

(12) 发明专利申请

(10) 申请公布号 CN 102834508 A

(43) 申请公布日 2012. 12. 19

(21) 申请号 201180016104. 6

(22) 申请日 2011. 01. 21

(30) 优先权数据

61/299, 794 2010. 01. 29 US

(85) PCT申请进入国家阶段日

2012. 09. 26

(86) PCT申请的申请数据

PCT/US2011/022046 2011. 01. 21

(87) PCT申请的公布数据

W02011/094131 EN 2011. 08. 04

(71) 申请人 基因组股份公司

地址 美国加利福尼亚州

(72) 发明人 罗宾·E·奥斯特豪特

(74) 专利代理机构 北京安信方达知识产权代理

有限公司 11262

代理人 武晶晶 郑霞

(51) Int. Cl.

C12N 9/00 (2006. 01)

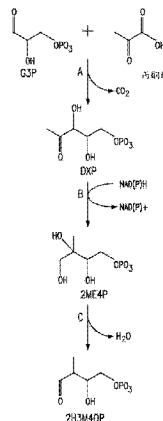
权利要求书 4 页 说明书 37 页 附图 3 页

(54) 发明名称

用于生物合成对甲苯甲酸和对苯二甲酸的微生物和方法

(57) 摘要

本发明提供具有 (2- 羟基 -3- 甲基 -4- 氧代丁氧基) 磷酸酯途径、对甲苯甲酸途径和 / 或对苯二甲酸途径的非天然存在的微生物。本发明另外还提供使用这类生物体来产生 (2- 羟基 -3- 甲基 -4- 氧代丁氧基) 磷酸酯途径、对甲苯甲酸途径或对苯二甲酸途径的方法。



1. 非天然存在的微生物,包括一种具有(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯途径的微生物,该微生物包含至少一种编码以足以产生(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯的量表达的(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯途径酶的外源核酸,所述(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯途径包含2-C-甲基-D-赤藓糖醇-4-磷酸脱水酶。

2. 权利要求1的非天然存在的微生物,其中所述(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯途径还包含1-脱氧木酮糖-5-磷酸合酶或1-脱氧-D-木酮糖-5-磷酸还原异构酶。

3. 权利要求1的非天然存在的微生物,其中所述(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯途径还包含1-脱氧木酮糖-5-磷酸合酶和1-脱氧-D-木酮糖-5-磷酸还原异构酶。

4. 权利要求1的非天然存在的微生物,其中所述微生物包含三种外源核酸,各自编码(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯途径酶。

5. 权利要求4的非天然存在的微生物,其中所述三种外源核酸编码2-C-甲基-D-赤藓糖醇-4-磷酸脱水酶、1-脱氧木酮糖-5-磷酸合酶和1-脱氧-D-木酮糖-5-磷酸还原异构酶。

6. 权利要求1的非天然存在的微生物,其中所述至少一种外源核酸是异源核酸。

7. 权利要求1的非天然存在的微生物,其中所述非天然存在的微生物是在基本上厌氧的培养基中。

8. 用于生产(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯的方法,包括将权利要求1的非天然存在的微生物在足以产生(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯的条件下培养足够的时间周期。

9. 权利要求8的方法,其中所述非天然存在的微生物是在基本上厌氧的培养基中。

10. 权利要求8的方法,其中所述微生物包含三种外源核酸,各自编码(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯途径酶。

11. 权利要求10的方法,其中所述三种外源核酸编码2-C-甲基-D-赤藓糖醇-4-磷酸脱水酶、1-脱氧木酮糖-5-磷酸合酶和1-脱氧-D-木酮糖-5-磷酸还原异构酶。

