattcgtggaatcagccgacttcatcctgatgcttggagttaaactcacagactcttcaaca
ggagccttcactcatcatttaaatgaaaataaaatgatttcactgaatatagatgaaggaaaaatatttaacgaaagcatccaaaattttgatttt
gaatccctcatctcctctctcttagacctaagcgaaatagaatacaaaggaaaatatatcgataaaaagcaagaagactttgttccatcaaatgcg
cttttatcacaagaccgcctatggcaagcagttgaaaacctaactcaaagcaatgaaacaatcgttgctgaacaagggacatcattctttggcgct
tcatcaattttcttaaaaccaaagagtcattttattggtcaacccttatggggatcaattggatatacattcccagcagcattaggaagccaaatt
gcagataaagaaagcagacacctttttatttattggtgatggttcacttcaacttacggtgcaagaattaggattagcaatcagagaaaaaattaat
ccaatttgctttattatcaataatgatggttatacagtcgaaagagaaattcatggaccaaatcaaagctacaatgatattccaatgtggaattac
tcaaaattaccagaatcatttggagcaacagaagaacgagtagtctcgaaaatcgttagaactgaaaatgaatttgtgtctgtcatgaaagaagct
caagcagatccaaatagaatgtactggattgagttaattttggcaaaagaagatgcaccaaaagtactgaaaaaaatgggcaaactatttgctgaa
caaaataaatcataa

图 15

PDC6（*酿酒酵母*）（序列 ID 号：29）

atgtctgaaattactcttggaaaatacttattgaaagattgaagcaagttaatgttaacaccatttttgggctaccaggcgacttcaacttgtcc
ctattggacaagatttacgaggtagatggattgagatgggctggtaatgcaaatgagctgaacgccgcctatgccgccgatggttacgcacgcatc
aagggtttatctgtgctggtaactacttttggcgtaggtgaattatccgccttgaatggtattgcaggatcgtatgcagaacacgtcggtgtactg
catgttgttggtgtcccctctatctccgctcaggctaagcaattgttgttgcatcataccttgggtaacggtgatttaccgtttttcacagaatg
tccgccaatatctcagaaactacatcaatgattacagacattgctacagcccctcagaaatcgataggttgatcaggacaacatttataacacaa
aggcctagctacttggggttgccagcgaatttggtagatctaaaggttcctggttctcttttggaaaaaccgattgatctatcattaaaacctaac
gatcccgaagctgaaaaggaagttattgataccgtactagaattgatccagaattcgaaaaaccctgttatactatcggatgcctgtgcttctagg
cacaacgttaaaaaagaaacccagaagttaattgatttgacgcaattcccagcttttgtgacacctctaggtaaagggtcaatagatgaacagcat
cccagatatggcggtgtttatgtgggaacgctgtccaaacaagacgtgaaacaggccgttgagtcggctgatttgatcctttcggtcggtgctttg
ctctctgattttaacacaggttcgttttcctactcctacaagactaaaaatgtagtggagtttcattccgattacgtaaaggtgaagaacgctacg
ttcctcggtgtacaaatgaaatttgcactacaaaacttactgaaggttattcccgatgttgttaagggctacaagagcgttcccgtaccaaccaaa
actcccgcaaacaaaggtgtacctgctagcacgcccttgaaacaagagtggttgtggaacgaattgtccaaattcttgcaagaaggtgatgttatc
atttccgagaccggcacgtctgccttcggtatcaatcaaactatctttcctaaggacgcctacggtatctcgcaggtgttgtgggggtccatcggt
tttacaacaggagcaactttaggtgctgcctttgccgctgaggagattgaccccaacaagagagtcatcttattcataggtgacgggtctttgcag
ttaaccgtccaagaaatctccaccatgatcagatgggggttaaagccgtatcttttgtccttaacaacgacggctacactatcgaaaagctgatt
catgggcctcacgcagagtacaacgaaatccagacctgggatcacctcgccctgttgcccgcatttggtgcgaaaaagtacgaaaatcacaagatc
gccactacgggtgagtgggatgccttaaccactgattcagagttccagaaaaactcggtgatcagactaattgaactgaaactgcccgtctttgat
gctccggaaagtttgatcaaacaagcgcaattgactgccgctacaaatgccaaacaataa

图 16

ARO10（酿酒酵母）（序列 ID 号：31）

atggcacctgttacaattgaaaagttcgtaaatcaagaagaacgacaccttgtttccaaccgatcagcaacaattccgtttggtgaatacatattt
aaaagattgttgtccatcgatacgaaatcagttttcggtgttcctggtgacttcaacttatctctattagaatatctctattcacctagtgttgaa
tcagctggcctaagatgggtcggcacgtgtaatgaactgaacgccgcttatgcggccgacggatattcccgttactctaataagattggctgttta
ataaccacgtatggcgttggtgaattaagcgccttgaacggtatagccggttcgttcgctgaaaatgtcaaagttttgcacattgttggtgtggcc
aagtccatagattcgcgttcaagtaactttagtgatcggaacctacatcatttggtcccacagctacatgattcaaattttaaagggccaaatcat
aaagtatatcatgatatggtaaaagatagagtcgcttgctcggtagcctacttggaggatattgaaactgcatgtgaccaagtcgataatgttatc
cgcgatatttacaaggattctaaacctggttatatttttgttcctgcagattttgcggatatgtctgttacatgtgataatttggttaatgttcca
cgtatatctcaacaagattgtatagtatacccttctgaaaaccaattgtctgacataatcaacaagattactagttggatatattccagtaaaaca
cctgcgatccttggagacgtactgactgataggtatggtgtgagtaacttttgaacaagcttatctgcaaaactgggatttggaatttttccact
gttatgggaaaatctgtaattgatgagtcaaacccaacttatatgggtcaatataatggtaaagaaggtttaaaacaagtctatgaacattttgaa
ctgtgcgacttggtcttgcattttggagtcgacatcaatgaaattaataatgggcattatacttttacttataaaccaaatgctaaaatcattcaa
tttcatccgaattatattcgccttgtggacactaggcagggcaatgagcaaatgttcaaaggaatcaattttgcccctattttaaaagaactatac
aagcgcattgacgtttctaaactttctttgcaatatgattcaaatgtaactcaatatacgaacgaaacaatgcggttagaagatcctaccaatgga
caatcaagcattattacacaagttcacttacaaaagacgatgcctaaatttttgaaccctggtgatgttgtcgtttgtgaaacaggctcttttcaa
ttctctgttcgtgatttcgcgtttccttcgcaattaaaatatatatcgcaaggatttttcctttccattggcatggcccttcctgccgccctaggt
gttggaattgccatgcaagaccactcaaacgctcacatcaatggtggcaacgtaaaagaggactataagccaagattaattttgtttgaaggtgac
ggtgcagcacagatgacaatccaagaactgagcaccattctgaagtgcaatattccactagaagttatcatttggaacaataacggctacactatt
gaaagagccatcatgggccctaccaggtcgtataacgacgttatgtcttggaaatggaccaaactatttgaagcattcggagacttcgacggaaag
tatactaatagcactctcattcaatgtccctctaaattagcactgaaattggaggagcttaagaattcaaacaaaagaagcgggatagaactttta
gaagtcaaattaggcgaattggatttccccgaacagctaaagtgcatggttgaagcagcggcacttaaaagaaataaaaaatag

图 17

THI3（酿酒酵母）（序列 ID 号：33）

atgaattctagctatacacagagatatgcactgccgaagtgtatagcaatatcagattatcttttccatcggctcaaccagctgaacatacatacc
atatttggactctccggagaatttagcatgccgttgctggataaactatacaacattccgaacttacgatgggccggtaattctaatgagttaaat
gctgcctacgcagcagatggatactcacgactaaaaggcttgggatgtctcataacaacctttggtgtaggcgaattatcggcaatcaatggcgtg
gccggatcttacgctgaacatgtaggaatacttcacatagtgggtatgccgccaacaagtgcacaaacgaaacaactactactgcatcatactctg
ggcaatggtgatttcacggtatttcatagaatagccagtgatgtagcatgctatacaacattgattattgacictgaattatgtgccgacgaagtc
gataagtgcatcaaaaaggcttggatagaacagaggccagtatacatgggcatgcctgtcaaccaggtaaatctcccgattgaatcagcaaggctt
aatacacctctggatttacaattgcataaaaacgacccagacgtagagaaagaagttatttctcgaatattgagttttatatacaaaagccagaat
ccggcaatcatcgtagatgcatgtactagtcgacagaatttaatcgaggagactaaagagctttgtaataggcttaaatttccagttttgttaca
cctatgggtaagggtacagtaaacgaaacagacccgcaatttgggggcgtattcacgggctcgatatcagccccagaagtaagagaagtagttgat
tttgccgattttatcatcgtcattggttgcatgctctccgaattcagcacgtcaactttccacttccaatataaaactaagaattgtgcgctacta
tattctacatctgtgaaattgaaaaatgccacatatcctgacttgagcattaaattactactacagaaaatattagcaaatcttgatgaatctaaa
ctgtcttaccaaccaagcgaacaacccagtatgatggttccaagaccttacccagcaggaaatgtcctcttgagacaagaatgggtctggaatgaa
atatcccattggttccaaccaggtgacataatcataacagaaactggtgcttctgcatttggagttaaccagaccagatttccggtaaatacacta
ggtatttcgcaagctctttggggatctgtcggatatacaatgggggcgtgtcttggggcagaatttgctgttcaagagataaacaaggataaattc
cccgcaactaaacatagagttattctgtttatgggtgacggtgctttccaattgacagttcaagaattatccacaattgttaagtggggattgaca
ccttatatttttgtgatgaataaccaaggttactctgtggacaggtttttgcatcacaggtcagatgctagttattacgatatccaaccttggaac
tacttgggattattgcgagtatttggttgcacgaactacgaaacgaaaaaaattattactgttggagaattcagatccatgatcagtgacccaaac
tttgcgaccaatgacaaaattcggatgatagagattatgctaccaccaagggatgttccacaggctctgcttgacaggtgggtggtagaaaaagaa
cagagcaaacaagtgcaagaggagaacgaaaattctagcgcagtaaatacgccaactccagaattccaaccacttctaaaaaaaaatcaagttgga
tactga

图 18

pdc (丙酮丁醇梭菌) (序列 ID 号：35)

ttgaagagtgaatacacaattggaagatatttgttagaccgtttatcagagttgggtattcggcatatctttggtgtacctggagattacaatcta
tccttttttagactatataatggagtacaaagggatagattgggttggaaattgcaatgaattgaatgctgggtatgctgctgatggatatgcaaga
ataaatggaattggagccatacttacaacatttggtgttggagaattaagtgccattaacgcaattgctggggcatacgctgagcaagttccagtt
gttaaaattacaggtatccccacagcaaaagttagggacaatggattatatgtacaccacacattaggtgacggaaggtttgatcacttttttgaa
atgtttagagaagtaacagttgctgaggcattactaagcgaagaaaatgcagcacaagaaattgatcgtgttcttatttcatgctggagacaaaaa
cgtcctgttcttataaatttaccgattgatgtatatgataaaccaattaacaaaccattaaagccattactcgattatactatttcaagtaacaaa
gaggctgcatgtgaatttgttacagaaatagtacctataataaatagggcaaaaaagcctgttattcttgcagattatggagtatatcgttaccaa
gttcaacatgtgcttaaaaacttggccgaaaaaaccggatttcctgtggctacactaagtatgggaaaaggtgttttcaatgaagcacaccctcaa
tttattggtgtttataatggtgatgtaagttctccttatttaaggcagcgagttgatgaagcagactgcattattagcgttggtgtaaaattgacg
gattcaaccacaggggattttctcatggattttctaaaaggaatgtaattcacattgatccttttttcaataaaggcaaaaggtaaaaaatatgca
cctattacgatgaaagatgctttaacagaattaacaagtaaaattgagcatagaaactttgaggatttagatataaagccttacaaatcagataat
caaaagtattttgcaaaagagaagccaattacacaaaaacgttttttgagcgtattgctcactttataaaagaaaaagatgtattattagcagaa
cagggtacatgctttttggtgcgtcaaccatacaactacccaaagatgcaacttttattggtcaacctttatggggatctattggatacacactt
cctgctttattaggttcacaattagctgatcaaaaaaggcgtaatattcttttaattggggatggtgcatttcaaatgacagcacaagaaatttca
acaatgcttcgtttacaaatcaaacctattatttttttaattaataacgatggttatacaattgaacgtgctattcatggtagagaacaagtatat
aacaatattcaaatgtggcgatatcataatgttccaaaggttttaggtcctaaagaatgcagcttaacctttaaagtacaaagtgaaactgaactt
gaaaaggctcttttagtggcagataaggattgtgaacatttgatttttatagaagttgttatggatcgttatgataaacccgagcctttagaacgt
ctttcgaaacgttttgcaaatcaaaataattag

