



(12) 发明专利申请

(10) 申请公布号 CN 105008514 A

(43) 申请公布日 2015. 10. 28

(21) 申请号 201480012326. 4

(51) Int. Cl.

(22) 申请日 2014. 03. 06

C12N 1/21(2006. 01)

C12N 15/52(2006. 01)

(30) 优先权数据

61/774, 342 2013. 03. 07 US

61/928, 333 2014. 01. 16 US

(85) PCT国际申请进入国家阶段日

2015. 09. 06

(86) PCT国际申请的申请数据

PCT/US2014/021258 2014. 03. 06

(87) PCT国际申请的公布数据

W02014/138419 EN 2014. 09. 12

(71) 申请人 凯利斯塔公司

地址 美国加利福尼亚州

(72) 发明人 E·伦纳德 J·米舒尔 J·内斯

T·J·珀塞尔

(74) 专利代理机构 北京市铸成律师事务所

11313

代理人 孟锐

权利要求书4页 说明书43页

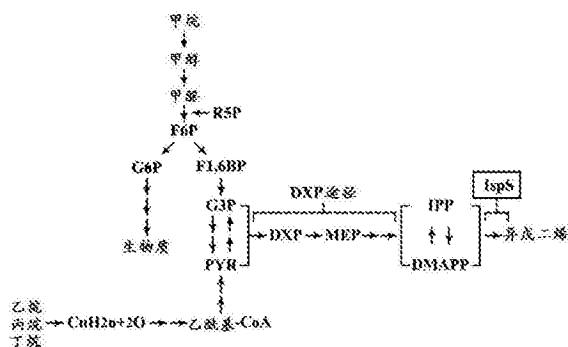
序列表25页 附图6页

(54) 发明名称

用于生物生产异戊二烯的组合物和方法

(57) 摘要

本公开提供使用利用碳原料诸如甲烷或者天然气的嗜甲烷细菌生物生产异戊二烯的组合物和方法。



1. 一种非天然存在的嗜甲烷细菌,其包含编码异戊二烯合成酶的外源性核酸,其中所述嗜甲烷细菌能够将碳原料转化成异戊二烯。

2. 根据权利要求 1 所述的非天然存在的嗜甲烷细菌,其中编码所述异戊二烯合成酶的所述核酸源于银白杨、毛果杨、美洲山杨、黑杨、银白杨 x 欧洲山杨、银灰杨、山葛、葛根、夏栎、蝶形花亚科、褪色柳、无毛柳、五蕊柳或百里香叶柳。

3. 根据权利要求 1 或 2 所述的非天然存在的嗜甲烷细菌,其中编码所述异戊二烯合成酶的所述外源性核酸序列是为在所述嗜甲烷细菌中表达而优化的密码子。

4. 根据权利要求 1 所述的非天然存在的嗜甲烷细菌,其中所述异戊二烯合成酶不包含 N- 末端质体-靶向序列。

5. 根据权利要求 2 所述的非天然存在的嗜甲烷细菌,其中所述核酸编码 SEQ ID NO:1-6 中任一项阐述的氨基酸序列。

6. 根据权利要求 2 所述的非天然存在的嗜甲烷细菌,其中所述核酸包含如 SEQ ID NO:14-19 中任一项阐述的序列。

7. 根据权利要求 1 所述的非天然存在的嗜甲烷细菌,其中所述编码异戊二烯合成酶的外源性核酸与表达控制序列可操作地连接。

8. 根据权利要求 7 所述的非天然存在的嗜甲烷细菌,其中所述表达控制序列是选自由甲醇脱氢酶启动子、己酮糖 6- 磷酸合成酶启动子、核糖体蛋白 S16 启动子、丝氨酸羟甲基转移酶启动子、丝氨酸乙醛酸氨基转移酶启动子、磷酸烯醇丙酮酸羧化酶启动子、T5 启动子和 Trc 启动子组成的组中的启动子。

9. 根据权利要求 1 所述的非天然存在的嗜甲烷细菌,其中所述嗜甲烷细菌与亲代嗜甲烷细菌的内源性 DXP 途径酶的正常表达水平相比能够过表达内源性 DXP 途径酶,用编码 DXP 途径酶的外源性核酸进行转化并能够表达所述编码 DXP 途径酶的外源性核酸,或它们的组合。

10. 根据权利要求 1 所述的非天然存在的嗜甲烷细菌,其中所述嗜甲烷细菌与亲代嗜甲烷细菌的内源性甲羟戊酸途径酶的正常表达水平相比能够过表达内源性甲羟戊酸途径酶,用编码甲羟戊酸途径酶的外源性核酸进行转化并能够表达所述编码甲羟戊酸途径酶的外源性核酸,或它们的组合。

11. 根据权利要求 9 所述的非天然存在的嗜甲烷细菌,其中所述 DXP 途径酶是 DXS、DXR、IDI、IspD、IspE、IspF、IspG、IspH 或它们的组合。

12. 根据权利要求 10 所述的非天然存在的嗜甲烷细菌,其中所述甲羟戊酸途径酶是乙酰乙酰辅酶 A 硫解酶、3- 羟基-3- 甲基戊二酰辅酶 A 合酶、3- 羟基-3- 甲基戊二酰辅酶 A 还原酶、甲羟戊酸激酶、磷酸甲羟戊酸激酶、甲羟戊酸焦磷酸脱羧酶、异戊烯基二磷酸异构酶或它们的组合。

13. 根据前述权利要求中任一项所述的非天然存在的嗜甲烷细菌,其中所述嗜甲烷细菌用至少一种编码变异体 DXP 途径酶的外源性核酸进行转化。

14. 根据权利要求 13 所述的非天然存在的嗜甲烷细菌,其中所述至少一种变异体 DXP 途径酶包括两种变异体 DXP 途径酶,其中所述两种变异体 DXP 途径酶包括突变型丙酮酸脱氢酶 (PDH) 和突变型 3,4 二羟基-2- 丁酮 4- 磷酸合成酶 (DHBPS)。

15. 根据前述权利要求中任一项所述的非天然存在的嗜甲烷细菌,其中所述嗜甲烷细

菌能够生产约 1mg/L 至约 500g/L 的异戊二烯。

16. 根据权利要求 1-15 中任一项所述的非天然存在的嗜甲烷细菌,其中所述嗜甲烷细菌是甲基单胞菌属、甲基杆菌属、甲基球菌属、甲基弯曲菌属、甲基孢囊菌属、甲基微菌属、甲烷单胞菌属、甲基细胞菌属、甲基帽菌属。

17. 根据权利要求 1-16 中任一项所述的非天然存在的嗜甲烷细菌,其中所述嗜甲烷细菌是荚膜甲基球菌 Bath 菌株、甲烷甲基单胞菌 16a(ATCC PTA 2402)、发泡甲基弯曲菌 OB3b(NRRL B-11, 196)、生孢甲基弯曲菌(NRRL B-11, 197)、小甲基孢囊菌(NRRL B-11, 198)、甲烷甲基单胞菌(NRRL B-11, 199)、白色甲基单胞(NRRL B-11, 200)、荚膜甲基杆菌(NRRL B-11, 201)、嗜有机物甲基杆菌(ATCC 27, 886)、甲基单胞菌 sp AJ-3670(FERM P-2400)、*Methylocella silvestris*、沼泽甲基细胞菌(ATCC 700799)、冻原甲基细胞菌、*Methylocystis daltona* 菌株 SB2、*Methylocystis bryophila*、金黄甲基帽菌 KYG、极端嗜酸甲烷氧化菌、*Methylacidiphilum fumariolicum*、*Methyloacida kamchatkensis*、甲基叔丁基醚降解菌或嗜碱甲基微菌。

18. 根据权利要求 1-17 中任一项所述的非天然存在的嗜甲烷细菌,其中所述碳原料是甲烷、甲醇、天然气或非常规天然气。

19. 一种生产异戊二烯的方法,其包括在碳原料存在下在足以生产异戊二烯的条件下培养包含编码异戊二烯合成酶的外源性核酸的非天然存在的嗜甲烷细菌。

20. 根据权利要求 19 所述的方法,其中编码所述异戊二烯合成酶的所述核酸源于银白杨、毛果杨、美洲山杨、黑杨、银白杨 x 欧洲山杨、银灰杨、山葛、葛根、夏栎、蝶形花亚科、褪色柳、无毛柳、五蕊柳或百里香叶柳。

21. 根据权利要求 19 或 20 所述的方法,其中编码所述异戊二烯合成酶的所述外源性核酸序列是为在所述宿主嗜甲烷细菌中表达而优化的密码子。

22. 根据权利要求 19-21 中任一项所述的方法,其中编码所述异戊二烯合成酶的所述外源性核酸不包含 N-末端质体-靶向序列。

23. 根据权利要求 19-22 中任一项所述的方法,其中所述外源性核酸编码 SEQ ID NO:1-6 中任一项阐述的氨基酸序列。

24. 根据权利要求 19-23 中任一项所述的方法,其中所述外源性核酸包含如 SEQ ID NO:14-19 中任一项阐述的序列。

25. 根据权利要求 19-24 中任一项所述的方法,其中所述编码异戊二烯合成酶的外源性核酸与一种表达控制序列可操作地连接。

26. 根据权利要求 25 所述的方法,其中所述表达控制序列是选自由甲醇脱氢酶启动子、己酮糖 6-磷酸合成酶启动子、核糖体蛋白 S16 启动子、T5 启动子和 Trc 启动子组成的组中的启动子。

27. 根据权利要求 19 所述的方法,其中所述嗜甲烷细菌与亲代嗜甲烷细菌的内源性 DXP 途径酶的正常表达水平相比能够过表达内源性 DXP 途径酶,用编码 DXP 途径酶的外源性核酸进行转化并能够表达所述编码 DXP 途径酶的外源性核酸,或它们的组合。

28. 根据权利要求 27 所述的方法,其中所述 DXP 途径酶是 DXS、DXR、IDI、IspD、IspE、IspF、IspG、IspH 或它们的组合。

29. 根据权利要求 19-28 中任一项所述的方法,其中所述嗜甲烷细菌用至少一种编码

变异体 DXP 途径酶的外源性核酸进行转化。

30. 根据权利要求 29 所述的方法,其中所述至少一种变异体 DXP 途径酶包括两种变异体 DXP 途径酶,其中所述两种变异体 DXP 途径酶包括突变型丙酮酸脱氢酶 (PDH) 和突变型 3,4 二羟基-2-丁酮 4-磷酸合成酶 (DHBPS)。

31. 根据权利要求 19-30 中任一项所述的方法,其中所述嗜甲烷细菌是甲基单胞菌属、甲基杆菌属、甲基球菌属、甲基弯曲菌属、甲基孢囊菌属、甲基微菌属、甲烷单胞菌属、甲基细胞菌属、甲基帽菌属。

32. 根据权利要求 31 所述的方法,其中所述嗜甲烷细菌是荚膜甲基球菌 Bath 菌株、甲烷甲基单胞菌 16a(ATCC PTA 2402)、发泡甲基弯曲菌 OB3b(NRRL B-11, 196)、生孢甲基弯曲菌(NRRL B-11, 197)、小甲基孢囊菌(NRRL B-11, 198)、甲烷甲基单胞菌(NRRL B-11, 199)、白色甲基单胞菌(NRRL B-11, 200)、荚膜甲基杆菌(NRRL B-11, 201)、嗜有机物甲基杆菌(ATCC 27, 886)、甲基单胞菌 sp AJ-3670(FERM P-2400)、Methylocella silvestris、沼泽甲基细胞菌(ATCC 700799)、冻原甲基细胞菌、Methylocystis daltona 菌株 SB2、Methylocystis bryophila、金黄甲基帽菌 KYG、极端嗜酸甲烷氧化菌、Methylacidiphilum fumariolicum、Methyloacida kamchatkensis、甲基叔丁基醚降解菌或嗜碱甲基微菌。

33. 根据权利要求 19-32 中任一项所述的方法,其中所述碳原料是甲烷、甲醇、天然气或非常规天然气。

34. 根据权利要求 19-33 中任一项所述的方法,其中所述嗜甲烷细菌能够生产约 1g/L 至约 500g/L 的异戊二烯。

35. 根据权利要求 19-34 中任一项所述的方法,其中所述嗜甲烷细菌通过发酵进行培养并且将从所述发酵生产的异戊二烯回收作为废气。

36. 根据权利要求 35 所述的方法,其中将所述回收的异戊二烯进一步修饰成二聚体(10 个碳)烃、三聚体(15 个碳)烃,或它们的组合。

37. 根据权利要求 36 所述的方法,其中将所述二聚体烃、三聚体烃或它们的组合氢化成为长链支链烷烃。

38. 根据权利要求 35 所述的方法,其中将所述回收的异戊二烯进一步修饰成类异戊二烯产品。

39. 一种筛选突变型嗜甲烷细菌的方法,其包括:

- (a) 使所述嗜甲烷细菌暴露于诱变剂以生产突变型嗜甲烷细菌;
- (b) 用编码香叶基香叶基二磷酸合成酶 (GGPPS)、八氢番茄红素合成酶 (CRTB) 和八氢番茄红素脱氢酶 (CRTI) 的外源性核酸转化所述突变型嗜甲烷细菌;和
- (c) 在足够生长的条件下培养所述转化的突变型嗜甲烷细菌;

其中与包含编码 GGPPS、CRTB 和 CRTI 的外源性核酸的参考嗜甲烷细菌相比呈现出增强的红色着色的突变型嗜甲烷细菌表明异戊二烯前体合成的增高。

40. 根据权利要求 39 所述的方法,其中所述诱变剂是辐射物、化学品、质粒或转座子。

41. 根据权利要求 39 所述的方法,其中步骤 (b) 在步骤 (a) 之前实施,然后在足够生长的条件下培养所述转化的突变型嗜甲烷细菌。

42. 根据权利要求 39-41 中任一项所述的方法,其中具有增强的红色着色的所述突变型嗜甲烷细菌或它们的克隆细胞用编码异戊二烯合成酶的外源性核酸进行转化。

43. 根据权利要求 42 所述的方法, 其中将编码 GGPPS、CRTB 和 CRTI 的所述核酸中的至少一种从所述突变型嗜甲烷细菌中除去或灭活。

44. 一种在嗜甲烷细菌中筛选异戊二烯途径基因的方法, 其包括:

(a) 用以下转化所述嗜甲烷细菌:

(i) 至少一种编码异戊二烯途径酶的外源性核酸;

(ii) 编码香叶基香叶基二磷酸合成酶 (GGPPS)、八氢番茄红素合成酶 (CRTB) 和八氢番茄红素脱氢酶 (CRTI) 的外源性核酸; 和

(b) 在足够生长的条件下培养来自步骤 (a) 的所述嗜甲烷细菌;

其中与包含编码 GGPPS、CRTB 和 CRTI 的外源性核酸并且不包含编码异戊二烯途径酶的外源性核酸的参考嗜甲烷细菌相比, 呈现出增强的红色着色的转化的嗜甲烷细菌表明异戊二烯前体合成的增高。

45. 根据权利要求 44 所述的方法, 其中所述异戊二烯途径酶是 DXP 途径酶或甲羟戊酸途径酶。

46. 根据权利要求 44 或 45 所述的方法, 其中所述至少一种编码异戊二烯途径酶的外源性核酸是异源性核酸。

47. 根据权利要求 44-46 中任一项所述的方法, 其中所述至少一种编码异戊二烯途径酶的外源性核酸是为在所述宿主嗜甲烷细菌中表达而优化的密码子。

48. 根据权利要求 44 或 45 所述的方法, 其中所述至少一种编码异戊二烯途径酶的外源性核酸是同源性核酸。

49. 根据权利要求 48 所述的方法, 其中所述同源性核酸在所述嗜甲烷细菌中过表达。

50. 根据权利要求 44-49 中任一项所述的方法, 其中所述至少一种编码异戊二烯途径酶的外源性核酸是非天然存在的变异体。

51. 根据权利要求 50 所述的方法, 其中所述非天然存在的变异体通过随机诱变、定向诱变或它们的组合产生。

52. 根据权利要求 49 所述的方法, 其中所述非天然存在的变异体被合成。

53. 根据权利要求 49 所述的方法, 其中所述非天然存在的变异体与编码异戊二烯途径酶的参考核酸相比包括至少一个氨基酸取代。

54. 一种异戊二烯组合物, 其中所述异戊二烯具有小于约 -30‰的 $\delta^{13}\text{C}$ 分布。

55. 根据权利要求 53 所述的异戊二烯组合物, 其中所述异戊二烯具有小于约 -40‰的 $\delta^{13}\text{C}$ 分布。

56. 根据权利要求 53 所述的异戊二烯组合物, 其中所述异戊二烯具有小于约 -50‰的 $\delta^{13}\text{C}$ 分布。

57. 根据权利要求 53 所述的异戊二烯组合物, 其中所述异戊二烯具有约 -30‰至约 -50‰范围中的 $\delta^{13}\text{C}$ 分布。

58. 根据权利要求 53 所述的异戊二烯组合物, 其中所述异戊二烯具有约 -30‰至约 -40‰范围中的 $\delta^{13}\text{C}$ 分布。

59. 根据权利要求 53 所述的异戊二烯组合物, 其中所述异戊二烯具有约 -40‰至约 -50‰范围中的 $\delta^{13}\text{C}$ 分布。

用于生物生产异戊二烯的组合物和方法

[0001] 关于序列表的说明

[0002] 以文本格式替代纸印本提供与本申请相关的序列表,并在此通过引用并入本说明书中。包括该序列表的文本文件的名称是 200206_411WO_SEQUENCE_LISTING.txt。该文本文件 58.5KB,创建于 2014 年 3 月 4 日,并通过 EFS-Web 电子提交。

[0003] 背景

技术领域

[0004] 本公开提供用于生物生产异戊二烯的组合物和方法,并且更具体地,使用嗜甲烷细菌从碳底物诸如甲烷或天然气生产异戊二烯。

[0005] 相关技术的描述

[0006] 异戊二烯也称为 2-甲基-1,3-丁二烯,是挥发性 5-碳烃。异戊二烯由各种生物体,包括微生物、植物和动物物种产生 (Kuzuyama, 2002, Biosci Biotechnol. Biochem. 66:1619-1627)。有两种异戊二烯生物合成的途径:甲羟戊酸 (MVA) 途径和非-甲羟戊酸 (或甲羟戊酸-非依赖性) 途径,也称为 1-脱氧-D-木酮糖-5-磷酸 (DXP) 途径。MVA 途径存在于真核细胞、古细菌和高等植物的细胞溶质中 (Kuzuyama, 2002, Biosci. Biotechnol. Biochem. 66:1619-1627)。DXP 途径发现于大多数细菌、绿藻以及高等植物的叶绿体中 (Kuzuyama, 2002, Biosci Biotechnol. Biochem. 66:1619-1627)。

[0007] 异戊二烯是一种重要的平台化学品,用于生产在轮胎和橡胶工业中使用的聚异戊二烯;在鞋类、医疗用品、胶乳、体育用品中使用的弹性体;粘合剂;以及医药用的类异戊二烯。异戊二烯还可以用作替代燃料。利用催化剂可以将异戊二烯化学修饰成二聚体 (10 个碳) 和三聚体 (15 个碳) 烃以制备烯烃 (Clement 等, 2008, Chem. Eur. J. 14:7408-7420; Gordillo 等, 2009, Adv. Synth. Catal. 351:325-330)。在被氢化以制备长链支链烷烃后这些分子可以适合用作柴油或喷气燃料替代品。

[0008] 目前,异戊二烯的工业用途受限于它的紧缺供应。大多数合成橡胶是基于丁二烯聚合物,其毒性远远高于异戊二烯。天然橡胶从橡胶树或来自中南美洲和非洲雨林的植物获得。异戊二烯还可以从石油最普遍地通过裂解存在于精炼石油的石脑油部分中的烃进行制备。制备一加仑基于化石的异戊二烯需要约七加仑的原油。自天然生产的生物体的异戊二烯产率没有商业吸引力。

[0009] 已经作出越来越多的努力以使得能够或增强从丰富且划算的可再生资源微生物生产异戊二烯。具体而言,重组微生物诸如大肠杆菌 (E. coli)、藻类和蓝细菌已被用来将源于生物质的原料转化成异戊二烯。然而,即使使用相对廉价的纤维素生物质作为原料,多于一半质量的糖类原料由氧构成,其呈现出对转化效率的显著限制。异戊二烯及它的衍生物 (诸如类异戊二烯) 具有比该原料显著更低的氧含量,该原料由于氧必须作为废物消除而限制理论产率。因此,从糖类原料生产异戊二烯及它的衍生物的经济情况是极其昂贵的。

[0010] 鉴于与用于生产异戊二烯及相关化合物的基于糖类的发酵方法关联的限制,在本领域中需要用于生产异戊二烯的替代的划算且环境友好的方法。本公开满足此类需要,并

进一步提供其它相关的优点。

[0011] 概述

[0012] 简言之,本公开提供包括编码异戊二烯合成酶(例如,IspS)的外源性核酸的非天然存在的嗜甲烷细菌,其中该嗜甲烷细菌能够将碳原料转化成异戊二烯。

[0013] 编码异戊二烯合成酶的核酸可以源于包含内源性异戊二烯合成酶的任何生物体,诸如银白杨(*Populus alba*)、毛果杨(*Populus trichocarpa*)、美洲山杨(*Populus tremuloides*)、黑杨(*Populus nigra*)、银白杨 x 欧洲山杨(*Populus alba* x *Populus tremula*)、银灰杨(*Populus*×*canescens*)、山葛(*Pueraria montana*)、葛根(*Pueraria lobata*)、夏栎(*Quercus robur*)、蝶形花亚科(*Faboideae*)、褪色柳(*Salix discolor*)、无毛柳(*Salix glabra*)、五蕊柳(*Salix pentandra*)或百里香叶柳(*Salix serpyllifolia*)。编码 IspS 的外源性核酸还可以是为在该嗜甲烷细菌中表达而优化的密码子。异戊二烯合成酶还可以包含包含 SEQ ID NO:1-6 中任一项的氨基酸序列。该异戊二烯合成酶还可以不包括 N-末端质体靶向序列。该编码异戊二烯合成酶的核酸还可以包含 SEQ ID NO:14-19 中的任一项。

[0014] 编码异戊二烯合成酶的外源性核酸还可以与表达控制序列可操作地连接。该表达控制序列还可以是选自由甲醇脱氢酶启动子、己酮糖-6-磷酸合成酶启动子、核糖体蛋白 S16 启动子、丝氨酸羟甲基转移酶启动子、丝氨酸乙醛酸氨基转移酶启动子、磷酸烯醇丙酮酸羧化酶启动子、T5 启动子和 Trc 启动子组成的组中的启动子。

[0015] 该非天然存在的嗜甲烷细菌还可以包括嗜甲烷细菌,与内源性 DXP 途径酶的正常表达水平相比,该嗜甲烷细菌过度表达过表达内源性 DXP 途径酶,用编码 DXP 途径酶的外源性核酸进行转化,或它们的组合。该 DXP 途径酶可以是 DXS、DXR、IDI、IspD、IspE、IspF、IspG、IspH 或它们的组合。该非天然存在的嗜甲烷细菌还可以包括表达经转化的编码甲羟戊酸途径酶的外源性核酸的嗜甲烷细菌。该甲羟戊酸途径酶可以是乙酰乙酰辅酶 A 硫解酶、3-羟基-3-甲基戊二酰辅酶 A 合酶、3-羟基-3-甲基戊二酰辅酶 A 还原酶、甲羟戊酸激酶、磷酸甲羟戊酸激酶、甲羟戊酸焦磷酸脱羧酶、异戊烯基二磷酸异构酶或它们的组合。该非天然存在的嗜甲烷细菌还可以包括至少一种编码变异体 DXP 途径酶的外源性核酸。该变异体 DXP 途径酶可以包括突变型丙酮酸脱氢酶(PDH)和突变型 3,4-二羟基-2-丁酮-4-磷酸合成酶(DHBPS)。

[0016] 该嗜甲烷细菌还可以生产约 1mg/L 至约 500g/L 的异戊二烯。

[0017] 可以将编码异戊二烯合成酶的外源性核酸导入嗜甲烷细菌中,诸如甲基单胞菌属(*Methylomonas*)、甲基杆菌属(*Methylobacter*)、甲基球菌属(*Methylococcus*)、甲基弯曲菌属(*Methylosinus*)、甲基孢囊菌属(*Methylocystis*)、甲基微菌属(*Methylomicrobium*)、甲烷单胞菌属(*Methanomonas*)、甲基细胞菌属(*Methylocella*)、甲基帽菌属(*Methylocapsa*)。在某些实施方案中,该嗜甲烷细菌是荚膜甲基球菌 Bath 菌株(*Methylococcus capsulatus* Bath strain)、甲烷甲基单胞菌(*Methylomonas methanica*) 16a(ATCC PTA 2402)、甲烷氧化菌(*Methylosinus trichosporium*) 0B3b(NRRL B-11, 196)、生孢甲基弯曲菌(*Methylosinus sporium*) (NRRL B-11, 197)、小甲基孢囊菌(*Methylocystis parvus*) (NRRL B-11, 198)、甲烷甲基单胞菌(*Methylomonas methanica*) (NRRL B-11, 199)、白色甲基单胞菌(*Methylomonas albus*) (NRRL B-11, 200)、荚膜甲基杆

菌 (*Methylobacter capsulatus*) (NRRL B-11, 201)、嗜有机物甲基杆菌 (*Methylobacterium organophilum*) (ATCC 27, 886)、甲基单胞菌 sp AJ-3670 (FERM P-2400)、*Methylocella silvestris*、沼泽甲基细胞菌 (*Methylocella palustris*) (ATCC 700799)、冻原甲基细胞菌 (*Methylocella tundrae*)、*Methylocystis daltona* 菌株 SB2、*Methylocystis bryophila*、金黄甲基帽菌 (*Methylocapsa aurea*) KYG、极端嗜酸甲烷氧化菌 (*Methylacidiphilum infernorum*)、*Methylacidiphilum fumariolicum*、*Methyloacidia kamchatkensis*、甲基叔丁基醚降解菌 (*Methylibium petroleiphilum*) 或嗜碱甲基微菌 (*Methylomicrobium alcaliphilum*)。

[0018] 在某些实施方案中,该碳原料是甲烷、甲醇、天然气或非常规天然气。

[0019] 本文还提供用于生产异戊二烯的方法,其包括:在碳原料存在下在足以生产异戊二烯的条件下培养包含编码异戊二烯合成酶的外源性核酸的非天然存在的嗜甲烷细菌。该方法包括使用对该非天然存在的嗜甲烷细菌所述的各种实施方案。在某些实施方案中,该方法还包括回收由发酵废气生产的异戊二烯。可以将回收的异戊二烯进一步修饰成二聚体(10个碳)烃、三聚体(15个碳)烃,或它们的组合。可以将该二聚体烃、三聚体烃或它们的组合进一步氢化成为长链支链烷烃。在其它实施方案中,可以将回收的异戊二烯进一步修饰成类异戊二烯产品。

[0020] 在另一方面,本公开提供用于筛选突变型嗜甲烷细菌的方法,其包括:a)使该嗜甲烷细菌暴露于诱变剂以生产突变型嗜甲烷细菌;b)用编码香叶基香叶基二磷酸合成酶(GGPPS)、八氢番茄红素合成酶(CRTB)和八氢番茄红素脱氢酶(CRTI)的外源性核酸转化该突变型嗜甲烷细菌;和c)在足够生长的条件下培养来自步骤b)的突变型嗜甲烷细菌;其中与已暴露于诱变剂并已用编码GGPPS、CRTB和CRTI的外源性核酸转化的参考嗜甲烷细菌相比,呈现出增强的红色着色的突变型嗜甲烷细菌,表明具有增强的红色着色的该突变型嗜甲烷细菌呈现出与该参考嗜甲烷细菌相比增高的异戊二烯前体合成。在某些实施方案中,该诱变剂是辐射物(radiation)、化学品、质粒或转座子。在某些实施方案中,将具有增强的红色着色的突变型嗜甲烷细菌或它的克隆细胞用编码IspS的外源性核酸进行转化。在其它实施方案中,将至少一种编码GGPPS、CRTB和CRTI的核酸从具有增强的红色着色的突变型嗜甲烷细菌中除去或灭活。

[0021] 在又一方面,本公开提供用于在嗜甲烷细菌中筛选异戊二烯途径基因的方法,其包括:a)用以下转化该嗜甲烷细菌:i)至少一种编码异戊二烯途径酶的外源性核酸;ii)编码香叶基香叶基二磷酸合成酶(GGPPS)、八氢番茄红素合成酶(CRTB)和八氢番茄红素脱氢酶(CRTI)的外源性核酸;和b)在足够生长的条件下培养来自步骤a)的嗜甲烷细菌;其中与已用编码GGPPS、CRTB和CRTI的外源性核酸转化并且不包含至少一种编码异戊二烯途径酶的外源性核酸的参考嗜甲烷细菌相比,呈现出增强的红色着色的转化的嗜甲烷细菌,表明该至少一种编码异戊二烯途径酶的外源性核酸赋予与该参考嗜甲烷细菌相比增高的异戊二烯前体合成。该异戊二烯途径酶包括DXP途径酶或甲羟戊酸途径酶。该至少一种编码异戊二烯途径酶的外源性核酸可以包括异源性或同源性核酸。该至少一种编码异戊二烯途径酶的外源性核酸可以是为在宿主嗜甲烷细菌中表达而优化的密码子。该同源性核酸可以在该嗜甲烷细菌中过表达。该至少一种编码异戊二烯途径酶的外源性核酸可以包括非天然存在的变异体。

[0022] 本公开还提供一种异戊二烯组合物,其中该异戊二烯具有约-30%至约-50%范围内的 $\delta^{13}\text{C}$ 分布。

[0023] 附图简述

[0024] 图 1 显示用于异戊二烯合成的 1-脱氧-D-木酮糖-5-磷酸 (DXP) 途径。所用的缩写:DXS = 1-脱氧-D-木酮糖-5-磷酸 (DXP) 合成酶;DXR = 1-脱氧-D-木酮糖-5-磷酸 (DXP) 还原异构酶,也称为 IspC;IspD = 4-二磷酸胞苷基-2-C-甲基-D-赤藓糖醇 (CDP-ME) 合成酶;IspE = 4-二磷酸胞苷基-2-C-甲基-D-赤藓糖醇 (CDP-ME) 激酶;IspF = 2-C-甲基-D-赤藓糖醇 2,4-环二磷酸 (ME-cPP) 合成酶;IspG = 1-羟基-2-甲基-2-(E)-丁烯基 4-二磷酸 (HMBPP) 合成酶;IspH = 1-羟基-2-甲基-丁烯基 4-二磷酸 (HMBPP) 还原酶;IDI = 异戊烯基二磷酸 (IPP) 异构酶,也称为 IPI;IspS = 异戊二烯合成酶。