12. 权利要求8的方法,其中所述至少一种外源核酸是异源核酸。

13. 非天然存在的微生物,包括一种具有对甲苯甲酸途径的微生物,该微生物包含至少一种编码以足以产生对甲苯甲酸的量表达的对甲苯甲酸途径酶的外源核酸,所述对甲苯甲酸途径包含2-脱氢-3-脱氧磷酸庚糖酸合酶、3-脱氢奎尼酸合酶、3-脱氢奎尼酸脱水酶、莽草酸脱氢酶、莽草酸激酶、3-磷酸莽草酸-2-羧基乙烯基转移酶、分支酸合酶或分支酸裂合酶。

14. 权利要求13的非天然存在的微生物,其中所述微生物包含两种外源核酸,各自编码对甲苯甲酸途径酶。

15. 权利要求13的非天然存在的微生物,其中所述微生物包含三种外源核酸,各自编码对甲苯甲酸途径酶。

16. 权利要求13的非天然存在的微生物,其中所述微生物包含四种外源核酸,各自编码对甲苯甲酸途径酶。

17. 权利要求13的非天然存在的微生物,其中所述微生物包含五种外源核酸,各自编码对甲苯甲酸途径酶。

18. 权利要求13的非天然存在的微生物,其中所述微生物包含六种外源核酸,各自编

码对甲苯甲酸途径酶。

19. 权利要求 13 的非天然存在的微生物,其中所述微生物包含七种外源核酸,各自编码对甲苯甲酸途径酶。

20. 权利要求 13 的非天然存在的微生物,其中所述微生物包含七种外源核酸,各自编码对甲苯甲酸途径酶。

21. 权利要求 19 的非天然存在的微生物,其中所述八种外源核酸编码 2-脱氢-3-脱氧磷酸庚糖酸合酶、3-脱氢奎尼酸合酶、3-脱氢奎尼酸脱水酶、莽草酸脱氢酶、莽草酸激酶、3-磷酸莽草酸-2-羧基乙烯基转移酶、分支酸合酶和分支酸裂合酶。

22. 权利要求 13 的非天然存在的微生物,其中所述微生物还包含(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径。

23. 权利要求 22 的非天然存在的微生物,其中所述(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径包含 2-C-甲基-D-赤藓糖醇-4-磷酸脱水酶、1-脱氧木酮糖-5-磷酸合酶或 1-脱氧-D-木酮糖-5-磷酸还原异构酶。

24. 权利要求 23 的非天然存在的微生物,其中所述(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径包含 2-C-甲基-D-赤藓糖醇-4-磷酸脱水酶、1-脱氧木酮糖-5-磷酸合酶和 1-脱氧-D-木酮糖-5-磷酸还原异构酶。

25. 权利要求 13 的非天然存在的微生物,其中至少一种外源核酸是异源核酸。

26. 权利要求 13 的非天然存在的微生物,其中所述非天然存在的微生物是在基本上厌氧的培养基中。

27. 用于生产对甲苯甲酸的方法,包括将权利要求 13 的非天然存在的微生物在足以产生对甲苯甲酸的条件培养足够的时间周期。

28. 权利要求 27 的方法,其中所述非天然存在的微生物是在基本上厌氧的培养基中。

29. 权利要求 27 的方法,其中所述微生物包含七种外源核酸,各自编码对甲苯甲酸途径酶。

30. 权利要求 29 的方法,其中所述七种外源核酸编码 2-脱氢-3-脱氧磷酸庚糖酸合酶、3-脱氢奎尼酸合酶、3-脱氢奎尼酸脱水酶、莽草酸脱氢酶、莽草酸激酶、3-磷酸莽草酸-2-羧基乙烯基转移酶、分支酸合酶和分支酸裂合酶。

31. 权利要求 27 的方法,其中所述微生物还包含(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径。

32. 权利要求 31 的方法,其中所述(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径包含 2-C-甲基-D-赤藓糖醇-4-磷酸脱水酶、1-脱氧木酮糖-5-磷酸合酶或 1-脱氧-D-木酮糖-5-磷酸还原异构酶。