图 19

ADH2: 醇脱氢酶(*酿酒酵母*) (序列 ID 号：37)

atgccttcgcaagtcattcctgaaaaacaaaaggctattgtctttttatgagacagatggaaaattggaatataaagacgtcacagttccggaacct
aagcctaacgaaatttttagtccacgttaaatattctggtgtttgtcatagtgacttgcacgcgtggcacggtgattggccatttcaattgaaatttt
ccattaatcggtggtcacgaaggtgctggtgttgttgttaagttgggatctaacgttaagggctggaaagtcggtgattttgcaggtataaaatgg
ttgaatgggacttgcatgtcctgtgaatattgtgaagtaggtaatgaatctcaatgtccttatttggatggtactggcttcacacatgatggtact
tttcaagaatacgcaactgccgatgccgttcaagctgcccatattccaccaaacgtcaatcttgctgaagttgccccaatcttgtgtgcaggtatc
actgtttataaggcgttgaaaagagccaatgtgataccaggccaatgggtcactatatccggtgcatgcggtggcttgggttctctggcaatccaa
tacgcccttgctatgggttacagggtcattggtatcgatggtggtaatgccaagcgaaagttatttgaacaattaggcggagaaatattcatcgat
ttcacggaagaaaaagacattgttggtgctataataaaggccactaatggcggttctcatggagttattaatgtgtctgtttctgaagcagctatc
gaggcttctacgaggtattgtaggcccaatggtactgtcgtcctggttggtatgccagctcatgcttactgcaattccgatgttttcaatcaagtt
gtaaaatcaatctccatcgttggatcttgtgttggaaatagagctgatacaagggaggctttagatttcttcgccagaggtttgatcaaatctccg
atccacttagctggcctatcggatgttcctgaaatttttgcaaagatggagaagggtgaaattgttggtagatatgttgttgagacttctaaatga

图 20

ilvI: (*大肠杆菌*) (序列 ID 号：39)

atggagatgttgtctggagccgagatggtcgtccgatcgcttatcgatcagggcgttaaacaagtattcggttatcccggaggcgcagtccttgat
atttatgatgcattgcataccgtgggtggtattgatcatgtattagttcgtcatgagcaggcggcggtgcatatggccgatggcctggcgcgcgcg
accggggaagtcggcgtcgtgctggtaacgtcgggtccaggggcgaccaatgcgattactggcatcgccaccgcttatatggattccattccatta
gttgtccttccgggcaggtagcgacctcgttgataggttacgatgcctttcaggagtgcgacatggtggggatttcgcgaccggtggttaaacac
agttttctggttaagcaaacggaagacattccgcaggtgctgaaaaaggctttctggctggcggcaagtggtcgcccaggaccagtagtcgttgat
ttaccgaaagatattcttaatccggcgaacaaattaccctatgtcctggccggagtcggtcagtatgcgttcttacaatcccactactaccggacat
aaagggcaaattaagcgtgctctgcaaacgctggtagcggcaaaaaaaccggttgtctacgtaggcggtggggcaatcacggcgggctgccatcag
cagttgaaagaaacggtggaggcgttgaatctgcccgttgtttgctcattgatggggctgggggcgtttccggcaacgcatcgtcaggcactgggc
atgctgggaatgcacggtacctacgaagccaatatgacgatgcataacgcggatgtgatttcgccgtcggggtacgatttgatgaccgaacgacg
aacaatctggcaaagtactgcccaaatgccactgttctgcatatcgatattgatcctacttccatttctaaaaccgtgactgcggatatcccgatt
gtgggggatgctcgccaggtcctcgaacaaatgcttgaactcttgtcgcaagaatccgcccatcaaccactggatgagatccgcgactggtggcag
caaattgaacagtggcgcgctcgtcagtgcctgaaatatgacactcacagtgaaaagattaaaccgcaggcggtgatcgagactctttggcggttg
acgaagggagacgcttacgtgacgtccgatgtcgggcagcaccagatgtttgctgcactttattatccattcgacaaaccgcgtcgctggatcaat
tccggtggcctcggcacgatgggttttggtttacctgcggcactgggcgtcaaaatggcgttgccagaagaaaccgtggtttgcgtcactggcgac
ggcagtattcagatgaacatccaggaactgtctaccgcgttgcaatacgagttgcccgtactggtggtgaatctcaataaccgctatctggggatg
gtgaagcagtggcaggacatgatctattccggccgtcattcacaatcttatatgcaatcgctacccgatttcgtccgtctggcggaagcctatggg
catgtcgggatccagatttctcatccgcatgagctggaaagcaaacttagcgaggcgctggaacaggtgcgcaataatcgcctggtgtttgttgat
gttaccgtcgatggcagcgagcacgtctacccgatgcagattcgcgggggcggaatggatgaaatgtggttaagcaaaacggagagaacctga

图 21

ilvH: (*大肠杆菌*) (序列 ID 号：41)

atgcgccggatattatcagtcttactcgaaaatgaatcaggcgcgttatcccgcgtgattggccttttttcccagcgtggctacaacattgaaagc
ctgaccgttgcgccaaccgacgatccgacattatcgcgtatgaccatccagaccgtgggcgatgaaaaagtacttgagcagatcgaaaagcaatta
cacaaactggtcgatgtcttgcgcgtgagtgagttggggcagggcgcgcatgttgagcgggaaatcatgctggtgaaaattcaggccagcggttac
gggcgtgacgaagtgaaacgtaatacggaaatattccgtgggcaaattatcgatgtcacaccctcgctttataccgttcaattagcaggcaccagc
ggtaagcttgatgcatttttagcatcgattcgcgatgtggcgaaaattgtggaggttgctcgctctggtgtggtcggactttcgcgcggcgataaa
ataatgcgttga

图 22

ilvC: (*大肠杆菌*) (序列 ID 号：43)

atggctaactacttcaatacactgaatctgcgccagcagctggcacagctgggcaaatgtcgctttatgggccgcgatgaattcgccgatggcgcg
agctaccttcagggtaaaaaagtagtcatcgtcggctgtggcgcacagggtctgaaccagggcctgaacatgcgtgattctggtctcgatatctcc
tacgctctgcgtaaagaagcgattgccgagaagcgcgcgtcctggcgtaaagcgaccgaaaatggttttaaagtgggtacttacgaagaactgatc
ccacaggcggatctggtgattaacctgacgccggacaagcagcactctgatgtagtgcgcaccgtacagccactgatgaaagacggcgcggcgctg
ggctactcgcacggtttcaacatcgtcgaagtgggcgagcagatccgtaaagatatcaccgtagtgatggttgcgccgaaatgcccaggcaccgaa
gtgcgtgaagagtacaaacgtgggttcggcgtaccgacgctgattgccgttcacccggaaaacgatccgaaaggcgaaggcatggcgattgccaaa
gcctgggcggctgcaaccggtggtcaccgtgcgggtgtgctggaatcgtccttcgttgcggaagtgaaatctgacctgatgggcgagcaaaccatc
ctgtgcggtatgttgcaggctggctctctgctgtgcttcgacaagctggtggaagaaggtaccgatccagcatacgcagaaaaactgattcagttc
ggttgggaaaccatcaccgaagcactgaaacagggcggcatcaccctgatgatggaccgtctctctaacccggcgaaactgcgtgcttatgcgctt
tctgaacagctgaaagagatcatggcacccctgttccagaaacatatggacgacatcatctccggcgaattctcttccggtatgatggcggactgg
gccaacgatgataagaaactgctgacctggcgtgaagagaccggcaaaaccgcgtttgaaaccgcgccgcagtatgaaggcaaaatcggcgagcag
gagtacttcgataaaggcgtactgatgattgcgatggtgaaagcgggcgttgaactggcgttcgaaaccatggtcgattccggcatcattgaagag
tctgcatattatgaatcactgcacgagctgccgctgattgccaacaccatcgcccgtaagcgtctgtacgaaatgaacgtggttatctctgatacc
gctgagtacggtaactatctgttctcttacgcttgtgtgccgttgctgaaaccgtttatggcagagctgcaaccgggcgacctgggtaaagctatt
ccggaaggcgcggtagataacgggcaactgcgtgatgtgaacgaagcgattcgcagccatgcgattgagcaggtaggtaagaaactgcgcggctat
atgacagatatgaaacgtattgctgttgcgggttaa

图 23

ilvD: *(大肠杆菌)* (序列 ID 号: 45)

atgcctaagtaccgttccgccaccaccactcatggtcgtaatatggcgggtgctcgtgcgctgtggcgcgccaccggaatgaccgacgccgatttc
ggtaagccgattatcgcggttgtgaactcgttcacccaatttgtaccgggtcacgtccatctgcgcgatctcggtaaactggtcgccgaacaaatt
gaagcggctggcggcgttgccaaagagttcaacaccattgcggtggatgatgggattgccatgggccacgggggggatgctttattcactgccatct
cgcgaactgatcgctgattccgttgagtatatggtcaacgcccactgcgccgacgccatggtctgcatctctaactgcgacaaaatcaccccgggg
atgctgatggcttccctgcgcctgaatattccggtgatctttgtttccggcggcccgatggaggccgggaaaaccaaactttccgatcagatcatc
aagctcgatctggttgatgcgatgatccagggcgcagacccgaaagtatctgactcccagagcgatcaggttgaacgttccgcgtgtccgacctgc
ggttcctgctccgggatgtttaccgctaactcaatgaactgcctgaccgaagcgctgggcctgtcgcagccgggcaacggctcgctgctggcaacc
cacgccgaccgtaagcagctgttccttaatgctggtaaacgcattgttgaattgaccaaacgttattacgagcaaaacgacgaaagtgcactgccg
cgtaatatcgccagtaaggcggcgtttgaaaacgccatgacgclggatatcgcgatgggtggatcgactaacaccgtacttcacctgctggcggcg
gcgcaggaagcggaaatcgacttcaccatgagtgatatcgataagctttcccgcaaggttccacagctgtgtaaagttgcgccgagcacccagaaa
taccatatggaagatgttcaccgtgctggtggtgttatcggtattctcggcgaactggatcgcgcggggttactgaaccgtgatgtgaaaaacgta
cttggcctgacgttgccgcaaacgctggaacaatacgacgttatgctgacccaggatgacgcggtaaaaaatatgttccgcgcaggtcctgcaggc
attcgtaccacacaggcattctcgcaagattgccgttgggatacgctggacgacgatcgcgccaatggctgtatccgctcgctggaacacgcctac
agcaaagacggcggcctggcggtgctctacggtaactttgcggaaaacggctgcatcgtgaaaacggcaggcgtcgatgacagcatcctcaaattc
accggcccggcgaaagtgtacgaaagccaggacgatgcggtagaagcgattctcggcggtaaagttgtcgccggagatgtggtagtaattcgctat
gaaggcccgaaaggcggtccggggatgcaggaaatgctctacccaaccagcttcctgaaatcaatgggtctcggcaaagcctgtgcgctgatcacc
gacggtcgtttctctggtggcacctctggtctttccatcggccacgtctcaccggaagcggcaagcggcggcagcattggcctgattgaagatggt
gacctgatcgctatcgacatcccgaaccgtggcattcagttacaggtaagcgatgccgaactggcggcgcgtcgtgaagcgcaggacgctcgaggt
gacaaagcctggacgccgaaaaatcgtgaacgtcaggtctcctttgcccigcgtgcttatgccagcctggcaaccagcgccgacaaaggcgcggtg
cgcgataaatcgaaactgggggttaa