[0025] 图 2 显示用于异戊二烯合成的甲羟戊酸 (MVA) 途径。所用的缩写:AACT = 乙酰乙酰辅酶 A 硫解酶;HMGS = 羟甲基戊二酰辅酶 A (HMG) 合成酶;HMGR = 羟甲基戊二酰辅酶 A (HMG) 还原酶;MK = 甲羟戊酸 (MVA) 激酶;PMK = 磷酸甲羟戊酸激酶;MPD = 甲羟戊酸焦磷酸脱羧酶,也称为二磷酸甲羟戊酸脱羧酶 (DPMDC);IDI = 异戊烯基二磷酸 (IPP) 异构酶;IspS = 异戊二烯合成酶。

[0026] 图 3 以实例的方式显示如本公开中提供的嗜甲烷细菌如何可以通过用编码 IspS 的外源性核酸转化嗜甲烷菌而利用低碳烷烃(甲烷、乙烷、丙烷、丁烷)生产异戊二烯。

[0027] 图 4 显示各种碳源的 $\delta^{13}\text{C}$ 分布。

[0028] 图 5 显示源于用 (A) pMS3 载体和 (B) pMS3[Pmdh+ 柳树 (*Salix* sp.) IspS] 转化的荚膜甲基球菌 Bath 菌株的封闭培养的顶空样品的 GC/MS 色谱。箭头表明对应于异戊二烯的峰。通过在 A 中峰面积的量化的异戊二烯产率低于检测限。异戊二烯产率在 B 中是约 10mg/L。

[0029] 图 6 显示蕃茄红素途径的较低部分,该较低部分可以被转化成嗜甲烷宿主细菌并用来筛选突变型细菌菌株用于异戊二烯前体代谢物的改进生产。所用的缩写:GGPPS = 香叶基香叶基二磷酸 (GGPP) 合成酶。

[0030] 图 7A 和 7B 显示在来自用包含 pLacIq-山葛 ispS 的表达载体转化并在存在或不存在 IPTG 下生长的荚膜甲基球菌 Bath 菌株的封闭培养的顶空样品中通过 GC/MS 检测的异戊二烯的量。

[0031] 详述

[0032] 本公开提供用于从存在于天然气诸如低碳烷烃(甲烷、乙烷、丙烷和丁烷)中的碳原料生物合成异戊二烯的组合物和方法。例如,将嗜甲烷细菌用编码异戊二烯合成酶(例如,IspS)的外源性核酸转化并用碳原料(例如,天然气)培养以产生异戊二烯。本文所述的重组嗜甲烷细菌及相关的方法允许将碳原料嗜甲烷菌生物转化成异戊二烯用于轮胎或橡胶工业、制药中或用作替代燃料。

[0033] 作为背景,甲烷(尤其以天然气的形式)代表着廉价且丰富的天然资源。如前述,基于糖类的原料包含占它们的质量多于一半的氧,由于异戊二烯不包含任何氧分子,并且类异戊二烯具有比此类原料远远更低的氧含量,所以这显著限制转化效率。一种针对目前生物合成系统的限制的解决方案是利用天然气中的甲烷或其它低碳烷烃作为原料用于转

化。来自天然气的甲烷及其它低碳烷烃（例如，甲烷、乙烷、丙烷和丁烷）不具有氧，这允许显著改进转化效率。此外，与糖类原料相比天然气廉价且丰富，促成异戊二烯生产的改进的经济性。

[0034] 在本公开中，碳原料生物转化为异戊二烯是通过将编码异戊二烯合成酶（例如，IspS）的外源性核酸导入宿主嗜甲烷细菌中实现。此外，宿主嗜甲烷细菌的代谢工程可以通过过表达天然基因或与异戊二烯途径关联的外源性基因（例如，DXS、DXR、IspD、IspE、IspF、IspG、IspH 或 IDI）以增加异戊二烯前体用来增加异戊二烯产率。还提供用于通过将蕃茄红素途径工程改造入细菌中以提供比色读数筛选用于增加异戊二烯前体产量的突变型嗜甲烷细菌的方法。

[0035] 在更详细地阐述本公开之前，提供本文所用的某些术语的定义可以有助于对它的理解。附加的定义在整个本公开中进行了阐述。

[0036] 在本说明书中，除非另有指明，术语“约”意指所指范围、值或结构的 $\pm 20\%$ 。术语“基本上由……组成”将要求保护的的范围限制于特定的材料或步骤以及不实质性影响要求保护的发明的基本的和新颖的特征的那些。应理解如本文所用的术语“一种”是指“一种或多种”列举的组分。替代（例如，“或”）的使用应理解意指替代物中的一个、两个或它们的任何组合。如本文所用的术语“包括”和“具有”同义地使用，它们的术语和变型旨在被解释为非限制性。术语“包含”意指如权利要求书中提及的所述特征、整数、步骤或组分的存在，但是它不排除一个或多个其它特征、整数、步骤、组分或它们的组的存在或增加。

[0037] 如本文所用的术语“异戊二烯”，也称为“2-甲基-1,3-丁二烯”，是指具有式 $\text{CH}_2=\text{C}(\text{CH}_3)\text{CH}=\text{CH}_2$ 的有机化合物。异戊二烯是由各种植物、微生物和动物物种生产的无色、疏水性、挥发性液体。异戊二烯是用于各种合成聚合物（包括合成橡胶）的关键起始材料，并且还可以用作燃料。

[0038] 如本文所用的术语“异戊二烯合成途径”、“异戊二烯生物合成途径”或“异戊二烯途径”是指用于生产异戊二烯的任何生物合成途径。异戊二烯生物合成通常通过两种途径完成：存在于真核细胞、古细菌和高等植物的细胞溶质中的甲羟戊酸（MVA）途径，以及非-甲羟戊酸途径，也称为甲基-赤藓糖醇-4-磷酸（MEP）或（1-脱氧-D-木酮糖-5-磷酸）DXP 途径，其可以是原核来源或来自植物质体。一种异戊二烯途径还可以包括已知生物合成途径的途径变型或修改或工程改造的生物合成途径。

[0039] 如本文所用的术语“类异戊二烯”是指从异戊二烯单元（5 个碳支链异戊二烯结构）合成的或包含异戊二烯单元（5 个碳支链异戊二烯结构）的任何化合物。类异戊二烯可以包括萜类、银杏苦内酯类、甾醇类和类胡萝卜素类，

[0040] 如本文所用的术语“甲羟戊酸途径”或“MVA 途径”是指通常存在于真核细胞和古细菌中的异戊二烯生物合成途径。该甲羟戊酸途径包括如图 2 中所述的经典途径，以及修改的 MVA 途径，诸如一种通过磷酸甲羟戊酸脱羧酶（PMD）将甲羟戊酸磷酸转化成异戊烯基磷酸的途径，异戊烯基磷酸通过异戊烯基磷酸激酶（IPK）被转化成异戊烯基二磷酸。

[0041] 如本文所用的术语“非甲羟戊酸途径”或“1-脱氧-D-木酮糖-5-磷酸（DXP）途径”是指通常存在于细菌和植物质体中的异戊二烯生物合成途径。一种示例性 DXP 途径示于图 1 中。

[0042] 如本文所用的术语“DXP”是指 1-脱氧-D-木酮糖-5-磷酸。1-脱氧-D-木酮

糖-5-磷酸合成酶(DXS)催化甘油醛与丙酮酸的缩合从而形成DXP,其在DXP途径中是异戊二烯的前体分子。

[0043] 如本文所用的术语“异戊二烯合成酶”(例如,IspS)是指催化二甲基烯丙基二磷酸(DMAPP)转化成异戊二烯的酶。

[0044] 如本文所用的术语“蕃茄红素途径”是指用于生产蕃茄红素的生物合成途径。蕃茄红素是通常存在于番茄及其它红色水果和蔬菜中的鲜红色类胡萝卜素色素。蕃茄红素途径的一个实例示于图6中。一般而言,在真核植物和原核生物中蕃茄红素生物合成是相似的,起始于甲羟戊酸,其被转化成二甲基烯丙基二磷酸(DMAPP)。二甲基烯丙基二磷酸与三分子的IPP缩合而产生香叶基香叶基焦磷酸(GGPP)。两分子的GGPP以尾-尾的方式缩合而产生八氢蕃茄红素,其进行若干去饱和步骤以生产蕃茄红素。

[0045] 如本文所用的术语“宿主”是指可以用异戊二烯生物合成途径组分(例如,IspS)进行基因修饰从而将碳底物原料(例如,甲烷、天然的、低碳烷烃)转化为异戊二烯的微生物(例如,嗜甲烷细菌)。宿主细胞可以包含用于异戊二烯前体合成的内源性途径(例如,DMAPP或IPP)或可以进行基因修饰以允许或增高前体产量。此外,宿主细胞可以已经具有赋予与本文公开的异戊二烯生物合成途径无关的期望性质的其它基因修饰。例如,宿主细胞可以具有赋予高增长、对污染物或特定培养条件的耐受性、代谢附加的碳底物的能力或合成期望的产品或中间体的能力的基因修饰。

[0046] 如本文所用的术语“嗜甲烷菌”、“嗜甲烷细菌(methanotrophic bacterium; methanotrophic bacteria)”是指能够利用C₁底物诸如甲烷或非常规天然气作为主要的或仅有的碳和能量来源的甲基营养菌。如本文所用的“嗜甲烷细菌”包括可以仅仅利用C₁底物作为碳和能量来源的“专性嗜甲烷细菌”,以及除了C₁底物之外还天然地能够使用多-碳底物诸如乙酸盐、丙酮酸盐、琥珀酸盐、苹果酸盐或乙醇作为它们的主要的或仅有的碳和能量来源的“兼性嗜甲烷细菌”。兼性嗜甲烷菌包括甲基细胞菌属、甲基孢囊菌属和甲基帽菌属中的一些物种(例如,Methylocella silvestris、沼泽甲基细胞菌、冻原甲基细胞菌、Methylocystis daltona SB2、Methylocystis bryophila 和金黄甲基帽菌 KYG),以及嗜有机物甲基杆菌(ATCC 27,886)。

[0047] 如本文所用的术语“C₁底物”或“C₁原料”是指缺少碳-碳键的任何含碳的分子。实例包括:甲烷、甲醇、甲醛、甲酸、甲酸盐、甲基化胺(例如,单甲基胺、二甲基胺和三甲基胺)、甲基化硫醇和二氧化碳。

[0048] 如本文所用的术语“低碳烷烃”是指甲烷、乙烷、丙烷或丁烷,或它们的任何组合。低碳烷烃可以包括基本上纯化的组合物,诸如为95-98%甲烷的“管道质量天然气”或“干天然气”,或未纯化的组合物,诸如“湿天然气”,其中其它烃(例如,乙烷、丙烷和丁烷)尚未被除去并且甲烷占该组合物的大于60%。低碳烷烃还可以“天然气液体”(也称为“天然气关联的烃”)提供,其是指在加工以生产管道质量干天然气期间从湿天然气分离的各种烃(例如,乙烷、丙烷、丁烷)。“湿天然气的部分分离的衍生物”包括天然气液体。

[0049] 如本文所用的术语“天然气”是指主要由甲烷组成的天然存在的烃气体混合物,其可以具有一种或多种其它烃(例如,乙烷、丙烷和丁烷)、二氧化碳、氮气和硫化氢。天然气包括常规天然气和非常规天然气(例如,致密砂岩气、气岩页、天然气水合物和煤层甲烷)。天然气包括干天然气(或管道质量天然气)或湿(未经加工的)天然气。

[0050] 如本文所用的术语“非天然存在的”，也称为“重组的”或“转基因的”，在提及微生物使用时，意指具有至少一个在参考物种的天然存在的菌株（包括该参考物种的野生型菌株）中通常不存在的遗传学改变的微生物。遗传学改变包括：例如，导入编码蛋白的可表达的核酸分子的修饰、其它核酸添加、核酸缺失、核酸取代，或细菌的遗传物质的其它功能性破坏。此类修饰包括，例如，对于参考物种的对异源性或同源性多肽的编码区及它们的功能性片段。附加的修饰包括，例如，其中该修饰改变基因或操纵子的表达的非编码调控区。示例性蛋白包括在异戊二烯途径（例如，IspS）内的蛋白。对编码酶的核酸分子或其功能性片段的基因修饰可以将生物化学反应能力或代谢途径能力或此类能力的改进赋予从它的天然存在的状态改变的非天然存在的微生物。

[0051] 如本文所用的“外源性”意指将提及的分子（例如，核酸）或提及的活性（例如，异戊二烯合成酶活性）导入宿主微生物中。该分子可以例如，通过将核酸导入宿主遗传物质中，诸如通过整合入宿主染色体中或通过将核酸作为非染色体遗传物质导入诸如在质粒上，进行导入。在提及编码核酸的表达使用该术语时，它是指将该编码核酸以可表达的形式导入宿主微生物中。在提及酶活性或蛋白活性使用时，该术语是指被导入宿主参考微生物中的活性。因此，术语“内源性”或“原生的”是指存在于宿主微生物中的提及的分子或活性。在提及核酸使用时，术语“嵌合”是指非内源性的包括自然界中不共同存在的序列的任何核酸。例如，嵌合核酸可以包括源于不同来源的调控序列和编码序列，或源于相同来源但以不同于自然界中存在的形式排列的调控序列和编码序列。术语“异源性”是指源于不同于参考物种或菌株的来源的分子或活性，而“同源性”是指源于宿主微生物的分子或活性。因此，包括如本公开中提供的外源性核酸的微生物可以使用异源性核酸或同源性核酸或二者兼有。

[0052] 应当理解的是当外源性核酸被包括在微生物中时，该外源性核酸是指如上所讨论提及的编码核酸或蛋白活性。还应当理解的是可以将此类外源性核酸导入宿主微生物的独立的核酸分子上，多顺反子核酸分子上，编码融合蛋白的单个核酸分子上，或其组合，并且仍然被视为多于一种外源性核酸。例如，如本文公开的，可以修饰微生物以表达一种或多种编码来自异戊二烯途径的酶（例如，异戊二烯合成酶）的外源性核酸。在将两种编码来自异戊二烯途径的酶的外源性核酸导入宿主微生物的情况中，应当理解的是可将这两种外源性核酸以单个核酸分子导入在例如单个质粒上，在独立的质粒上，可整合入宿主染色体的单个位点或多个位点处，并且仍然被视为两种外源性核酸。相似地，应当理解的是多于两种外源性核酸分子可以以任何期望的组合被导入宿主微生物中，例如，在单个质粒上，在独立的质粒上，可以被整合入宿主染色体中在单个位点或多个位点处，并且仍然被视为两种或多种外源性核酸。因此，提及的外源性核酸或酶活性的数量是指编码核酸的数量或蛋白活性的数量，不是被导入宿主微生物中的独立的核酸分子的数量。

[0053] 如本文所用的“核酸”，也称为多聚核苷酸，是指由共价连接的称为核苷酸的亚单位构成的聚合化合物。核酸包括多聚核糖核酸（RNA）、多聚脱氧核糖核酸（DNA），二者可以是单链或双链。DNA 包括 cDNA、基因组 DNA、合成 DNA 和半合成 DNA。

[0054] 如本文所用的“过表达的”在提及基因或蛋白使用时是指基因或蛋白的表达或活性增高。增高的表达或活性包括当基因或蛋白的表达或活性增高超出野生型（非基因工程的）对照或参考微生物中的水平时。若该表达或活性是在它的通常非表达或活性的微生物

中,则基因或蛋白被过表达。若该表达或活性在该微生物中存在时间段比在野生型对照或参考微生物中更长,则基因或蛋白被过表达。

[0055] 宿主嗜甲烷细菌

[0056] 转化是指核酸(例如,外源性核酸)转移入宿主微生物的基因组中,导致遗传上稳定的遗传。包含转化的核酸的宿主微生物称为“非天然存在的”或“重组的”或“转化的”或“转基因的”微生物。宿主微生物可以选自嗜甲烷细菌,或非天然存在的微生物从嗜甲烷细菌生成,嗜甲烷细菌通常包括具有氧化作为碳和能量来源的甲烷的能力的细菌。

[0057] 嗜甲烷细菌根据它们的碳同化途径及内部膜结构被分类成三组:I型(γ 变形杆菌(*proteobacteria*))、II型(α 变形杆菌,和X型(γ 变形杆菌)。为了碳同化,I型嗜甲烷菌使用单磷酸核酮糖(RuMP)途径而II型嗜甲烷菌使用丝氨酸途径。X型嗜甲烷菌使用RuMP途径但也表达低水平的丝氨酸途径的酶。嗜甲烷细菌被划分成若干属:甲基单胞菌属、甲基杆菌属、甲基球菌属、甲基孢囊菌属、甲基弯曲菌属、甲基微菌属、甲烷单胞菌属和甲基细胞菌属。

[0058] 嗜甲烷细菌包括专性嗜甲烷菌和兼性嗜甲烷菌,其天然地具有使用一些多-碳底物作为仅有的碳和能量来源的能力。兼性嗜甲烷菌包括甲基细胞菌属、甲基孢囊菌属和甲基帽菌属中的一些物种(例如,*Methylocella silvestris*、沼泽甲基细胞菌、冻原甲基细胞菌、*Methylocystis daltona*菌株SB2、*Methylocystis bryophila*和金黄甲基帽菌KYG)。示例性嗜甲烷细菌物种包括:荚膜甲基球菌Bath菌株、甲基单胞菌16a(ATCC PTA 2402)、发孢甲基弯曲菌OB3b(NRRL B-11, 196)、生孢甲基弯曲菌(NRRL B-11, 197)、小甲基孢囊菌(NRRL B-11, 198)、甲烷甲基单胞菌(NRRL B-11, 199)、白色甲基单胞菌(NRRL B-11, 200)、荚膜甲基杆菌(NRRL B-11, 201)、嗜有机物甲基杆菌(ATCC27, 886)、甲基单胞菌sp AJ-3670(FERM P-2400)、*Methylocella silvestris*、沼泽甲基细胞菌(ATCC 700799)、冻原甲基细胞菌、*Methylocystis daltona*菌株SB2、*Methylocystis bryophila*、金黄甲基帽菌KYG、极端嗜酸甲烷氧化菌、*Methyloacidiphilum fumariolicum*、*Methyloacida kamchatkensis*、甲基叔丁基醚降解菌和嗜碱甲基微菌。

[0059] 选定的嗜甲烷宿主细菌还可以经受在选择条件下菌株适应以鉴定具有用于生产的改进的性质的变异体。改进的性质可以包括:增加的生长率、期望产品的产率以及对可能的工艺污染物的耐受性(参见,例如,美国6,689,601)。在具体的实施方案中,高生长变异体嗜甲烷细菌是一种能够在作为仅有的碳和能量来源的甲烷上生长,并且具有与它的亲代细菌、参考细菌或野生型细菌相比更快的指数期生长率(即,较短的倍增时间)的生物体。

[0060] 异戊二烯合成途径、核酸和多肽

[0061] 本公开提供已经工程改造而具有生产异戊二烯的能力的嗜甲烷细菌。包括DXP途径的上部分的酶存在于许多嗜甲烷细菌中。然而,在(E)-4-羟基-3-甲基丁-2-烯基-二磷酸(HMBPP)转化成异戊烯基二磷酸(IPP)和二甲基烯丙基二磷酸(DMAPP)之后,目前已知的嗜甲烷菌缺少用于将DMAPP转化成异戊二烯的异戊二烯合成酶(例如,IspS)。相反,嗜甲烷菌通过香叶基转移酶和法尼基二磷酸合成酶(IspA)将DMAPP转化成法尼基二磷酸,其随后被转化成类胡萝卜素(参见,例如,美国专利7,105,634)。

[0062] 在某些实施方案中,本公开提供包括编码异戊二烯合成酶(例如,IspS)的外源性核酸的非天然存在的嗜甲烷细菌,其中该嗜甲烷细菌能够将碳原料转化成异戊二烯。用编

码异戊二烯合成酶的外源性核酸转化的嗜甲烷细菌通常能够利用如图 1 中所示的 DXP 途径将丙酮酸和甘油醛-3-磷酸转化成异戊二烯。

[0063] 异戊二烯合成酶核酸和多肽序列是本领域中已知的,并且可以从天然具有异戊二烯合成酶的任何生物体获得。IspS 基因已从多种植物(包括例如杨树、山杨和野葛)中分离出并克隆。虽然多种细菌具有 DXP 途径,但在目前的任何数据库中不能获得来自原核生物的 ispS 基因的序列(参见,例如,Xue 等,2011,Appl. Environ. Microbiol. 77:2399-2405)。在某些实施方案中,编码异戊二烯合成酶的核酸源于银白杨、毛果杨、美洲山杨、黑杨、银白杨 x 欧洲山杨、银灰杨、山葛、葛根、夏栎、蝶形花亚科、褪色柳、无毛柳、五蕊柳或百里香叶柳。在 NCBI 数据库中可获得的异戊二烯合成酶的核酸序列的实例包括:登录号 AB198180(银白杨)、AY341431(美洲山杨)、AJ294819(银白杨 x 欧洲山杨)、AY316691(葛麻姆 (Montana var. lobata))、HQ684728(黑杨)和 EU693027(毛果杨)。在表 1 中提供异戊二烯合成酶多肽的实例。带下划线的序列代表 N-末端质体靶向序列,其在截短形式中被除去。在某些实施方案中,该外源性核酸编码具有如 SEQ ID NO:1-6 中任一项阐述的氨基酸序列的异戊二烯合成酶多肽。

[0064] 表 1. 异戊二烯合成酶多肽的实例

[0065]

物种	氨基酸序列	SEQ ID NO.
银白杨	MATELLCLHRPTSLTHKLFNRNPLPKVIQATPLTLKLRC SVSTENVSTETETETEARRSANYEPNSWDYDYLSSDT ESIEVYKDKAKKLEAEVRREINNEKAEFLTLELIDNV QRLGLGYRFESDIRGALDRFVSSGGFDAVTKTSLHGTA LSFRLLRQHGFVVSQEAFFSGFKDQNGNFLENLKEDIKA ILSLYEASFLALEGENILDEAKVFAISHLKELSEEKIG KELAEQVNHAELEPLHRRRTQRLEAVWSIEAYRKKEAN QVLELAILDYNMIQSVYQRDLRETSRWRRVGLATKL HFARDRLIESFYWAVGVAFEPQYSDCRNSVAKMFSFVT IIDDIDYDVGTLDELELFTDAVERWDVNAINDLDPYMK LCFLALYNTINEIAYDNLKDKGENILPYLTKAWADLCN AFLQEAQWLYNKSTPTFDDYFGNAWKSSSGPLQLVPAY PAVVQNIKKEEIEENLQKYHDTISRPSHIFRLCNDLASA SAEIARGETANSVSCYMRKGISEELATESVMNLIDET WKKMNKEKLGGSFLFAKPFVETAINLARQSHCTYHNGDA HTSPDELTRKRVLSVITEPILPFER	1
银白杨 (截短的)	MCSVSTENVSTETETETEARRSANYEPNSWDYDYLSSD TDESIEVYKDKAKKLEAEVRREINNEKAEFLTLELID NVQRLGLGYRFESDIRGALDRFVSSGGFDAVTKTSLHG TALSFRLLRQHGFVVSQEAFFSGFKDQNGNFLENLKEDI KAILSLYEASFLALEGENILDEAKVFAISHLKELSEEK IGKELAEQVNHAELEPLHRRRTQRLEAVWSIEAYRKKEA ANQVLELAILDYNMIQSVYQRDLRETSRWRRVGLAT KLHFARDRLIESFYWAVGVAFEPQYSDCRNSVAKMFSF VTIIDDIDYDVGTLDELELFTDAVERWDVNAINDLDPY MKLCFLALYNTINEIAYDNLKDKGENILPYLTKAWADL CNAFLQEAQWLYNKSTPTFDDYFGNAWKSSSGPLQLVF AYFAVVQNIKKEEIEENLQKYHDTISRPSHIFRLCNDLA SASAEIARGETANSVSCYMRKGISEELATESVMNLID ETWKKMNKEKLGGSFLFAKPFVETAINLARQSHCTYHNG DAHTSPDELTRKRVLSVITEPILPFER	2
葛麻姆	MATNLLCLSNKLSSPTPTPSTRFPQSKNFITQKTSAN PKPWRVICATSSQFTQITEHNSRRSANYQPNLWNFEFL	3

[0066]

物种	氨基酸序列	SEQ ID NO.
	QSLNDLKVEKLEEKATKLEEEVRCMINRVDTQPLSLL ELIDDVQRLGLTYKFEKDI I KALENIVLLDENKKNKSD LHATALSFRLLRQHGFVVSQDVFERFKD KEGGFSGELKGDVQGLLSLYEASYLGFEGENLLEEART FSITHLKNNLKEGINTKVAEQVSHALELPYHQRHLRLE ARWFLDKYEPKEPHHQLLLELAKLDFNMVQTLHQKELQ DLSRWWTENGLASKLDFVRDRLMEVYFWALGMAPDPQF GECRAVTKMFGLVTIIDDVYDVYGTLDLDELQFLTDAVE RWDVNAINTLPDYMKL CFLALYNTVNDTSYSILKEKGH NNLSYLT KSWRELCKAFLQEA KWSNNKI I PAFSKYLEN ASVSSSGVALLAPSYF SVCQQQEDISDHALRSLTDFHG LVRSSCVIFRLCNDLATSAAELERGETTNSIISYMHEN DGTSEEQAREELRKLIDA EWKKMNRERVSDSTLLPKAF MEIAVN MARVSHCTYQYGDGLGRPDYATENRIKLLID PFPINQLMYV	
葛麻姆 (截短的)	MCATSSQFTQITEHNSRRSANYQPNLWNFEFLQSLEND LKVEKLEEKATKLEEEVRCMINRVDTQPLSLLLELIDDV QRLGLTYKFEKDI I KALENIVLLDENKKNKSD LHATAL SFRLLRQHGFVVSQDVFERFKD KEGGFSGELKGDVQGL LSLYEASYLGFEGENLLEEARTFSITHLKNNLKEGINT KVAEQVSHALELPYHQRHLRLE ARWFLDKYEPKEPHH QLLLELAKLDFNMVQTLHQKELQDLSRWWTENGLASKL D FVRDRLMEVYFWALGMAPDPQFGECAVTKMFGLVTI IDDVYDVYGTLDLDELQFLTD AVERWDVNAINTLPDYM KLCFLALYNTVNDTSYSILKEKGHNNLSYLT KSWRELCK A FLQEA KWSNNKI I PAFSKYLENASVSSSGVALLAPSY F SVCQQQEDISDHALRSLTDFHGLVRSSCVIFRLCNDL A TSAAELEERGETTNSIISYMHENDGTSEEQAREELRKL I DA EWKKMNRERVSDSTLLPKAFMEIAVN MARVSHCTY QYGDGLGRPDYATENRIKLLIDFPINQLMYV	4
柳属 DG- 2011	MATELLCLHRPISLTPKLFNRNPLPKVILATPLTLKLRC SVSTENVSFTETETETETRRSANYEPNSWDYDYLLSSD TD ESIEVYKDKAKKLEAEVRREINNEKAEFLTLELIDNV QRLGLGYRFESDIRRALDRFVSSGGFDAVTKTS LHATA LSFRFLRQHGFVVSQEA FGGFKDQNGNFLENLKEDIKA ILSLYEASFLALEGENILDEAKVFAISHLKELSEEKIG KDLAEQVNHAELEPLHRRRTQRLEAVWSIEAYRKKE DAN QVLELAILDYNMIQSVYQRDLRETSRWWRVGLATKL HFARDRLIESFYWAVGVAFEPQYSDCRNSVAKMFSFVT IIDDYDVYGTLDLELEFLTD AVERWDVNAINDLPDYM KLCFLALYNTINEIAYDNLKEKGENILPYLTKAWADLCN AFLQEA KWLYNKSTPTFDDYFGNAWKSSSGPLQLVPAY FAVVQNIKKEETENLQKYHDIISRPSHIFRLCNDLASA SAEIARGETANSVSCYMRKGI SEELATESVMNLIDET WKKMNKEKLGGSLFPKPFVETAINLARQSHCTYHNGDA HTSFDLTKRKRVLSVITEPILPFER	5

[0067]

物种	氨基酸序列	SEQ ID NO.
柳属 DG-2011 (截短的)	MCSVSTENVSFTEETETETETRRSANYEPNSWDYDYLLSSD TDESIEVYKDKAKKLEAEVRREINNEKAEFLTLLELID NVQRLGLGYRFESDIRRALDRFVSSGGFDAVTXTSLHA TALSFRFLRQHGFEVSQEAFFGGFKDQNGNFLENLKEDI KAILSLYEASFLALEGENILDEAKVFAISHLKELSEEK IGKDLAEQVNHAELEPLHRRRTQRLEAVWSIEAYRKED ANQVLLLELAILDYNNMIQSVYQRDLRETSRWRRVGLAT KLHFARDRLIESFYWAVGVAFEPQYSDCRNSVAKMPSF VTIIDDIDYDVYGTLDLELELFTDAVERWDVNAINDLPDY MKLCFLALYNTINEIAYDNLKEKGENILPYLTAKAWADL CNAFLQEAKWLYNKSTFTFDDYFGNAWKSSSGPLQLVF AYFAVVQNIKKKEELENLQKYHDIISRPSHIFRLCNDLA SASAEIARGETANSVSCYMRITGISEELATESVMNLID ETWKKMNKEKLGGSLFFPKPFVETAINLARQSHCTYHNG DAHTSPDELTRKRVLSVITEPILPFR	6

[0068] 用于本文所述的组合物和方法中的异戊二烯合成酶核酸和多肽序列包括具有改进的溶解性、表达、稳定性、催化活性和转换率的变异体。例如,美国专利 8,173,410(在此通过引用整体并入)公开了具有增强的溶解性、表达和活性的特定的异戊二烯合成酶氨基酸取代。

[0069] 在某些实施方案中,可以期望的是过表达内源性 DXP 途径酶或将外源性 DXP 途径基因导入宿主嗜甲烷菌中以增加 IPP 和 DMAPP 产量和异戊二烯产率。在某些实施方案中,如本文提供的包括编码异戊二烯合成酶(例如, IspS)的外源性核酸的非天然存在的嗜甲烷细菌,与内源性 DXP 途径酶的正常表达水平相比进一步过表达内源性 DXP 途径酶,用编码 DXP 途径酶的外源性核酸进行转化,或二者兼具。“内源性”或“原生的”是指存在于宿主嗜甲烷细菌中的提及的分子或活性。在其它实施方案中,如本文提供的包括编码异戊二烯合成酶(例如, IspS)的外源性核酸的非天然存在的嗜甲烷细菌,与两种、三种、四种、五种、六种、七种、八种或更多种内源性 DXP 途径酶的正常表达水平相比,过表达所述的两种、三种、四种、五种、六种、七种、八种或更多种内源性 DXP 途径酶;用编码两种、三种、四种、五种、六种、七种、八种或更多种 DXP 途径酶的外源性核酸进行转化;或其任何组合。来自 DXP 途径的内源性酶诸如 DXS 的过表达,可以增加异戊二烯产量(Xue 和 Ahring, 2011, Applied Environ. Microbiol. 77:2399-2405)。不希望受理论的束缚,可认为增加 DXS 的量使经过 DXP 途径的碳的流量增大,从而导致异戊二烯产量增加。在某些实施方案中,采用液相色谱法-质谱法的代谢物谱分析(metabolite profiling)用来鉴定异戊二烯合成途径中的瓶颈和待过表达的酶(参见,例如, Pitera 等, 2007, Metabolic Engineering 9:193-207)。