33. 权利要求 32 的方法,其中所述(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径包含 2-C-甲基-D-赤藓糖醇-4-磷酸脱水酶、1-脱氧木酮糖-5-磷酸合酶和 1-脱氧-D-木酮糖-5-磷酸还原异构酶。

34. 权利要求 27 的方法,其中所述至少一种外源核酸是异源核酸。

35. 非天然存在的微生物,包括一种具有对苯二甲酸途径的微生物,该微生物包含至少一种编码以足以产生对苯二甲酸的量表达的对苯二甲酸途径酶的外源核酸,所述对苯二甲酸途径包含对甲苯甲酸甲基-单加氧酶还原酶、4-羧基苯甲醇脱氢酶或 4-羧基苯甲醛

脱氢酶;和其中所述微生物还包含对甲苯甲酸途径,其中所述对甲苯甲酸途径包含 2-脱氢-3-脱氧磷酸庚糖酸合酶、3-脱氢奎尼酸合酶、3-脱氢奎尼酸脱水酶、莽草酸脱氢酶、莽草酸激酶、3-磷酸莽草酸-2-羧基乙烯基转移酶、分支酸合酶或分支酸裂合酶。

36. 权利要求 35 的非天然存在的微生物,其中所述微生物包含两种外源核酸,各自编码对苯二甲酸途径酶。

37. 权利要求 35 的非天然存在的微生物,其中所述微生物包含三种外源核酸,各自编码对苯二甲酸途径酶。

38. 权利要求 37 的非天然存在的微生物,其中所述三种外源核酸编码对甲苯甲酸甲基-单加氧酶还原酶、4-羧基苯甲醇脱氢酶和 4-羧基苯甲醛脱氢酶。

39. 权利要求 35 的非天然存在的微生物,其中所述对甲苯甲酸途径包含 2-脱氢-3-脱氧磷酸庚糖酸合酶、3-脱氢奎尼酸合酶、3-脱氢奎尼酸脱水酶、莽草酸脱氢酶、莽草酸激酶、3-磷酸莽草酸-2-羧基乙烯基转移酶、分支酸合酶和分支酸裂合酶。

40. 权利要求 35 的非天然存在的微生物,其中所述微生物还包含(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径。

41. 权利要求 40 的非天然存在的微生物,其中所述(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径包含 2-C-甲基-D-赤藓糖醇-4-磷酸脱水酶、1-脱氧木酮糖-5-磷酸合酶或 1-脱氧-D-木酮糖-5-磷酸还原异构酶。

42. 权利要求 41 的非天然存在的微生物,其中所述(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径包含 2-C-甲基-D-赤藓糖醇-4-磷酸脱水酶、1-脱氧木酮糖-5-磷酸合酶和 1-脱氧-D-木酮糖-5-磷酸还原异构酶。

43. 权利要求 35 的非天然存在的微生物,其中所述至少一种外源核酸是异源核酸。

44. 权利要求 35 的非天然存在的微生物,其中所述非天然存在的微生物是在基本上厌氧的培养基中。

45. 用于生产对苯二甲酸的方法,包括将权利要求 35 的非天然存在的微生物在足以产生对苯二甲酸的条件培养足够的时间周期。

46. 权利要求 45 的方法,其中所述非天然存在的微生物是在基本上厌氧的培养基中。

47. 权利要求 45 的方法,其中所述微生物包含三种外源核酸,各自编码对苯二甲酸途径酶。

48. 权利要求 47 的方法,其中所述三种外源核酸编码对甲苯甲酸甲基-单加氧酶还原酶、4-羧基苯甲醇脱氢酶或 4-羧基苯甲醛脱氢酶。

49. 权利要求 45 的方法,其中所述对甲苯甲酸途径包含 2-脱氢-3-脱氧磷酸庚糖酸合酶、3-脱氢奎尼酸合酶、3-脱氢奎尼酸脱水酶、莽草酸脱氢酶、莽草酸激酶、3-磷酸莽草酸-2-羧基乙烯基转移酶、分支酸合酶和分支酸裂合酶。