图 24

ilvA: *(大肠杆菌)* (序列 ID 号: 47)

atggctgactcgcaacccctgtccggtgctccggaaggtgccgaatatttaagagcagtgctgcgcgcgccggtttacgaggcggcgcaggttacg
ccgctacaaaaaaatggaaaaactgtcgtcgcgtcttgataacgtcattctggtgaagcgcgaagatcgccagccagtgcacagctttaagctgcgc
ggcgcatacgccatgatggcgggcctgacggaagaacagaaagcgcacggcgtgatcactgcttctgcgggtaaccacgcgcagggcgtcgcgttt
tcttctgcgcggttaggcgtgaaggccctgatcgttatgccaaccgccaccgccgacatcaaagtcgacgcggtgcgcggcttcggcggcgaagtg
ctgctccacggcgcgaactttgatgaagcgaaagccaaagcgatcgaactgtcacagcagcaggggttcacctgggtgccgccgttcgaccatccg
atggtgattgccgggcaaggcacgctggcgctggaactgctccagcaggacgcccatctcgaccgcgtatttgtgccagtcggcggcggcggtctg
gctgctggcgtggcggtgctgatcaaacaactgatgccgcaaatcaaagtgatcgccgtagaagcggaagactccgcctgcctgaaagcagcgctg
gatgcgggtcatccggttgatctgccgcgcgtagggctatttgctgaaggcgtagcggtaaaacgcatcggtgacgaaaccttccgtttatgccag
gagtatctcgacgacatcatcaccgtcgatagcgatgcgatctgtgcggcgatgaaggatttattcgaagatgtgcgcgcggtggcggaaccctct
ggcgcgctggcgctggcgggaatgaaaaaaatatatcgccctgcacaacattcgcggcgaacggctggcgcatattctttccggtgccaacgtgaac
ttccacggcctgcgctacgtctcagaacgctgcgaactgggcgaacagcgtgaagcgttgttggcggtgaccattccggaagaaaaaggcagcttc
ctcaaattctgccaactgcttggcgggcgttcggtcaccgagttcaactaccgttttgccgatgccaaaaacgcctgcatctttgtcggtgtgcgc
ctgagccgcggcctcgaagagcgcaaagaaattttgcagatgctcaacgacggcggctacagcgtggttgatctctccgacgacgaaatggcgaag
ctacacgtgcgctatatggtcggcggacgtccatcgcatccgttgcaggaacgcctctacagcttcgaattcccggaatcaccgggcgcgctgctg
cgcttcctcaacacgctgggtacgtactggaacatttctttgttccactatcgcagccatggcaccgactacgggcgcgtactggcggcgttcgaa
cttggcgaccatgaaccggatttcgaaacccggctgaatgagctgggctacgattgccacgacgaaaccaataacccggcgttcaggttctttttg
gcgggttag

图 25

leuA: (序列 ID 号：49)

atgagccagcaagtcattattttcgataccacattgcgcgacggtgaacaggcgttacaggcaagcttgagtgtgaaagaaaaactgcaaattgcg
ctggcccttgagcgtatgggtgttgacgtgatggaagtcggtttccccgtctcttcgccgggcgattttgaatcggtgcaaaccatcgcccgccag
gttaaaaacagccgcgtatgtgcgttagctcgctgcgtggaaaaagatatcgacgtggcggccgaatccctgaaagtcgccgaagccttccgtatt
catacctttattgccacttcgccaatgcacatcgccaccaagctgcgcagcacgctggacgaggtgatcgaacgcgctatctatatggtgaaacgc
gcccgtaattacaccgatgatgttgaattttcttgcgaagatgccgggcgtacacccattgccgatctggcgcgagtggtcgaagcggcgattaat
gccggtgccaccaccatcaacattccggacaccgtgggctacaccatgccgtttgagttcgccggaatcatcagcggcctgtatgaacgcgtgcct
aacatcgacaaagccattatctccgtacatacccacgacgatttgggcctggcggtcggaaactcactggcggcggtacatgccggtgcacgccag
gtggaaggcgcaatgaacgggatcggcgagcgtgccggaaactgttccctggaagaagtcatcatggcgatcaaagttcgtaaggatattctcaac
gtccacaccgccattaatcaccaggagatatggcgcaccagccagttagttagccagatttgtaatatgccgatcccggcaaacaaagccattgtt
ggcagcggcgcattcgcacactcctccggtatacaccaggatggcgtgctgaaaaaccgcgaaaactacgaaatcatgacaccagaatctattggt
ctgaaccaaatccagctgaatctgacctctcgttcggggcgtgcggcggtgaaacatcgcatggatgagatggggtataaagaaagtgaatataat
ttagacaatttgtacgatgctttcctgaagctggcggacaaaaaaggtcaggtgtttgattacgatctggaggcgctggccttcatcggtaagcag
caagaagagccggagcatttccgtctggattacttcagcgtgcagtctggctctaacgatatcgccaccgccgccgtcaaactggcctgtggcgaa
gaagtcaaagcagaagccgccaacggtaacggtccggtcgatgccgtctatcaggcaattaaccgcatcactgaatataacgtcgaactggtgaaa
tacagcctgaccgccaaaggccacggtaaagatgcgctgggtcaggtggatatcgtcgctaactacaacggtcgccgcttccacggcgtcggcctg
gctaccgatattgtcgagtcatctgccaaagccatggtgcacgttctgaacaatatctggcgtgccgcagaagtcgaaaaagagttgcaacgcaaa
gctcaacacaacgaaaacaacaaggaaaccgtgtga

图 26

leuB: (序列 ID 号：51)

gtgatgtcgaagaattaccatattgccgtattgccgggggacggtattggtccggaagtgatgacccaggcgctgaaagtgctggatgccgtgcgc
aaccgctttgcgatgcgcatcaccaccagccattacgatgtaggcggcgcagccattgataaccacgggcaaccactgccgcctgcgacggttgaa
ggttgtgagcaagccgatgccgtgctgtttggctcggtaggcggcccgaagtgggaacatttaccaccagaccagcaaccagaacgcggcgcgctg
ctgcctctgcgtaagcacttcaaattattcagcaacctgcgcccggcaaaactgtatcaggggctggaagcattctgtccgctgcgtgcagacatt
gccgcaaacggcttcgacatcctgtgtgtgcgcgaactgaccggcggcatctatttcggtcagccaaaaggccgcgaaggtagcggacaatatgaa
aaagcctttgataccgaggtgtatcaccgttttgagatcgaacgtatcgcccgcatcgcgtttgaatctgctcgcaagcgtcgccacaaagtgacg
tcgatcgataaagccaacgtgctgcaatcctctatttttatggcgggagatcgttaacgagatcgccacggaatacccggatgtcgaactggcgcat
atgtacatcgacaacgccaccatgcagctgattaaagatccatcacagtttgacgttctgctgtgctccaacctgtttggcgacattctgtctgac
gagtgcgcaatgatcactggctcgatggggatgttgccttccgccagcctgaacgagcaaggttttggactgtatgaaccggcgggcggctcggca
ccagatatcgcaggcaaaaacatcgccaacccgattgcacaaatcctttcgctggcactgctgctgcgttacagcctggatgccgatgatgcggct
tgcgccattgaacgcgccattaaccgcgcattagaagaaggcattcgcaccggggatttagcccgtggcgctgccgccgttagtaccgatgaaatg
ggcgatatcattgcccgctatgtagcagaaggggtgtaa

图 27

leuC: (序列 ID 号：53)

atggctaagacgttatacgaaaaattgttcgacgctcacgttgtgtacgaagccgaaaacgaaaccccactgttatatatcgaccgccacctggtg
catgaagtgacctcaccgcaggcgttcgatggtctgcgcgcccacggtcgcccggtacgtcagccgggcaaaaccttcgctaccatggatcacaac
gtctctacccagaccaaagacattaatgcctgcggtgaaatggcgcgtatccagatgcaggaactgatcaaaaactgcaaagaatttggcgtcgaa
ctgtatgacctgaatcacccgtatcaggggatcgtccacgtaatggggccggaacagggcgtcaccttgccggggatgaccattgtctgcggcgac
tcgcataccgccacccacggcgcgtttggcgcactggcctttggtatcggcacttccgaagttgaacacgtactggcaacgcaaaccctgaaacag
ggccgcgcaaaaaccatgaaaattgaagtccagggcaaagccgcgccgggcattaccgcaaaagatatcgtgctggcaattatcggtaaaaccggt
agcgcaggcggcaccgggcatgtggtggagtttgcggcgaagcaatccgtgatttaagcatggaaggtcgtatgaccctgtgcaatatggcaatc
gaaatgggcgcaaaagccggtctggttgcaccggacgaaaccaccttaactatgtcaaaggccgtctgcatgcgccgaaaggcaaagatttcgac
gacgccgttgcctactggaaaaccctgcaaaccgacgaaggcgcaactttcgataccgttgtcactctgcaagcagaagaaatttcaccgcaggtc
acctggggcaccaatcccggccaggtgatttccgtgaacgacaatattcccgatccggcttcgtttgccgatccggttgaacgcgcgtcggcagaa
aaagcgctggcctatatggggctgaaaccgggtattccgctgaccgaagtggctatcgacaaagtgtttatcggttcctgtaccaactcgcgcatt
gaagatttacgcgcggcagcggagatcgccaaagggcgaaaagtcgcgccaggcgtgcaggcactggtggttcccggctctggcccggtaaaagcc
caggcggaagcggaaggtctggataaaatctttattgaagccggttttgaatggcgcttgcctggctgctcaatgtgtctggcgatgaacaacgac
cgtctgaatccgggcgaacgttgtgcctccaccagcaaccgtaactttgaaggccgccaggggcgcggcgggcgcacgcatctggtcagcccggca
atggctgccgctgctgctgtgaccggacatttcgccgacattcgcaacattaaataa

图 28

leuD: (序列 ID 号：55)

atggcagagaaatttatcaaacacacaggcctggtggttccgctggatgccgccaatgtcgataccgatgcaatcatcccgaaacagtttttgcag
aaagtgacccgtacgggttttggcgcgcatctgtttaacgactggcgttttctggatgaaaaaggccaacagccaaacccggacttcgtgctgaac
ttcccgcagtatcagggcgcttccattttgctggcacgagaaaacttcggctgtggctcttcgcgtgagcacgcgccctgggcattgaccgactac
ggttttaaagtggtgattgcgccgagttttgctgacatcttctacggcaatagctttaacaaccagctgctgccggtgaaattaagcgatgcagaa
gtggacgaactgtttgcgctggtgaaagctaatccggggatccatttcgacgtggatctggaagcgcaagaggtgaaagcgggagagaaaacctat
cgctttaccatcgatgccttccgccgccactgcatgatgaacggtctggacagtattgggcttaccttgcagcacgacgacgccattgccgcttat
gaagcaaaacaacctgcgtttatgaattaa

图 29

cimA (詹氏甲烷暖球菌): (序列 ID 号：57)

atgatggtaaggatatttgatacaacacttagagatggagagcaaacaccaggagtttctttaacaccaaatgataagttagagatagcaaaaaaa
ttggatgagcttggagttgatgttatagaggcaggttcagctataacttcaaaaggagagagagaaggaataaaattaataacaaaagaaggttta
aatgcagaaatctgctcatttgttagagcttacctgtagatattgatgctgccttagaatgtgatgtagatagtgtccatttagtagtgccaaca
tctccaatacacatgaaatataagcttagaaaaacagaagatgaggttttagagacagctttaaaggctgtagagtatgctaaagaacatggattg
attgttgagttatctgcagaggatgcaacaagaagtgatgtaaatttcttaataaaactatttaatgaaggggaaaaggttggagcagacagagtt
tgtgtttgtgacacagtaggagttttaactccacaaaagagtcaggaattatttaaaaaaataactgaaaatgttaatttaccggtctcagttcat
tgccacaacgactttggaatggctactgctaatacttgctcagcagttttaggtggagctgttcagtgccacgtaacagttaatggtattggagag
agagcaggaaatgcctcattggaagaggttgttgctgctttaaaaatactctatggctatgatactaagataaagatggaaaagttatatgaggtt
tcaagaattgtctcaagattgatgaaacttcctgttccaccaaataaagcaattgttggggacaatgcatttgctcatgaagcaggaatacatgtt
gatggattaataaaaaatactgaaacctatgagccaataaaaccagaaatggttgggaatagaagaagaattattttgggtaagcattctggtaga
aaagctttaaaatacaaacttgatttgatgggcataaacgttagtgatgagcaattaaataaaatatatgaaagagttaaagaatttggggattg
ggtaaatacatttcagacgctgatttgttggctatagttagagaagttactggaaaattggtagaagagaaaatcaaattagatgaattaactgtt
gtatctggaaataaaataacaccaattgcatctgttaaactccattataaaggagaagatataactttaatagaaactgcttatggtgttggaccg
gtagatgcagcaataaatgctgtgagaaaggcaataagtggagttgcagatattaagttggtagagtatagagttgaagcaattggtggaggaact
gatgcgttaatagaggttgttgttaaattaagaaaaggaactgaaattgttgaagttagaaaatcagacgctgatataataagggcttctgtagat
gctgtaatggaaggaatcaatatgttattgaattaa