[0070] 用于在宿主生物体中过表达核酸的方法是本领域中已知的。过表达可以通过将编码内源性 DXP 途径酶的核酸或编码 DXP 途径酶的外源性(例如, 异源性)核酸的拷贝导入宿主嗜甲烷细菌中来实现。作为实例,可以将编码内源性 DXS 酶的核酸与编码异戊二烯合成酶(例如, IspS)的外源性核酸一起转化入宿主嗜甲烷细菌中,或可以将源于非-宿主嗜甲烷物种的编码 DXS 酶的外源性核酸与编码异戊二烯合成酶(例如, IspS)的外源性核酸一起转化入宿主嗜甲烷细菌中。内源性 DXP 途径酶的过表达还可以通过用导致增强的转录的启动子或调控区替换内源性启动子或调控区来实现。

[0071] 在某些实施方案中,在宿主嗜甲烷细菌中过表达的 DXP 途径酶是 DXS、DXR、IDI、IspD、IspE、IspF、IspG、IspH 或它们的组合。在一些实施方案中,在宿主嗜甲烷菌中过表达的 DXP 途径酶是 DXS、IDI、IspD、IspF 或它们的组合。

[0072] DXP 途径酶的来源是本领域中已知的,并且可以来自天然具有 DXP 途径的任何生物体,包括多种植物和细菌物种。例如,DXP 途径酶可以存在于炭疽杆菌 (*Bacillus anthracis*)、幽门螺旋杆菌 (*Helicobacter pylori*)、鼠疫杆菌 (*Yersinia pestis*)、结核分枝杆菌 (*Mycobacterium tuberculosis*)、恶性疟原虫 (*Plasmodium falciparum*)、海分枝杆菌 (*Mycobacterium marinum*)、枯草芽孢杆菌 (*Bacillus subtilis*)、大肠杆菌 (*Escherichia coli*)、超嗜热菌 (*Aquifex aeolicus*)、鼠衣原体 (*Chlamydia muridarum*)、空肠弯曲杆菌 (*Campylobacter jejuni*)、沙眼衣原体 (*Chlamydia trachomatis*)、肺炎嗜衣原体 (*Chlamydophila pneumoniae*)、流感嗜血杆菌 (*Haemophilus influenzae*)、脑膜炎奈瑟菌 (*Neisseria meningitidis*)、集胞藻属 (*Synechocystis*)、极端嗜酸甲烷氧化菌 V4、甲基孢囊菌 sp. SC2、甲基单胞菌菌株 16A、荚膜甲基球菌 Bath 菌株、一些单细胞藻类,包括斜生栅藻 (*Scenedesmus oliquus*),以及在大多数植物物种的质体中,包括拟南芥 (*Arabidopsis thaliana*)、银白杨、毛果杨、美洲山杨、黑杨、银白杨 x 欧洲山杨、银灰杨、山葛、葛根、夏栎、蝶形花亚科、褪色柳、无毛柳、五蕊柳或百里香叶柳中。

[0073] 在 NCBI 数据库中可获得的 DXS 的核酸序列的实例包括登录号:AF035440(大肠杆菌);Y18874(聚球藻 (*Synechococcus*)PCC6301);AB026631(链霉菌 (*Streptomyces* sp.)CL190);AB042821(灰孢链霉菌 (*Streptomyces griseolosporeus*));AF11814(恶性疟原虫);AF143812(番茄 (*Lycopersicon esculentum*));AJ279019(假水仙 (*Narcissus pseudonarcissus*));AJ291721(烟草 (*Nicotiana tabacum*));AX398484.1(甲基单胞菌菌株 16A);NC_010794.1(区 1435594..1437486, 补体)(极端嗜酸甲烷氧化菌 V4);和 NC_018485.1(区 2374620..2376548)(甲基孢囊菌 sp. SC2)。

[0074] 在 NCBI 数据库中可获得的 DXR 的核酸序列的实例包括登录号:AB013300(大肠杆菌);AB049187(灰孢链霉菌 (*Streptomyces griseolosporeus*));AF111813(恶性疟原虫);AF116825(胡椒薄荷 (*Mentha x piperita*));AF148852(拟南芥);AF182287(黄花蒿 (*Artemisia annua*));AF250235(长春花 (*Catharanthus roseus*));AF282879(铜绿假单胞菌 (*Pseudomonas aeruginosa*));AJ242588(拟南芥);AJ250714(运动发酵单胞菌 (*Zymomonas mobilis*) 菌株 ZM4);AJ292312(肺炎克雷伯杆菌 (*Klebsiella pneumoniae*));AJ297566(玉米 (*Zea mays*));和 AX398486.1(甲基单胞菌菌株 16A)。在 NCBI 数据库中可获得的 IspD 的核酸序列的实例包括登录号:AB037876(拟南芥);AF109075(难辨梭状芽孢杆菌 (*Clostridium difficile*));AF230736(大肠杆菌);AF230737(拟南芥);和 AX398490.1(甲基单胞菌菌株 16A)。

[0075] 在 NCBI 数据库中可获得的 IspE 的核酸序列的实例包括登录号:AF216300(大肠杆菌);AF263101(番茄);AF288615(拟南芥);和 AX398496.1(甲基单胞菌菌株 16A)。

[0076] 在 NCBI 数据库中可获得的 IspF 的核酸序列的实例包括登录号:AF230738(大肠杆菌);AB038256(大肠杆菌);AF250236(长春花);AF279661(恶性疟原虫);AF321531(拟南芥);和 AX398488.1(甲基单胞菌菌株 16A)。

[0077] 在 NCBI 数据库中可获得的 IspG 的核酸序列的实例包括登录号:AY033515(大

肠杆菌)YP_005646(嗜热栖热菌(*Thermus thermophilus*))和YP_475776.1(聚球藻)。在NCBI数据库中可获得的IspH的核酸序列的实例包括登录号:AY062212(大肠杆菌)、YP_233819.1(丁香假单胞菌(*Pseudomonas syringae*))和YP_729527.1(聚球藻)。在NCBI数据库中可获得的IDI的核酸序列的实例包括登录号:AF119715(大肠杆菌)、P61615(芝田硫化叶菌(*Sulfolobus shibatae*))和042641(红发夫酵母(*Phaffia rhodozyme*))。

[0078] 在表2中提供可用于各种实施方案中的来自荚膜甲基球菌Bath菌株(ATCC 33009)的DXP途径酶的氨基酸序列。

[0079] 表2. 荚膜甲基球菌Bath菌株的DXP途径酶

[0080]

基因名	氨基酸序列	SEQ ID NO.
DXS	MTETKRYALLEAADHPAALRNLPEDLPELAEELRGYLLLESVS RSGGHLAAGLGTVELTIALHYVFNTPEDKLVWDVGHQAYPHKI LTGRRARLPTIRKKGGLSAFPNRAESFYDCFCVGHSSSTSISAA LGMVAAAALERRPIHAVAIIGDGGLTGGMAFEALNHAGTLDAN LLIIILNDNEMSISPNVGALNNYLAKILSGKFYSSVRESGKHLL GRHMPGVWELARRAEHVKGMPVAPGTLFEELGFNYFGPIDGHD LDTLITTLRNLRDQKGPRFLHVVTRKGKGYAPAEKDPVAYHGV GAFDLDADELPKSKPGTPSYTEVFGQWLCDMAARDRLLGITP AMREGSGLVEFSQRFPDRYFDVGIAEQHAVTFAAGQASEGYKP VVAIYSTFLQRAYDQLIHDVALQNLFVLFADIRAGLVGPDGPT HAGSFDLSFMRCIPNMLIMAPSDENECRQMLYTGFIHDGPAAV RYPRGRGPGVRPEETMTAFPPVGKGEVLRGKGTAIFAFTPLA	7

[0081]

基因名	氨基酸序列	SEQ ID NO.
	AALAVGERIGATVANMRFVKPLDE ALILELAATHDRIVTVEENAIAGGAGSAVGEFLAAQHCGIFVC HIGLKDEFLDQGTREELLAIAGLDQAGIARSIDAFIQATAAAD KPRRARGQAKDKH	
DXR	MKGICILGSTGSGISGVSTLDVLAHPDRYRVVALSANGNVDRLEF EQCRAHRPRYAAVIRAEAAACLRERLMAAGLGGIEVLAGEAL EQIASLPEVDSVMAAIVGAAGLLPTLAAARAGKDVLLANKEAL VMSGPLPMAEVARSGARLLPIDSEHNAVFCMPAAYRAGSRAV GVRRILLTASGGPFLHTPLAELETVTPEQAVAHPNWVMGRKIS VDSATMMNKGLEVIEACLLFNAPDDVQVVVHRQSVIHSMDY VDGTVLAQMGTDPDMRIPIAHALAWPDRFESGAESLDLFAVRQL NFERPDLARFFCLRLAYEAVGAGGTAPAILNAANETAVAAFLD RRLAFTGIPREVIENCMARVAPNAADATIESVLQADAETRKVAQK YIDDLRV	8
IspD	MSTDARFWIVVPAAGVGKRMGADIPKQYLDVAGKPVLOHTLER LLSVRRVTAVMVALGANDEFWPELPCSRPRVLATTGGRERAD SVLSALTALAGRAADGDWVLVHDAARLCVTRDDVERLMETLED DPVGGILALPVTDTLKTVENTIQGSADRSRVWRALTFQMFRY RALKEALEAAARRGLTVDASALELAGLSPRVVEGRFDNIKI TRPEDLPLAAFYLERQCFE	9
IspE	MDRESSVMKSPSLRLPAPAKLNLTLRITGRRPDGYHDLQTVF QFVDVCDWLEFRADASGEIRLQTSLAGVPAERNLIVRAARLLK EYAGVAAGADIVLEKNLPMGGGLGGSSNAATTLVALNRLWDL GLDRQTLNMLGLRLGADVPIFVFGGAWAEGVGERLQVLELPE FWYVIVVPPCHVSTAEIFNAPDLTRNDPITIAADFLAGSHQNH CLDAVVRYPVVGGEAMCVLGRYSRDVRLTGTGACVYSVHGSEE EAKAACDDLSRDWVAIVASGRNLSPLYEALNER	10
IspF	MFRIGQGYDAHRFKEGDHIVLCGVKIPFGRGFAAHSDGDVALH ALCDALLGAAALGDIGRHFPDTDARYKGIDSRVLLREVRQRIA SLGYTVGNVDVTVVAQAPRLAAHIQAMRENLAQDLEIPFDCVN VKATTTEGMGFEGRGEGISAHAVALLARR	11
IspG	MMNRKQTVGVRVGSVRIGGAPIVVQSMNTNTDTADVAGTVRQV IDLARAGSELVRITVNNEEAAEAVPRIREELDRQGCNVPLVGD FHFNGHKLLDKYPACAEALGKFRINPENVGRGSKRDPQFAQMI EFACRYDKFVRIGVNWGSLDQSVLARLLDENARLAEPRPLPEV MREAVITSALESABKAQGLGLPKDRIVLSCKMSGVQELISVYE ALSSRCDHALHLGLTEAGMGSKGIVASTAALSPLLQGGIGDTI RISLTPEPGADRSLEVIVAQEILQTMGLRSFTPMVISCPGCCR TTSDYFQKLAQQIQTHLRHKMPFWRRRYRGVEDMHVAVMGCVV NGPGESKNANIGISLPGTGEQFVAPVPFEDGVKTVTLKGDRIAE EFQELVERYIETHYGSRAEA	12

[0082]

基因名	氨基酸序列	SEQ ID NO.
IspH	MEIILANPRGFCAGVDRAIEIVDRAIEVFGAPIYVRHEVVHNE YVVDGLRERGAVFVEELSEVPENSTVIFSAHGVSKQIQEEARE RGLQVFDATCPLVTKVHIEVHQHASEGREIVFIGHAGHFEEVG TMGQYDNPAGGIYLVESPEDVEMLOVKNPDLAYVTQTTLSD DTGAVVEALKMRFPKILGPRKDDICYATQNRQDAVKKLAAQCD TILVVGSPNSSNSNRLREIADKLGRKAFLIDNAAQLTRDMVAG AQRIGVTAGASAPEILVQQVIAQLKEWGGRTATETQGIEEKVV FSLPKELRRLNA	13

[0083] 本领域技术人员应该理解的是被导入宿主嗜甲烷细菌中的每种 DXP 途径酶的来源可以是相同的,被导入宿主嗜甲烷细菌中的两种或多种 DXP 途径酶的来源可以是相同的,或被导入宿主嗜甲烷细菌中的每种 DXP 途径酶的来源可以彼此不同。DXP 途径酶的来源可以与 IspS 的来源相同或不同。在某些实施方案中,具有源于两种或多种来源的核酸的混合途径用来增大异戊二烯产量(参见,例如, Yang 等, 2012, PLoS ONE 7:e33509)。

[0084] 还可以期望的是通过凭借替代途径使异戊二烯前体 DMAPP 和 IPP 的合成增大来提高异戊二烯产量。作为实例, DMAPP 和 IPP 还可以通过甲羟戊酸途径进行合成(参见图 2)。不希望受理论的束缚,认为增加在细胞中 DMAPP 和 IPP 多肽的量可以增加生产的异戊二烯的量。目前,在已经完全被测序的几种嗜甲烷细菌中,尚未鉴定内源性甲羟戊酸途径。然而,在几种细菌物种中,已经鉴定甲羟戊酸途径。若在宿主嗜甲烷菌中不存在甲羟戊酸途径,可以期望的是导入构建用于生产 DMAPP 和 IPP 前体的甲羟戊酸途径所需的基因。若在宿主嗜甲烷菌中存在甲羟戊酸途径,可以期望的是导入或过表达某些甲羟戊酸途径基因以增大 DMAPP 和 IPP 的产量。在某些实施方案中,包括编码 IspS 的外源性核酸的非天然存在的嗜甲烷细菌,与原生的甲羟戊酸途径酶的正常表达水平相比,过表达内源性甲羟戊酸途径酶,表达转化的编码甲羟戊酸途径酶的外源性核酸,或其组合。在其它实施方案中,如本文提供的包括编码 IspS 的外源性核酸的非天然存在的嗜甲烷细菌,与内源性甲羟戊酸途径酶各自的正常表达水平相比,过表达一种、两种、三种、四种、五种、六种或更多种内源性甲羟戊酸途径酶;用编码一种、两种、三种、四种、五种、六种或更多种甲羟戊酸途径酶的外源性核酸进行转化;或二者兼具。

[0085] 将甲羟戊酸途径工程改造入嗜甲烷菌中或增强内源性甲羟戊酸途径可以通过增加 DMAPP 和 IPP 前体的供应来增大异戊二烯产量(参见,例如, Martin 等, 2003, Nature Biotechnol. 21:796-802)。在某些实施方案中,采用液相色谱法-质谱法的代谢物谱分析用来鉴定异戊二烯合成途径中的瓶颈和待过表达的酶(参见,例如, Pitera 等, 2007, 代谢工程 9:193-207)。过表达可以通过将编码内源性甲羟戊酸途径酶的核酸或编码甲羟戊酸途径酶的外源性(即,异源性)核酸导入宿主嗜甲烷细菌中来实现。作为实例,可以将编码内源性甲羟戊酸(mevlaonate)酶的核酸的拷贝与编码 IspS 的外源性核酸一起转化入宿主嗜甲烷细菌中,或可以将源于非-宿主嗜甲烷物种的编码甲羟戊酸酶的外源性核酸与编码 IspS 的外源性核酸一起转化入宿主嗜甲烷细菌中。内源性甲羟戊酸途径酶的过表达还可以通过用导致增强的转录的启动子或调控区替换内源性启动子或调控区来

实现。在某些实施方案中,在宿主嗜甲烷细菌中过表达的甲羟戊酸途径酶是 AACT、HMGS、HMGR、MK、PMK、MPD、IDI,或它们的组合。在一些实施方案中,将编码具有 Ala110Gly 取代(以增加反应速率)的多肽的突变型 HMGS 核酸导入宿主嗜甲烷细菌中(Steussy 等,2006, *Biochem.* 45:14407-14)。

[0086] 甲羟戊酸途径酶的来源是本领域中已知的,并且可以是来自天然具有甲羟戊酸途径的任何生物体,包括多种植物、动物、真菌、古细菌和细菌物种。例如,甲羟戊酸途径酶可以存在于 *Caldariella acidophilus*、红皮盐杆菌 (*Halobacterium cutirubrum*)、橙色粘球菌 (*Myxococcus fulvus*)、橙色绿屈挠菌 (*Chloroflexus aurantiacus*)、酿酒酵母 (*Saccharomyces cerevisiae*)、粟酒裂殖酵母 (*Schizosaccharomyces pombe*)、秀丽隐杆线虫 (*Caenorhabditis elegans*)、拟南芥、植物乳杆菌 (*Lactobacillus plantarum*)、金黄色葡萄球菌 (*Staphylococcus aureus*)、肉葡萄球菌 (*Staphylococcus carnosus*)、溶血性葡萄球菌 (*Staphylococcus haemolyticus*)、表皮葡萄球菌 (*Staphylococcus epidermidis*)、变异链球菌 (*Streptococcus mutans*)、肺炎链球菌 (*Streptococcus pneumoniae*)、化脓性链球菌 (*Streptococcus pyogenes*)、*Streptomyces aerioouifer*、伯氏疏螺旋体 (*Borrelia burgdorferi*)、*Chloropseudomonas ethylica*、橙色粘球菌、小眼虫 (*Euglena gracilis*)、粪肠球菌 (*Enterococcus faecalis*)、尿肠球菌 (*Enterococcus faecium*)、闪烁古生球菌 (*Archaeoglobus fulgidus*)、热自养甲烷杆菌 (*Methanobacterium thermoautotrophicum*)、詹氏甲烷球菌 (*Methanococcus jannaschii*)、智人 (*Homo sapiens*)、鸸鹋肠球菌 (*Enterococcus gallinarum*)、铅黄肠球菌 (*Enterococcus casseliflavus*)、格氏李斯特菌 (*Listeria grayi*)、马氏甲烷八叠球菌 (*Methanosarcina mazei*)、*Methanococcoides buronii*、清酒乳杆菌 (*Lactobacillus sakei*) 和链霉菌属 CL190 (*Streptomyces* CL190) 中。本领域技术人员应当理解的是被导入嗜甲烷宿主细菌中的每种甲羟戊酸途径酶的来源可以是相同的,两种或多种甲羟戊酸途径酶的来源可以是相同的,或每种甲羟戊酸途径酶的来源可以彼此不同。甲羟戊酸途径酶的来源可以与异戊二烯合成酶(例如, *IspS*) 的来源相同或不同。在某些实施方案中,具有源于两种或多种来源的核酸的混合途径用来增大异戊二烯产量(Yang 等,2012, *PLoS ONE* 7:e33509)。

[0087] 在某些实施方案中,包括编码 *IspS* 的外源性核酸的非天然存在的嗜甲烷细菌还可以包括如本文所述的基因修饰的 DXP 和甲羟戊酸途径。例如,如本文所述的非天然存在的嗜甲烷细菌,与内源性 DXP 途径酶的正常表达水平相比,可以过表达内源性 DXP 途径酶,表达转化的编码 DXP 途径酶的外源性核酸,或二者兼具;并且与原生的甲羟戊酸途径酶的正常表达水平相比,过表达内源性甲羟戊酸途径酶,表达转化的编码甲羟戊酸途径酶的外源性核酸,或二者兼具;或其任何组合。如前述,所有 DXP 和甲羟戊酸途径酶的来源可以是相同的,一些 DXP 或甲羟戊酸途径酶的来源可以是相同的,或 DXP 和甲羟戊酸途径酶的来源可以彼此都不同。

[0088] 还可以工程改造本公开的非天然存在的嗜甲烷细菌以包括变异体异戊二烯生物合成途径或酶。异戊二烯合成途径中的变化可以发生在途径的一个或多个的单个步骤或涉及整个新的途径。具体的途径反应可以被不同类别的酶催化,这些酶可能不具有与已知异戊二烯酶的序列、结构或催化相似性。例如,流产布鲁氏菌 (*Brucella abortus*) 2308 包含用于除了 DXR 之外的 DXP 途径的基因。相反,流产布鲁氏菌 2308 使用 DXR-样基因 (DRL)

来催化从 DXP 形成 2-C- 甲基 -D- 赤藓糖醇 -4- 磷酸 (MEP) (Sangari 等, 2010, Proc. Natl. Acad. Sci. USA 107:14081-14086)。在另一个实例中, 分别编码丙酮酸脱氢酶和 3, 4- 二羟基 -2- 丁酮 4- 磷酸合成酶的催化 E 亚单位的突变型 aceE 和 ribE 基因, 已经被鉴定各自能够救援 DXS- 缺陷的突变型细菌并且通过变异体 DXP 途径生产 DXP (Perez-Gil 等, 2012, PLoS ONE 7:e43775)。在又一个实例中, 也已经鉴定各种类型的异戊烯基二磷酸异构酶 (Kaneda 等, 2001, Proc. Natl. Acad. Sci. USA 98:932-7; Laupitz 等, 2004, Eur. J. Biochem. 271:2658-69)。除了 DXP 和甲羟戊酸途径之外, 还可以存在着可替代的异戊二烯合成途径 (参见, Poliquin 等, 2004, J. Bacteriol. 186:4685-4693; Ershov 等, 2002, J. Bacteriol. 184:5045-5051)。在某些实施方案中, 具体的途径反应被变异体或可替代的异戊二烯酶诸如 DRL、丙酮酸脱氢酶的催化 E 亚单位、3, 4- 二羟基 -2- 丁酮 4- 磷酸合成酶、变异体异戊烯基二磷酸异构酶, 或它们的任何组合催化。

[0089] 编码异戊二烯途径组分的核酸 (例如, 编码异戊二烯合成酶 (例如, IspS) 的核酸) 包括编码具有编码的异戊二烯途径多肽的至少一种活性 (例如, 将 DMAPP 转化成异戊二烯的能力) 的多肽、多肽片段、肽或融合多肽的核酸。本领域中已知的方法可以用来通过测量多肽将底物转化成产品的能力确定该多肽是否具有具体的活性 (参见, 例如, Silver 等, 1995, J. Biol. Chem. 270:13010-13016)。

[0090] 利用可获得的数以百计的生物体的全基因组序列, 在相关的或远系物种中编码异戊二烯合成酶及其它异戊二烯途径酶的基因包括例如同源基因 (homolog)、直向同源基因、横向同源基因等的鉴定是本领域中公知的。因此, 关于来自具体的生物体的具体的核酸, 本文所述的编码异戊二烯合成酶、DXS、DXR、IDI 等的外源性核酸可以容易地包括来自其它生物体的编码异戊二烯合成酶、DXS、DXR、IDI 等的其它核酸。

[0091] 本文所述的多肽序列以及对蛋白、蛋白结构域及它们的片段的编码核酸, 诸如异戊二烯合成酶及其它异戊二烯途径酶, 可以包括天然和重组工程的变异体。核酸变异体是指可以包含一种或多种取代、添加、缺失、插入的核酸, 或可以是或包含参考核酸的一个或多个片段的核酸。参考核酸是指编码具体的异戊二烯途径酶 (例如, IspS) 的选定的野生型或亲代核酸。变异体核酸与参考核酸可以具有 40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98% 或 99% 的序列同一性, 只要该变异体核酸编码仍然可以执行它的必要功能或生物活性 (例如, 对于 IspS, 将 DMAPP 转化成异戊二烯) 的多肽。变异体多肽与参考蛋白可以具有 50%、55%、60%、65%、70%、75%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98% 或 99% 的序列同一性, 只要该变异体多肽仍然可以执行它的必要功能或生物活性 (例如, 对于 IspS, 将 DMAPP 转化成异戊二烯)。在某些实施方案中, 被导入如本文提供的非天然存在的嗜甲烷细菌中的异戊二烯合成酶 (例如, IspS) 包含氨基酸序列, 该氨基酸序列与在 SEQ ID NO:1-6 中提供的氨基酸序列具有至少 50%、55%、60%、65%、70%、75%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98% 或 99% 的同一性。这些变异体与亲代 (或野生型) 蛋白相比可以具有改进的功能和生物活性 (例如, 较高的酶活性、对底物改进的特异性或较高的转换率)。由于遗传密码中的冗余, 核酸变异体可能影响或不影响氨基酸序列。

[0092] 核酸变异体还可以编码氨基酸序列, 该氨基酸序列与参考氨基酸序列相比包括一

种或多种保守取代。保守取代可以天然存在于多肽中（例如，天然存在的遗传变异体）或可以在重组生产多肽时被导入。保守取代是其中一个氨基酸被具有相似性质的另一个氨基酸取代，从而本领域技术人员将预期该多肽的二级结构及亲疏水性（hydropathic）基本上没有改变。氨基酸取代通常可以根据在残基的极性、电荷、溶解性、疏水性或两亲性方面的相似性进行，并且是本领域中已知的。

[0093] 利用公知的和常规操作的诱变方法可以将氨基酸取代、缺失和添加导入多肽中（参见，例如，Sambrook等Molecular Cloning:A Laboratory Manual, 第3版,Cold Spring Harbor Laboratory Press, NY2001）。寡核苷酸-定向的位点特异性（或区段特异性）诱变程序可以用来提供改变的多聚核苷酸，该多聚核苷酸具有根据期望的取代、缺失或插入而改变的具体密码子。蛋白的缺失或截短变异体还可以通过使用与期望的缺失相邻的方便的限制性内切核酸酶位点进行构建。或者，随机诱变技术诸如丙氨酸扫描诱变、易错聚合酶链反应诱变和寡核苷酸-定向诱变可以用来制备多肽变异体（参见，例如，Sambrook等，同上）。

[0094] 野生型（或亲代或参考）核酸或多肽与它们的变异体之间的差异可以通过确定同一性的已知方法进行确定，该方法设计用来给出在测试序列之间最大的匹配。可以从公众可获得的计算机程序应用确定序列同一性的方法。确定两个序列之间同一性的计算机程序方法包括，例如，BLASTP-BLASTN(Altschul, S. F. 等, J. Mol. Biol. 215:403-410(1990)，以及FASTA(Pearson 和 Lipman Proc. Natl. Acad. Sci. USA 85 ;2444-2448(1988)，并使用默认的参数。

[0095] 用于确定多肽变异体是否折叠成与非-变异体多肽或片段可比的构象的测定包括，例如，蛋白与对原生的或未折叠的表位特异性的单克隆抗体或多克隆抗体反应的能力、配体-结合功能的保留、酶活性的保留（若可适用），以及突变型蛋白对蛋白酶消化的敏感性或抗性（参见 Sambrook 等，同上）。可以制备多肽、变异体及它们的片段而不改变所得蛋白分子的生物活性（即，不得以统计学显著的或生物学显著的方式改变一种或多种功能活性）。例如，此类取代通常通过用包括在相同组诸如极性残基、带电荷残基、疏水性残基或小残基等的组内的另一种氨基酸互换氨基酸进行。任何氨基酸取代的效果可以凭经验仅仅通过测试所得修饰的蛋白在生物测定中的行使功能的能力或与同源配体或靶分子结合的能力进行确定。

[0096] 在某些实施方案中，被导入宿主嗜甲烷细菌中的编码 IspS 或其它异戊二烯途径酶的外源性核酸不包含 N-末端质体-靶向序列。通常叶绿体蛋白诸如许多植物异戊二烯合成酶及其它异戊二烯途径酶在细胞核中被编码并在细胞溶质中作为前体被合成。N-末端质体-靶向序列（也称为信号肽或过渡肽）编码靶向叶绿体被膜所需的信号，其在叶绿体输入后被肽酶切除。N-末端靶向序列的除去可以增强异源性核酸的表达。N-末端质体-靶向序列可以利用本领域中已知的预测程序包括 ChloroP(Emanuelsson 等, 1999, Protein Sci. 8:978-984) ;PLCR(Schein 等, 2001, Nucleic Acids Res. 29:e82) ;MultiP(<http://sbi.postech.ac.kr/MultiP/>) 进行确定。在导入嗜甲烷细菌中之前可以通过重组方法从核酸除去 N-末端质体靶向序列。在某些实施方案中，缺少 N-末端质体靶向序列的 IspS 的氨基酸序列在 SEQ ID NO:2、4 和 6 中的任一项中提供。在其它实施方案中，被导入宿主嗜甲烷细菌中的编码异戊二烯合成酶（例如，IspS）或其它异戊二烯途径酶的外源性核酸不包

括对其它细胞器例如质体样细胞器或内质网的靶向序列。

[0097] 在某些实施方案中,编码异戊二烯合成酶或其它异戊二烯途径酶的外源性核酸与表达控制序列可操作地连接。表达控制序列意指引导与它可操作连接的核酸的转录的核酸序列。表达控制序列包括启动子(例如,组成型启动子、渗漏型启动子或诱导型启动子)或增强子。在某些实施方案中,该表达控制序列是选自由甲醇脱氢酶启动子(MDH)、己酮糖 6-磷酸合成酶启动子、核糖体蛋白 S16 启动子、丝氨酸羟甲基转移酶启动子、丝氨酸乙醛酸氨基转移酶启动子、磷酸烯醇丙酮酸羧化酶启动子、T5 启动子和 Trc 启动子组成的组中的启动子。不希望受理论的束缚,甲醇脱氢酶启动子、己酮糖 6-磷酸合成酶启动子、核糖体蛋白 S16 启动子、丝氨酸羟甲基转移酶启动子、丝氨酸乙醛酸氨基转移酶启动子、磷酸烯醇丙酮酸羧化酶启动子、T5 启动子和 Trc 启动子提供允许异源性多肽在嗜甲烷细菌中的表达的不同强度的启动子。

[0098] 在某些实施方案中,编码 IspS 的核酸与诱导型启动子可操作地连接。诱导型启动子系统是本领域中已知的,并且包括四环素诱导型启动子系统;IPTG/lac 操纵子诱导型启动子系统、热激诱导型启动子系统;金属应答性启动子系统;硝酸盐诱导型启动子系统;光诱导型启动子系统;蜕皮激素诱导型启动子系统等。例如,非天然存在的嗜甲烷菌可以包括与侧接 lacO 操纵序列的启动子可操作地连接的编码异戊二烯合成酶(例如,IspS)的外源性核酸,并且还包括与组成型启动子(例如,己酮糖-6-磷酸合成酶启动子)可操作地连接的编码 lacI 阻遏蛋白的外源性核酸。LacI 阻遏蛋白与 IspS 启动子侧面的 lacO 操纵序列结合,以阻止转录。IPTG 结合 lacI 阻遏物并将它从 lacO 序列释放,以允许转录。通过利用诱导型启动子系统,可以通过添加诱导物控制异戊二烯合成。编码 IspS 或其它异戊二烯途径酶的核酸还可以与其它核酸序列、多聚腺苷酸化信号、限制性酶位点、多克隆位点、其它编码区段等组合。