50. 权利要求 45 的方法,其中所述微生物还包含(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径。

51. 权利要求 50 的方法,其中所述(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径包含 2-C-甲基-D-赤藓糖醇-4-磷酸脱水酶、1-脱氧木酮糖-5-磷酸合酶或 1-脱氧-D-木酮糖-5-磷酸还原异构酶。

52. 权利要求 51 的方法,其中所述(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径包

含 2-C- 甲基 -D- 赤藓糖醇 -4- 磷酸脱水酶、1- 脱氧木酮糖 -5- 磷酸合酶和 1- 脱氧 -D- 木酮糖 -5- 磷酸还原异构酶。

53. 权利要求 45 的方法,其中所述至少一种外源核酸是异源核酸。

## 用于生物合成对甲苯甲酸和对苯二甲酸的微生物和方法

### [0001] 发明背景

[0002] 本发明总的来讲涉及生物合成过程,更具体地讲,本发明涉及具有对甲苯甲酸(p-toluic acid)对苯二甲酸(terephthalic acid)或(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯生物合成能力的生物体。

[0003] 对苯二甲酸(也称为对苯二甲酸(terephthalic acid)和PTA)是聚对苯二甲酸乙二酯(polyethylene terephthalate, PET)的直接(immediate)前体,用于制造衣物、树脂、塑料瓶,甚至用作禽饲料添加剂。几乎所有的PTA都是在称为中世纪工艺(the Mid Century Process)的工艺中在空气中氧化而由对二甲苯生产。这个氧化作用在高温下、在乙酸溶剂中用由钴和/或锰盐组成的催化剂来进行。对二甲苯来源于石油化学来源并由石脑油的高苛刻度催化重整而形成。二甲苯也从石脑油蒸汽裂化器中的热解汽油气流和通过甲苯歧化获得。

[0004] 用于生产可再生PTA的成本-有效方法迄今尚未开发出来。PTA、甲苯和其它的芳族前体能被有些细菌自然地降解。然而,这些降解途径通常包括以可降解方向不可逆操作的单加氧酶。因此,PTA的生物合成途径受到迄今已知酶特性的严格限制。

[0005] 一种颇有前景的PTA前体是对甲苯甲酸,也称为对甲基苯甲酸酯。对甲苯甲酸是在某些工业过程中将对二甲苯氧化成PTA的中间体。它也是聚合物稳定剂、杀虫剂、光敏化合物、动物饲料补充剂和其它有机化学品的中间体。仅在水溶液中微溶,对甲苯甲酸在生理温度下为固体,其熔点为275°C。用于从糖贮存物(feedstocks)合成该化合物的微生物催化剂迄今为止还没有描述过。

[0006] 因此,存在有效生产商用量化合物例如对甲苯甲酸或对苯二甲酸的替代方法的需求。本发明满足了这种需求,也提供了有关优势。

### [0007] 发明概述

[0008] 本发明提供具有(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径、对甲苯甲酸途径和/或对苯二甲酸途径的非天然存在的微生物。本发明另外还提供使用这类生物体来产生(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径、对甲苯甲酸途径或对苯二甲酸途径的方法。

### [0009] 附图简述

[0010] 图1显示由3-磷酸甘油醛和丙酮酸生成(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯(2H3M4OP)的示例性途径的示意图。G3P为3-磷酸甘油醛,DXP为1-脱氧-D-木酮糖-5-磷酸,和2ME4P为C-甲基-D-赤藓糖醇-4-磷酸。酶是:(A)DXP合酶;(B)DXP还原异构酶(reductoisomerase);和(C)2ME4P脱水酶。