图 30

ilvM: (序列 ID 号:59)

atgatgcaacatcaggtcaatgtatcggctcgcttcaatccagaaaccttagaacgtgttttacgcgtggtgcgtcatcgtggtttccacgtctgc
tcaatgaatatggccgccgccagcgatgcacaaaatataaatatcgaattgaccgttgccagcccacggtcggtcgacttactgtttagtcagtta
ataaaactggtggacgtcgcacacgttgccatctgccagagcacaaccacatcacaacaaatccgcgcctga

图 31

ilvG: (序列 ID 号：61)

ttgttgttaaaacaactgtcggatcgtaaacctgcggattgcgtcgtgaccacagatgtggggcagcaccagatgtgggctgcgcagcacatcgcc
cacactcgcccggaaaatttcatcacctccagcggtttaggtaccatgggtttggtttaccggcggcggttggcgcacaagtcgcgcgaccgaac
gataccgttgtctgtatctccggtgacggctctttcatgatgaatgtgcaagagctgggcaccgtaaaacgcaagcagttaccgttgaaaatcgtc
ttactcgataaccaacggttagggatggttcgacaatggcagcaactgtttttttcaggaacgatacagcgaaaccacccttactgataaccccgat
ttcctcatgttagccagcgccttcggcatccatggccaacacatcacccggaaagaccaggttgaagcggcactcgacaccatgctgaacagtgat
gggccatacctgcttcatgtctcaatcgacgaacttgagaacgtctggccgctggtgccgcctggcgccagtaattcagaaatgttggagaaatta
tcatga

图 32

ilvN: (序列 ID 号：63)

atgcaaaacacaactcatgacaacgtaattctggagctcaccgttcgcaaccatccgggcgtaatgacccacgtttgtggccttttgcccgccgc
gcttttaacgttgaaggcattctttgtctgccgattcaggacagcgacaaaagccatatctggctactggtcaatgacgaccagcgtctggagcag
atgataagccaaatcgataagctggaagatgtcgtgaaagtgcagcgtaatcagtccgatccgacgatgtttaacaagatcgcggtgttttttcag
taa

图 33

ilvB: (序列 ID 号: 65)

atggcaagttcgggcacaacatcgacgcgtaagcgctttaccggcgcagaatttatcgttcatttcctggaacagcagggcattaagattgtgaca
ggcattccgggcggttctatcctgcctgtttacgatgccttaagccaaagcacgcaaatccgccatattctggcccgtcatgaacagggcgcgggc
tttatcgctcagggaatggcgcgcaccgacggtaaaccggcggtctgtatggcctgtagcggaccgggtgcgactaacctggtgaccgccattgcc
gatgcgcggctggactccatcccgctgatttgcatcactggtcaggttcccgcctcgatgatcggcaccgacgccttccaggaagtggacacctac
ggcatctctatccccatcaccaaacacaactatctggtcagacatatcgaagaactcccgcaggtcatgagcgatgccttccgcattgcgcaatca
ggccgcccaggcccggtgtggatagacattcctaaggatgtgcaaacggcagtttttgagattgaaacacagcccgctatggcagaaaaagccgcc
gccccccgcctttagcgaagaaagcattcgtgacgcagcggcgatgattaacgctgccaaacgcccggtgctttatctgggcggcggtgtgatcaat
gcgcccgcacgggtgcgtgaactggcggagaaagcgcaactgcctaccaccatgactttaatggcgctgggcatgttgccaaaagcgcatccgttg
tcgctgggtatgctggggatgcacggcgtgcgcagcaccaactatattttgcaggaggcggatttgttgatagtgctcggtgcgcgttttgatgac
cgggcgattggcaaaaccgagcagttctgtccgaatgccaaaatcattcatgtcgatatcgaccgtgcagagctgggtaaaatcaagcagccgcac
gtggcgattcaggcggatgttgatgacgtgctggcgcagttgatcccgctggtggaagcgcaaccgcgtgcagagtggcaccagttggtagcggat
ttgcagcgtgagtttccgtgtccaatcccgaaagcgtgcgatccgttaagccattacggcctgatcaacgccgttgccgcctgtgtcgatgacaat
gcaattatcaccaccgacgttggtcagcatcagatgtggaccgcgcaagcttatccgctcaatcgcccacgccagtggctgacctccggtgggctg
ggcacgatgggttttggcctgcctgcggcgattggcgctgcgctggcgaacccggatcgcaaagtgttgtgtttctccggcgacggcagcctgatg
atgaatattcaggagatggcgaccgccagtgaaaatcagctggatgtcaaaatcattctgatgaacaacgaagcgctggggctggtgcatcagcaa
cagagtctgttctacgagcaaggcgtttttgccgccacctatccgggcaaaatcaactttatgcagattgccgccggattcggcctcgaaacctgt
gatttgaataacgaagccgatccgcaggcttcattgcaggaaatcatcaatcgccctggcccggcgctgatccatgtgcgcattgatgccgaagaa
aaagtttacccgatggtgccgccaggtgcggcgaatactgaaatggtgggggaataa

图 34

图 36 *adhE2* (*丙酮丁醇梭菌*): (序列 ID 号: 67)

atgaaagttacaaatcaaaaagaactaaaacaaaagctaaatgaattgagagaagcgcaaaagaagtttgcaacctatactcaagagcaagttgat
aaaatttttaaacaatgtgccatagccgcagctaaagaaagaataaacttagctaaattagcagtagaagaaacaggaaiaggtcttgtagaagat
aaaattataaaaaatcattttgcagcagaatatatatacaataaatataaaaatgaaaaaacttgtggcataatagaccatgacgattctttaggc
ataacaaaggttgctgaaccaattggaattgttgcagccatagttcctactactaatccaacttccacagcaattttcaaatcattaatttcttta
aaaacaagaaacgcaatattcttttcaccacatccacgtgcaaaaaaatctacaattgctgcagcaaaattaattttagatgcagctgttaaagca
ggagcacctaaaaatataataggctggatagatgagccatcaatagaactttctcaagatttgatgagtgaagctgatataatattagcaacagga
ggtccttcaatggttaaagcggcctattcatctggaaaacctgcaattggtgttggagcaggaaatacaccagcaataatagatgagagtgcagat
atagatatggcagtaagctccataattttatcaaagacttatgacaatggagtaatatgcgcttctgaacaatcaatattagttatgaattcaata
tacgaaaaagttaaagaggaatttgtaaaacgaggatcatatatactcaatcaaaatgaaatagctaaaataaaagaaactatgtttaaaaatgga
gctattaatgctgacatagttggaaaatctgcttatataattgctaaaatggcaggaattgaagttcctcaaactacaaagatacttataggcgaa
gtacaatctgttgaaaaaagcgagctgttctcacatgaaaaactatcaccagtacttgcaatgtataaagttaaggattttgatgaagctctaaaa
aaggcacaaaggctaatagaattaggtggaagtggacacacgtcatctttatatatagattcacaaaacaataaggataaagttaaagaatttgga
ttagcaatgaaaacttcaaggacatttattaacatgccttcttcacagggagcaagcggagatttatacaattttgcgatagcaccatcatttact
cttggatgcggcacttggggaggaaactctgtatcgcaaaatgtagagcctaaacatttattaaatattaaaagtgttgctgaaagaagggaaaat
atgctttggtttaaagtgccacaaaaaatatattttaaatatggatgtcttagatttgcattaaaagaattaaaagatatgaataagaaaagagcc
tttatagtaacagataaagatctttttaaacttggatatgttaataaaataacaaaggtactagatgagatagatattaaatacagtatatttaca
gatattaaatctgatccaactattgattcagtaaaaaaaggtgctaaagaaatgcttaactttgaacctgatactataatctctattggtggtgga
tcgccaatggatgcagcaaaggttatgcacttgttatatgaatatccagaagcagaaattgaaaatctagctataaactttatggatataagaaag
agaatatgcaatttccctaaattaggtacaaaggcgatttcagtagctattcctacaactgctggtaccggttcagaggcaacacctttgcagtt
ataactaatgatgaaacaggaatgaaataccctttaacttcttatgaattgaccccaaacatggcaataatagatactgaattaatgttaaatatg
cctagaaaattaacagcagcaactggaatagatgcattagttcatgctatagaagcatatgtttcggttatggctacggattatactgatgaatta
gccttaagagcaataaaaatgatatttaaatatttgcctagagcctataaaaatgggactaacgacattgaagcaagagaaaaaatggcacatgcc
tctaatattgcggggatggcatttgcaaatgctttcttaggtgtatgccattcaatggctcataaacttggggcaatgcatcacgttccacatgga
attgcttgtgctgtattaatagaagaagttattaaatataacgctacagactgtccaacaaagcaaacagcattccctcaatataaatctcctaat
gctaagagaaaatatgctgaaattgcagagtatttgaatttaaagggtactagcgataccgaaaaggtaacagccttaatagaagctatttcaaag
ttaaagatagatttgagtattccacaaaatataagtgccgctggaataaataaaaaagatttttataatacgctagataaaatgtcagagcttgct
tttgatgaccaatgtacaacagctaatcctaggtatccacttataagtgaacttaaggatatctatataaaatcattttaa

Li-*cimA (问号钩端螺旋体)*: (序列 ID 号: 69)

atgacaaaagtagaaactcgattggaaattttagacgtaactttgagagacggggagcagaccagaggggtcagtttttccacttccgaaaaactaaatatcgcaaaatttctattacaaaaactaaatgtagatcgggtagagattgcgtctgcaagagtttctaaaggggaattggaaacggtccaaaaaaatcatggaatgggctgcaacagaacagcttacggaaagaatcgaaatcttaggttttgtagacgggaataaaaccgtagattggatcaaagatagtggggctaaggttttaaatcttttgactaagggatcgcttcatcatttagaaaaacaattaggcaaaactccgaaagaattctttacagacgtttcttttgtaatagaatacgcgatcaaaagcggacttaaaataaacgtatatttagaagattggtccaacggtttcagaaacagtccagattacgtcaaatcgctcgtagaacatctaagtaaagaacatatagaaagaatttttcttccagacacgttaggcgttcttcgccagaagagacgtttcaaggagtggattcactcattcaaaaatacccggatattcattttgaatttcacggacataacgactacgatctttccgtggcaaatagtcttcaagcgattcgtgccggagtcaaaggtcttcacgcttctataaatggtctcggagaaagagccggaaatactccgttggaagcactcgtaaccacgattcatgataagtctaactctaaaacgaacataaacgaaattgcaattacggaagcaagccgtcttgtagaagtattcagcggaaaaagaatttctgcaaatagaccgatcgtaggagaagacgtgtttactcagaccgcgggagtacacgcagacggagacaaaaaaggaaatttatacgcaaatcctattttaccggaaagatttggtaggaaaagaagttacgcgttaggcaaacttgcaggtaaggcgagtatctccgaaaatgtaaaacaactcggaatggttttaagtgaagtggttttacaaaaggttttagaaagggtgatcgaattaggagatcagaataaactagtgacacctgaagatcttccatttatcattgcggacgtttctggaagaaccggagaaaaggtacttacaatcaaatcttgtaatattcattccggaattggaattcgtcctcacgcacaaattgaattggaatatcaggggaaagattcataaggaaatttctgaaggagacggagggtatgatgcgtttatgaatgcacttactaaaattacgaatcgcctcggtattagtattcctaaattgatagattacgaagtaaggattcctcctggtggaaaaacagatgcacttgtagaaactaggatcacctggaacaagtccttagatttagaagaggaccagactttcaaaacgatgggagttcatccggatcaaacggttgcagcggttcatgcaactgaaaagatgctcaatcaaattctacaaccatggcaaatctaa