[0099] 在某些实施方案中,利用本领域中已知的方法可以调节表达 DXP 途径酶和异戊二烯合成酶(例如,IspS)或甲羟戊酸途径酶和异戊二烯合成酶(例如,IspS)的强度和时机以提高异戊二烯产量。例如,变化的启动子强度或基因拷贝数可以用来调节表达水平。在另一个实例中,通过利用诱导型启动子系统或具有排列的基因顺序的多顺反子操纵子可以调节表达的时机。例如,DXP 途径酶和异戊二烯合成酶(例如,IspS)或甲羟戊酸途径酶和异戊二烯合成酶(例如,IspS)的表达可以在培养的生长期和静止期期间或仅在静止期期间进行。在另一个实例中,异戊二烯 DXP 途径酶和 IspS 或甲羟戊酸途径酶和 IspS 可以进行有序的共表达。已发现编码各种 DXP 途径酶的核酸的有序的共表达增加异戊二烯产量(Lv 等,2012,Appl. Microbiol. Biotechnol., “Significantly enhanced production of isoprene by ordered coexpression of genes dxs, dxr, and idi in Escherichia coli,” 网上发表于 2012 年 11 月 10 日)。

[0100] 密码子优化

[0101] 重组蛋白在它们的原宿主的外部通常难以表达。例如,在细菌的不同物种中已经观察到密码子使用偏向性的变化(Sharp 等,2005,Nucl. Acids. Res. 33:1141-1153)。重组蛋白的过表达即使在它们的原生宿主内也可能是困难的。在本发明的某些实施方案中,要被导入本发明的微生物中的核酸(例如,编码异戊二烯合成酶的核酸)可以进行密码子优化以增强蛋白表达。密码子优化是指改变用于转化生物体的基因中的密码子或核

酸的编码区从而反映宿主生物体的典型的密码子使用而不改变该 DNA 编码的多肽。在某些实施方案中,为在宿主嗜甲烷细菌中表达对编码 IspS、其它异戊二烯途径组分或蕃茄红素途径组分的外源性核酸进行密码子优化。用于在异源性生物体中最佳的基因表达的密码子优化方法是本领域中已知的并且以前已有描述(参见,例如, Welch 等, 2009, PLoS One 4:e7002 ;Gustafsson 等, 2004, Trends Biotechnol. 22:346-353 ;Wu 等, 2007, Nucl. Acids Res. 35:D76-79 ;Villalobos 等, 2006, BMC Bioinformatics 7:285 ;美国专利公布 2011/0111413 ;和美国专利公布 2008/0292918)。

[0102] 在表 3 中提供为在荚膜甲基球菌 Bath 菌株中表达而密码子优化的异戊二烯合成酶(例如, IspS)多聚核苷酸序列的实例。SEQ ID NO:15、17 和 19 是分别来自银白杨、山葛和柳树的没有它们的 N-末端质体-靶向序列的截短的 IspS 序列。SEQ ID NO:14、16 和 18 是分别来自银白杨、山葛和柳树的全长 IspS 序列(具有 N-末端质体-靶向序列)。

[0103] 表 3. 为在荚膜甲基球菌 Bath 菌株中表达而密码子-优化的 IspS 多聚核苷酸序列

[0104]

物种	核苷酸序列	SEQ ID NO.
银白杨	ATGGCCACTGAACTTCTTTGTTTGCACCGCCCGATTTCCTCC TGACCCATAAGCTGTTTCGCAACCTCTGCCCCAAAGTTAT CCAGGCAACCCCGCTGACGCTCAAGCTCCGGTGCAGCGTA TCCACCGAAAATGTATCGTTACCGAAACCGAAACTGAAG CCCGTGCGAGCGCGAACTACGAGCCCAACTCGTGGGATTA CGACTATCTGCTGAGCTCGGATACCGACGAATCCATCGAA GTCTATAAGGACAAAGCCAAGAAGCTCGAAGCCGAGGTGC GCCGTGAGATCAACAACGAGAAGGCCGAGTTCTCTGACCCT GTTGGAAGTGAATCGACAACGTCCAGCGCCTGGGCCCTCGGC TACCGGTTTCGAGAGCGATATCCGGGGTGCCCTGGACCGTT TCGTCAGCTCGGGCGGATTTCGACGCACTGACCAAAACGTC GCTGCATGGGACGGCCCTGTCTTCCGTCTGCTGCGCCAG CATGGCTTCGAGGTGTCCAGGAAGCCTTCAGCGGCTTCA AGGATCAGAACGGAACTTTCTGGAAAAGTTGAAAGAGGA CATCAAGGCCATCTCTCAGCCTGTACGAGGCGTCTCTCTG GCCCTCGAAGGTGAAAACATCTCTGATGAAGCCAAGGTGT TCGCAATCTCGCATCTTAAAGAGCTGTCCGAAGAGAAGAT TGGCAAAGAGCTGGCCGAACAAGTCAACCACGCGTTGGAG CTGCCGCTCCACCGCGCACCCAGCGGCTGGAAGCGGTCT GGTGATCGAAGCCTACCGCAAGAAAGAGGACGCCAATCA GGTCTGCTGGAGCTCGCGATTCTGGATTACAATATGATC CAGTCGGTCTATCAGCGCGATCTGCGCGAAACGTCCCGGT GGTGGCGGCGTGTGCGCTTGGCGACCAAGTTGCACTTCGC GCGTGACCGCTTGATCGAGAGCTTCTATTTGGGCCGTGGG GTGGCCTTTTGAGCCCCAGTACTCCGACTGCCGCAATAGCG TGGCGAAGATGTTTCAGCTTCGTTACCATCATCGACGACAT CTACGACGTGTATGGCACGCTCGACGAGCTCGAACTGTTTC ACCGACGCGGTGGAACGTTGGGACGTCAACGCCATCAATG ATCTCCCCGACTACATGAAGCTGTGCTTCTGGCGTTGTA TAACACCATCAACGAGATTGCCTACGATAACCTCAAGGAC AAGGGCGAGAACATCCTGCCGTACTTGACCAAGGCCCTGGG CCGATTTGTGCAACGCTTTCTGCGAGGAAGCAAAGTGGCT GTACAACAAATCCACGCCGACGTTTCGACGACTATTTCCGGC AATGCATGGAAATCGAGCTCGGGTCTCTGCAACTTGTGT TCGCGTACTTCGCCGTGTCGAGAAATATCAAGAAAGAAGA AATCGAGAACCTTCAGAAATATCATGACACCATCAGCCGT CCATCGCACATCTTTGCGCTGTGCAACGACCTCGCGTCCG CATCCGCCGAGATCGCACGCGGCGAAACGGCCAATTCGGT GTCTTGCTACATGCGGACCAAGGGCATCTCGGAAGAGCTG GCGACGGAATCCGTGATGAACCTGATCGATGAAACCTGGA AGAAGATGAACAAAGAGAAGCTCGGCCGGGAGCCTGTTCCG	14

[0105]

物种	核苷酸序列	SEQ ID NO.
	GAAGCCCTTCGTCGAAACCGCAATTAACCTGGCACGCCAA TCCCACTGTACCTACCATAACGGAGATGCCACACGAGCC CGGACGAGCTGACTCGCAAGCGCGTCCTTTTCGGTCATCAC CGAGCCGATCCTGDCGTTTCGAGCGGTAA	
银白杨 (截短的)	ATGTGCAGCGTATCCACCGAAAATGTATCGTTCACCGAAA CCGAAACTGAAGCCCGTTCGAGCGCGAACTACGAGCCCAA CTCGTGGGATTACGACTATCTGCTGAGCTCGGATACCGAC GAATCCATCGAAGTCTATAAGGACAAAGCCAAGAAGCTCG AAGCCGAGGTGCGCCGTGAGATCAACAACGAGAAGGCCGA GTTCTTGACCTGTGGAACGATCGACAACGTCCAGCGC CTGGGCCCTCGGCTACCGGTTTCGAGAGCGATATCCGGGGTG CCCTGGACCGTTTCGTCAGCTCGGGCGGATTTCGACGCAGT GACCAAAACGTGCTGTCATGGGACGGCCCTGTCTTCCTCGT CTGCTGCGCCAGCATGGCTTCGAGGTGTCCAGGAAGCCT TCAGCGGCTTCAAGGATCAGAACGGAACTTTCTGGAAAA CTTGAAAGAGGACATCAAGGCCATCCTCAGCCTGTACGAG GCGTCCTTCCTGGCCCTCGAAGGTGAAAACATCCTCGATG AAGCCAAGGTGTTTCGCAATCTCGCATCTTAAAGAGCTGTC CGAAGAGAAGATTGGCAAAGAGCTGGCCGAACAAGTCAAC CACGCGTTGGAGCTGCCGCTCCACCGGCGCAACCAGCGGC TGGAAGCGGTCTGGTCGATCGAAGCCTACCGCAAGAAAGA GGACGCCAATCAGGTCTCTGCTGGAGCTCGCGATTCTGGAT TACAATATGATCCAGTCCGCTATCAGCSCGATCTGCGCG AAACGTCCCGGTGSGTGGCGGCGTGTGCGCTTGGCGACCAA GTTGCACCTTCGCGCGTGACCGCTTGATCGAGAGCTTCTAT TGGGCCGTCGGGGTGGCCTTTGAGCCCCAGTACTCCGACT GCCGCAATAGCGTGGCGAAGATGTTTCAGCTTCGTTACCAT CATCGACGACATCTACGACGTGTATGGCACGCTCGACGAG CTCGAACTGTTTCACCGACGCCGTTGGAAAGTTGGGACGTCA ACGCCATCAATGATCTCCCCGACTACATGAAGCTGTGCTT CCTGGCGTTGTATAACACCATCAACGAGATTGCCTACGAT AACCTCAAGGACAAGGGCGAGAACATCCTGCCGTACTTGA CCAAGSCCTGGGCCGATTGTTGCAACGCCCTTTCTGCAGGA AGCAAAGTGGCTGTACAACAAATCCAAGCCGACGTTTCGAC GACTATTTTCGGCAATGCATGGAAATCGAGCTCGGGTCCTC TGCAACTTGTGTTTCGCGTACTTCGCCGTCGTGCAGATAT CAAGAAAGAAGAAATCGAGAACCTTCAGAAATATCATGAC ACCATCAGCCGTCCATCGCACATCTTTCGCCTGTGCAACG ACCTCGCGTCCGCATCCGCCGAGATCGCACGCGCGGAAAC GGCCAATTCCGGTGTCTGTACATGCCGACCAAGGGCATC TCGGAAGAGCTGGCGACGGAATCCGTGATGAACCTGATCG ATGAAACCTGGAAGAAGATGAACAAAGAGAAGCTCGGCGG GAGCCTGTTTCGCGAAGCCCTTCGTCGAAACCGCAATTAAC CTGGCACGCCAATCCCACTGTACCTACCATAACGGAGATG CCCACACGAGCCCGGACGAGCTGACTCGCAAGCCGCTCCT	15

[0106]

物种	核苷酸序列	SEQ ID NO.
山萸	TTCGGTCATCACCGAGCCGATCCTGCCGTTTCGAGCGGTAA ATGGCCACCAATCTGCTCTGCCTGTGGAATAAACTGTCCA GCCCCACGCCCCACGCCGTCCACGCGGPTCCCGCAGTCCAA GAACTTCATTACCCAGAAAACCAGCCTCCGCCAACCCGAAG CCATGGCGCGTGATCTGCGCAACCTCGTCCCAATTCACCC AGATCACGGAACACAACCTCGCGTCGCTCGGCCAACTACCA GCCTAATTTGTGGAACCTTCGAGTTCTTCGAGAGCTTGGAG AACGATCTGAAGGTCGAGAAGCTGGAAGAGAAAGCCACCA AGCTCGAAGAAGAGGTCCGTTGCATGATCAACCGCGTCGA CACTCAGCCGCTCTCCCTGCTGGAGCTTATCGACGACGTC CAGCGCCTCGGCTTGACTTACAAGTTCGAGAAAGACATTA TCAAGGCCCTTGAGAATATCGTCTCTGCTGGATGAAAACAA AAAGAACAAGTCGGATCTGCATGCGACCGCCCTGAGCTTC CGGCTGCTGCGCCAGCACGGCTTTGAGGTGAGCCCAAGACG TATTTCGAACGCTTCAAGGATAAAGAAGGCGGGTTTTCCGG CGAATTGAAAGGCGACGTGCAGGGCTTGCTCTCGCTGTAC GAGGCCAGCTACCTGGGCTTTGAGGGTGAAAATCTGCTCG AAGAGGCGCGTACCTTCAGCATCACGCATCTGAAGAATAA CCTCAAAGAGGGCATCAACACCAAGGTGGCCGAACAAGTG TCCCACGCGCTGGAACTGCCATACCATCAACGGCTGCATC GCCTGGAAGCGCGCTGGTTCTTGGAACAAGTATGAACCCAA AGAACCTCACCATCAGCTGCTTCTGGAGCTCGCCAAGTTG GACTTCAACATGGTCCAGACCTTGACCCAGAAAGAACTGC AGGACTTGTCCCGGTGGTGGACCGAAATGGGACTGGCGTC CAAGCTTGACTTCGTCCGCGATCGCCTCATGGAAGTGTAC TTTTGGGCCCTCGGAATGGCACCGGACCCGCAGTTCGGCG AGTGCCCGCAAAGCAGTTACCAAGATGPTCGGCCCTGGTCAC CATTATCGACGATGTCTACGACGTATACGGGACGTTGGAT GAGCTGCAACTGTTACGGACGCGGTGGAGCGGTGGGACG TCAACGCCATCAACACGCTCCCCGACTATATGAAGCTCTG CTTCTGGCATTTGTACAATACCGTGAACGACACCTCGTAT TCCATTCTGAAAGAAAAAGGACACAATAACCTGTCTATC TGACCAAGTCCTGGCGTGAGCTGTGCAAGGCGTTCTTGCA AGAAGCCAAGTGGAGCAATAACAAGATCATCCCCGCGTTC TCGAAGTATCTTGAGAACGCATCCGTGTGAGCAGCGGGG TCGCCCTGCTGGCCCCGTCGTACTTCAGCGTATGTCAGCA GCAGGAAGATATCTCGGACCACGCGCTGCGTAGCCTTACG GACTTCCATGGCCTCGTCCGGTCCGAGCTGCGTGATCTTCC GTTTGTGCAACGACCTGGCGACCTCGGCCGAGAACTGGA GCGGGGTGAAACCACCAACAGCATCATCTCGTACATGCAC GAGAACGATGGCACGTGGGAAGAGCAGGCACGCGAAGAGC TGCGTAAGCTGATCAGCGCCGAGTGGGAAGAAATGAACCG CGAACGCGTCAGCGACTCCACCCTGCTGCCGAAGGCTTC ATGGAAATCGCCGTGAACATGGCACGTGTGTCCCATTTGTA CTTATCAGTACGGCGATGGCCTGGGTGCCCCGACTATGC	16

[0107]

物种	核苷酸序列	SEQ ID NO.
	CACGGAGAACCGGATCAAGCTCCTGTTGATCGATCCGTTCCGATCAACCAGCTGATGTACGTGTAA	
山萸 (截短的)	ATGTGCGCAACCTCGTCCCAATTCACCCAGATCACGGAAC ACAACTCGCGTCGCTCGGCCAACTACCAGCCTAATTTGTG GAACTTCGAGTTCCTGCAGAGCTTGGAGAACGATCTGAAG GTCGAGAAGCTGGAAGAGAAAGCCACCAAGCTCGAAGAAG AGGTCCGTTGCATGATCAACCGCGTCGACACTCAGCCGCT CTCCCTGCTGGAGCTTATCGACGACGTCAGCGCCTCGGC TTGACTTACAAGTTCGAGAAAGACATTATCAAGGCCCTTG AGAATATCGTCCCTGCTGGATGAAAACAAAAGAACAAGTC GGATCTGCATGCGACCGCCCTGAGCTTCCGGCTGCTGCGC CAGCACGGCTTTGAGGTGAGCCAAGACGTATTGGAACGCT TCAAGGATAAAGAAGGCGGGTTTTCCGGCGAATTGAAAGG CGACGTGCAAGGCTTGCTCTCGCTGTACGAGGCCAGCTAC CTGGGCTTTGAGGGTGAAAATCTGCTCGAAGAGGCGCGTA CCTTCAGCATCACGCATCTGAAGAATAACCTCAAAGAGGG CATCAACACCAAGGTGGCCGAACAAGTGTCCACGCGCTG GAACTGCCATACCATCAACGGCTGCATCGCCTGGAAGCGC GCTGGTTCTTGGAACAAGTATGAACCCAAAGAACCTCACCA TCAGCTGCTTCTGGAGCTCGCCAAGTTGGACTTCAACATG GTCCAGACCTTGCACCAGAAAGAACTGCAGGACTTGTCCC GGTGGTGGACCGAAATGGGACTGGCGTCCAAGCTTGACTT CGTCCGCGATCGCCTCATGGAAGTGACTTTTGGGCCCTC GGAATGGCACCGGACCCGCAAGTTCCGGCGAGTGCCGCAAAG CAGTTACCAAGATGTTCCGGCCTGGTCACCATTATCGACGA TGTCTACGACGTATACGGGACGTTGGATGAGCTGCAACTG TTCACGACGCGCTGGAGCGGTGGGACGTCAACGCCATCA ACACGCTCCCGACTATATGAAGCTCTGCTTCTTGGCATT GTACAATACCGTGAACGACACCTCGTATTCCATTCTGAAA GAAAAAGGACACAATAACCTGTCTATCTGACCAAGTCCT GGCGTGAGCTGTGCAAGGCGTTCTGCAAGAAGCCAAGTG GAGCAATAACAAGATCATCCCCGCGTTCTCGAAGTATCTT GAGAACGCATCCGTGTCGAGCAGCGGGTCCGCTGCTGG CCCCCTCGTACTTCAGCGTATGTCAGCAGCAGGAAGATAT CTCGGACACGCGCTGCGTAGCCTTACGGACTTCCATGGC CTCGTCCGGTTCGAGCTGCGTGATCTTCCGTTTGTGCAACG ACCTGGCGACCTCGGCCCGCAGAACTGGAGCGGGGTGAAAC CACCAACAGCATCATCTCGTACATGCACGAGAACGATGGC ACGTCGGAAGAGCAGGCACGCCAAGAGCTGCGTAAGCTGA TCGACGCGAGTGAAGAAAATGAACCGCAACGCGTCAG CGACTCCACCCTGCTGCCGAAGGCCTTCATGGAAATCGCC GTGAACATGSCACGTGTGTCCATTGTACTTATCAGTACG GCGATGGCCTGGGTGCCCCGACTATGCCACGGAGAACCG GATCAAGCTCCTGTTGATCGATCCGTTCCCGATCAACCAG CTGATGTACGTGTAA	17

[0108]

物种	核苷酸序列	SEQ ID NO.
柳树	ATGGCCACTGAACTTCTGTGCTTGCACCGTCCCATTTCGC TCACCCCTAAACTGTTCGCAACCCGCTCCCGAAGGTAAT CCTGGCGACGCCGCTGACCCCTGAAGCTGCGGTGCAGCGTA TCCACCGAAAACGTGAGCTTTACTGAAACCGAAAACCGAAA CGCGTCGCTCGGCGAACTACGAACCCAATTCTCTGGGATTA TGACTACCTTCTGTGCTCGACACGGACGAGTCGATCGAG GTGTATAAGGATAAGGCCAAGAAGCTTGAGGCGGAAGTCC GTCCGGAGATCAACAACGAGAAGGCGGAGTTCTTGACGCT GCTCGAACTGATTGACAACGTCCAGCGCCTCGGCCCTGGGC TATCGCTTTCGAGTCCGATATCCGTCCGCGACTCGACCGCT TCGTTTCGTCCGCTGGCTTCGACGCACTGACGAAAACCTC GCTGCATGCCACCGCGCTGTGCTTCGCTTCTGCGCCAG CACGGATTTCGAGGTCAGCCAGGAAGCGTTCGGCGGGTTCA AGGACCAGAACGSGAATTTCTTGGAAAATCTGAAAGAAGA TATCAAAGCCATCTTGTCGCTGTACGAGGCGTCGTTTCTC GCGCTCGAAGGCGAGAACATTTCTCGACGAAGCGAAGGTGT TCGCCATCTCGACCTGAAAGAGCTCTCCGAAGAGAAGAT CGGCAAAGACTTGGCCGAGCAAGTCAATCACGCCCTGGAG TTGCCCTTGCATCGCCGCAACCAGCGCTTGGAAGCCGTTT GGAGCATTGAAGCCTATCGTAAGAAAGAGGACGCCAACCA AGTCTTGTGCTGGAGCTGGCCATCCTGGACTACAACATGATC CAGTCCGTGTACCAGCGGGACTTGGCGCGAAACCAGCCGGT GGTGGCGTCGCGTCGGCCTCGCCACCAAGCTGCACTTCGC ACGCGACCGCCTGATCGAGTCCTTCTACTGGGCCGTGGGC GTGCGATTTCGAGCCGCAATATAGCGACTGCCGGAACAGCG TGGCAAAGATGTTTCAGCTTCGTGACCATCATCGACGATAT CTATGACGTGTATGGGACGCTTGACGAACTGGAGCTGTTT ACGGATGCCGTGAGCGGTGGGACGTCAATGCCATCAACG ATTTGCCGGACTACATGAAGCTGTGCTTCTTGGCCTTGTA TAACACTATCAACGAGATCGCCTACGATAACCTGAAAGAA AAGGGTGAGAACATCCTGCCCTACCTCACCAAGGCCTGGG CCGACCTGTGTAACGCCTTTCTGCAGGAAGCCAAGTGGCT CTACAACAAGTCCACCCCAACCTTCGACGATTACTTCGGA AATGCCTGGAAGAGCAGCTCCGGACCTCTCCAGCTGGTGT TCGCATACTTCGCGCTCGTGCAGAACATCAAGAAAGAAGA GATCGAAAACCTTGCAAGTACCACGATATCATCAGCCGT CCCTCGCACATCTTCCGGCTCTGCAACGACCTTGCAAGCG CGTCCGCGGAGATCGCACGGGGCGAAACGGCCAACTCGGT GAGCTGCTACATGCGCACCAAGGGCATCTCGGAAGAAGTT GCGACGGAGTCCGTGATGAACCTTGATCGACGAAACCTGGA AGAAAATGAATAAAGAGAAACTCGGCCGCGAGCCTGTTCCC GAAGCCATTTCGTGAAACCGCCATCAACCTGGCGCGTCAG TCGCATTGCACCTACCATAATGGCGATGCCATACGTGCG CGGATGAACTGACCCGTAAGCGGGTCTGTCCGTGATCAC CGAGCCGATTCTGCGGTTTCGAGCGCTAA	18

[0109]

物种	核苷酸序列	SEQ ID NO.
柳树 (截短的)	ATGTGCAGCGTATCCACCGAAAAACGTGAGCTTTACTGAAA CCGAAACCGAAACGCGTCGCTCGGCGAACTACGAACCCAA TTCCTGGGATTATGACTACCTTCTGTCTGTCGACACGCGAC GAGTCGATCGAGGTGTATAAGGATAAGGCCAAGAAGCTTG AGGCGGAAGTCCGTCGGGAGATCAACAACGAGAAGGCGGA GTTCTTGACGCTGCTCGAACTGATTGACAACTCCAGCGC CTCGGCCTGGGCTATCGCTTCGAGTCCGATATCCGTCGCG CACTCGACCGCTTCGTTTTCGTCCGGTGGCTTCGACGCGAGT GACGAAAACCTCGCTGCATGOCACCGCGCTGTCTTCGCGC TTCCTGCGCCAGCAACGGATTTCGAGGTCAGCCAGGAAGCGT TCGGCGGGTTCAAGGACCAGAACGGGAATTTCTTGAAAA TCTGAAAGAAGATATCAAAGCCATCTTGTCTGCTGTACGAG GCGTCGTTTCTCGCGCTCGAAGGCGAGAACATTCTCGACG AAGCGAAGGTGTTTCGCCATCTCGCACCTGAAAGAGCTCTC CGAAGAGAAGATCGGCAAAGACTTGGCCGAGCAAGTCAAT CACGCCCTGGAGTTGCCCTGCATCGCCGCACCCAGCGCT TGGAAGCCGTTTGGAGCATTGAAGCCTATCGTAAGAAAGA GGACGCCAACCAAGTCTCTGCTGGAGCTGGCCATCCTGGAC TACAAATGATCCAGTCCGTGTACAGCGGGACTTGCCTCG AAACCAGCCGGTGGTGGCGTCGCGTCGGCCTCGCCACCAA GCTGCACCTTCGCACGCGACCGCCTGATCGAGTCCCTTCTAC TGGGCCGTGGGCGTCGCATTTCGAGCCGCAATATAGCGACT GCCGGAACAGCGTGGCAAAGATGTTTCAGCTTCGTGACCAT CATCGACGATATCTATGACGTGTATGGGACGCTTGACGAA CTGGAGCTGTTTACGGATGCCGTTCGAGCGGTGGGACGTCA ATGCCATCAACGATTTGCCGGACTACATGAAGCTGTGCTT CCTGGCCTTGTATAACACTATCAACGAGATCGCCTACGAT AACCTGAAAGAAAAGGGTGAGAACATCCTGCCCTACCTCA CCAAGGCCTGGGCGGACCTGTGTAAACGCCTTTCTGCAGGA AGCCAAGTGGCTCTACAACAAGTCCACCCCAACCTTCGAC GATTACTTCGGAAATGCCTGGAAGAGCAGCTCCGGACCTC TCCAGCTGGTGTTCGCATACTTCGCCGTTCGTGCAGAACAT CAAGAAAGAAGAGATCGAAAACCTTCAGAAAGTACCAAGAT ATCATCAGCCGTCCCTCGCACATCTTCCGGCTCTGCAACG ACCTTCGAAGCGCGTCCGCGGAGATCGCACGGGGCGGAAAC GGCCAACTCGGTGAGCTGCTACATGCGCACCAAGGGCATC TCGGAAGAACTTCGACGAGTCCGTCATGAACCTTGATCG ACGAAACCTGGAAGAAAATGAATAAAGAGAACTCGGCGG CAGCCTGTTCCCGAAGCCATTCGTGAAACCGCCATCAAC CTGGCGCGTCAGTCGCATTGCACCTACCATAATGGCGATG CCCATACGTCCCGGATGAACGTACCCGTAAGCGGGTCCT GTCCGTCATCACCGAGCCGATTCTGCCGTTTCGAGCGCTAA	19

[0110] 嗜甲烷菌的示例性培养

[0111] 利用本领域中公知的材料和方法可以培养如本文所述的非天然存在的嗜甲烷细菌。在某些实施方案中,在允许表达被导入宿主嗜甲烷菌细胞中的一种或多种核酸(例如, IspS)的条件下培养非天然存在的嗜甲烷细菌。

[0112] 经典的分批培养方法是封闭的系统,其中培养基的组成在培养开始时设定并且在

培养过程期间不受外部改变。因此,在培养过程开始时,使培养基接种期望的生物体或生物体,并且在不向该系统添加任何物质的情况下允许发生生长或代谢活动。然而,通常“分批”培养是相对于添加碳源而言的分批,并且通常试图控制诸如 pH 和氧浓度之类的因素。在分批系统中,该系统的代谢物和生物质组成不断变化直至培养终止之时。在分批培养物内,细胞经由静滞期至高对数生长期并最终至静止期变缓和,在静止期中生长速率减小或停止。若不处理,处于静止期的细胞最终将死亡。处于对数期的细胞通常负责一些系统中的最终产品或中间体的批量生产。在其它系统中可以获得静止期或指数期后的生产。

[0113] 补料分批系统是对标准分批系统的变型。补料分批培养过程包括具有随着培养进程递增地添加底物的修改的典型分批系统。当分解代谢物阻遏作用倾向于抑制细胞的代谢时以及在期望在培养基中具有有限量的底物的情况中,补料分批系统是有用的。在补料分批系统中实际底物浓度的测量是困难的,并因此根据可测量因素诸如 pH、溶解氧和废气诸如 CO₂ 的分压的变化进行估计。分批培养方法和补料分批培养方法在本领域中是普遍的并且已知的(参见,例如,Thomas D. Brock, *Biotechnology: A Textbook of Industrial Microbiology*, 第 2 版 (1989) Sinauer Associates, Inc., Sunderland, MA; Deshpande, 1992, *Appl. Biochem. Biotechnol.* 36:227, 其通过引用整体并入)。

[0114] 连续培养是“开放”系统,其中将限定的培养基连续地添加到生物反应器并同时为了加工除去等量的条件培养基。连续培养通常使细胞维持在恒定的高液相密度,其中细胞主要处于对数期生长。或者,可以用固定化细胞实施连续培养,其中连续地添加碳和营养物并从细胞群连续地除去有价值的产品、副产物和废品。利用由天然材料或合成材料构成的各种固体支持体可以执行细胞固定化。

[0115] 连续培养或半连续培养允许调节影响细胞生长或最终产品浓度的一个因素或任何数量的因素。例如,一种方法可使有限的营养物诸如碳源或氮水平维持在固定比率并允许调节所有其它参数。在其它系统中,可以连续地改变影响生长的多种因素同时使通过培养基浊度测得的细胞浓度保持恒定。连续系统力求维持稳定的状态生长条件,因此必须使因培养基被取出所致的细胞损失与培养物中细胞生长速率保持平衡。调节用于连续培养过程的营养物和生长因素的方法以及用于最大化产品形成的速率的技术是本领域中公知的,并且 Brock 详述了各种方法,参见同上。

[0116] 还可以将嗜甲烷细菌固定化在固体底物上作为全细胞催化剂并经受用于异戊二烯生产的发酵条件。

[0117] 本公开中提供的嗜甲烷细菌可以作为分离的纯培养物生长,可以与可辅助生长的异源性非嗜甲烷生物体或嗜甲烷细菌的一种或多种不同的菌株 / 或物种合并以产生混合的培养物。

[0118] 任何碳源,能够被嗜甲烷细菌代谢的含碳化合物(也称为碳原料),可以用来培养本文所述的非天然存在的嗜甲烷细菌。碳原料可以用于维持活力,使嗜甲烷细菌生长,或被转化成异戊二烯。

[0119] 在某些实施方案中,如本文所述用一种或多种异戊二烯途径酶基因工程改造的非天然存在的嗜甲烷细菌能够将碳原料转化成异戊二烯,其中该碳原料是 C₁ 底物。C₁ 底物包括但不限于甲烷、甲醇、天然气和非常规天然气。非天然存在的嗜甲烷细菌还可以将非-C₁ 底物诸如多-碳底物转化成异戊二烯。一旦已经将异戊二烯生物合成能力导入细