[0011] 图2显示对甲苯甲酸的示例性替代莽草酸途径的示意图。酶是:(A)2-脱氢-3-脱氧磷酸庚糖酸合酶;(B)3-脱氢奎尼酸合酶;(C)3-脱氢奎尼酸脱水酶;(D)莽草酸脱氢酶;(E)莽草酸激酶;(F)3-磷酸莽草酸-2-羧基乙基转移酶;(G)分支酸合酶;和(H)分支酸裂合酶。化合物是:(1)(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯;(2)2,4-二羟基-5-甲基-6-[(膦酰氧基)甲基]噁烷-2-羧酸酯;(3)1,3-二羟基-4-甲基-5-氧代

环己烷-1-羧酸酯；(4)5-羟基-4-甲基-3-氧代环己-1-烯-1-羧酸酯；(5)3,5-二羟基-4-甲基环己-1-烯-1-羧酸酯；(6)5-羟基-4-甲基-3-(磷酸氧基)环己-1-烯-1-羧酸酯；(7)5-[(1-羧基乙-1-烯-1-基)氧基]-4-甲基-3-(磷酸氧基)环己-1-烯-1-羧酸酯；(8)3-[(1-羧基乙-1-烯-1-基)氧基]-4-甲基环己-1,5-二烯-1-羧酸酯；和(9)对甲苯甲酸。

[0012] 图3显示将对甲苯甲酸(p-toluic acid)转变成对苯二甲酸(PTA)的示例性途径。反应A、B和C分别由对甲苯甲酸甲基-单加氧酶还原酶、4-羧基苯甲醇脱氢酶和4-羧基苯甲醛脱氢酶催化。所示化合物是：(1)对甲苯甲酸(p-toluic acid)；(2)4-羧基苯甲醇；(3)4-羧基苯甲醛；和(4)对苯二甲酸。

[0013] 发明详述

[0014] 本发明涉及具有对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯的生物合成生产能力的细胞和生物体的设计和生产。本文所述结果表明代谢途径可以进行设计和重组工程以实现对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯在大肠杆菌(*Escherichia coli*)及其它细胞或生物体中的生物合成。对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯的生物合成性生产可以通过构建具有设计的代谢基因型的菌株来证实。这些代谢工程改造的细胞或生物体也可以经受适应性进化以进一步增强对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯的生物合成，包括在接近理论最大生长的条件下。

[0015] 大肠杆菌中的莽草酸生物合成途径将赤藓糖-4-磷酸转变成分支酸，一种重要中间体，导致生物合成许多必需代谢物，包括4-羟基苯甲酸酯。4-羟基苯甲酸酯在结构上类似于对甲苯甲酸，即对苯二甲酸的一种工业前体。正如本文所公开的，莽草酸途径酶可用于接受替代底物(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯(2H3M4OP)并将其转化成对甲苯甲酸。另外，利用来自类异戊二烯(isoprenoid)生物合成的非甲羟戊酸途径的酶，该途径用于合成2H3M4OP前体。

[0016] 本文公开的是对微生物进行工程改造以从碳水化合物贮存物生产可再生对甲苯甲酸或对苯二甲酸(PTA)的策略。首先，在三个酶促步骤中，将3-磷酸甘油醛(G3P)和丙酮酸转变成2-羟基-3-甲基-4-氧代丁氧基)磷酸酯(2H3M4OP)(参见实施例I和图1)。随后，2H3M4OP中间体通过莽草酸途径中的酶转化成对甲苯甲酸(参见实施例II和图2)。对甲苯甲酸可以通过微生物进一步转变成PTA(参见实施例III和图3)。

[0017] G3P转变成对甲苯甲酸需要一个ATP、两个还原当量的(NAD(P)H)和两分子磷酸烯醇丙酮酸，按照以下的净反应式。

[0018]  $G3P+2PEP+ATP+2NAD(P)H+2H^+ \rightarrow \text{对甲苯甲酸} +4Pi+ADP+2NAD(P)^++CO_2+H_2O$