图 36

Li-*leuC (问号钩端螺旋体)*: (序列 ID 号:71)

atgaagacaatgttcgaaaaaatttgggaagatcatctagtcggagaactagatgctggatcctatctaatctatatagatcgccatctcattcatgaagttacaagtcctcaggcgtttgaaggacttaaacttgcaggcagaaaggttcgtcgtcctgaagctacttttgccacaatggatcataacgtttctactagaacacgtgatttaagtctggccgatcctgtttccgcaattcaaatgcagactttaaaaaaagaactgcgacgaaaacggaatccgcgtttatgattttcaaaaccctgaccaaggaatcattcacgtaatcgctcctgaaatgggactgactcatcctggaatgacaatcgtatgcggagattctcatacttctacacacggtgcgtttggtgcgcttgctttcgggatcggaaccagcgaagtagagcacgttcttgcgactcaaaccttagttcaaaaaagagcaaaaacaatggagattagagtcgatggaaaactttccgataaggtcacagcaaaagacatcattcttgcgatcattggaaaaattggaaccgcaggtgcgacaggttatgtgatcgaatatagaggttctgcaattcaagccctcagtatggaagctagaatgactatttgtaatatgtctatcgaagcgggagctagagcaggtttaatcgcaccagatgaaactacttttaattatattcaaggaaaggacttttctccaaaaggagtcgaatgggatcttgcggtcaaaaaatggaaacactatgtaacggacgaaggtgctaaatttgatagaaccgtaattcttcatgcagatgaaatcgctcctatggtaacttggggaacttctcccagtcaggttgtttcgataaaaggagtcgttccagatccaaaagatgcaaatgatccggtggaaaaaattggaattgagtctgcgcttaaatatatggatctcaaatcgggccagaagatagaagacatttcaattaataaagtgtttatcggttcctgtactaattctagaatcgaagatttaagagcggccgctgctaccgtaaaaggaaaaaaagtttcctctaaggttcaggcgattgtggttcccggttcaggcagagtcaaacgtcaggcggaacaagaaggtctggataaaatttttaccgcggccggttttgaatggagaaatccaggctgttctatgtgtcttgcgatgaacgacgacgtattagaaccgggagatcgttgtgcttctacttctaaccgaaactttgaaggtcgtcaaggaaaaggtggaagaacccatctagtaggaccggaaatggccgccgccgcggctatcgaaggccattttgtggatattcgaaactggaaataa

图 37

Li-*leuD* (*问号钩端螺旋体*): (序列 ID 号: 73)

atgaaacccttactatattaaatggaattgccgccttactggacagacccaacgtggatacggatcagatcattccaaaacaatttttacggaagatagaacgaaccggtttcggagttcatctgtttcacgattggagatacttagacgacgcgggtaccaaactcaatcctgatttttccctcaatcaagaacgatataagggagcttctatccttatcaccagagataactttggttgtggatcttccagagaacacgctccttgggctttagaagactacgggtttagggcaatcattgctccttcttacgcggatatttttttcaacaactgctttaaaaacggaatgcttccagtcattttaaaatcggaagaagtagaagagctgttccatttggtttcgactaacgtaggagcgaaagtcatagtggatctggacaaacaaactgtaaccggaccgactggaaaaatatattatttttgaagtggattcttttcgtaaatactgtctttataacggacttgatgacataggtctaactctaaaacaagaaagtaaaattggagagtttgaaaaaaagcagaaagaagttgaaccttggttatacgccatataa

图 38

Li-*leuB*（*向号钩端螺旋体*）：（序列 ID 号：75）

atgaagaatgtagcagtactttcaggagacggaatcggaccggaagtcatggagatagccatctccgttttgaaaaaggctctcggtgcaaaagtt
tccgagtttcaatttaaagaaggatttgtaggtggaatcgcaatcgataaaactggacacccacttccaccggaaactcttaaactatgtgaagaa
tcttccgcaattcttttcggaagtgtgggaggtcctaaatgggaaacactccctccggaaaaacaaccggaacgaggggcacttctacctttgaga
aaacattttgatctatttgcaaacttaagacctgcgatcatttatccagagttgaaaaatgcttctccagttcgttctgatattattggaaacgga
ttagatattctcatattaagagagttaaccggaggaatttattttggacaaccaaaaggaagagaaggatcaggtcaggaagaatttgcatacgac
acgatgaagtattccagaagagaaatcgaaaggattgctaaagtcgcattccaggcggccagaaaaagaaataataaagtgactagtatcgataaa
gcaaacgtcttgactacttccgttttttggaaggaagtagtaatcgaattgcataagaaagaattttcagacgtccaattgaatcatctttatgtg
gacaatgcggcgatgcagttaatcgtaaatccgaaacaattcgacgtggttctttgtgagaatatgtttggtgatattctttcggacgaggcttcc
atcattacgggttcaatcggaatgcttccttctgcctctctttccgaatctggatttggattgtatgaaccttctggtggttctgcgccggacata
gccggaaaaggagtggcaaatccgattgctcaagtattgagtgcggcgttgatgttacgttattctttttctatggaagaagaagcaaacaagata
gaaaccgccgtgcgtaaaacgattgcctccggaaaaagaaccagagacatagcggaagtaggatctacgatcgtaggaactaaagaaatcggtcaa
ttgatcgaatcctttctctaa

图 39

pheA：（序列 ID 号：77）

atgacatcggaaaacccgttactggcgctgcgagagaaaatcagcgcgctggatgaaaaattattagcgttactggcagaacggcgcgaactggcc
gtcgaggtgggaaaagccaaactgctctcgcatcgcccggtacgtgatattgatcgtgaacgcgatttgctggaaagattaattacgctcggtaaa
gcgcaccatctggacgcccattacattactcgcctgttccagctcatcattgaagattccgtattaactcagcaggctttgctccaacaacatctc
aataaaattaatccgcactcagcacgcatcgcttttctcggccccaaaggttcttattcccatcttgcggcgcgccagtatgctgcccgtcacttt
gagcaattcattgaaagtggctgcgccaaatttgccgatatttttaatcaggtggaaaccggccaggccgactatgccgtcgtaccgattgaaaat
accagctccggtgccataaacgacgtttacgatctgctgcaacataccagcttgtcgattgttggcgagatgacgttaactatcgaccattgtttg
ttggtctccggcactactgatttatccaccatcaatacggtctacagccatccgcagccattccagcaatgcagcaaattccttaatcgttatccg
cactggaagattgaatataccgaaagtacgtctgcggcaatggaaaaggttgcacaggcaaaatcaccgcatgttgctgcgttgggaagcgaagct
ggcggcactttgtacggtttgcaggtactggagcgtattgaagcaaatcagcgacaaaacttcacccgatttgtgggtgttggcgcgtaaagccatt
aacgtgtctgatcaggttccggcgaaaaccacgttgttaatggcgaccgggcaacaagccggtgcgctggttgaagcgttgctggtactgcgcaac
cacaatctgattatgacccgtctggaatcacgcccgattcacggtaatccatgggaagagatgttctatctggatattcaggccaatcttgaatca
gcggaaatgcaaaaagcattgaaagagttaggggaaatcacccgttcaatgaaggtattgggctgttacccaagtgagaacgtagtgcctgttgat
ccaacctga

图 40

TyrA：（序列 ID 号：79）

atggttgctgaattgaccgcattacgcgatcaaattgatgaagtcgataaagcgctgctgaatttattagcgaagcgtctggaactggttgctgaa
gtgggcgaggtgaaaagccgctttggactgcctatttatgttccggagcgcgaggcatctatgttggcctcgcgtcgtgcagaggcggaagctctg
ggtgtaccgccagatctgattgaggatgttttgcgtcgggtgatgcgtgaatcttactccagtgaaaacgacaaaggatttaaaacactttgtccg
tcactgcgtccggtggttatcgtcggcggtggcggtcagatgggacgcctgttcgagaagatgctgacctctcgggttatcaggtgcggattctg
gagcaacatgactgggatcgagcggctgatattgttgccgatgccggaatggtgattgttagtgtgccaatccacgttactgagcaagttattggc
aaattaccgcctttaccgaaagattgtattctggtcgatctggcatcagtgaaaaatgggccattacaggccatgctggtggcgcatgatggtccg
gtgctggggctacacccgatgttcggtccggacagcggtagcctggcaaagcaagttgtggtctggtgtgatggacgtaaaccggaagcataccaa
tggtttctggagcaaattcaggtctggggcgctcggctgcatcgtattagcgccgtcgagcacgatcagaatatggcgtttattcaggcactgcgc
cactttgctacttttgcttacgggctgcacctggcagaagaaaatgttcagcttgagcaacttctggcgctctcttcgccgatttaccgccttgag
ctggcgatggtcgggcgactgtttgctcaggatccgcagctttatgccgacatcattatgtcgtcagagcgtaatctggcgttaatcaaacgttac
tataagcgtttcggcgaggcgattgagttgctggagcagggcgataagcaggcgtttattgacagttccgcaaggtggagcactggttcggcgat
tacgcacagcgttttcagagtgaaagccgcgtgttattgcgtcaggcgaatgacaatcgccagtaa

图 41

alsS: 乙酰乳酸合成酶 *(枯草芽孢杆菌 枯草亚种)* (序列 ID 号: 81)

atgttgacaaaagcaacaaaagaacaaaaatcccttgtgaaaaacagaggggcggagcttgttgttgattgcttagtggagcaaggtgtcacacat
gtatttggcattccaggtgcaaaaattgatgcggtatttgacgctttacaagataaaggacctgaaattatcgttgcccggcacgaacaaaacgca
gcattcatggcccaagcagtcggccgtttaactggaaaaccgggagtcgtgttagtcacatcaggaccgggtgcctctaacttggcaacaggcctg
ctgacagcgaacactgaaggagaccctgtcgttgcgcttgctggaaacgtgatccgtgcagatcgtttaaaacggacacatcaatctttggataat
gcggcgctattccagccgattacaaaatacagtgtagaagttcaagatgtaaaaaatataccggaagctgttacaaatgcatttaggatagcgtca
gcagggcaggctggggccgcttttgtgagctttccgcaagatgttgtgaatgaagtcacaaatacgaaaaacgtgcgtgctgttgcagcgccaaaa
ctcggtcctgcagcagatgatgcaatcagtgcggccatagcaaaaatccaaacagcaaaacttcctgtcgttttggtcggcatgaaaggcggaaga
ccggaagcaattaaagcggttcgcaagctttgaaaaaggttcagcttccatttgttgaaacatatcaagctgccggtacccttctagagattta
gaggatcaatattttggccgtatcggtttgttccgcaaccagcctggcgatttactgctagagcaggcagatgttgttctgacgatcggctatgac
ccgattgaatatgatccgaaattctggaatatcaatggagaccggacaattatccatttagacgagattatcgctgacattgatcatgcttaccag
cctgatcttgaattgatcggtgacattccgtccacgatcaatcatatcgaacacgatgctgtgaaagtggaatttgcagagcgtgagcagaaaatc
ctttctgatttaaaacaatatatgcatgaaggtgagcaggtgcctgcagattggaaatcagacagagcgcaccctcttgaaatcgttaaagagttg
cgtaatgcagtcgatgatcatgttacagtaacttgcgatatcggttcgcacgccatttggatgtcacgttatttccgcagctacgagccgttaaca
ttaatgatcagtaacggtatgcaaacactcggcgttgcgcttccttgggcaatcggcgcttcattggtgaaaccgggagaaaaagtggtttctgtc
tctggtgacggcggtttcttattctcagcaatggaattagagacagcagttcgactaaaagcaccaattgtacacattgtatggaacgacagcaca
tatgacatggttgcattccagcaattgaaaaaatataaccgtacatctgcggtcgatttcggaaatatcgatatcgtgaaatatgcggaaagcttc
ggagcaactggcttgcgcgtagaatcaccagaccagctggcagatgttctgcgtcaaggcatgaacgctgaaggtcctgtcatcatcgatgtcccg
gttgactacagtgataacattaatttagcaagtgacaagcttccgaaagaattcggggaactcatgaaaacgaaagctctctag