菌中,则非天然存在的嗜甲烷细菌可以内源地具有将多-碳底物诸如低碳烷烃(乙烷、丙烷和丁烷)转化成异戊二烯的能力(参见图3)。或者,非天然存在的嗜甲烷细菌可以需要附加的基因工程以使用替代碳原料(参见,例如,2012年10月24日提交的美国临时申请61/718,024,“Engineering of Multi-Carbon Substrate Utilization Pathways in Methanotrophic Bacteria”,其通过引用整体并入),然后可以根据本公开将其转化成异戊二烯。可以向嗜甲烷细菌提供纯的或相对纯的主要由单碳底物诸如甲烷或干天然气组成的碳原料。还可以向嗜甲烷细菌提供混合的碳原料诸如湿天然气,其包括甲烷和低碳烷烃。

[0120] 非天然存在的嗜甲烷细菌的构建

[0121] 本文使用的重组DNA和分子克隆技术是本领域中公知的,在Sambrook等, *Molecular Cloning: A Laboratory Manual*, 第3版, Cold Spring Harbor Laboratory, New York (2001); 和 Ausubel 等, *Current Protocols in Molecular Biology*, John Wiley and Sons, Baltimore, MD (1999) 中有描述。

[0122] 用于将异源性核酸导入嗜甲烷细菌中的重组方法是本领域中已知的。用于在嗜甲烷细菌中表达异源性核酸的表达系统和表达载体是已知的。用于转化嗜甲烷细菌的载体或盒是已知的。

[0123] 代谢 C1 细菌的电穿孔之前已经在 Toyama 等, 1998, *FEMS Microbiol. Lett.* 166:1-7 (扭脱甲基杆菌 (*Methylobacterium extorquens*)); Kim 和 Wood, 1997, *Appl. Microbiol. Biotechnol.* 48:105-108 (甲基营养型嗜甲基菌 (*Methylophilus methylotrophus*) AS1); Yoshida 等, 2001, *Biotechnol. Lett.* 23:787-791 (甲基杆菌 (*Methylobacillus* sp.) 菌株 12S), 和 US2008/0026005 (扭脱甲基杆菌) 中有描述。

[0124] 细菌接合, 其是指涉及供体细胞与受体细胞的直接接触的特定类型的转化, 更频繁地用于将核酸转移入嗜甲烷细菌中。细菌接合涉及将“供体”细胞与“受体”细胞以彼此紧密接触方式混合在一起。通过在供体细菌与受体细菌之间形成细胞质连接发生接合, 并且新合成的供体核酸单向转移入受体细胞中。在接合反应中受体是可以通过水平转移从供体细菌接受核酸的任何细胞。在接合反应中供体是包含接合质粒、接合转座子或移动的质粒的细菌。通过自我转移性 (self-transmissible) 质粒或借助于“辅助”质粒可以发生供体质粒的物理转移。涉及代谢 C1 的细菌包括嗜甲烷细菌的接合之前已经在 Stolyar 等, 1995, *Mikrobiologiya* 64:686-691; Martin 和 Murrell, 1995, *FEMS Microbiol. Lett.* 127:243-248; Motoyama 等, 1994, *Appl. Micro. Biotech.* 42:67-72; Lloyd 等, 1999, *Archives of Microbiology* 171:364-370; 和 Odom 等, PCT 公布 WO 02/18617; Ali 等, 2006, *Microbiol.* 152:2931-2942 中有描述。

[0125] 如本文所述, 可以期望的是过表达各种上游异戊二烯途径基因以增大产量。利用本领域中已知的方法, 诸如多拷贝质粒或强启动子, 可以实现内源性核酸或异源性核酸的过表达。在嗜甲烷菌中使用多拷贝表达系统是本领域中已知的(参见, 例如, Cardy 和 Murrell, 1990 *J. Gen. Microbiol.* 136:343-352; Sharpe 等, 2007, *Appl. Environ. Microbiol.* 73:1721-1728)。例如, 在甲基杆菌中异源性基因的基于转座子的多拷贝表达已有描述(参见, 例如美国专利公布 2008/0026005)。还可以使用对外源性核酸的高表达而言适合的同源性启动子或异源性启动子。例如, 美国专利 7,098,005 描述在甲烷或甲醇的存在下高度表达的启动子用于在嗜甲烷细菌中异源性基因表达。可以使用的附加的启

动子包括：脱氧-木酮糖磷酸合成酶甲醇脱氢酶操纵子启动子 (Springer 等, 1998, FEMS Microbiol. Lett. 160:119-124)；用于 PHA 合成的启动子 (Foellner 等 1993, Appl. Microbiol. Biotechnol. 40:284-291)；或从甲基营养生物 (methylophil) 中的原生质粒鉴定的启动子 (EP296484)。可以使用的非-原生的启动子包括：lac 操纵子 Plac 启动子 (Toyama 等, 1997, Microbiology 143:595-602) 或杂合启动子诸如 Ptrc (Brosius 等, 1984, Gene 27:161-172)。可以使用的附加的启动子包括渗漏型启动子或诱导型启动子系统。例如，在甲基营养型细菌和嗜甲烷细菌中重组蛋白表达的阻遏物 / 操纵基因系统在美国专利 8, 216, 821 中已有描述。

[0126] 或者，某些基因的中断可以适合于消除竞争能量或碳汇，增加异戊二烯途径前体的累积，或防止异戊二烯的进一步代谢。用于中断的基因的选择可以根据经验证据进行确定。用于中断的候选基因可以包括 IspA。已知嗜甲烷细菌具有可以为异戊二烯前体 DMAPP 和 IPP 而竞争的类胡萝卜素生物合成途径（参见，美国专利 6, 969, 595）。IspA 是指催化三个异戊二烯基转移酶反应的序列的香叶基转移酶或法尼基二磷酸合成酶，其中从 DMAPP 和 IPP 形成香叶基二磷酸 (GPP)、法尼基二磷酸 (FPP) 和香叶基香叶基二磷酸 (GGPP)。用于在嗜甲烷细菌中将内源性基因功能下调、失活、敲除或缺失的各种方法是本领域中已知的。例如，靶向基因中断是一种用于基因下调的有效方法，其中将外源 DNA 插入结构基因中以中断转录。将基因盒导入宿主嗜甲烷细菌中，该基因盒包括侧接与要被中断的靶宿主基因的部分具有高度同源性的序列的外源插入 DNA（通常遗传标记）。外源 DNA 通过原生 DNA 复制机制中断靶宿主基因。等位交换从而在代谢 C₁ 的细菌（包括嗜甲烷细菌）中构建缺失 / 插入突变型已经在 Toyama 和 Lidstrom, 1998, Microbiol. 144:183-191；Stolyar 等, 1999, Microbiol. 145:1235-1244；Ali 等, 2006, Microbiology 152:2931-2942；Van Dien 等, 2003, Microbiol. 149:601-609；Martin 和 Murrell, 2006, FEMS Microbiol. Lett. 127:243-248 中有描述。

[0127] 被转化入宿主嗜甲烷细菌中的核酸，诸如编码 IspS、DXP 途径酶、甲羟戊酸途径酶或蕃茄红素途径酶的核酸，可以作为独立的核酸分子被导入在多顺反子核酸分子上，在编码融合蛋白的单个核酸分子上，或其组合。若将多于一种核酸分子导入宿主嗜甲烷细菌中，可以根据相关的代谢途径将它们以各种顺序（包括随机的顺序或相继的顺序）导入。在某些实施方案中，在将编码来自选定生物合成途径的多种酶的多种核酸转化入宿主嗜甲烷细菌中时，以保留与选定的生物合成途径一致的相继的顺序的方式将它们转化。

[0128] 生产异戊二烯的方法

[0129] 本文提供用于生产异戊二烯的方法，其包括：在碳原料存在下在足以生产异戊二烯的条件下培养包含编码异戊二烯合成酶的外源性核酸的非天然存在的嗜甲烷细菌。用于生长和维持嗜甲烷细菌培养物的方法是本领域中公知的。本文所述的非天然存在的嗜甲烷细菌的各种实施方案可以用于生产异戊二烯的方法中。

[0130] 在某些实施方案中，在细胞生长的特定时期（例如，延滞期、对数期、静止期或死亡期）期间生产异戊二烯。可以期望的是来自原料的碳被转化成异戊二烯而不是嗜甲烷细菌的生长和维持。在一些实施方案中，将本文提供的非天然存在的嗜甲烷细菌培养为低至中等的细胞密度 (OD₆₀₀)，然后开始异戊二烯的生产。在一些实施方案中，生产异戊二烯同时嗜甲烷细菌不再分裂或非常缓慢地分裂。在一些实施方案中，仅在静止期期间生产异戊二

烯。在一些实施方案中,在对数期和静止期期间生产异戊二烯。

[0131] 包括由本文提供的非天然存在的嗜甲烷细菌生产的异戊二烯的发酵器废气还可以包括与生物发酵过程关联的其它有机化合物。例如,发酵的生物副产物可以包括下列中的一种或多种:醇类、环氧化物类、醛类、酮类和酯类。在某些实施方案中,发酵器废气可以包含以下醇类中的一种或多种:甲醇、乙醇、丁醇或丙醇。在某些实施方案中,发酵器废气可以包含以下环氧化物类中的一种或多种:环氧乙烷、环氧丙烷或环氧丁烷。其它化合物诸如 H_2O 、 CO 、 CO_2 、 CO 、 N_2 、 H_2 、 O_2 以及未使用的碳原料诸如甲烷、乙烷、丙烷和丁烷也可以存在于发酵器废气中。

[0132] 在某些实施方案中,本文提供的非天然存在的嗜甲烷细菌以约 0.001g/L 的培养物至约 500g/L 的培养物生产异戊二烯。在一些实施方案中,生产的异戊二烯的量是约 1g/L 的培养物至约 100g/L 的培养物。在一些实施方案中,生产的异戊二烯的量是约 0.001g/L、0.01g/L、0.025g/L、0.05g/L、0.1g/L、0.15g/L、0.2g/L、0.25g/L、0.3g/L、0.4g/L、0.5g/L、0.6g/L、0.7g/L、0.8g/L、0.9g/L、1g/L、2.5g/L、5g/L、7.5g/L、10g/L、12.5g/L、15g/L、20g/L、25g/L、30g/L、35g/L、40g/L、45g/L、50g/L、60g/L、70g/L、80g/L、90g/L、100g/L、125g/L、150g/L、175g/L、200g/L、225g/L、250g/L、275g/L、300g/L、325g/L、350g/L、375g/L、400g/L、425g/L、450g/L、475g/L 或 500g/L。

[0133] 通过碳指纹图谱,可以将利用本文提供的组合物和方法生产的异戊二烯与由石油化学品生产的异戊二烯或由非嗜甲烷细菌生物合成的异戊二烯区分开。作为背景,稳定的同位素测量法和质量平衡方法广泛用来评价甲烷的全球来源和吸收汇 (sink) (参见 Whiticar 和 Faber, *Org. Geochem.* 10:759, 1986; Whiticar, *Org. Geochem.* 16:531, 1990)。由甲烷的微生物氧化所致的碳同位素分馏的程度的测量可以通过测量残余甲烷的同位素特征值 (isotopic signature) (即,稳定的同位素之比 $^{13}C:^{12}C$) 进行确定。例如,需氧嗜甲烷菌可以通过特定的酶,甲烷单加氧酶 (monooxygenase) (MMO) 代谢甲烷。嗜甲烷菌将甲烷转化成甲醇并随后转化为甲醛。可以将甲醛进一步氧化成 CO_2 从而以还原当量 (NADH) 的形式向细胞提供能量,或通过 RuMP 或丝氨酸循环被并入生物质中 (Hanson 和 Hanson, *Microbiol. Rev.* 60:439, 1996), 其与光合生物体中的碳同化途径直接类似。更具体地, I 型嗜甲烷菌使用 RuMP 途径用于生物质合成并完全从 CH_4 产生生物质,而 II 型嗜甲烷菌使用丝氨酸途径,同化来自 CH_4 的细胞碳的 50-70% 和来自 CO_2 的细胞碳的 30-50% (Hanson 和 Hanson, 1996)。用于测量碳同位素组成的方法在例如, Templeton 等 (*Geochim. Cosmochim. Acta* 70:1739, 2006) 中提供,其方法在此通过引用整体并入。异戊二烯的 $^{13}C/^{12}C$ 稳定的碳同位素比 (以千分比‰报告为 $\delta^{13}C$ 值) 随着所用的 C_1 底物的来源和纯度而改变 (参见, 例如, 图 4)。

[0134] 例如,源于石油的异戊二烯具有约 -22‰至约 -24‰的 $\delta^{13}C$ 分布。主要从玉米衍生的葡萄糖 ($\delta^{13}C$ -10.73‰) 生物合成的异戊二烯具有约 -14.66‰至 -14.85‰的 $\delta^{13}C$ 。从可再生的碳源生物合成的异戊二烯预期具有与源于石油的异戊二烯相比较小负值的 $\delta^{13}C$ 值。然而,来自天然气的甲烷的 $\delta^{13}C$ 分布有别于大多数碳源,并具有比原油更大负值的 $\delta^{13}C$ 分布。嗜甲烷细菌表现出在过量甲烷的条件下优选使用 ^{12}C 并减少它们摄取的 ^{13}C , 从而导致 $\delta^{13}C$ 值进一步负迁移。如本文所述的嗜甲烷细菌生产的异戊二烯具有与来自原油或可再生的碳源的异戊二烯相比更大负值的 $\delta^{13}C$ 分布,介于约 -30‰至约 -50‰的范围内。

在某些实施方案中,异戊二烯组合物具有小于约-30%、-40%或-50%的 $\delta^{13}\text{C}$ 分布。在某些实施方案中,异戊二烯组合物具有约-30%至约-40%或约-40%至约-50%的 $\delta^{13}\text{C}$ 分布。

[0135] 在某些实施方案中,异戊二烯组合物具有小于约-30%、-40%或-50%的 $\delta^{13}\text{C}$ 分布。在某些实施方案中,异戊二烯组合物具有约-30%至约-40%,或约-40%至约-50%的 $\delta^{13}\text{C}$ 分布。在其它实施方案中,异戊二烯组合物具有小于-30%、小于-31%、小于-32%、小于-33%、小于-34%、小于-35%、小于-36%、小于-37%、小于-38%、小于-39%、小于-40%、小于-41%、小于-42%、小于-43%、小于-44%、小于-45%、小于-46%、小于-47%、小于-48%、小于-49%、小于-50%、小于-51%、小于-52%、小于-53%、小于-54%、小于-55%、小于-56%、小于-57%、小于-58%、小于-59%、小于-60%、小于-61%、小于-62%、小于-63%、小于-64%、小于-65%、小于-66%、小于-67%、小于-68%、小于-69%或小于-70%的 $\delta^{13}\text{C}$ 分布。

[0136] 测量异戊二烯产量

[0137] 利用本领域中已知的方法可以测量异戊二烯产量。例如,来自发酵器气体的废气的样品可以通过配有火焰离子化检测器和为检测短链烃而选择的柱的气相色谱进行分析(Lindberg等,2010, Metabolic Eng. 12:70-79)。生产的异戊二烯的量可以通过与纯的异戊二烯标准比较进行估计。Silver等, J. Biol. Chem. 270:13010, 1995, 美国专利 5,849,970 以及其中引用的文献描述了利用配有氧化汞气体检测器的气相色谱测量异戊二烯产量的方法,该方法在此通过引用整体并入。

[0138] 异戊二烯的回收和纯化

[0139] 在某些实施方案中,本文所述的方法中的任一种方法还可以包括回收或纯化由宿主嗜甲烷细菌生产的异戊二烯。虽然以下描述的示例性回收和纯化方法是指异戊二烯,但是它们也可以应用于类异戊二烯或源于异戊二烯的其它化合物。

[0140] 利用本公开中提供的组合物和方法生产的异戊二烯可以通过使气流(例如,氮气、空气)起泡经过生产异戊二烯的嗜甲烷菌的培养物从发酵系统进行回收。改变发酵培养基的气体鼓泡速率以使异戊二烯在发酵废气中的浓度增加的方法是本领域中已知的。利用本领域中已知的技术,诸如气提(gas stripping)、蒸馏、聚合物膜增强分离、分馏、全蒸发、吸附/解吸(例如,硅胶、碳筒(carbon cartridge))、从固相热解吸或真空解吸异戊二烯,或用溶剂萃取被固定化或吸附于固相的异戊二烯,进一步回收和纯化异戊二烯(参见,例如,美国专利 4,703,007、美国专利 4,570,029、美国专利 4,147,848、美国专利 5,035,794、PCT 公布 W02011/075534,来自其中每篇的方法在此通过引用整体并入)。用醇(例如,乙醇、甲醇、丙醇或它们的组合)萃取蒸馏可以用来回收异戊二烯。异戊二烯回收可以涉及以液体形式(例如,异戊二烯的纯溶液或异戊二烯与溶剂的溶液)分离异戊二烯。以气体形式回收异戊二烯可以包括气提,其中以连续的方式从发酵废气除去异戊二烯蒸气。利用各种方法,包括,例如,吸附至固相、分配入液相或直接冷凝,可以实现气提。高于蒸气露点的稀释异戊二烯蒸汽流的膜富集也可以用来冷凝液体异戊二烯。异戊二烯气体还可以被压缩和冷凝。

[0141] 异戊二烯的回收和纯化可以包括一个或多个步骤。回收和纯化方法可以单独地或联合使用以获得高纯度异戊二烯。在一些实施方案中,从发酵废气除去异戊二烯气体和转

化成液相同时执行。例如,可以将异戊二烯从废气流直接冷凝成液体。在其它实施方案中,从发酵废气除去异戊二烯气体和转化成液相相继地执行(例如,可以将异戊二烯吸附至固相,然后用溶剂萃取)。

[0142] 在某些实施方案中,利用本文所述的组合物和方法从培养系统回收的异戊二烯受到进一步纯化(例如,从一种或多种在异戊二烯生产期间存在于异戊二烯液体或蒸气中的非异戊二烯组分分离)。在某些实施方案中,异戊二烯是基本上纯化的液体。纯化方法是本领域中已知的,并且包括萃取蒸馏和色谱法,并且纯度可以通过诸如柱色谱法、HPLC 或 GC-MS 分析的方法进行评估。在某些实施方案中,异戊二烯具有按重量计至少 20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95% 或 99% 的纯度。

[0143] 在某些实施方案中,将在异戊二烯回收的一个或多个步骤后剩余的气相的至少一部分反向再循环入发酵系统中。

[0144] 异戊二烯的进一步加工

[0145] 利用本领域中已知的方法可以将利用本文所述的组合物和方法生产的异戊二烯进一步加工成其它高价值产品。在回收或纯化后,利用各种催化剂可以使异戊二烯聚合而形成各种聚异戊二烯异构体(Senyei, "Isoprene Polymers", Encyclopedia of Polymer Science and Technology, 2002, John Wiley & Sons, Inc.)。还可以使异戊二烯与苯乙烯或丁二烯聚合而形成各种弹性体。由过氧化氢引发的异戊二烯的光化聚合形成羟基封端的聚异戊二烯,其可以用作压敏粘合剂。异戊二烯调聚反应产品也可用作燃料(Clement 等, 2008, Chem. Eur. J. 14:7408-7420 ; Jackstell 等, 2007, J. Organometallic Chem. 692:4737-4744)。利用催化剂还可以将异戊二烯化学修饰成二聚体(10 个碳)和三聚体(15 个碳)烯烃(Clement 等, 2008, Chem. Eur. J. 14:7408-7420 ; Gordillo 等, 2009, Adv. Synth. Catal. 351:325-330)。可以将烯烃氢化而形成长链支链烷烃,其可以用作燃料或溶剂。可以将异戊二烯转化成类异戊二烯化合物,诸如萜、银杏苦内酯、甾醇或类胡萝卜素。还可以将异戊二烯转化成基于类异戊二烯的生物燃料,诸如法呢烷、没药烷、松油萜、异戊醇或它们的任何组合(Peralta-Yahya 等, 2012, Nature 488:320-328)。

[0146] 筛选具有增加的异戊二烯途径前体的突变型的方法

[0147] 可以在宿主嗜甲烷细菌中诱导基因组或基因特异性突变以试图提高异戊二烯前体的产量。诱发基因组突变的方法是本领域中已知的(参见,例如,Thomas D. Brock, Biotechnology: A Textbook of Industrial Microbiology, 第 2 版 (1989) Sinauer Associates, Inc., Sunderland, MA ; Deshpande, 1992, Appl. Biochem. Biotechnol. 36:227),并且包括例如,UV 辐射、化学诱变(例如,吡啶染料、 HNO_2 、 NH_2OH)和转座子诱变(例如, Ty1、Tn7、Tn5)。随机诱变技术,例如易错 PCR、滚环易错 PCR 或突变基因菌株,可以用来创建特定基因或基因集的随机突变型库。定向诱变还可以用来创建特定基因或基因集的突变型库。

[0148] 本公开提供用于通过将蕃茄红素途径工程改造入嗜甲烷细菌中来筛选具有提高的异戊二烯前体产量的突变型嗜甲烷菌菌株的方法。蕃茄红素和异戊二烯合成途径使用相同的通用前体,异戊烯基二磷酸(IPP)和二甲基烯丙基二磷酸(DMAPP)(参见图 1 和 6);蕃茄红素和异戊二烯生物合成共同具有大部分的 DXP 途径。根据细菌的增强的红色着色测量

导致提高的番茄红素 (lycoprene) 产量的有益基因组突变还可以通过增加 IPP 和 DMAPP 产量导致提高的异戊二烯合成 (若该突变影响重叠的途径组分)。

[0149] 在某些实施方案中,用于筛选突变型嗜甲烷细菌的方法包括:(a) 使嗜甲烷细菌暴露于诱变剂以生产突变型嗜甲烷细菌;(b) 用编码香叶基香叶基二磷酸合成酶 (GGPPS)、八氢番茄红素合成酶 (CRTB) 和八氢番茄红素脱氢酶 (CRTI) 的外源性核酸转化该突变型嗜甲烷细菌;和 (c) 在足够生长的条件下培养该突变型嗜甲烷细菌;其中与已用 GGPPS、CRTB 和 CRTI 转化并尚未暴露于诱变剂的参考嗜甲烷细菌相比,呈现出增强的红色着色的突变型嗜甲烷细菌,表明具有增强的红色着色的该突变型嗜甲烷细菌呈现出与该参考嗜甲烷细菌相比增高的异戊二烯前体合成。在某些实施方案中,异戊二烯前体是 IPP 或 DMAPP。在一些实施方案中,该诱变剂是辐射物、化学品、质粒或转座子。然后可以用如本文所述的异戊二烯生物合成途径对通过增加的蕃茄红素途径活性被鉴定为具有增加的异戊二烯前体产量的突变型嗜甲烷细菌进行工程改造。在一些实施方案中,将具有增强的红色着色的突变型嗜甲烷细菌或它的克隆细胞用编码异戊二烯合成酶 (例如, IspS) 的外源性核酸进行转化。在某些实施方案中,在用编码 IspS 的核酸转化之前或之后,从被鉴定为具有增加的异戊二烯前体产量的突变型嗜甲烷细菌将蕃茄红素途径基因 (GGPPS、CRTB 和 CRTI) 中的至少一种、两种或全部除去或灭活。功能性蕃茄红素途径与功能性异戊二烯途径的共表达可竞争共同具有的前体 DMAPP 和 IPP,并且可能降低异戊二烯产量。然后可以将通过本文所述的筛选方法鉴定的突变型嗜甲烷细菌中的异戊二烯产量与具有确认增加异戊二烯水平的异戊二烯生物合成能力的参考嗜甲烷细菌进行比较。对本领域技术人员而言显而易见的是在该方法期间的每一个步骤可以保存克隆细菌存储料以便后续使用。例如,对于已被鉴定为具有增强的红色着色的特定细菌,在用蕃茄红素途径转化之前保存的该细菌的克隆存储料 (即,具有通过蕃茄红素筛选被鉴定对异戊二烯合成潜在有益的突变但没有外源性蕃茄红素途径的细菌) 可以用异戊二烯合成酶 (例如, IspS) 进行转化。

[0150] 本公开中还提供用于在嗜甲烷细菌中筛选异戊二烯途径基因的方法。这些筛选方法通过将蕃茄红素途径工程改造入嗜甲烷细菌中作为比色读数可以用来鉴定导致异戊二烯前体 DMAPP 和 IPP 的合成增高的异戊二烯途径基因。蕃茄红素和异戊二烯合成途径使用相同的通用前体 IPP 和 DMAPP (参见图 1 和 6)。可以用异源性异戊二烯途径基因、同源性异戊二烯途径基因、变异体异戊二烯途径基因的过表达或其任何组合修饰嗜甲烷细菌来鉴定根据细菌的增强的红色着色测量的具有提高的蕃茄红素产量的细菌。由于提高的 IPP 和 DMAPP 产量,被鉴定为具有增加的蕃茄红素产量的细菌还可以呈现出提高的异戊二烯合成。

[0151] 在某些实施方案中,用于在嗜甲烷细菌中筛选异戊二烯途径基因的方法包括:(a) 用 (i) 至少一种编码异戊二烯途径酶的外源性核酸, (ii) 编码香叶基香叶基二磷酸合成酶 (GGPPS)、八氢番茄红素合成酶 (CRTB) 和八氢番茄红素脱氢酶 (CRTI) 的外源性核酸转化该嗜甲烷细菌;和 (b) 在足够生长的条件下培养来自步骤 (a) 的嗜甲烷细菌;其中与已用编码 GGPPS、CRTB 和 CRTI 的外源性核酸转化并且不包含至少一种编码异戊二烯途径酶的外源性核酸的参考嗜甲烷细菌相比,呈现出增强的红色着色的转化的嗜甲烷细菌表明该至少一种编码异戊二烯途径酶的外源性核酸赋予与该参考嗜甲烷细菌相比增高的异戊二烯前体合成。在某些实施方案中,该异戊二烯途径酶是 DXP 途径酶 (例如, DXS、DXR、IspD、IspE、IspF、IspG、IspH 或 IDI) 或甲羟戊酸途径酶 (例如, AACT、HMGS、HMGR、MK、PMK、MPD 或 IDI)。

该至少一种编码异戊二烯途径酶的外源性核酸可以是异源性核酸或同源性核酸。该异源性核酸可以是为该宿主嗜甲烷细菌中表达而优化的密码子。在一些实施方案中,该同源性核酸在嗜甲烷细菌中过表达。在本文所述的各种实施方案中,该至少一种编码异戊二烯途径酶的外源性核酸可以是非天然存在的变异体。该非天然存在的变异体可以通过随机诱变、定向诱变产生,或合成(全部或部分)。在某些实施方案中,该非天然存在的变异体与编码异戊二烯途径酶的参考核酸相比包括至少一个氨基酸取代。

[0152] 蕃茄红素途径酶的来源是本领域中已知的,并且可以是天然具有蕃茄红素途径的任何生物体,包括植物、光合细菌、真菌和藻类的物种。在 NCBI 数据库中可获得的香叶基香叶基二磷酸合成酶的核酸序列的实例包括登录号:AB000835(拟南芥);AB016043(智人);AB019036(智人);AB016044(小家鼠(*Mus musculus*));AB027705(胡萝卜(*Daucus carota*));AB034249(近琴巴豆(*Croton sublyratus*));AB034250(野甘草(*Scoparia dulcis*));AF049659(果蝇(*Drosophila melanogaster*));AF139916(亚麻短杆菌(*Brevibacterium linens*));AF279807(覃青霉(*Penicillium paxilli*));AJ010302(类球红细菌(*Rhodobacter sphaeroides*));AJ133724(金色分枝杆菌(*Mycobacterium aurum*));L25813(拟南芥);U44876(拟南芥);和 U15778(白羽扇豆(*Lupinus albus*))。在 NCBI 数据库中可获得的八氢番茄红素合成酶的核酸序列的实例包括登录号:AB001284(钝顶螺旋藻(*Spirulina platensis*));AB032797(胡萝卜(*Daucus carota*));AB034704(胶状红长命菌(*Rubrivivax gelatinosus*));AB037975(温州蜜柑(*Citrus unshui*));AF009954(拟南芥);AF139916(亚麻短杆菌);AF152892(葡萄柚(*Citrus x paradise*));AF218415(根瘤菌(*Bradyrhizobium* sp. ORS278));AF220218(温州蜜柑(*Citrus unshui*));AJ133724(金色分枝杆菌);和 AJ304825(向日葵(*Helianthus annuus*))。在 NCBI 数据库中可获得的八氢番茄红素脱氢酶的核酸序列的实例包括登录号:AB046992(温州蜜柑);AF139916(亚麻短杆菌);AF218415(根瘤菌 ORS278);AF251014(万寿菊(*Tagetes erecta*));L16237(拟南芥);L39266(玉米);M64704(大豆(*Glycine max*));AF364515(葡萄柚);D83514(长赤细菌(*Erythrobacter longus*));M88683(番茄);和 X55289(聚球藻)。

实施例

[0153] 实施例 1

[0154] 在荚膜甲基球菌(*Methanococcus capsulatus*) Bath 菌株中异戊二烯合成酶的克隆和表达

[0155] 为了创建生产异戊二烯的嗜甲烷菌菌株,通过接合交配将包含编码异戊二烯合成酶(IspS)的基因的嗜甲烷菌表达载体插入荚膜甲基球菌 Bath、发酵甲基弯曲菌 OB3b 和甲基单胞菌 16A 中。附加体(episomal)表达质粒(包含编码复制起点、转移起点、抗药性标记(卡那霉素)和多克隆位点的序列)用来克隆甲醇脱氢酶(MDH)启动子下游的密码子优化的柳属 IspS 多聚核苷酸序列(荚膜甲基球菌 Bath 的 SEQ ID NO:19),或 IPTG-诱导型(LacIq)启动子下游的山葛密码子优化的 IspS 多聚核苷酸序列(去除氨基-末端叶绿体靶向序列)(荚膜甲基球菌 Bath 的 SEQ ID NO:17)。将包含载有 IspS 的质粒的大肠杆菌菌株(供体菌株)和包含 pRK2013 质粒(ATCC)的大肠杆菌(辅助菌株)的菌落接种在包含卡那

霉素 (30 μ g/mL) 的液体 LB 中并在 37°C 下过夜生长。在将它们用来与受体嗜甲烷菌菌株交配之前, 将各一份的液体供体培养物和辅助培养物接种在 100 份包含卡那霉素 (30 μ g/mL) 的新鲜 LB 中 3-5h。在交配之前, 将嗜甲烷菌 (受体) 菌株接种在含有约 40mL 甲烷的液体 MM-W1 培养基中 (Pieja 等, 2011, *Microbial Ecology* 62:564-573) 1-2 天直至它们到达对数生长期 (OD600 约 0.3)。