[0019] 从葡萄糖合成G3P需要额外的ATP。最大理论对甲苯甲酸产量为0.67mol/mol(0.51g/g)(从葡萄糖中减去碳需要能量)。在假定每合成1分子对甲苯甲酸要消耗掉2个ATP的情况下，从葡萄糖到对甲苯甲酸的产量预计为0.62mol/mol(0.46g/g)对甲苯甲酸。

[0020] 如果对甲苯甲酸被实施例III中描述的酶进一步转变成PTA，那么从葡萄糖到PTA的产量预计为0.64mol/mol(0.58g/g)。在这种情形下，对甲苯甲酸氧化成PTA产生额外的净还原当量，按照以下的净反应式：

[0021] 对甲苯甲酸 + O<sub>2</sub> + NAD<sup>+</sup> → PTA + NADH + 2H<sup>+</sup>

[0022] 以下各小节描述了用于催化建议途径各步骤的候选酶。

[0023] 本文所用的术语“非天然存在的”当用于涉及本发明的微生物时，意指该微生物具有至少一个遗传改变，该遗传改变在正常情况下在参考菌种的天然存在的菌株（包括参考菌种的野生型菌株）中不存在。遗传改变包括例如微生物遗传材料的修饰，该修饰引入编码代谢多肽的可表达核酸、其它核酸添加、核酸缺失和 / 或其它功能破坏。这样的修饰包括例如参考菌种的异源多肽、同源多肽或者异源多肽和同源多肽二者的编码区及其功能片段。另外的修饰包括例如非编码调节区，其中该修饰改变了基因或操纵子的表达。示例性的代谢多肽包括对甲苯甲酸、对苯二甲酸或（2-羟基-3-甲基-4-氧代丁氧基）磷酸酯生物合成途径内的酶或蛋白质。

[0024] 代谢修饰是指自其天然存在的状态被改变的生物化学反应。因此，非天然存在的微生物可具有代谢多肽编码核酸或其功能片段的遗传修饰。在本文中公开了示例性的代谢修饰。

[0025] 本文所用的术语“分离的”当用于涉及微生物时，意指基本上不含随该参考微生物在自然界中发现的至少一种组分的生物体。该术语包括去除了随该微生物在天然环境中被发现的某些或所有组分的微生物。该术语也包括去除了随该微生物在非天然存在的环境中被发现的某些或所有组分的微生物。因此，分离的微生物与随其在自然界中被发现或者随其在非天然存在的环境中生长、贮存或生存时被发现的其他物质部分或完全分离开。分离的微生物的具体实例包括部分纯的微生物、相当纯的微生物和在培养基即非天然存在的环境中培养的微生物。

[0026] 本文所用的术语“微生物”意指作为用显微镜可看见的细胞存在的任何生物体，它被包括在古细菌 (archaea)、细菌 (bacteria) 或真核微生物 (eukarya) 的范围内。因此，该术语预计包括具有用显微镜可看见的大小的原核或真核细胞或者原核或真核生物体，并且包括所有菌种的细菌、古细菌和真细菌以及真核微生物（例如酵母和真菌）。该术语也包括可以进行培养以用于生物化学品生产的任何菌种的细胞培养物。

[0027] 本文所用的术语“CoA”或“辅酶 A”意指形成有活性的酶系统的有机辅因子或辅基 (prosthetic group) (酶的非蛋白质部分)，其存在对于许多酶（脱辅基酶）的活性是必需的。辅酶 A 在某些缩合酶中发挥作用，在乙酰基或其它酰基转移中和在脂肪酸合成和氧化、丙酮酸氧化中和在其它乙酰化中起作用。

[0028] 本文所用的术语“（2-羟基-3-甲基-4-氧代丁氧基）磷酸酯”，在本文中缩写为 2H3M40P，具有图 1 所示的化学结构式。这样一种化合物也可描述为 3-羟基-2-甲基丁醛-4-磷酸。