图 42

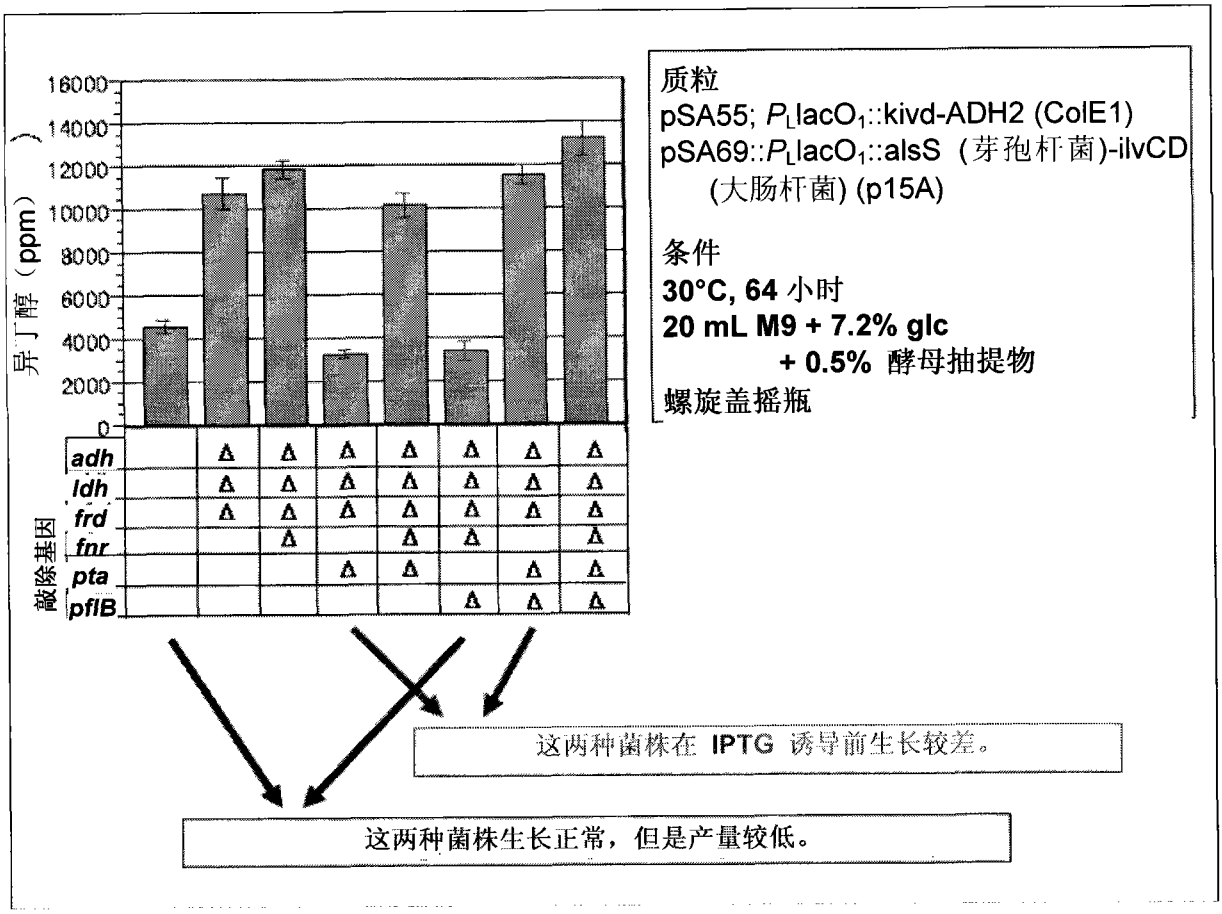
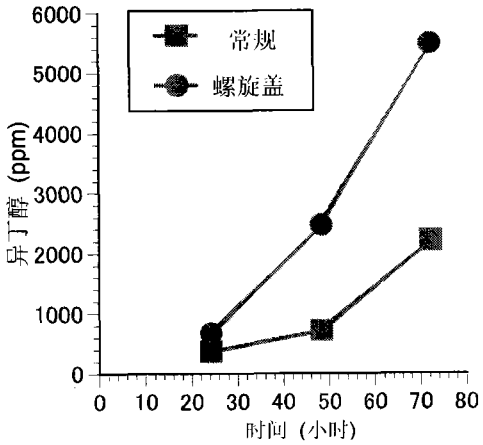
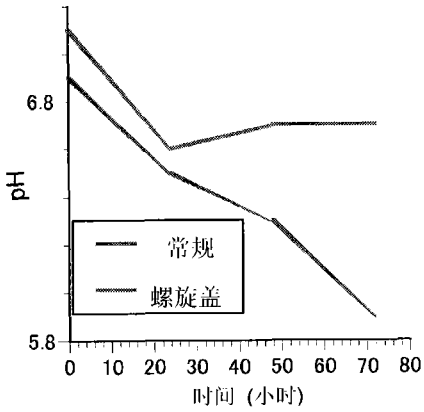
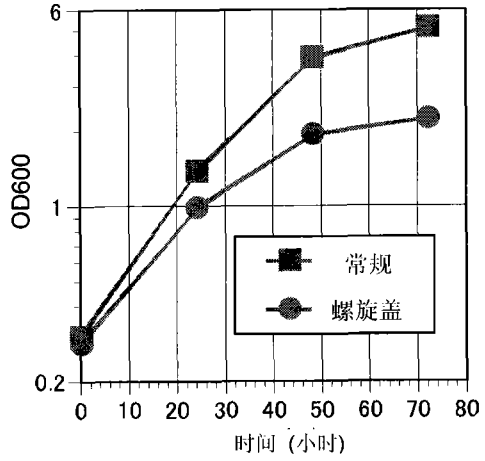


图 43







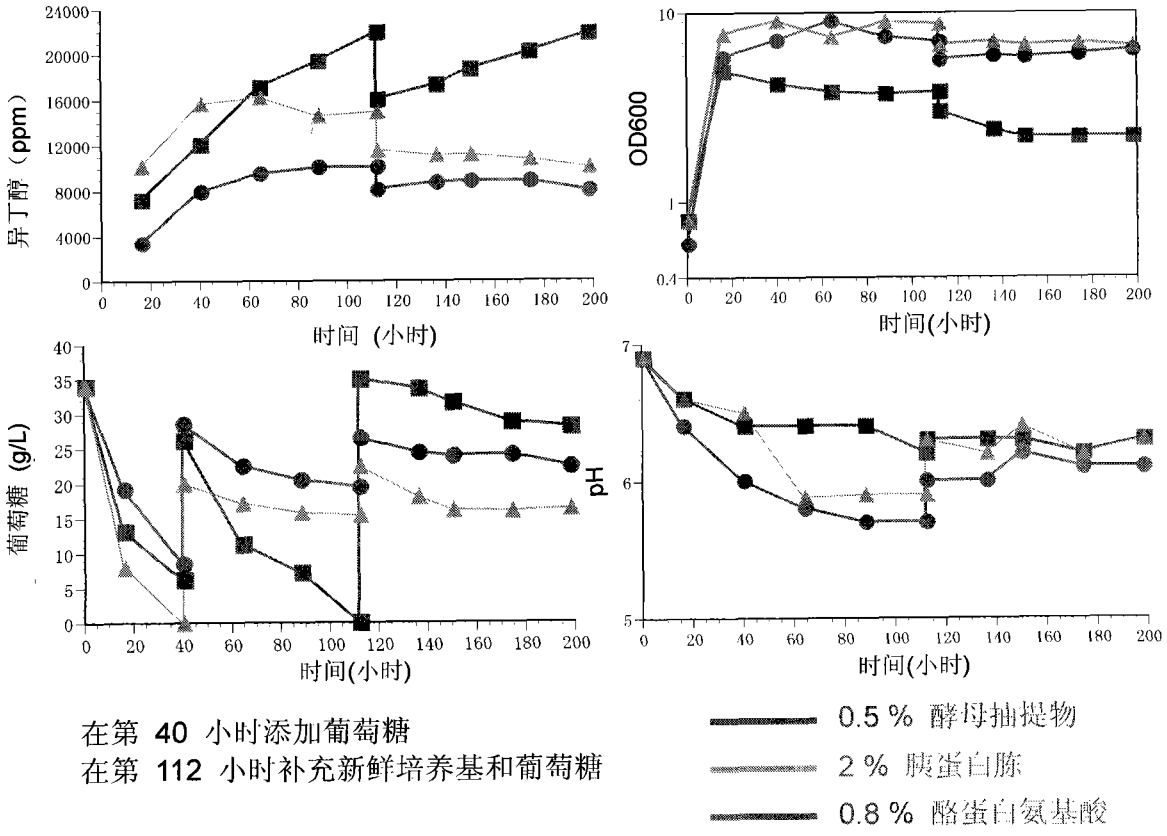
pSA55; $P_L lacO_1$::kivd-ADH2 (ColE1)
pSA69:: *P*LlacO1::alsS (芽孢杆菌)-ilvCD (*大肠杆菌*) (p15A)
JCL260:: JCL88+Δ*pflB*

SA237:: JCL260+pSA55&69

30°C，20 ml M9+3.6% glc

常规摇瓶
螺旋盖摇瓶

图 44



菌株

SA237:: JCL260+pSA55&69

pSA55; $P_L lacO_1$::kivd-ADH2 (ColE1)

pSA69:: $P_L$lacO1::alsS (芽孢杆菌)-ilvCD (*大肠杆菌*) (p15A)

JCL260:: JCL88+Δ*pflB*

条件

30°C, 20 ml M9+3.6% glc

图 45

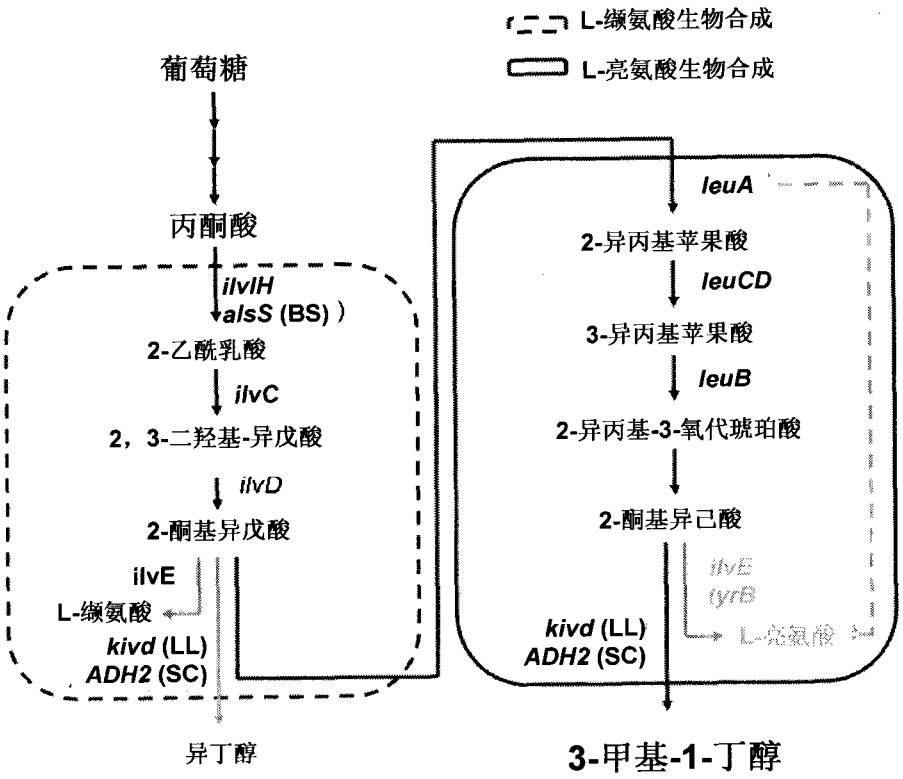
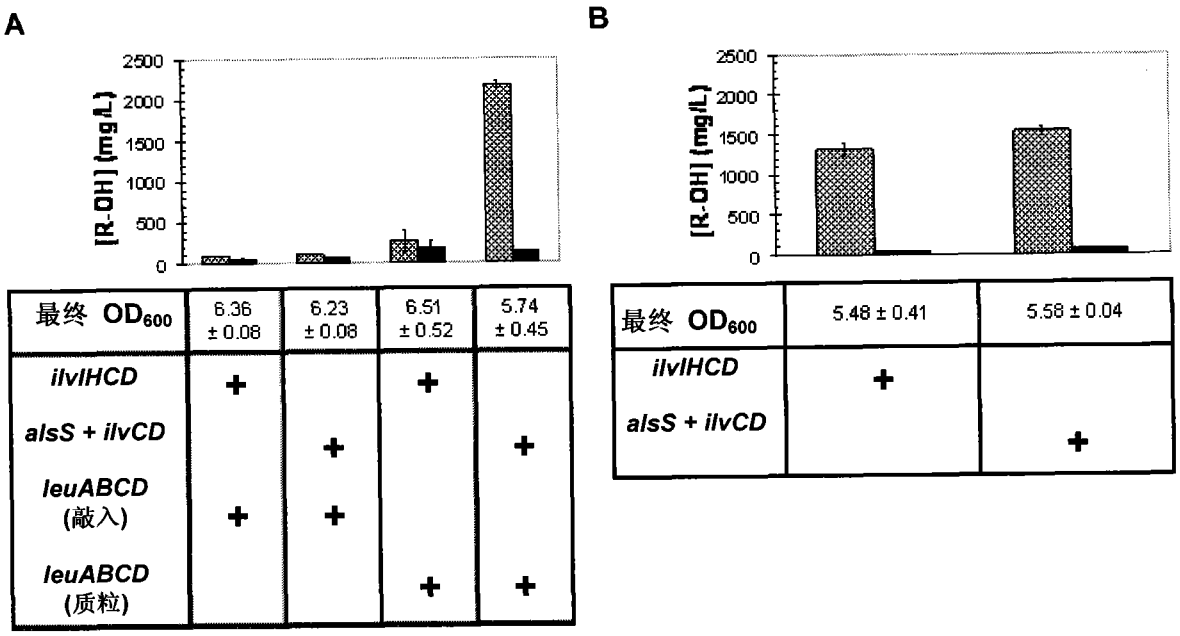