[0156] 通过以使 OD600 比率为 2:1:1 (例如, 1mL OD600 为 1.5 的嗜甲烷菌、1mL OD600 为 0.75 供体和 1mL OD600 为 0.75 的辅助菌株) 的体积制备受体菌株、供体菌株和辅助菌株来实施三亲交配。然后通过 25°C 下以 5,300rpm 离心 7min 来采集这些细胞。除去上清液, 并将细胞团块温和地再悬浮于 500 μ L MM-W1 中。对于大肠杆菌供体和辅助菌株而言, 将离心和再悬浮重复多于 2 次以确保除去抗生素。然后通过温和地吸移将等体积的受体菌株、供体菌株和辅助菌株的再悬浮细胞合并和混合。使交配组合物以 13.2k rpm 旋转减缓 (spun down) 30-60s, 并尽量除去上清液。然后将细胞团块温和地混合并以单个液滴沉积在交配琼脂上 (包含无菌的 0.5% 酵母提取物的完全 MM-W1 培养基)。在包含甲烷和空气的 oxid 室中, 在使用发孢甲基弯曲菌 OB3b 或甲基单胞菌 16a 作为受体的情况中在 30°C 下, 或在使用荚膜甲基球菌 Bath 作为受体菌株的情况中在 37°C 下, 使交配板孵育 48h。在 48h 孵育期后, 通过将 1mL MM-W1 培养基加在板上并将悬浮的细胞转移到 2mL Eppendorf 管中, 从交配板采集细胞。在平板接种于选择板 (包含卡那霉素 10 μ g/mL 的完全 MM-W1 琼脂培养基) 上以选择稳定维持构建体的细胞之前, 通过离心使细胞成团块, 并用 100 μ L 新鲜的 MM-W1 再悬浮。在包含甲烷-空气混合物的 oxid 室中, 在荚膜甲基球菌 Bath 菌株在 42°C 下孵育 1 周后, 或在甲基单胞菌 16a 和发孢甲基弯曲菌 OB3b 在 30°C 下孵育 1 周后, 携带质粒的嗜甲烷菌出现在这些板上。然后在 1mL 液体培养基中培养荚膜甲基球菌 Bath 菌株克隆, 并分析异戊二烯产量。

[0157] 实施例 2

[0158] 通过荚膜甲基球菌 Bath 菌株生产异戊二烯

[0159] 从包含具有组成型 MDH 启动子-柳属 *IspS* 的载体或具有 IPTG-诱导型 (LacIq) 启动子-山葛 *IspS* (在存在或不存在 0.1-10mM IPTG 下生长) 的载体的荚膜甲基球菌 Bath 菌株的过夜生长的封装的 5mL 培养物获得顶空气体样品 (250 μ L)。将气体样品注射到配有火焰离子化检测器的气相色谱 (Hewlett Packard 5890) 上。色谱条件包括: 安捷伦 CP-PoraBOND U (25m x 0.32mm 内径) 柱, 升温程序 50°C, 1.5min; 25°C, 1min; 300°C, 10min。洗脱峰通过火焰离子化进行检测, 并且通过与异戊二烯标准 (溶于去离子水中的纯异戊二烯) 比较来量化积分峰。

[0160] 与柳属 *IspS* 的表达相比, 当表达山葛 *IspS* 时, 荚膜甲基球菌 Bath 生产更多的异戊二烯。此外, 在表达山葛 *IspS* 的荚膜甲基球菌 Bath 中生产的异戊二烯的量与用 IPTG 诱导 LacIq 启动子直接相关 (参见图 7A)。图 5 和 7B 显示分别来自柳属和山葛变异体样品的顶空样品的 GC/MS 色谱。在图 5 中, 样品 A 是阴性对照, 其显示来自未转化的细胞的顶空的背景信号。图 5 的样品 B 中异戊二烯产率是约 10mg/L。图 7B 显示生产的异戊二烯的实质量 (substantial amount)。

[0161] 实施例 3

[0162] 工程改造具有提高的异戊二烯产量的 DXP 途径

[0163] 为了产生在途径内总体导致改进用于合成异戊二烯的定向 (committed) 前体 (IPP 和 DMAPP) 的酶的新的基因序列或调控元件, 将随机突变导入 DXP 途径操纵子 (即, DXS-DXR-IspD-IspE-IspF-IspG-IspH) 中。为了构建用于分离改进的 DXP 途径的容易的高通量筛选方法, 将包括 ggpps、crtB 和 crtI 的蕃茄红素合成途径用作比色报告物 (colorimetric reporter)。利用 GENEMORPH® II 随机诱变试剂盒 (Stratagene), 以低、中等和高突变率通过易错 PCR 创建 DXP 途径的随机诱变库。然后将该库克隆入包含 ggpps、crtB 和 crtI 基因序列的嗜甲烷菌表达质粒中, 其中它们的多顺反子表达由强嗜甲烷菌启动子序列 (例如, 甲醇脱氢酶启动子) 驱动。然后将包含质粒的库的池 (pool) 与大肠杆菌 DH10B 的多于大约 10^6 个转化体分离。然后该质粒库用来转化嗜甲烷菌菌株。在延长的孵育期后, 分离显现鲜红色着色的菌落 (如在 MM-WI 板上显现)。在质粒萃取和测序后, 将突变型 DXP 途径基因作为池用于下一轮的易错 PCR 中。包含野生型 DXP 途径基因的嗜甲烷菌菌株, 连同包含 ggpps、crtB、crtI 的质粒, 作为蕃茄红素形成的基线对比以分离突变型 DXP 途径基因。在鉴定没有额外的菌落显示增强的红色着色后, 停止突变和筛选的重复。然后从嗜甲烷菌宿主分离载有新的 DXP 途径基因的质粒。然后使这些新的 DXP 途径基因与 IspS 在嗜甲烷宿主细菌中共表达以确认异戊二烯产量的提高。

[0164] 实施例 4

[0165] 在代谢 C_1 的微生物中稳定的碳同位素分布

[0166] 利用与 IsoPrime100IRMS (Isoprime, Cheadle, UK) 偶联的 CHNOS 元素分析仪 (vario ISOTOPE cube, Elementar, Hanau, Germany), 通过元素分析仪 / 连续流同位素比率质谱法, 分析发孢甲基弯曲菌生物质的干样品的碳和氮含量 (% 干重) 以及碳 (^{13}C) 和氮 (^{15}N) 稳定同位素比率。将在发酵器或血清瓶中培养的嗜甲烷菌生物质的样品离心, 再悬浮于去离子水中, 并将与 0.2-2mg 碳 (约 0.5-5mg 干细胞重量) 对应的体积转移到 5x 9mm 锡囊中 (Costech Analytical Technologies, Inc., Valencia, CA), 并在 80°C 下干燥 24 小时。包含 0.1mg 碳的标准提供可靠的 $\delta^{13}C$ 值。

[0167] 同位素比率以 “ δ ” 符号 (‰) 表示, 其中相对于标准, 材料的同位素组成按照每百万偏差以 $\delta^{13}C$ (或 $\delta^{15}N$) = $(R_{\text{样品}} / R_{\text{标准}} - 1) \times 1,000$ 表示, 其中 R 是重同位素形式比轻同位素形式的分子比率。碳的标准是 Vienna Pee Dee Belemnite (V-PDB), 氮的标准是空气。NIST (国家标准与技术研究院) 提议的 SRM (标准参考材料) No. 1547 桃树叶片用作校正标准。所有的同位素分析在加利福尼亚大学伯克利分校稳定同位素生物地球化学中心进行。用于 C 和 N 同位素分析的长期外部精确度分别是 0.10‰ 和 0.15‰。

[0168] 在三个不同的发酵批次中, 使发孢甲基弯曲菌菌株 OB3b 在甲烷上生长, 在两个不同的发酵批次中, 使荚膜甲基球菌 Bath 在甲烷上生长, 并且在单个发酵批次中使甲基单胞菌 16a 在甲烷上生长。分析来自每个这些培养物的生物质的稳定碳同位素分布 ($\delta^{13}C$ 值; 参见表 4)。

[0169] 表 4. 在不同的嗜甲烷菌中稳定的碳同位素分布

[0170]

嗜甲烷菌	批次号	EFT (h)†	OD ₆₀₀	DCW*	$\delta^{13}\text{C}$ 细胞
Mt OB3b	68A	48	1.80	1.00	-57.9
		64	1.97	1.10	-57.8
		71	2.10	1.17	-58.0
		88	3.10	1.73	-58.1
		97	4.30	2.40	-57.8
		113	6.00	3.35	-57.0
		127	8.40	4.69	-56.3
Mt OB3b	68B	32	2.90	1.62	-58.3
		41	4.60	2.57	-58.4
		47	5.89	3.29	-58.0
		56	7.90	4.41	-57.5
Mt OB3b	68C	72	5.32	2.97	-57.9
		79.5	5.90	3.29	-58.0
		88	5.60	3.12	-57.8

[0171]

嗜甲烷菌	批次号	EFT (h) [†]	OD ₆₀₀	DCW*	δ ¹³ C 细胞
		94	5.62	3.14	-57.7
Mc Bath	62B	10	2.47	0.88	-59.9
		17.5	5.80	2.06	-61.0
		20	7.32	2.60	-61.1
		23	9.34	3.32	-60.8
		26	10.30	3.66	-60.1
Mc Bath	62A	10	2.95	1.05	-55.9
		13.5	3.59	1.27	-56.8
		17.5	5.40	1.92	-55.2
		23	6.08	2.16	-57.2
		26	6.26	2.22	-57.6
Mms 16a	66B	16	2.13	0.89	-65.5
		18	2.59	1.09	-65.1
		20.3	3.62	1.52	-65.5
		27	5.50	2.31	-66.2
		40.5	9.80	4.12	-66.3

[0172] *DCW, 干细胞重量是使用关于 1.0 至 0.558g/L 的 OD (对于 Mt OB3b), 1.0 至 0.355g/L 的 OD (对于 Mc Bath), 以及 1.0 至 0.42g/L 的 OD (对于 Mms 16a) 的特定的相关因子, 由测得的光密度 (OD₆₀₀) 计算, 以 g/L 进行报告。对于 Mt OB3b, 每次发酵所用的碳酸氢盐的初始浓度是 1.2mM 或 0.01% (批次号 68C) 和 0.1% 或 12mM (批次号 68A 和 68B)。

[0173] [†] EFT = 有效发酵时间 (小时)

[0174] 实施例 5

[0175] 甲烷来源和纯度对稳定碳同位素分布的影响

[0176] 为了检查在包含天然气组分的甲烷上的嗜甲烷菌生长, 将来自在供有甲烷和空气的 1:1 (v/v) 混合物的相同培养基中生长的血清瓶分批培养物 (5% v/v) 的发孢甲基弯曲菌 OB3b 或荚膜甲基球菌 Bath 接种在一系列包含 100mL 限定培养基 MMS1.0 的 0.5 升血清瓶中。培养基 MMS1.0 的组成如下: 0.8mM MgSO₄*7H₂O、30mM NaNO₃、0.14mM CaCl₂、1.2mM NaHCO₃、2.35mM KH₂PO₄、3.4mM K₂HPO₄、20.7 μM Na₂MoO₄*2H₂O、6 μM CuSO₄*5H₂O、10 μM Fe^{III}-Na-EDTA, 以及 1mL 每升的痕量金属溶液 (每 L 包含: 500mg FeSO₄*7H₂O、400mg ZnSO₄*7H₂O、20mg MnCl₂*7H₂O、50mg CoCl₂*6H₂O、10mg NiCl₂*6H₂O、15mg H₃BO₃、250mg EDTA)。在培养基被高压灭菌和冷却后, 添加磷酸盐、碳酸氢盐和 Fe^{III}-Na-EDTA。培养基的最终 pH 是 7.0±0.1。

[0177] 用橡胶塞密封接种瓶,并通过注射器经由无菌 0.45 μm 过滤器和无菌 27G 针头注射添加 60mL 甲烷气体。一式两份的培养物各被注射 60mL 体积的 (A) 99% 纯度的甲烷 (2.0 级, Praxair through Alliance Gas, San Carlos, CA)、(B) 代表天然气标准的 70% 纯度的甲烷 (Sigma-Aldrich ;还包含 9% 乙烷、6% 丙烷、3% 甲基丙烷、3% 丁烷及其它较少量的烃组分)、(C) 以甲烷来源 A 和 B 的 1:1 混合物递送的 85% 纯度的甲烷 ;以及 (D) >93% 甲烷 (1.3 级, Specialty Chemical Products, South Houston, TX ;内部分析显示组成 >99% 甲烷)。在 250rpm 旋转震动下,使培养物在 30°C (发孢甲基弯曲菌菌株 OB3b) 或 42°C (荚膜甲基球菌 Bath) 下孵育,并在大约 12 小时间隔通过抽取 1mL 样品确定 OD_{600} 来测量生长。这时,使该瓶通风,并用 60mL 各自的甲烷来源 (A、B、C 或 D) 和 60mL 浓缩氧气 (至少 85% 纯度) 置换顶空。在约 24 小时间隔,取出 5mL 样品,通过离心 (8,000rpm, 10 分钟) 回收细胞,并且然后在分析前存储在 -80°C 下。

[0178] 如实施例 4 中所述,对源于发孢甲基弯曲菌菌株 OB3b 和荚膜甲基球菌 Bath 的嗜甲烷菌生物质分析碳和氮含量 (% 干重量) 以及碳 (^{13}C) 和氮 (^{15}N) 稳定同位素比率。表 5 显示分析来自在具有不同纯度水平的甲烷上和在各种批次的瓶培养物中生长的荚膜甲基球菌 Bath 的生物质样品的稳定碳同位素的结果。

[0179] 表 5. 在具有不同纯度的不同甲烷来源上生长的荚膜甲基球菌 Bath 的稳定碳同位素分布

[0180]

甲烷*	批次号	时间(h)†	OD ₆₀₀	DCW (g/L)	$\delta^{13}\text{C}$ 细胞
A	62C	22	1.02	0.36	-40.3
		56	2.01	0.71	-41.7
		73	2.31	0.82	-42.5
	62D	22	1.14	0.40	-39.3
		56	2.07	0.73	-41.6
		73	2.39	0.85	-42.0
B	62E	22	0.47	0.17	-44.7
		56	0.49	0.17	-45.4
		73	0.29	0.10	-45.4
	62F	22	0.62	0.22	-42.3
		56	0.63	0.22	-43.6
		73	0.30	0.11	-43.7
C	62G	22	0.70	0.25	-40.7
		56	1.14	0.40	-44.8
		73	1.36	0.48	-45.8
	62H	22	0.62	0.22	-40.9
		56	1.03	0.37	-44.7
		73	1.23	0.44	-45.9

[0181] * 甲烷纯度 :A :99%甲烷, 2.0 级 (min. 99%) ;B :70%甲烷, 天然气标准 (包含 9% 乙烷、6%丙烷、3%甲基丙烷、3%丁烷) ;C :85%甲烷 (A 和 B 甲烷的 1:1 混合)

[0182] † 时间=瓶培养时间 (小时)

[0183] 在甲烷的一种来源 (A, 99%) 上生长的荚膜甲基球菌 Bath 的平均 $\delta^{13}\text{C}$ 是 -41.2 ± 1.2 , 而在甲烷的不同来源 (B, 70%) 上生长的荚膜甲基球菌 Bath 的平均 $\delta^{13}\text{C}$ 是 -44.2 ± 1.2 。当甲烷来源 A 和 B 混合时, 观察到 -43.8 ± 2.4 的中间平均 $\delta^{13}\text{C}$ 。这些数据表明, 由于输入甲烷的 $\delta^{13}\text{C}$ 的不同, 在甲烷来源 A 和 B 上生长的细胞材料的 $\delta^{13}\text{C}$ 彼此显著不同。但是, 在两种气体的混合物上生长的细胞优选使用 ^{12}C , 因此显示 $\delta^{13}\text{C}$ 更大负值的趋势。

[0184] 实施相似的试验以检查在不同甲烷来源上和在各种批次的瓶培养物中生长的两种不同的嗜甲烷菌 (荚膜甲基球菌 Bath 和发孢甲基弯曲菌 OB3b) 是否显示 $\delta^{13}\text{C}$ 分布的不同 (参见表 6)。

[0185] 表 6. 在不同纯度的不同甲烷来源上生长的不同嗜甲烷菌的稳定碳同位素分布

[0186]

菌株	甲烷*	批次号	时间(h)†	OD ₆₀₀	DCW (g/L)	$\delta^{13}\text{C}$ 细胞
Mc Bath	A	62I	18	0.494	0.18	-54.3
			40	2.33	0.83	-42.1
			48	3.08	1.09	-37.1
Mc Bath	D	62J	18	0.592	0.21	-38.3
			40	1.93	0.69	-37.8
			48	2.5	0.89	-37.8
Mc Bath	D	62K	18	0.564	0.20	-38.6
			40	1.53	0.54	-37.5
			48	2.19	0.78	-37.6
Mt OB3b	A	68D	118	0.422	0.24	-50.2
			137	0.99	0.55	-47.7
			162	1.43	0.80	-45.9
Mt OB3b	A	68E	118	0.474	0.26	-49.9
			137	1.065	0.59	-47.6
			162	1.51	0.84	-45.2

[0187]

菌株	甲烷*	批次号	时间(h)†	OD ₆₀₀	DCW (g/L)	$\delta^{13}\text{C}$ 细胞
Mt OB3b	D	68F	118	0.534	0.30	-45.6
			137	1.119	0.62	-38.7
			162	1.63	0.91	-36.4
Mt OB3b	D	68G	118	0.544	0.30	-44.8
			137	1.131	0.63	-39.1
			162	1.6	0.89	-34.2

[0188] * 甲烷来源和纯度 :A :99%甲烷 (2.0 级) ;D :>93%甲烷 (1.3 级)

[0189] † 时间=瓶培养时间 (小时)

[0190] 在第一甲烷来源 (A) 上生长的荚膜甲基球菌的平均 $\delta^{13}\text{C}$ 是 -44.5 ± 8.8 , 而在相同甲烷来源上生长的发孢甲基弯曲菌的平均 $\delta^{13}\text{C}$ 是 -47.8 ± 2.0 。在第二甲烷来源 (B) 上生长的荚膜甲基球菌的平均 $\delta^{13}\text{C}$ 是 -37.9 ± 0.4 , 而发孢甲基弯曲菌的平均 $\delta^{13}\text{C}$ 是 -39.8 ± 4.5 。这些数据表明,在甲烷来源上生长的细胞材料的 $\delta^{13}\text{C}$ 与来自在相同甲烷来源上生长的不同菌株的细胞材料的 $\delta^{13}\text{C}$ 高度相似。因此,观察到的细胞材料的 $\delta^{13}\text{C}$ 似

乎主要取决于输入气体的组成而不是被研究的具体细菌菌株的性质。

[0191] 可以将上述各种实施方案组合以提供其它实施方案。在本说明书中引用的或列于申请数据页中的所有的美国专利、美国专利申请公布、美国专利申请、外国专利、外国专利申请和非专利公布（包括但不限于美国专利申请 No. 61/774, 342 和美国专利申请 No. 61/928, 333）在此通过引用整体并入。若需要采用各种专利、申请和公布的构思，可以对所述实施方案的方面进行修改以提供另外更多的实施方案。

[0192] 根据以上详细描述，可以对实施方案作出这些及其它改变。一般而言，在以下权利要求书中，所用的术语不应解释为将本权利要求书限制于在本说明书和本权利要求书中公开的具体的实施方案，而应解释为包括有权利要求保护的所有可能的实施方案以及等同方案的全部范围。因此，本权利要求书不限于本公开。

[0001]

序列表

<110> 凯利斯塔公司

E·伦纳德

J·米舒尔

J·内斯

T·J·珀塞尔

<120> 用于生物生产异戊二烯的组合物和方法

<130> 200206.411WO

<140> PCT

<141> 2014-03-06

<150> US 61/928,333

<151> 2014-01-16

<150> US 61/774,342

<151> 2013-03-07

<160> 19

<170> FastSEQ Windows 4.0 版

<210> 1

<211> 595

<212> PRT

<213> 银白杨

<400> 1

Met	Ala	Thr	Glu	Leu	Leu	Cys	Leu	His	Arg	Pro	Ile	Ser	Leu	Thr	His
1				5					10					15	
Lys	Leu	Phe	Arg	Asn	Pro	Leu	Pro	Lys	Val	Ile	Gln	Ala	Thr	Pro	Leu
			20					25					30		
Thr	Leu	Lys	Leu	Arg	Cys	Ser	Val	Ser	Thr	Glu	Asn	Val	Ser	Phe	Thr
			35				40					45			
Glu	Thr	Glu	Thr	Glu	Ala	Arg	Arg	Ser	Ala	Asn	Tyr	Glu	Pro	Asn	Ser
		50				55					60				
Trp	Asp	Tyr	Asp	Tyr	Leu	Leu	Ser	Ser	Asp	Thr	Asp	Glu	Ser	Ile	Glu
65					70					75				80	
Val	Tyr	Lys	Asp	Lys	Ala	Lys	Lys	Leu	Glu	Ala	Glu	Val	Arg	Arg	Glu
				85					90				95		
Ile	Asn	Asn	Glu	Lys	Ala	Glu	Phe	Leu	Thr	Leu	Leu	Glu	Leu	Ile	Asp

[0002]

100	105	110
Asn Val Gln Arg Leu Gly Leu Gly Tyr Arg Phe Glu Ser Asp Ile Arg		
115	120	125
Gly Ala Leu Asp Arg Phe Val Ser Ser Gly Gly Phe Asp Ala Val Thr		
130	135	140
Lys Thr Ser Leu His Gly Thr Ala Leu Ser Phe Arg Leu Leu Arg Gln		
145	150	155
His Gly Phe Glu Val Ser Gln Glu Ala Phe Ser Gly Phe Lys Asp Gln		
165	170	175
Asn Gly Asn Phe Leu Glu Asn Leu Lys Glu Asp Ile Lys Ala Ile Leu		
180	185	190
Ser Leu Tyr Glu Ala Ser Phe Leu Ala Leu Glu Gly Glu Asn Ile Leu		
195	200	205
Asp Glu Ala Lys Val Phe Ala Ile Ser His Leu Lys Glu Leu Ser Glu		
210	215	220
Glu Lys Ile Gly Lys Glu Leu Ala Glu Gln Val Asn His Ala Leu Glu		
225	230	235
Leu Pro Leu His Arg Arg Thr Gln Arg Leu Glu Ala Val Trp Ser Ile		
245	250	255
Glu Ala Tyr Arg Lys Lys Glu Asp Ala Asn Gln Val Leu Leu Glu Leu		
260	265	270
Ala Ile Leu Asp Tyr Asn Met Ile Gln Ser Val Tyr Gln Arg Asp Leu		
275	280	285
Arg Glu Thr Ser Arg Trp Trp Arg Arg Val Gly Leu Ala Thr Lys Leu		
290	295	300
His Phe Ala Arg Asp Arg Leu Ile Glu Ser Phe Tyr Trp Ala Val Gly		
305	310	315
Val Ala Phe Glu Pro Gln Tyr Ser Asp Cys Arg Asn Ser Val Ala Lys		
325	330	335
Met Phe Ser Phe Val Thr Ile Ile Asp Asp Ile Tyr Asp Val Tyr Gly		
340	345	350
Thr Leu Asp Glu Leu Glu Leu Phe Thr Asp Ala Val Glu Arg Trp Asp		
355	360	365
Val Asn Ala Ile Asn Asp Leu Pro Asp Tyr Met Lys Leu Cys Phe Leu		
370	375	380
Ala Leu Tyr Asn Thr Ile Asn Glu Ile Ala Tyr Asp Asn Leu Lys Asp		
385	390	395
Lys Gly Glu Asn Ile Leu Pro Tyr Leu Thr Lys Ala Trp Ala Asp Leu		
405	410	415
Cys Asn Ala Phe Leu Gln Glu Ala Lys Trp Leu Tyr Asn Lys Ser Thr		
420	425	430
Pro Thr Phe Asp Asp Tyr Phe Gly Asn Ala Trp Lys Ser Ser Ser Gly		
435	440	445
Pro Leu Gln Leu Val Phe Ala Tyr Phe Ala Val Val Gln Asn Ile Lys		

[0003]

```

450          455          460
Lys Glu Glu Ile Glu Asn Leu Gln Lys Tyr His Asp Thr Ile Ser Arg
465          470          475          480
Pro Ser His Ile Phe Arg Leu Cys Asn Asp Leu Ala Ser Ala Ser Ala
          485          490          495
Glu Ile Ala Arg Gly Glu Thr Ala Asn Ser Val Ser Cys Tyr Met Arg
          500          505          510
Thr Lys Gly Ile Ser Glu Glu Leu Ala Thr Glu Ser Val Met Asn Leu
          515          520          525
Ile Asp Glu Thr Trp Lys Lys Met Asn Lys Glu Lys Leu Gly Gly Ser
          530          535          540
Leu Phe Ala Lys Pro Phe Val Glu Thr Ala Ile Asn Leu Ala Arg Gln
          545          550          555          560
Ser His Cys Thr Tyr His Asn Gly Asp Ala His Thr Ser Pro Asp Glu
          565          570          575
Leu Thr Arg Lys Arg Val Leu Ser Val Ile Thr Glu Pro Ile Leu Pro
          580          585          590
Phe Glu Arg
          595

```

<210> 2

<211> 559

<212> PRT

<213> 银白杨

<400> 2

```

Met Cys Ser Val Ser Thr Glu Asn Val Ser Phe Thr Glu Thr Glu Thr
  1              5              10              15
Glu Ala Arg Arg Ser Ala Asn Tyr Glu Pro Asn Ser Trp Asp Tyr Asp
          20              25              30
Tyr Leu Leu Ser Ser Asp Thr Asp Glu Ser Ile Glu Val Tyr Lys Asp
          35              40              45
Lys Ala Lys Lys Leu Glu Ala Glu Val Arg Arg Glu Ile Asn Asn Glu
          50              55              60
Lys Ala Glu Phe Leu Thr Leu Leu Glu Leu Ile Asp Asn Val Gln Arg
          65              70              75              80
Leu Gly Leu Gly Tyr Arg Phe Glu Ser Asp Ile Arg Gly Ala Leu Asp
          85              90              95
Arg Phe Val Ser Ser Gly Gly Phe Asp Ala Val Thr Lys Thr Ser Leu
          100             105             110
His Gly Thr Ala Leu Ser Phe Arg Leu Leu Arg Gln His Gly Phe Glu
          115             120             125
Val Ser Gln Glu Ala Phe Ser Gly Phe Lys Asp Gln Asn Gly Asn Phe
          130             135             140

```

[0004]

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

[0005]

Trp Lys Lys Met Asn Lys Glu Lys Leu Gly Gly Ser Leu Phe Ala Lys
 500 505 510
 Pro Phe Val Glu Thr Ala Ile Asn Leu Ala Arg Gln Ser His Cys Thr
 515 520 525
 Tyr His Asn Gly Asp Ala His Thr Ser Pro Asp Glu Leu Thr Arg Lys
 530 535 540
 Arg Val Leu Ser Val Ile Thr Glu Pro Ile Leu Pro Phe Glu Arg
 545 550 555

<210> 3

<211> 608

<212> PRT

<213> 葛麻姆

<400> 3

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

[0006]

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

[0007]

	580		585		590									
Lys	Leu	Leu	Ile	Asp	Pro	Phe	Pro	Ile	Asn	Gln	Leu	Met	Tyr	Val
	595		600		605									

<210> 4

<211> 564

<212> PRT

<213> 葛麻姆

<400> 4

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

[0008]

```

Met Glu Val Tyr Phe Trp Ala Leu Gly Met Ala Pro Asp Pro Gln Phe
    275                280                285
Gly Glu Cys Arg Lys Ala Val Thr Lys Met Phe Gly Leu Val Thr Ile
    290                295                300
Ile Asp Asp Val Tyr Asp Val Tyr Gly Thr Leu Asp Glu Leu Gln Leu
    305                310                315                320
Phe Thr Asp Ala Val Glu Arg Trp Asp Val Asn Ala Ile Asn Thr Leu
    325                330                335
Pro Asp Tyr Met Lys Leu Cys Phe Leu Ala Leu Tyr Asn Thr Val Asn
    340                345                350
Asp Thr Ser Tyr Ser Ile Leu Lys Glu Lys Gly His Asn Asn Leu Ser
    355                360                365
Tyr Leu Thr Lys Ser Trp Arg Glu Leu Cys Lys Ala Phe Leu Gln Glu
    370                375                380
Ala Lys Trp Ser Asn Asn Lys Ile Ile Pro Ala Phe Ser Lys Tyr Leu
    385                390                395                400
Glu Asn Ala Ser Val Ser Ser Ser Gly Val Ala Leu Leu Ala Pro Ser
    405                410                415
Tyr Phe Ser Val Cys Gln Gln Gln Glu Asp Ile Ser Asp His Ala Leu
    420                425                430
Arg Ser Leu Thr Asp Phe His Gly Leu Val Arg Ser Ser Cys Val Ile
    435                440                445
Phe Arg Leu Cys Asn Asp Leu Ala Thr Ser Ala Ala Glu Leu Glu Arg
    450                455                460
Gly Glu Thr Thr Asn Ser Ile Ile Ser Tyr Met His Glu Asn Asp Gly
    465                470                475                480
Thr Ser Glu Glu Gln Ala Arg Glu Glu Leu Arg Lys Leu Ile Asp Ala
    485                490                495
Glu Trp Lys Lys Met Asn Arg Glu Arg Val Ser Asp Ser Thr Leu Leu
    500                505                510
Pro Lys Ala Phe Met Glu Ile Ala Val Asn Met Ala Arg Val Ser His
    515                520                525
Cys Thr Tyr Gln Tyr Gly Asp Gly Leu Gly Arg Pro Asp Tyr Ala Thr
    530                535                540
Glu Asn Arg Ile Lys Leu Leu Leu Ile Asp Pro Phe Pro Ile Asn Gln
    545                550                555                560
Leu Met Tyr Val

```

<210> 5

<211> 595

<212> PRT

<213> 柳属 DG-2011

[0009]

<400> 5
 Met Ala Thr Glu Leu Leu Cys Leu His Arg Pro Ile Ser Leu Thr Pro
 1 5 10 15
 Lys Leu Phe Arg Asn Pro Leu Pro Lys Val Ile Leu Ala Thr Pro Leu
 20 25 30
 Thr Leu Lys Leu Arg Cys Ser Val Ser Thr Glu Asn Val Ser Phe Thr
 35 40 45
 Glu Thr Glu Thr Glu Thr Arg Arg Ser Ala Asn Tyr Glu Pro Asn Ser
 50 55 60
 Trp Asp Tyr Asp Tyr Leu Leu Ser Ser Asp Thr Asp Glu Ser Ile Glu
 65 70 75 80
 Val Tyr Lys Asp Lys Ala Lys Lys Leu Glu Ala Glu Val Arg Arg Glu
 85 90 95
 Ile Asn Asn Glu Lys Ala Glu Phe Leu Thr Leu Leu Glu Leu Ile Asp
 100 105 110
 Asn Val Gln Arg Leu Gly Leu Gly Tyr Arg Phe Glu Ser Asp Ile Arg
 115 120 125
 Arg Ala Leu Asp Arg Phe Val Ser Ser Gly Gly Phe Asp Ala Val Thr
 130 135 140
 Lys Thr Ser Leu His Ala Thr Ala Leu Ser Phe Arg Phe Leu Arg Gln
 145 150 155 160
 His Gly Phe Glu Val Ser Gln Glu Ala Phe Gly Gly Phe Lys Asp Gln
 165 170 175
 Asn Gly Asn Phe Leu Glu Asn Leu Lys Glu Asp Ile Lys Ala Ile Leu
 180 185 190
 Ser Leu Tyr Glu Ala Ser Phe Leu Ala Leu Glu Gly Glu Asn Ile Leu
 195 200 205
 Asp Glu Ala Lys Val Phe Ala Ile Ser His Leu Lys Glu Leu Ser Glu
 210 215 220
 Glu Lys Ile Gly Lys Asp Leu Ala Glu Gln Val Asn His Ala Leu Glu
 225 230 235 240
 Leu Pro Leu His Arg Arg Thr Gln Arg Leu Glu Ala Val Trp Ser Ile
 245 250 255
 Glu Ala Tyr Arg Lys Lys Glu Asp Ala Asn Gln Val Leu Leu Glu Leu
 260 265 270
 Ala Ile Leu Asp Tyr Asn Met Ile Gln Ser Val Tyr Gln Arg Asp Leu
 275 280 285
 Arg Glu Thr Ser Arg Trp Trp Arg Arg Val Gly Leu Ala Thr Lys Leu
 290 295 300
 His Phe Ala Arg Asp Arg Leu Ile Glu Ser Phe Tyr Trp Ala Val Gly
 305 310 315 320
 Val Ala Phe Glu Pro Gln Tyr Ser Asp Cys Arg Asn Ser Val Ala Lys
 325 330 335
 Met Phe Ser Phe Val Thr Ile Ile Asp Asp Ile Tyr Asp Val Tyr Gly

[0010]

<400> 6
Met Cys Ser Val Ser Thr Glu Asn Val Ser Phe Thr Glu Thr Glu Thr
1 5 10 15
Glu Thr Arg Arg Ser Ala Asn Tyr Glu Pro Asn Ser Trp Asp Tyr Asp
20 25 30

58

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

[0012]

```

Leu Gln Glu Ala Lys Trp Leu Tyr Asn Lys Ser Thr Pro Thr Phe Asp
385              390              395              400
Asp Tyr Phe Gly Asn Ala Trp Lys Ser Ser Ser Gly Pro Leu Gln Leu
              405              410              415
Val Phe Ala Tyr Phe Ala Val Val Gln Asn Ile Lys Lys Glu Glu Ile
              420              425              430
Glu Asn Leu Gln Lys Tyr His Asp Ile Ile Ser Arg Pro Ser His Ile
              435              440              445
Phe Arg Leu Cys Asn Asp Leu Ala Ser Ala Ser Ala Glu Ile Ala Arg
              450              455              460
Gly Glu Thr Ala Asn Ser Val Ser Cys Tyr Met Arg Thr Lys Gly Ile
465              470              475              480
Ser Glu Glu Leu Ala Thr Glu Ser Val Met Asn Leu Ile Asp Glu Thr
              485              490              495
Trp Lys Lys Met Asn Lys Glu Lys Leu Gly Gly Ser Leu Phe Pro Lys
              500              505              510
Pro Phe Val Glu Thr Ala Ile Asn Leu Ala Arg Gln Ser His Cys Thr
              515              520              525
Tyr His Asn Gly Asp Ala His Thr Ser Pro Asp Glu Leu Thr Arg Lys
              530              535              540
Arg Val Leu Ser Val Ile Thr Glu Pro Ile Leu Pro Phe Glu Arg
545              550              555

```

<210> 7

<211> 639

<212> PRT

<213> 荚膜甲基球菌

<400> 7

```

Met Thr Glu Thr Lys Arg Tyr Ala Leu Leu Glu Ala Ala Asp His Pro
1              5              10              15
Ala Ala Leu Arg Asn Leu Pro Glu Asp Arg Leu Pro Glu Leu Ala Glu
              20              25              30
Glu Leu Arg Gly Tyr Leu Leu Glu Ser Val Ser Arg Ser Gly Gly His
              35              40              45
Leu Ala Ala Gly Leu Gly Thr Val Glu Leu Thr Ile Ala Leu His Tyr
              50              55              60
Val Phe Asn Thr Pro Glu Asp Lys Leu Val Trp Asp Val Gly His Gln
65              70              75              80
Ala Tyr Pro His Lys Ile Leu Thr Gly Arg Arg Ala Arg Leu Pro Thr
              85              90              95
Ile Arg Lys Lys Gly Gly Leu Ser Ala Phe Pro Asn Arg Ala Glu Ser
              100             105             110
Pro Tyr Asp Cys Phe Gly Val Gly His Ser Ser Thr Ser Ile Ser Ala

```

[0013]

115	120	125
Ala Leu Gly Met Ala Val	Ala Ala Ala Leu Glu Arg Arg	Pro Ile His
130	135	140
Ala Val Ala Ile Ile Gly Asp Gly Gly	Leu Thr Gly Gly Met Ala Phe	
145	150	155
Glu Ala Leu Asn His Ala Gly Thr	Leu Asp Ala Asn Leu Leu Ile Ile	
165	170	175
Leu Asn Asp Asn Glu Met Ser Ile	Ser Pro Asn Val Gly Ala Leu Asn	
180	185	190
Asn Tyr Leu Ala Lys Ile Leu Ser Gly	Lys Phe Tyr Ser Ser Val Arg	
195	200	205
Glu Ser Gly Lys His Leu Leu Gly Arg	His Met Pro Gly Val Trp Glu	
210	215	220
Leu Ala Arg Arg Ala Glu Glu His Val	Lys Gly Met Val Ala Pro Gly	
225	230	235
Thr Leu Phe Glu Glu Leu Gly Phe Asn	Tyr Phe Gly Pro Ile Asp Gly	
245	250	255
His Asp Leu Asp Thr Leu Ile Thr Thr	Leu Arg Asn Leu Arg Asp Gln	
260	265	270
Lys Gly Pro Arg Phe Leu His Val Val	Thr Arg Lys Gly Lys Gly Tyr	
275	280	285
Ala Pro Ala Glu Lys Asp Pro Val Ala	Tyr His Gly Val Gly Ala Phe	
290	295	300
Asp Leu Asp Ala Asp Glu Leu Pro Lys	Ser Lys Pro Gly Thr Pro Ser	
305	310	315
Tyr Thr Glu Val Phe Gly Gln Trp Leu	Cys Asp Met Ala Ala Arg Asp	
325	330	335
Arg Arg Leu Leu Gly Ile Thr Pro Ala	Met Arg Glu Gly Ser Gly Leu	
340	345	350
Val Glu Phe Ser Gln Arg Phe Pro Asp	Arg Tyr Phe Asp Val Gly Ile	
355	360	365
Ala Glu Gln His Ala Val Thr Phe Ala	Ala Gly Gln Ala Ser Glu Gly	
370	375	380
Tyr Lys Pro Val Val Ala Ile Tyr Ser	Thr Phe Leu Gln Arg Ala Tyr	
385	390	395
Asp Gln Leu Ile His Asp Val Ala Leu	Gln Asn Leu Pro Val Leu Phe	
405	410	415
Ala Ile Asp Arg Ala Gly Leu Val Gly	Pro Asp Gly Pro Thr His Ala	
420	425	430
Gly Ser Phe Asp Leu Ser Phe Met Arg	Cys Ile Pro Asn Met Leu Ile	
435	440	445
Met Ala Pro Ser Asp Glu Asn Glu Cys	Arg Gln Met Leu Tyr Thr Gly	
450	455	460
Phe Ile His Asp Gly Pro Ala Ala Val	Arg Tyr Pro Arg Gly Arg Gly	

[0014]

```

465          470          475          480
Pro Gly Val Arg Pro Glu Glu Thr Met Thr Ala Phe Pro Val Gly Lys
          485          490          495
Gly Glu Val Arg Leu Arg Gly Lys Gly Thr Ala Ile Leu Ala Phe Gly
          500          505          510
Thr Pro Leu Ala Ala Ala Leu Ala Val Gly Glu Arg Ile Gly Ala Thr
          515          520          525
Val Ala Asn Met Arg Phe Val Lys Pro Leu Asp Glu Ala Leu Ile Leu
          530          535          540
Glu Leu Ala Ala Thr His Asp Arg Ile Val Thr Val Glu Glu Asn Ala
545          550          555          560
Ile Ala Gly Gly Ala Gly Ser Ala Val Gly Glu Phe Leu Ala Ala Gln
          565          570          575
His Cys Gly Ile Pro Val Cys His Ile Gly Leu Lys Asp Glu Phe Leu
          580          585          590
Asp Gln Gly Thr Arg Glu Glu Leu Leu Ala Ile Ala Gly Leu Asp Gln
          595          600          605
Ala Gly Ile Ala Arg Ser Ile Asp Ala Phe Ile Gln Ala Thr Ala Ala
610          615          620
Ala Asp Lys Pro Arg Arg Ala Arg Gly Gln Ala Lys Asp Lys His
625          630          635

```

<210> 8

<211> 394

<212> PRT

<213> 荚膜甲基球菌

<400> 8

```

Met Lys Gly Ile Cys Ile Leu Gly Ser Thr Gly Ser Ile Gly Val Ser
  1          5          10          15
Thr Leu Asp Val Leu Ala Arg His Pro Asp Arg Tyr Arg Val Val Ala
          20          25          30
Leu Ser Ala Asn Gly Asn Val Asp Arg Leu Phe Glu Gln Cys Arg Ala
          35          40          45
His Arg Pro Arg Tyr Ala Ala Val Ile Arg Ala Glu Ala Ala Ala Cys
          50          55          60
Leu Arg Glu Arg Leu Met Ala Ala Gly Leu Gly Gly Ile Glu Val Leu
65          70          75          80
Ala Gly Pro Glu Ala Leu Glu Gln Ile Ala Ser Leu Pro Glu Val Asp
          85          90          95
Ser Val Met Ala Ala Ile Val Gly Ala Ala Gly Leu Leu Pro Thr Leu
          100          105          110
Ala Ala Ala Arg Ala Gly Lys Asp Val Leu Leu Ala Asn Lys Glu Ala
          115          120          125

```

[0015]

Leu Val Met Ser Gly Pro Leu Phe Met Ala Glu Val Ala Arg Ser Gly
 130 135 140
 Ala Arg Leu Leu Pro Ile Asp Ser Glu His Asn Ala Val Phe Gln Cys
 145 150 155 160
 Met Pro Ala Ala Tyr Arg Ala Gly Ser Arg Ala Val Gly Val Arg Arg
 165 170 175
 Ile Leu Leu Thr Ala Ser Gly Gly Pro Phe Leu His Thr Pro Leu Ala
 180 185 190
 Glu Leu Glu Thr Val Thr Pro Glu Gln Ala Val Ala His Pro Asn Trp
 195 200 205
 Val Met Gly Arg Lys Ile Ser Val Asp Ser Ala Thr Met Met Asn Lys
 210 215 220
 Gly Leu Glu Val Ile Glu Ala Cys Leu Leu Phe Asn Ala Lys Pro Asp
 225 230 235 240
 Asp Val Gln Val Val Val His Arg Gln Ser Val Ile His Ser Met Val
 245 250 255
 Asp Tyr Val Asp Gly Thr Val Leu Ala Gln Met Gly Thr Pro Asp Met
 260 265 270
 Arg Ile Pro Ile Ala His Ala Leu Ala Trp Pro Asp Arg Phe Glu Ser
 275 280 285
 Gly Ala Glu Ser Leu Asp Leu Phe Ala Val Arg Gln Leu Asn Phe Glu
 290 295 300
 Arg Pro Asp Leu Ala Arg Phe Pro Cys Leu Arg Leu Ala Tyr Glu Ala
 305 310 315 320
 Val Gly Ala Gly Gly Thr Ala Pro Ala Ile Leu Asn Ala Ala Asn Glu
 325 330 335
 Thr Ala Val Ala Ala Phe Leu Asp Arg Arg Leu Ala Phe Thr Gly Ile
 340 345 350
 Pro Arg Val Ile Glu His Cys Met Ala Arg Val Ala Pro Asn Ala Ala
 355 360 365
 Asp Ala Ile Glu Ser Val Leu Gln Ala Asp Ala Glu Thr Arg Lys Val
 370 375 380
 Ala Gln Lys Tyr Ile Asp Asp Leu Arg Val
 385 390

<210> 9

<211> 234

<212> PRT

<213> 荚膜甲基球菌

<400> 9

Met Ser Thr Asp Ala Arg Phe Trp Ile Val Val Pro Ala Ala Gly Val
 1 5 10 15
 Gly Lys Arg Met Gly Ala Asp Ile Pro Lys Gln Tyr Leu Asp Val Ala

[0016]

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

<210> 10

<211> 291

<212> PRT

<213> 荚膜甲基球菌

<400> 10

Met Asp Arg Arg Glu Ser Ser Val Met Lys Ser Pro Ser Leu Arg Leu			
1	5	10	15
Pro Ala Pro Ala Lys Leu Asn Leu Thr Leu Arg Ile Thr Gly Arg Arg			
20	25	30	
Pro Asp Gly Tyr His Asp Leu Gln Thr Val Phe Gln Phe Val Asp Val			
35	40	45	
Cys Asp Trp Leu Glu Phe Arg Ala Asp Ala Ser Gly Glu Ile Arg Leu			
50	55	60	
Gln Thr Ser Leu Ala Gly Val Pro Ala Glu Arg Asn Leu Ile Val Arg			
65	70	75	80

[0017]

Ala Ala Arg Leu Leu Lys Glu Tyr Ala Gly Val Ala Ala Gly Ala Asp
 85 90 95
 Ile Val Leu Glu Lys Asn Leu Pro Met Gly Gly Gly Leu Gly Gly Gly
 100 105 110
 Ser Ser Asn Ala Ala Thr Thr Leu Val Ala Leu Asn Arg Leu Trp Asp
 115 120 125
 Leu Gly Leu Asp Arg Gln Thr Leu Met Asn Leu Gly Leu Arg Leu Gly
 130 135 140
 Ala Asp Val Pro Ile Phe Val Phe Gly Glu Gly Ala Trp Ala Glu Gly
 145 150 155 160
 Val Gly Glu Arg Leu Gln Val Leu Glu Leu Pro Glu Pro Trp Tyr Val
 165 170 175
 Ile Val Val Pro Pro Cys His Val Ser Thr Ala Glu Ile Phe Asn Ala
 180 185 190
 Pro Asp Leu Thr Arg Asp Asn Asp Pro Ile Thr Ile Ala Asp Phe Leu
 195 200 205
 Ala Gly Ser His Gln Asn His Cys Leu Asp Ala Val Val Arg Arg Tyr
 210 215 220
 Pro Val Val Gly Glu Ala Met Cys Val Leu Gly Arg Tyr Ser Arg Asp
 225 230 235 240
 Val Arg Leu Thr Gly Thr Gly Ala Cys Val Tyr Ser Val His Gly Ser
 245 250 255
 Glu Glu Glu Ala Lys Ala Ala Cys Asp Asp Leu Ser Arg Asp Trp Val
 260 265 270
 Ala Ile Val Ala Ser Gly Arg Asn Leu Ser Pro Leu Tyr Glu Ala Leu
 275 280 285
 Asn Glu Arg
 290

<210> 11

<211> 158

<212> PRT

<213> 荚膜甲基球菌

<400> 11

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

[0018]

65	70	75	80
Val Arg Gln Arg Ile Ala Ser Leu Gly Tyr Thr Val Gly Asn Val Asp			
	85	90	95
Val Thr Val Val Ala Gln Ala Pro Arg Leu Ala Ala His Ile Gln Ala			
	100	105	110
Met Arg Glu Asn Leu Ala Gln Asp Leu Glu Ile Pro Pro Asp Cys Val			
	115	120	125
Asn Val Lys Ala Thr Thr Thr Glu Gly Met Gly Phe Glu Gly Arg Gly			
	130	135	140
Glu Gly Ile Ser Ala His Ala Val Ala Leu Leu Ala Arg Arg			
145	150	155	

<210> 12

<211> 407

<212> PRT

<213> 荚膜甲基球菌

<400> 12

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

[0019]

Glu Leu Ile Ser Val Tyr Glu Ala Leu Ser Ser Arg Cys Asp His Ala
 210 215 220
 Leu His Leu Gly Leu Thr Glu Ala Gly Met Gly Ser Lys Gly Ile Val
 225 230 235 240
 Ala Ser Thr Ala Ala Leu Ser Val Leu Leu Gln Gln Gly Ile Gly Asp
 245 250 255
 Thr Ile Arg Ile Ser Leu Thr Pro Glu Pro Gly Ala Asp Arg Ser Leu
 260 265 270
 Glu Val Ile Val Ala Gln Glu Ile Leu Gln Thr Met Gly Leu Arg Ser
 275 280 285
 Phe Thr Pro Met Val Ile Ser Cys Pro Gly Cys Gly Arg Thr Thr Ser
 290 295 300
 Asp Tyr Phe Gln Lys Leu Ala Gln Gln Ile Gln Thr His Leu Arg His
 305 310 315 320
 Lys Met Pro Glu Trp Arg Arg Arg Tyr Arg Gly Val Glu Asp Met His
 325 330 335
 Val Ala Val Met Gly Cys Val Val Asn Gly Pro Gly Glu Ser Lys Asn
 340 345 350
 Ala Asn Ile Gly Ile Ser Leu Pro Gly Thr Gly Glu Gln Pro Val Ala
 355 360 365
 Pro Val Phe Glu Asp Gly Val Lys Thr Val Thr Leu Lys Gly Asp Arg
 370 375 380
 Ile Ala Glu Glu Phe Gln Glu Leu Val Glu Arg Tyr Ile Glu Thr His
 385 390 395 400
 Tyr Gly Ser Arg Ala Glu Ala
 405

<210> 13

<211> 313

<212> PRT

<213> 荚膜甲基球菌

<400> 13

Met Glu Ile Ile Leu Ala Asn Pro Arg Gly Phe Cys Ala Gly Val Asp
 1 5 10 15
 Arg Ala Ile Glu Ile Val Asp Arg Ala Ile Glu Val Phe Gly Ala Pro
 20 25 30
 Ile Tyr Val Arg His Glu Val Val His Asn Arg Tyr Val Val Asp Gly
 35 40 45
 Leu Arg Glu Arg Gly Ala Val Phe Val Glu Glu Leu Ser Glu Val Pro
 50 55 60
 Glu Asn Ser Thr Val Ile Phe Ser Ala His Gly Val Ser Lys Gln Ile
 65 70 75 80
 Gln Glu Glu Ala Arg Glu Arg Gly Leu Gln Val Phe Asp Ala Thr Cys

[0020]

85	90	95
Pro Leu Val Thr Lys Val His Ile Glu Val His Gln His Ala Ser Glu		
100	105	110
Gly Arg Glu Ile Val Phe Ile Gly His Ala Gly His Pro Glu Val Glu		
115	120	125
Gly Thr Met Gly Gln Tyr Asp Asn Pro Ala Gly Gly Ile Tyr Leu Val		
130	135	140
Glu Ser Pro Glu Asp Val Glu Met Leu Gln Val Lys Asn Pro Asp Asn		
145	150	155
Leu Ala Tyr Val Thr Gln Thr Thr Leu Ser Ile Asp Asp Thr Gly Ala		
165	170	175
Val Val Glu Ala Leu Lys Met Arg Phe Pro Lys Ile Leu Gly Pro Arg		
180	185	190
Lys Asp Asp Ile Cys Tyr Ala Thr Gln Asn Arg Gln Asp Ala Val Lys		
195	200	205
Lys Leu Ala Ala Gln Cys Asp Thr Ile Leu Val Val Gly Ser Pro Asn		
210	215	220
Ser Ser Asn Ser Asn Arg Leu Arg Glu Ile Ala Asp Lys Leu Gly Arg		
225	230	235
Lys Ala Phe Leu Ile Asp Asn Ala Ala Gln Leu Thr Arg Asp Met Val		
245	250	255
Ala Gly Ala Gln Arg Ile Gly Val Thr Ala Gly Ala Ser Ala Pro Glu		
260	265	270
Ile Leu Val Gln Gln Val Ile Ala Gln Leu Lys Glu Trp Gly Gly Arg		
275	280	285
Thr Ala Thr Glu Thr Gln Gly Ile Glu Glu Lys Val Val Phe Ser Leu		
290	295	300
Pro Lys Glu Leu Arg Arg Leu Asn Ala		
305	310	

<210> 14

<211> 1788

<212> DNA

<213> 银白杨

<400> 14

atggccactg aacttctttg tttgcaccgc ccgatttccc tgacccataa gctgtttcgc	60
aacctctctgc ccaaagttat ccaggcaacc ccgctgacgc tcaagctccg gtgcagcgta	120
tccaccgaaa atgtatcggt caccgaaacc gaaactgaag cccgtcgcag cgcgaactac	180
gagcccaact cgtgggatta cgactatctg ctgagctcgg ataccgacga atccatcgaa	240
gtctataagg acaaagccaa gaagctcgaa gccgaggtgc gccgtgagat caacaacgag	300
aaggccgagt tectgacctt gttggaactg atcgacaacg tccagcgcct gggcctcggc	360
taccggttcg agagcgatat ccggggtgcc ctggaccgtt tegtcagctc gggcggattc	420
gacgcagtga ccaaaacgtc gctgcatggg acggccctgt ccttcctctt gctgcgcag	480

[0021]

```

catggcttcg aggtgtccca ggaagccttc agcggcttca aggatcagaa cggaaacttt 540
ctgaaaaact tgaaagagga catcaaggcc atcttcagcc tgtacgagge gtccttcctg 600
gccctcgaag gtgaaaacat cctcgatgaa gccaaagggt tgcgaatctc gcattctaaa 660
gagctgtccg aagagaagat tggcaaagag ctggcgaac aagtcaacca cgcgttgag 720
ctgcccctcc accggcgcac ccagcggctg gaagcggctt ggtegatcga agcctaccgc 780
aagaaagagg acgccaatca ggctctgctg gagetcgga ttctggatta caatatgac 840
cagtcggctt atcagcgcga tcctgcgcga acgtcccggt ggtggcggcg tctcggcttg 900
gegaccaagt tgcacttcgc gcgtgaccgc ttgatcgaga gcttctattg ggccgtcggg 960
gtggcctttg agccccagta ctccgactgc cgcaatagcg tggcgaagat gtacagcttc 1020
gttaccatca tegacgacat ctacgacgtg tatggcacgc tcgacgagct cgaactgttc 1080
accgacgccg tggaaacttg ggaactcaac gccatcaatg atctccccga ctacatgaag 1140
ctgtgcttcc tggcggtgta taacaccatc aacgagattg cctacgataa cctcaaggac 1200
aaggcgaga acatcctgcc gtacttgacc aagcctggg ccgatttgtg caacgccttt 1260
ctgcaggaag caaagtggct gtacaacaaa tccacgccga cgttcgacga ctatttcggc 1320
aatgcatgga aatcgagctc gggctctcgt caacttgtgt tcgctactt cgcctcgtg 1380
cagaatatca agaaagaaga aatcgagaac cttcagaaat atcatgacac caccagcgt 1440
ccatcgaca tctttcgct gtgcaacgac ctgcgtccg catccgccga gatcgacgc 1500
ggcgaacgg ccaattcggg gtctgctac atcgcgacca aggcctctc ggaagagctg 1560
gcgacggaat ccgtgatgaa cctgatcgat gaaacctgga agaagatgaa caaagagaag 1620
ctcggcgga gcctgttcgc gaagcccttc gtcgaaaccg caattaacct ggcacgcaa 1680
tcccactgta cctaccataa cggagatgcc cacacgagcc cggacgagct gactcgcaag 1740
cgcgtccttt cggctcatcac cgagccgac ctgccgttcg agcggtaa 1788

```

<210> 15

<211> 1680

<212> DNA

<213> 银白杨

<400> 15

```

atgtgcagcg tatccaccga aaatgtatcg ttaccgaaa ccgaactga agcccgctgc 60
agcgcgaact acgagcccaa ctcgigggat tacgaactat tctgagctc ggataccgac 120
gaatccatcg aagtctataa ggacaaagcc aagaagctcg aagccgaggt gcgcccgtgag 180
atcaacaacg agaaggccga gtctctgacc ctgttggaac tgatcgacaa cgtccagcgc 240
ctgggcctcg gctaccggtt cgagagcgat atccgggggt cctggaccg ttctgtcagc 300
tcgggcggat tcgacgcagt gacaaaaacg tcgtgcatg ggacggccct gtccttcctg 360
ctgtgcgcc agcatggett cgaggtgtcc caggaaacct tcagoggctt caaggatcag 420
aacggaaaact ttctgaaaaa cttgaaagag gacatcaagg ccctctcag ctgtacgag 480
gcgtcccttc tggccctcga aggigaaaac atctctgatg aagccaaggt gtctgcaatc 540
tcgcatctta aagagctgtc cgaagagaag attggcaaag agctggccga acaagtcaac 600
cacgcgttgg agctgcgct ccaccggcgc acccagcggc tggaaagcgt ctggctgac 660
gaagcctacc gcaagaaaga ggacgccaat caggctctgc tggagctcgc gattctggat 720
tacaatatga tccagtcggg ctatcagcgc gatctgcgcg aaacgtcccg gtggtggcgg 780
cgtgtcggct tggcgaccaa gttgcacttc gcgctgacc gcttgatcga gagctctat 840
tgggcgctcg ggggtgacct tgagcccag tactccgact gccgcaatag cgtggcgaag 900

```

[0022]

atgttcagct	tgtttaccat	catcgacgac	atctacgacg	tgtatggcac	gctcgacgag	960
ctcgaactgt	tcaccgacgc	cgtggaacgt	tgggacgtca	acgccatcaa	tgatctcccc	1020
gactacatga	agctgtgctt	cctggcggtg	tataacacca	tcaacgagat	tgcctacgat	1080
aacctcaagg	acaagggcga	gaacatcctg	ccgtacttga	ccaaggcctg	ggcgcatttg	1140
tgcacgcct	ttctgcagga	agcaaagtgg	ctgtacaaca	aatccacgcc	gacgttcgac	1200
gactatttcg	gcaatgcctg	gaaatcgagc	tcgggtcctc	tgcaccttgt	gttcgcgtac	1260
ttcgccgtcg	tgcagaatat	caagaaagaa	gaaatcgaga	accttcagaa	atatcatgac	1320
accatcagcc	gtccatcgca	catctttcgc	ctgtgcaacg	acctcgcgtc	cgcctccgcc	1380
gagatcgac	gcggcgaaac	ggccaattcg	gtgtcctgct	acatgcggac	caagggcac	1440
tcggaagagc	tggcgacgga	atccgtgatg	aacctgatcg	atgaaacctg	gaagaagatg	1500
aacaagagga	agctcggcgg	gagcctgttc	gcgaagccct	tcgtcgaaac	cgcaattaac	1560
ctggcacgcc	aatcccactg	tacctaccat	aacggagatg	cccacacgag	cccggacgag	1620
ctgactcgca	agcgcgtect	ttcggtcctc	accgagccga	tcctgcctgt	cgagcggtaa	1680

<210> 16

<211> 1827

<212> DNA

<213> 山葛

<400> 16

atggccacca	atctgctctg	cctgtcgaat	aaactgtcca	gccccacgcc	cacgccgtcc	60
acggggttcc	cgcagtcctc	gaacttcatt	acccagaaaa	ccagcctcgc	caaccggaag	120
ccatggcgcg	tgatctgcgc	aacctcgtcc	caattcaccc	agatcacgga	acacaactcg	180
cgtegetcgg	ccaactacca	gcctaatttg	tggaaattcg	agttcctgca	gagcttgtag	240
aacgatctga	aggtcgagaa	gctggaagag	aaagccacca	agctcgaaga	agaggctcgt	300
tgcattgaca	accgcgtcga	cactcagcgc	ctctccctgc	tggagcttat	cgcacagctc	360
cagcgcctcg	gcttgactta	caagttcgag	aaagacatta	tcaaggccct	tgagaatata	420
gtctgtctgg	atgaaaacaa	aaagaacaag	tcggatctgc	atgcgaccgc	cctgagcttc	480
cggctgtctg	gccagcacgg	ctttgaggtc	agccaagacg	tattcgaacg	cttcaaggat	540
aaagaaggcg	ggttttcggg	cgaattgaaa	ggcgacgtgc	agggettgtc	ctcgtgttac	600
gaggccagct	acctgggctt	tgagggtgaa	aatctgctcg	aagaggcgcg	tacctcagc	660
atcacgcctc	tgaagaataa	cctcaagag	ggcatcaaca	ccaagggtgc	cgaacaagtg	720
tcaccaagcg	tggaaactgc	ataccatcaa	cggctgcctc	gcctggaagc	gcgtgtgttc	780
ttggacaagt	atgaacccaa	agaacctcac	catacgtcgc	ttctggagct	cgcgaagttg	840
gacttcaaca	tggctccagc	cttgccaccg	aaagaactgc	aggacttgte	ccggtgtgtg	900
accgaaatgg	gactggcgct	caagcttgac	ttcgtccgcg	atcgccctat	ggaagtgtac	960
ttttgggccc	tcggaatggc	accggaccgc	cagttcggcg	agtgcgcgaa	agcagttacc	1020
aagatgttcg	gcctgggtc	cattatcgac	gaigtctacg	acgtatacgg	gacgttggtg	1080
gagctgaac	tgttcacgga	cgcgttgtag	cgttggtgac	tcaacgccat	caacacgctc	1140
cccgactata	tgaagctctg	cttccttgca	ttgtacaata	ccgtgaacga	cacctcgtat	1200
tcattcttga	aagaaaaagg	acacaataac	ctgtcctatc	tgaccaagtc	ctggcgtgag	1260
ctgtgcaagg	cgttctctga	agaagccaag	tggagcaata	acaagatcat	ccccgcgttc	1320
tcgaagtatc	ttgagaacgc	atccgtgtcg	agcagcgggg	tcgccctgct	ggccccgtcg	1380
tacttcagcg	tatgtcagca	gcaggaagat	atctcggacc	acgcgctgcg	tagccttacg	1440

[0023]

gacttccatg gccctcgtccg gtcgagctgc gtgatcttcc gtttgtgcaa cgacctggcg 1500
 acctcggccg cagaactgga gcggggtgaa accaccaaca gcatcatctc gtacatgcac 1560
 gagaacgatg gcacgtcgga agagcaggca cgcgaagagc tgcgtaagct gatcgacgcc 1620
 gagtgggaaga aaatgaaccg cgaacgcgtc agcgactcca ccctgctgcc gaaggccttc 1680
 atggaaatcg ccgtgaacat ggcacgtgtg tcccattgta cttatcagta cggcgatggc 1740
 ctgggtcgcc ccgactatgc caccggagaac cggatcaagc tectgttgat cgatccgttc 1800
 ccgatecaacc agctgatgta cgtgtaa 1827