图 46

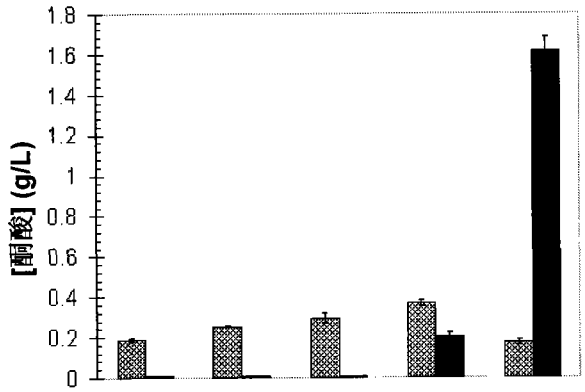


最终 OD_{600}	6.36 ± 0.08	6.23 ± 0.08	6.51 ± 0.52	5.74 ± 0.45
ilvIHCD	+		+	
alsS + ilvCD		+		+
leuABCD (敲入)	+	+		
leuABCD (质粒)			+	+

最终 OD_{600}	5.48 ± 0.41	5.58 ± 0.04
ilvIHCD	+	
alsS + ilvCD		+

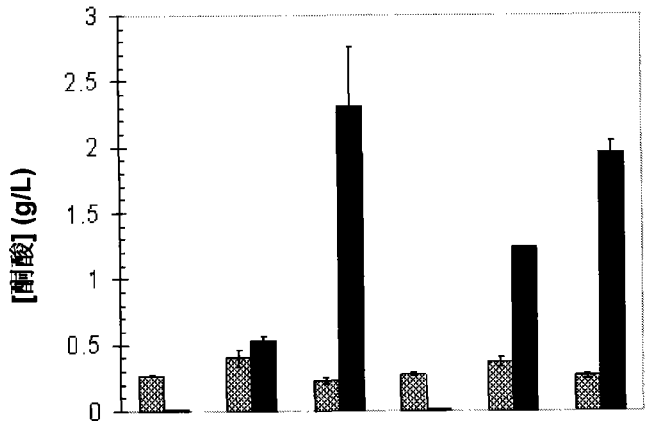
图 47

(A)



最终 OD_{600}	5.51 ± 0.65	4.53 ± 0.06	6.51 ± 0.62	6.32 ± 0.19	6.04 ± 0.20
O_2	高	低	高	高	高
alsS + ilvCD	+	+	+	+	+
WT IPMS		+	+	+	
IPMS *(G462D)*					+
天然 RBS	+	+	+		
合成 RBS				+	+

(B)



最终 OD_{600}	5.24 ± 0.13	6.32 ± 0.42	6.00 ± 0.30	6.37 ± 0.59	6.11 ± 0.11	6.29 ± 0.83
WT IPMS	+	+		+	+	
IPMS *(G462D)*			+			+
天然 RBS	+			+		
合成 RBS		+	+		+	+
ilvE	Δ	Δ	Δ	Δ	Δ	Δ
tyrB				Δ	Δ	Δ

图 48

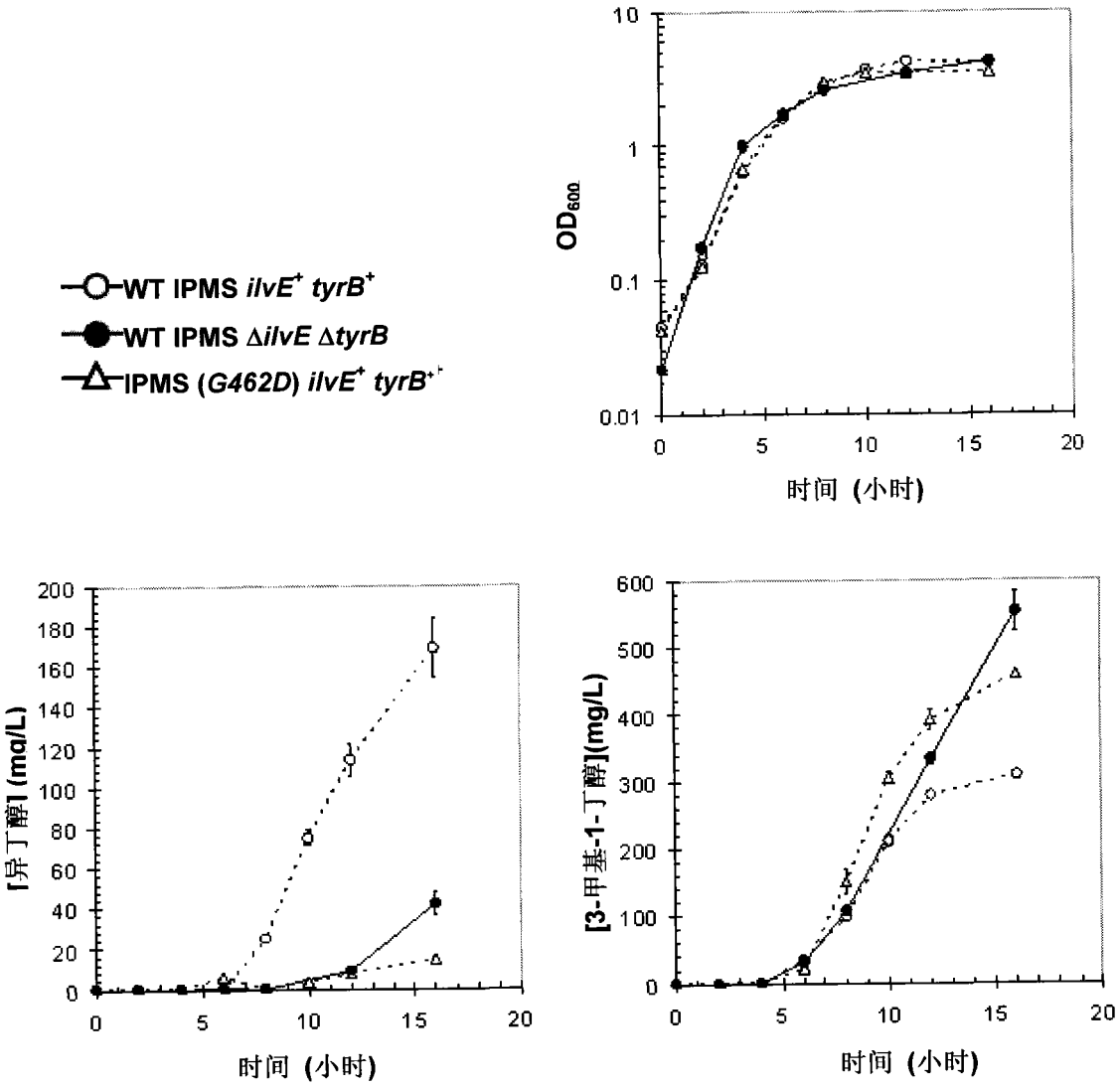

WT IPMS ilvE+ tyrB+
WT IPMS ΔilvE ΔtyrB
IPMS (G462D) ilvE+ tyrB+
OD600
时间 (小时)
[异丁醇] (mg/L)
时间 (小时)
[3-甲基-1-丁醇](mg/L)
时间 (小时)

图 49

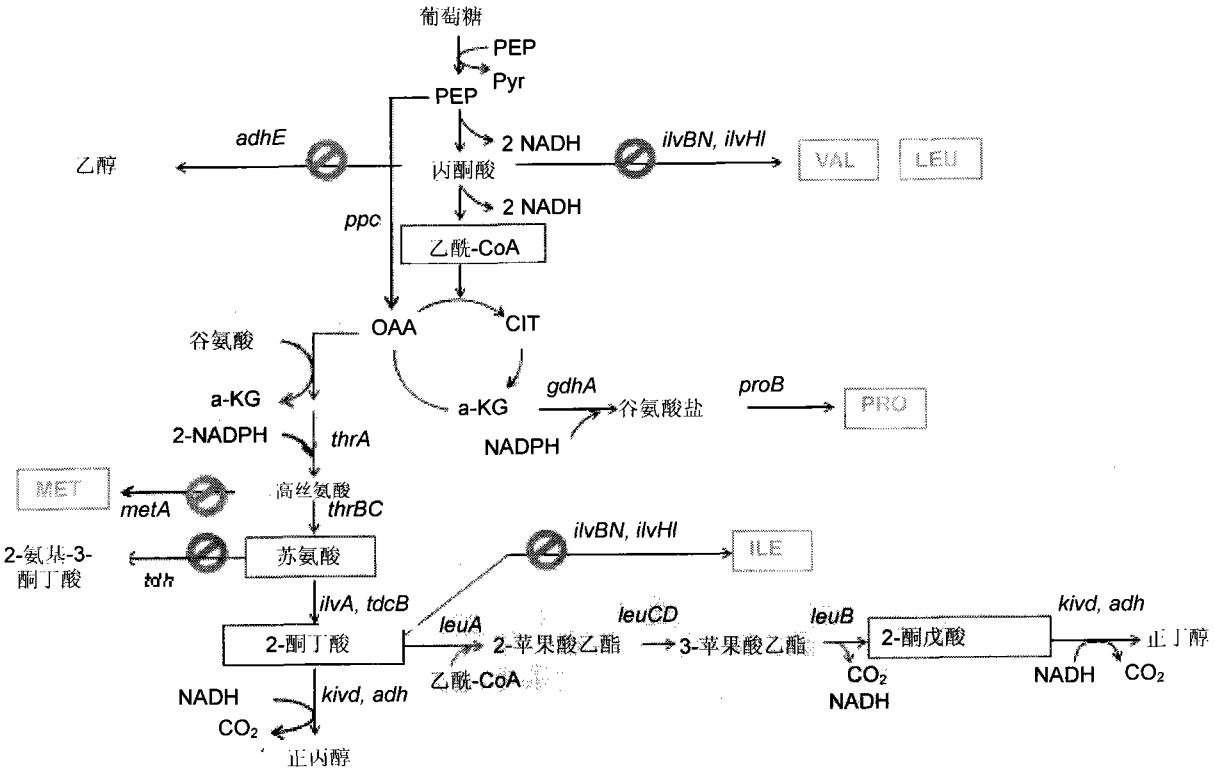

葡萄糖
PEP
Pyr
PEP
adhE
乙醇
2 NADH
丙酮酸
ilvBN, ilvHI
VAL
LEU
2 NADH
ppc
乙酰-CoA
OAA
CIT
谷氨酸
a-KG
gdhA
谷氨酸盐
proB
PRO
a-KG
NADPH
2-NADPH
thrA
MET
metA
高丝氨酸
thrBC
2-氨基-3-酮丁酸
tdh
苏氨酸
ilvBN, ilvHI
ILE
ilvA, tdcB
leuA
leuCD
leuB
kivd, adh
2-酮丁酸
2-苹果酸乙酯
3-苹果酸乙酯
2-酮戊酸
正丁醇
乙酰-CoA
CO2
NADH
NADH
kivd, adh
CO2
NADH
CO2
正丙醇

图 50

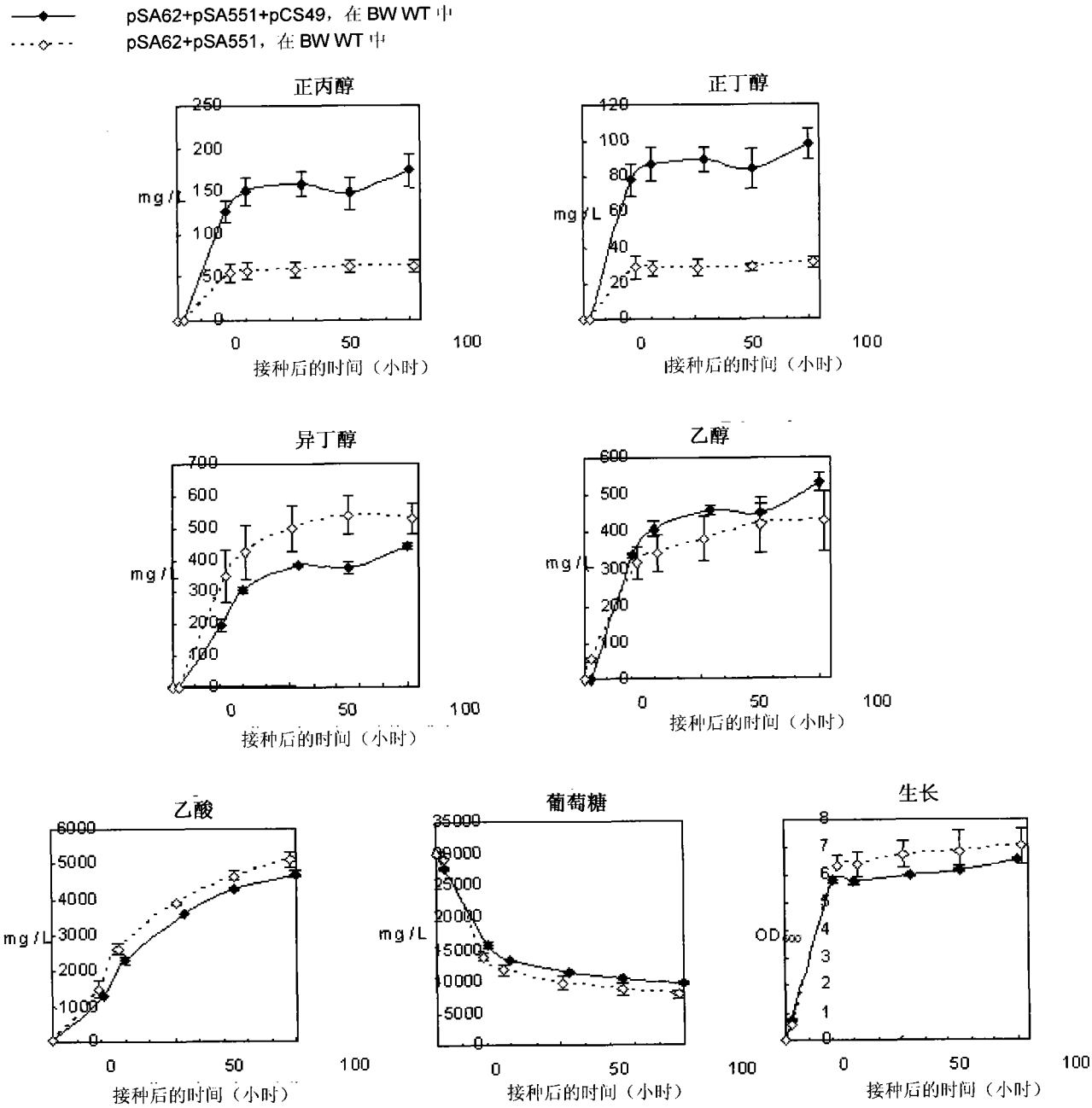

pSA62+pSA551+pCS49，在 BW WT 中
pSA62+pSA551，在 BW WT 中
正丙醇
mg/L
接种后的时间（小时）
正丁醇
mg/L
接种后的时间（小时）
异丁醇
mg/L
接种后的时间（小时）
乙醇
mg/L
接种后的时间（小时）
乙酸
mg/L
接种后的时间（小时）
葡萄糖
mg/L
接种后的时间（小时）
生长
OD
接种后的时间（小时）

图 51

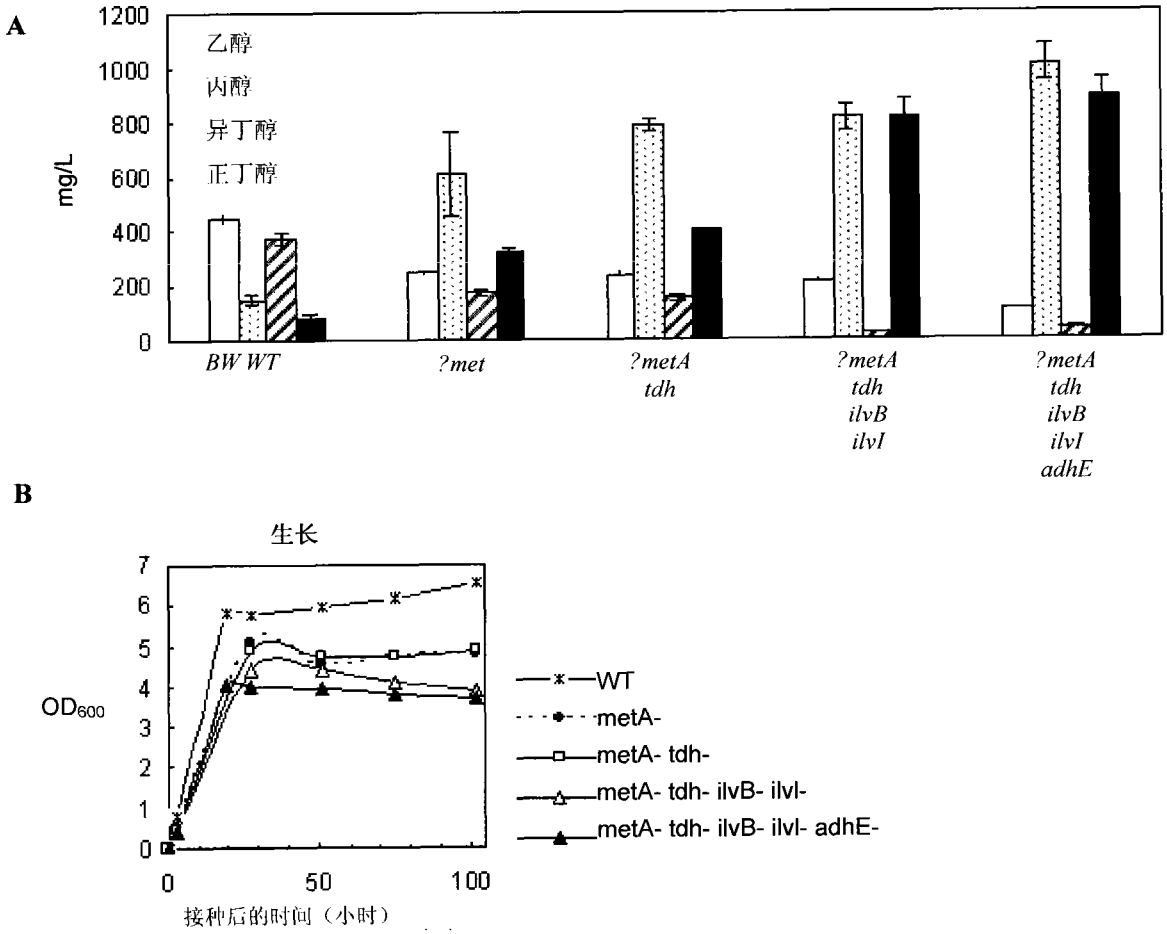

A
1200
1000
800
600
400
200
0
mg/L
乙醇
丙醇
异丁醇
正丁醇
BW WT
?met
?metA tdh
?metA tdh ilvB ilvI
?metA tdh ilvB ilvI adhE
B
生长
OD600
7
6
5
4
3
2
1
0
0
50
100
接种后的时间（小时）
WT
metA-
metA- tdh-
metA- tdh- ilvB- ilvI-
metA- tdh- ilvB- ilvI- adhE-

图 52

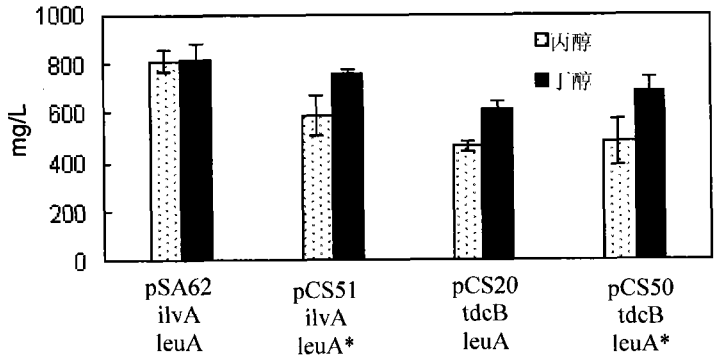

1000
800
600
400
200
0
mg/L
丙醇
丁醇
pSA62 ilvA leuA
pCS51 ilvA leuA*
pCS20 tdcB leuA
pCS50 tdcB leuA*

图 53

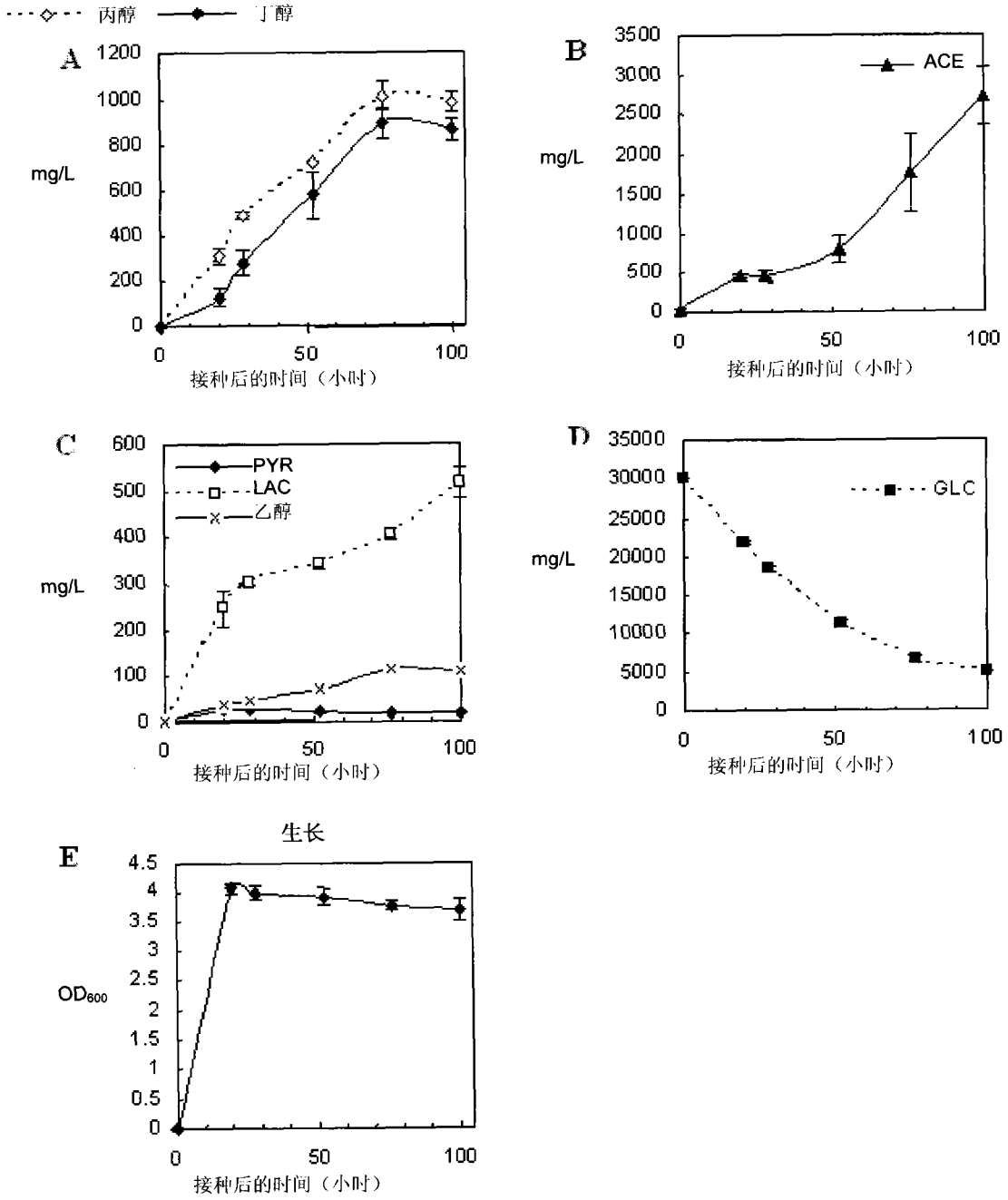

丙醇
丁醇
A
mg/L
接种后的时间（小时）
B
ACE
mg/L
接种后的时间（小时）
C
PYR
LAC
乙醇
mg/L
接种后的时间（小时）
D
GLC
mg/L
接种后的时间（小时）
生长
E
OD600
接种后的时间（小时）

图 54