<210> 17

<211> 1695

<212> DNA

<213> 山葛

<400> 17

atgtgcgcaa cctcgtccca attcaccag atcaccgaac acaactcgcg tcgtcggcc 60
 aactaccagc ctaatttgtg gaacttcgag ttcttcgaga gcttgagaa cgatctgaag 120
 gtcgagaagc tggaaagaga agccaccaag ctcgagaag aggtccgttg catgatcaac 180
 cgcgtcgaca ctacgccgt ctccctgtcg gagcttatcg acgacgtcca gcgcctcggc 240
 ttgacttaca agttcgagaa agacattatc aaggcccttg agaatatcgt cctgctggat 300
 gaaaacaaaa agaacaangtc ggatctgcac gcgaccgccc tgagcttcgg gctgctgcgc 360
 cagcacggtt ttgaggtcag ccaagacgta ttcgaaagct tcaaggataa agaaggcggg 420
 ttttcggcg aattgaaagg cgacgtcgag ggcttgcctt cgctgtacga ggccagctac 480
 ctgggctttg aggggtgaaa tctgctcgaa gaggcgcgta ccttcagcat cacgeatctg 540
 aagaataacc tcaagagggg catcaacacc aagggtggcg aacaagtgtc ccacgcgctg 600
 gaactgccat accatcaacg gctgcacgct ctggaagcgc gctggttctt ggacaagtat 660
 gaacccaaag aacctcacca tcagctgctt ctggagctcg ccaagttgga cttaacatg 720
 gtccagacct tgcaccagaa agaactgcag gacttgccc ggtggtggac cgaaatggga 780
 ctggcgcca agcttgactt cgtcccgcat cgcctcatgg aagtgtactt ttgggccctc 840
 ggaatggcac cggaccgca gttcggcgag tgcgcgaaag cagttacca gatgttcggc 900
 ctggtcacca ttatcgacga tgtctacgac gtatcggga cgttggaiga gctgcaactg 960
 ttcacggacg ccgtggagcg gtgggacgct aacgccatca acacgtccc cgaatatatg 1020
 aagctctgct tctggcatt gtacaatacc gtgaacgaca cctcgtatc cattctgaaa 1080
 gaaaaaggac acaataacct gtctatctg accaagtcct ggctgagct gtgcaaggcg 1140
 ttctgcaag aagccaagt gagcaataac aagatcatcc ccgcgtctc gaagtatctt 1200
 gagaacgcat ccgtgtcgag cagcggggtc gccctgctgg ccccgctgta cttcagcgta 1260
 igtacgagc aggaagatat ctcgaccac gcgttcgta gccttacgga ctccatggc 1320
 ctgctccggt cgagctcgct gatcttcgt ttgtgcaacg acctggcgac ctggcgcca 1380
 gaactggagc ggggtgaaac caccaacagc atcatctcgt acatgcacga gaacgatggc 1440
 acgtcggag agcaggcacg cgaagagctg cgtgaagctga tcgacgccga gtggaagaaa 1500
 atgaaccgag aacgcgtcag cgactccacc ctgctgcga aggccttcac ggaatcgcc 1560
 gtgaacatgg cacgtgtgtc ccattgtact tatcagtacg gcgatggcct gggtcgcccc 1620
 gactatgcca cggagaaccg gatcaagctc ctgttgatcg atccgtccc gatcaaccag 1680
 ctgatgtacg tgtaa 1695

[0024]

<210> 18
 <211> 1788
 <212> DNA
 <213> 柳属

<400> 18

```

atggccactg aacttcctgtg cttgcaccgt cccatttcgc tcaccacctaa actgttccgc      60
aaccgcctcc cgaaggtaat cctggcgacg ccgtgaccc tgaagctgcg gtgcagcgta      120
tccaccgaaa acgtgagctt tactgaaacc gaaaccgaaa cgcgtcgcctc ggcgaaactac      180
gaacccaatt cctgggatta tgactacctt ctgtcgtccg acacggacga gtgatcgag      240
gtgtataagg ataaggccaa gaagcttgag gcggaagtcc gtcgggagat caacaacgag      300
aaggcggagt tectgacgt gctcgaaactg attgacaacg tccagcgctt cggcctgggc      360
tategcttcg agtcgatata ccgtcgcgca ctgaccgct tegtctcgtc cgggtggettc      420
gacgcagtga cgaaaacctc gctgcattgc accgcgctgt egttccgctt cctgcgccag      480
cacggattcg aggtcagcca ggaagcgttc ggcggggttc aggaccagaa cgggaatttc      540
ctggaanaac tgaagaaga tatcaaaccc atcttgctgc tgtacgagc gtctttctc      600
gcgtcgaag gcgagaacat tctcgacgaa gcgaaggtgt tcgccatctc gcacctgaaa      660
gagctctccg aagagaagat cggcaaagac ttggccgagc aagtcataca cgccttgag      720
ttgcccctgc atgcgcgcac ccagcgcttg gaagccgttt ggagcattga agcctatcgt      780
aagaaagagg acgccaacca agtctgtctg gagetggcca tcttgacta caacatgatc      840
cagtcctgtg accagcggga cttgcgcgaa accagccggt ggtggcgtcg cgtcggcctc      900
gccaccaagc tgcacttcgc acgcgaccgc ctgatcgagt ccttctaetg ggccgtgggc      960
gtgcatttcg agccgcaata tagcgactgc cggaaacagc tggeaaagat gttcagcttc      1020
gtgaccatca tcgacgatat ctatgacgtg tatgggacgc ttgacgaact ggagctgttt      1080
atggatgccg tcgagcggtg ggacgtcaat gccatcaacg atttgccgga ctacatgaag      1140
ctgtgcttcc tggccttgta taacctatc aacgagatcg cctacgataa cctgaaagaa      1200
aagggtgaga acatcctgcc ctacctacc aaggcctggg ccgacctgtg taacgccttt      1260
ctgcaggaag ccaagtggct ctacaacaag tccaccctaa ccttcgacga ttacttcgga      1320
aatgcctgga agagcagctc cggacctctc cagctgggtg tgcatactt cgccgtcgtg      1380
cagaacatca agaaagaaga gatcgaaaac ttgcagaagt accacgatat catcagcgt      1440
ccctcgcaca tcttcgggt ctgcaacgac cttgcaagcg cgtccgcgga gatcgcaagg      1500
ggcgaaacgg ccaactcgtt gagctgtctc atgcgcacca agggcatctc ggaagaactt      1560
gcgacggagt ccgtcatgaa cttgatcgac gaaacctgga agaaaatgaa taaagagaaa      1620
ctcggcgga gccgtgtccc gaagccattc gtcgaaaccg ccatcaacct ggccgctcag      1680
tcgcatigca cclaccataa tggcgatgcc catacgtcgc cggatgaact gaccgtaag      1740
cgggtcctgt ccgtcateac cgagccgatt ctgccgttcg agcgttaa      1788

```

<210> 19
 <211> 1680
 <212> DNA
 <213> 柳属

<400> 19

```

atgtgcagcg tatccaccga aaacgtgagc ttactgaaa ccgaaaccga aacgcgtcgc      60

```

[0025]

tcggcgaaact	acgaacccaa	ttcctgggat	taigactacc	ttctgtcgtc	cgacacggac	120
gagtcgatcg	aggtgtataa	ggataaggcc	aagaagcttg	aggcggaagt	ccgtcgggag	180
atcaacaacg	agaaggcgga	gttctgacg	ctgctcgaa	tgattgacaa	cgtccagcgc	240
ctcggcctgg	gctatcgctt	cgagtcgat	atccgtcgcg	cactcgaccg	cttcgtttcg	300
tccggtggct	tcgacgcagt	gacgaaaacc	tcgtcgcatg	ccaccgcgct	gtcgttccgc	360
ttcctgcgcc	agcacggatt	cgaggtcagc	caggaagcgt	tcggcggtt	caaggaccag	420
aacgggaatt	tcctggaaaa	tctgaaagaa	gatataaag	ccatcttgtc	gctgtacgag	480
gcgtcgtttc	tcgcgctcga	aggcgagaac	attctcgacg	aagcgaaggt	gttcgccatc	540
tcgcacctga	aagagctctc	cgaagagaag	atcggaag	acttgccga	gcaagtcaat	600
cacgccttgg	agttgccctt	gcacgcgcgc	acccagegct	tggagccgt	ttggagcatt	660
gaagcctatc	gtaagaaaga	ggacgccaac	caagtcctgc	tggagctggc	catcctggac	720
tacaacatga	tccagtcctg	gtaccagcgg	gaacttgcgcg	aaaccagccg	gtggtggcgt	780
cgcgtcggcc	tcgccaccaa	gctgcacttc	gcacgcgacc	gcctgatcga	gtccttctac	840
tgggccttgg	gcgtcgcatt	cgagccgcaa	tatagcgact	gccggaacag	cgtggcaaaag	900
atgttcagct	tcgtgaccat	catcgacgat	atctatgacg	tgtatgggac	gcttgacgaa	960
ctggagctgt	ttacggatgc	cgctcgagcg	tgggacgtca	atgccatcaa	cgatttgcgc	1020
gactacatga	agctgtgctt	cctggccttg	tataaacta	tcaacgagat	cgcctacgat	1080
aaactgaaag	aaaagggtga	gaacatcctg	ccctacctca	ccaaggcctg	ggccgacctg	1140
tgtaacgcct	ttctgcagga	agccaagtgg	ctctacaaca	agtccacccc	aaccttcgac	1200
gattacttgc	gaaatgcctg	gaagagcagc	tccggacctc	tccagctggg	gttcgcatac	1260
ttcgccgtcg	tgcagaacat	caagaaagaa	gagatcgaaa	acttgcagaa	gtaccacgat	1320
atcatcagcc	gtccctcgca	catcttcggg	ctctgcaacg	accttgcaag	cgcgtccgcg	1380
gagatcgcac	ggggcgaaac	ggccaactcg	gtgagctgct	acatgcgcac	caagggcatac	1440
tcggaagaac	ttgcgacgga	gtccgtcatg	aacttgatcg	acgaaacctg	gaagaaaatg	1500
aataaagaga	aactcggcgg	cagcctgttc	ccgaagccat	tcgtcgaaac	cgccatcaac	1560
ctggcgcgtc	agtcgcattg	cacctaccat	aatggcgatg	cccatacgtc	gccggatgaa	1620
ctgacccgta	agcgggtcct	gtccgtcctc	accgagccga	ttctgccgtt	cgagcgctaa	1680

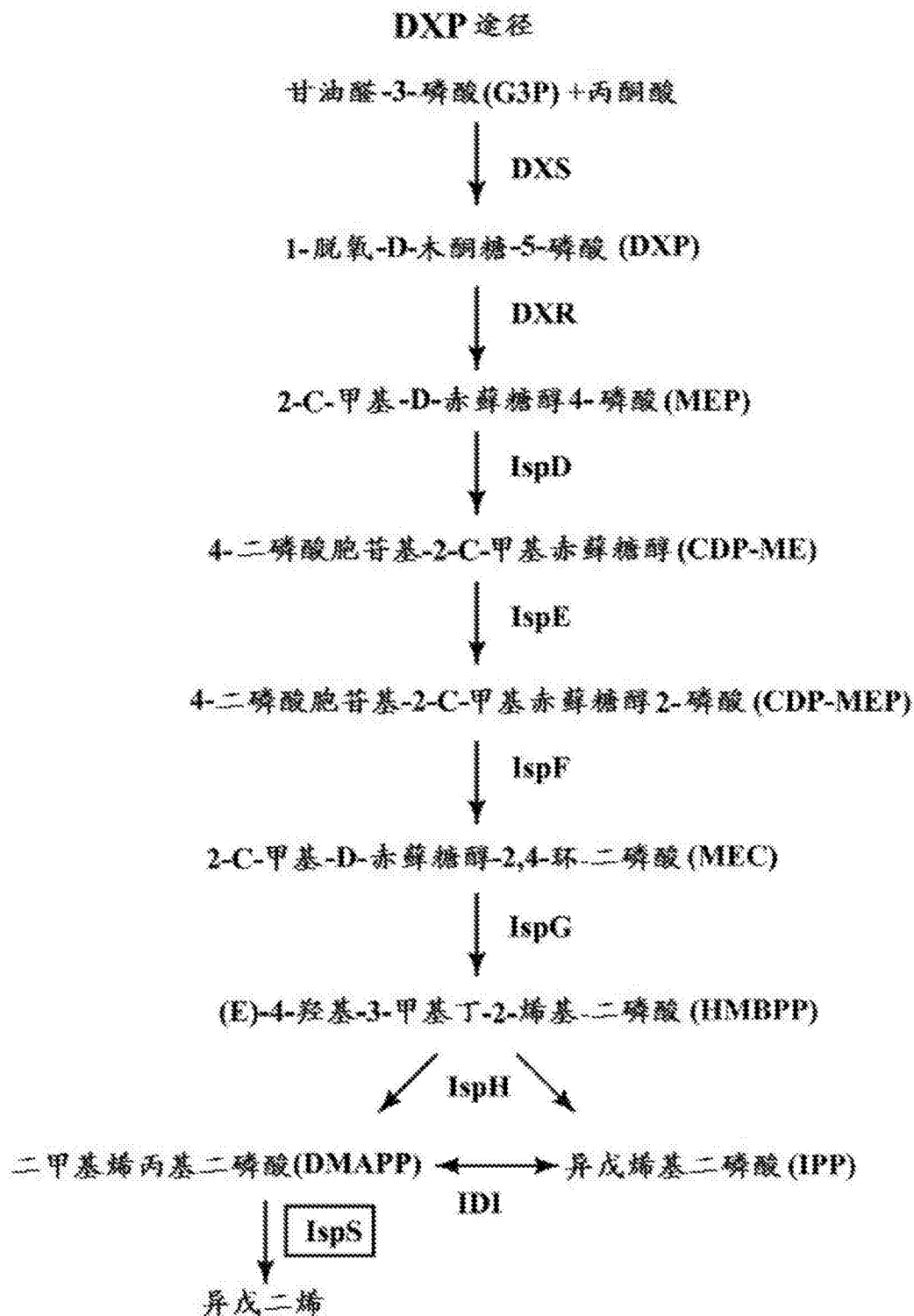


图 1

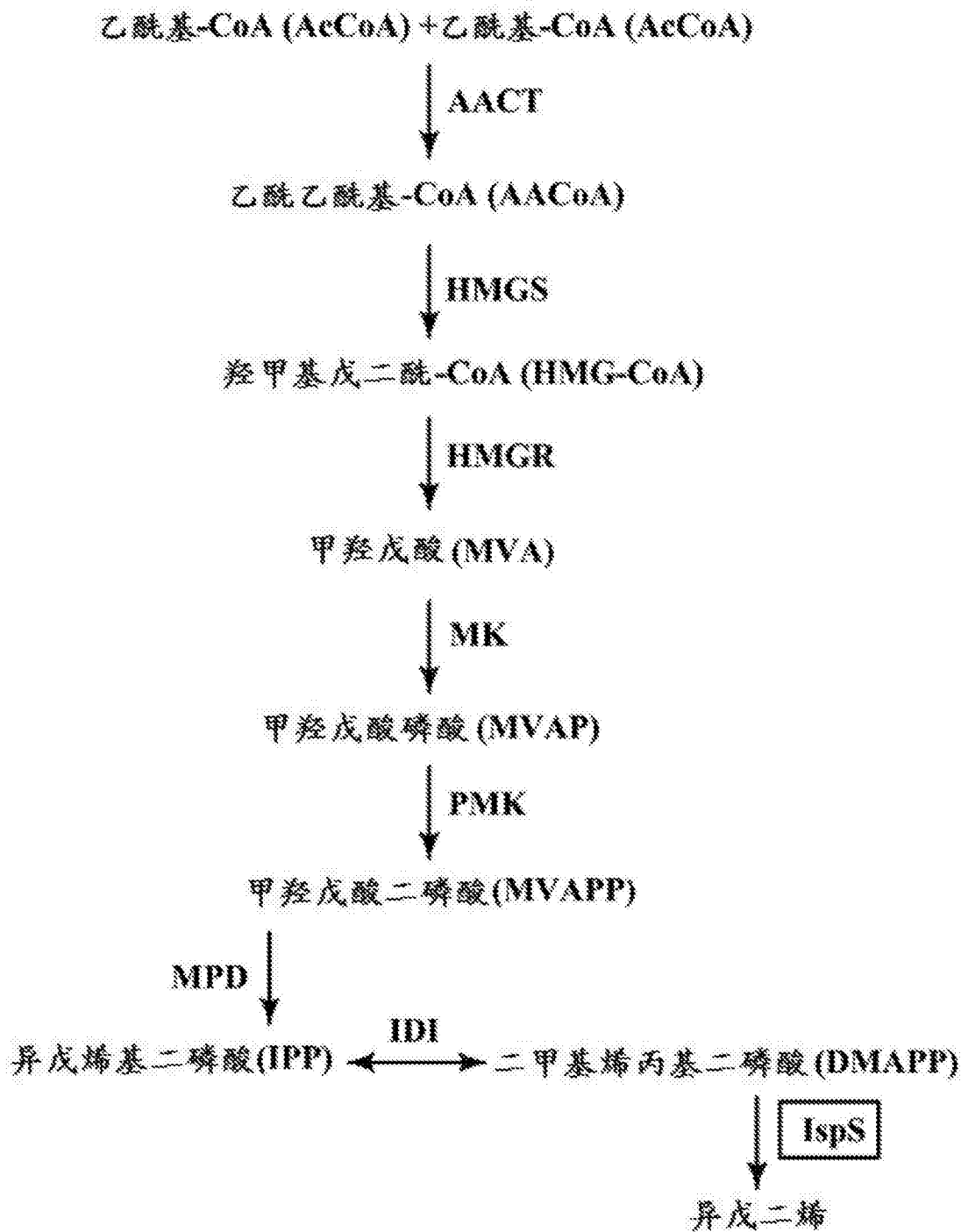
MVA 途径

图 2

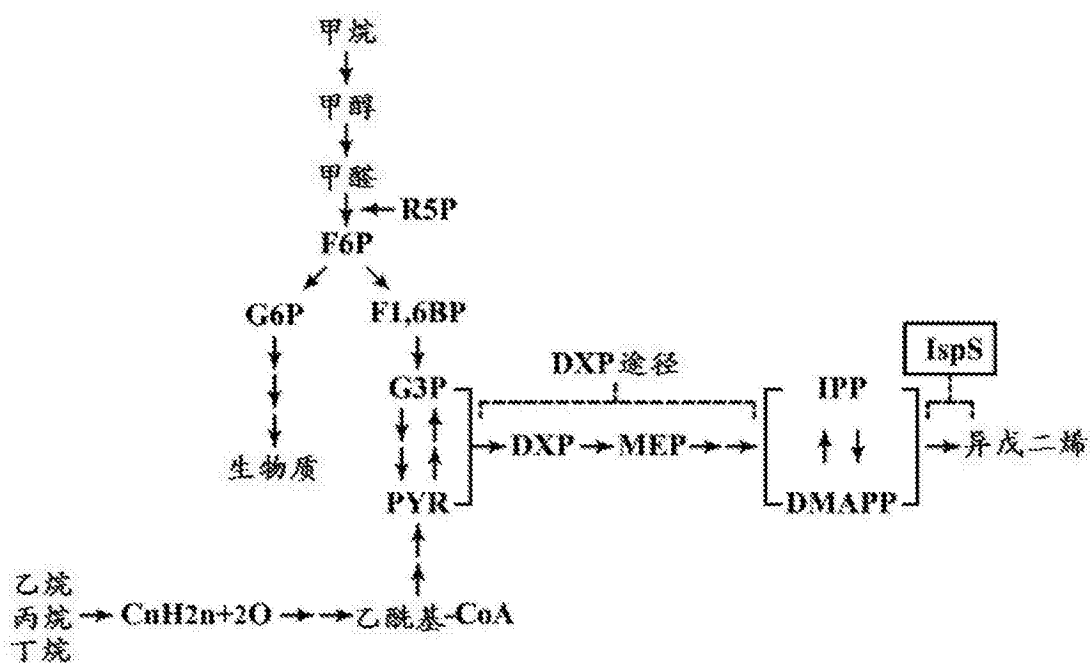


图 3

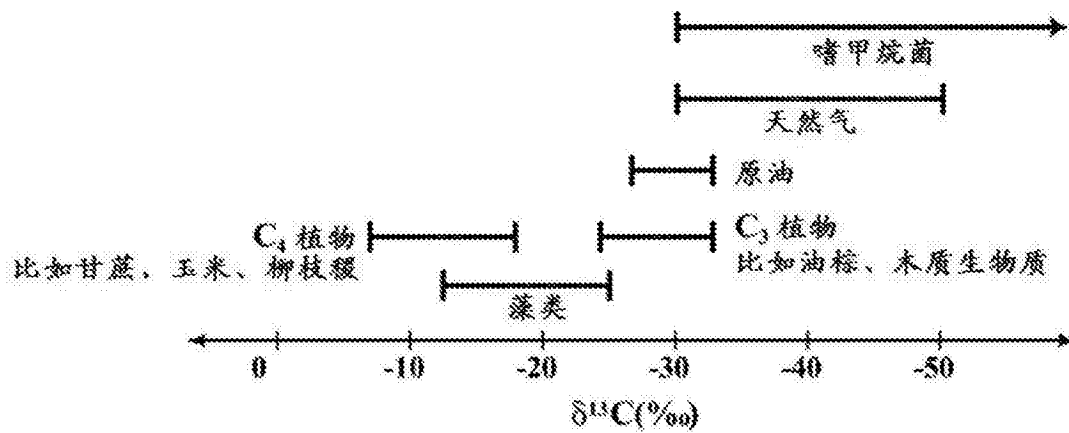


图 4

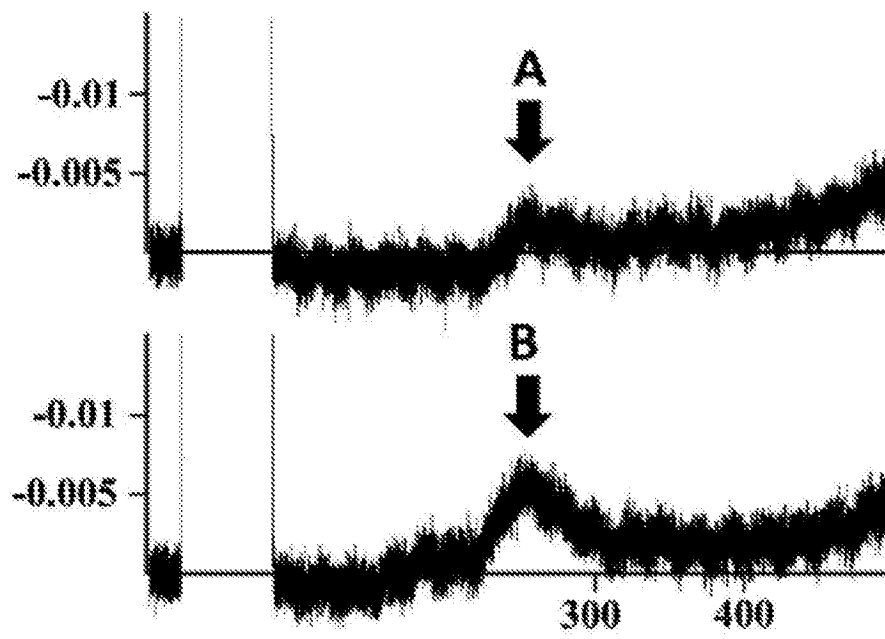


图 5

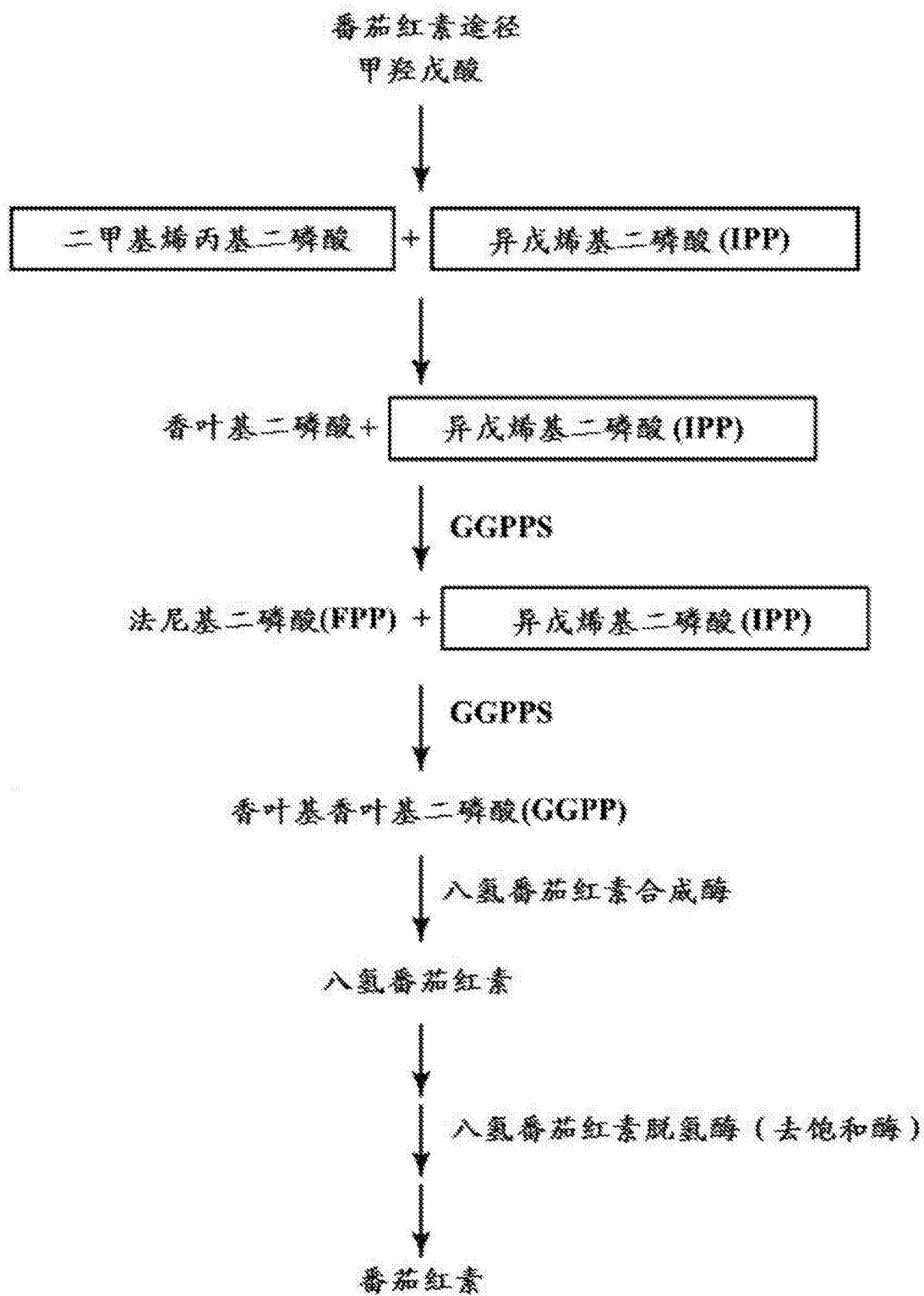


图 6

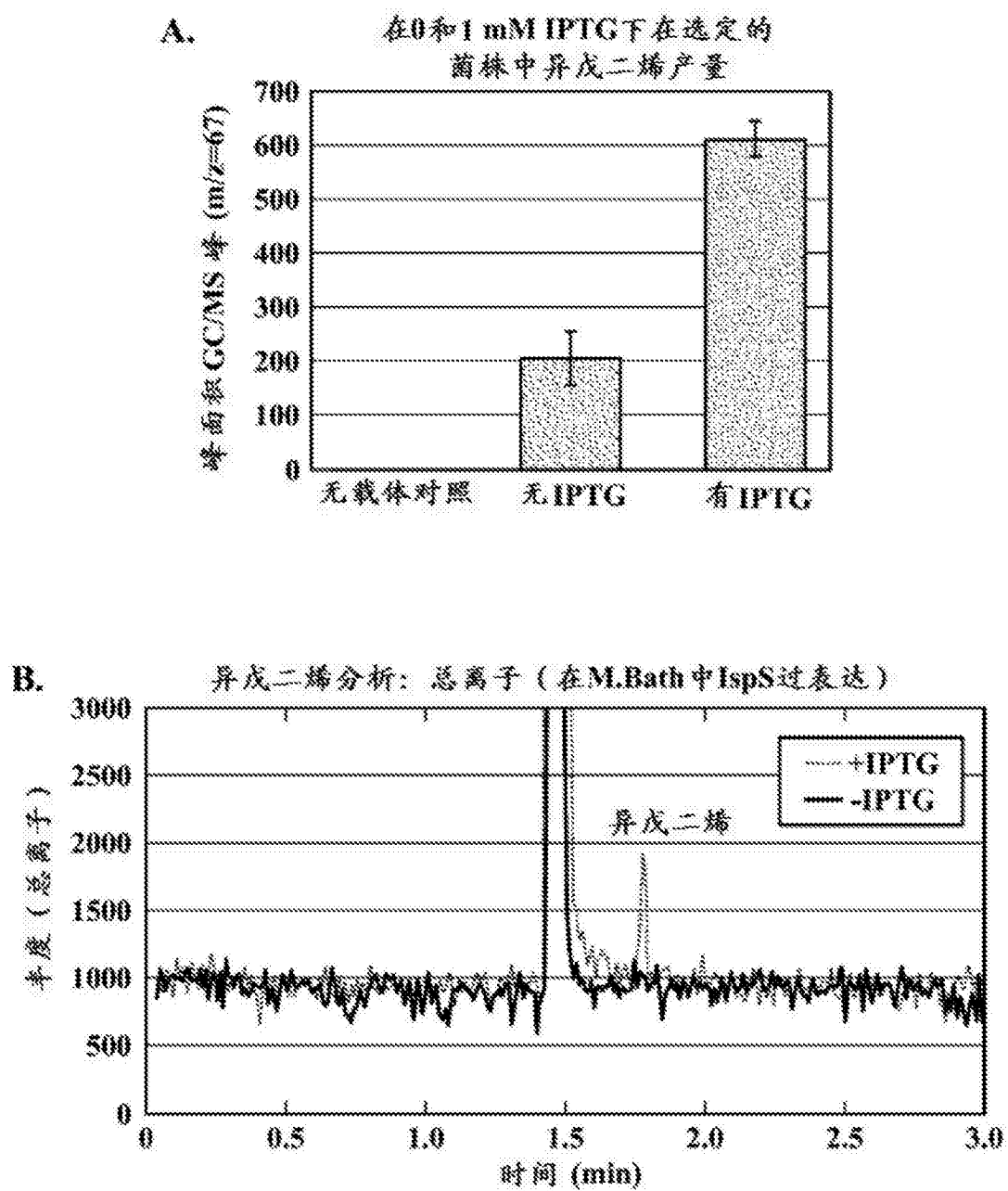


图 7