[0029] 本文所用的术语“对甲苯甲酸 (p-toluate)”，分子式为 C<sub>8</sub>H<sub>7</sub>O<sub>2</sub><sup>-</sup> (参见图 2，化合物 9) (IUPAC 命名为 4-甲基苯甲酸酯)，是对甲苯甲酸 (p-toluic acid) 的离子化形式。应当理解，对甲苯甲酸 (p-toluate) 和对甲苯甲酸 (p-toluic acid) 在涉及化合物呈其中性或离子化形式的任何一种（包括其任何盐形式）时可互换使用。技术人员全都知道，具体的形式将取决于 pH。

[0030] 本文所用的术语“对苯二甲酸 (terephthalate)”，分子式为 C<sub>8</sub>H<sub>4</sub>O<sub>4</sub><sup>-2</sup> (参见图 3，化合物 4) (IUPAC 命名为对苯二甲酸)，是对苯二甲酸 (terephthalic acid)，也称为对-苯二

甲酸 (p-phthalic acid) 或 PTA 的离子化形式。应当理解,对苯二甲酸 (terephthalate) 和对苯二甲酸 (p-phthalic acid) 在涉及化合物呈其中性或离子化形式的任何一种 (包括其任何盐形式) 时可互换使用。技术人员全都知道,具体的形式将取决于 pH。

[0031] 本文所用的术语“基本上厌氧的”当用于涉及培养或生长条件时,意指氧的量低于液体培养基中溶解氧饱和度的约 10%。该术语也预计包括用小于约 1% 氧的气氛维持的液体或固体培养基的密闭室。

[0032] 本文所用的“外源的”意指参考分子或参考活性被引入到宿主微生物中。该分子可以例如通过将编码核酸引入到宿主遗传材料中,例如通过整合到宿主染色体或作为非染色体遗传材料例如质粒而引入。因此,该术语当用于涉及编码核酸的表达时,是指将编码核酸以可表达形式引入到微生物中。当用于涉及生物合成活性时,该术语是指被引入到宿主参比生物体中的活性。活性来源可以是例如在引入到宿主微生物之后表达参考活性的同源或异源编码核酸。因此,术语“内源的”是指宿主体内存在的参考分子或活性。同样地,该术语当用于涉及编码核酸的表达时,是指微生物内含有的编码核酸的表达。术语“异源的”是指从不是参考物种的来源获得的分子或活性,而“同源的”是指来源于宿主微生物的分子或活性。因此,本发明的编码核酸的外源表达可以利用异源或同源编码核酸中的任何一个或二者。

[0033] 应当理解,当在微生物中包括超过一种外源核酸时,超过一种外源核酸是指如上所论述的参考编码核酸或生物合成活性。还应当理解,正如本文所公开的,这样的超过一种外源核酸可以在独立的核酸分子上、在多顺反子核酸分子或其组合上引入到宿主微生物中,并且仍然被视为超过一种外源核酸。例如,本文所公开的微生物可以经过工程改造以表达两种或更多种编码所需途径酶或蛋白质的外源核酸。在将编码所需活性的两种外源核酸引入到宿主微生物的情况下,应当理解,这两种外源核酸可以作为一种核酸 (例如在单个质粒上、在独立的质粒上) 被引入,可以整合到宿主染色体的单个位点或多个位点上,并且仍然被视为两种外源核酸。同样地,应当理解,超过两种外源核酸可以以任何所需组合 (例如在单个质粒上、在独立的质粒上) 引入到宿主生物中,可以整合到宿主染色体的单个位点或多个位点上,并且仍然被视为两种或更多种外源核酸,例如三种外源核酸。因此,参考外源核酸或生物合成活性的数目是指编码核酸的数目或生物合成活性的数目,而不是被引入到宿主生物中的独立核酸的数目。

[0034] 本发明的非天然存在的微生物可含有稳定的遗传改变,它是指可培养 5 代以上而没有遗传改变损失的微生物。一般而言,稳定的遗传改变包括持续 10 代以上的修饰,具体地说,稳定的修饰将持续约 25 代以上,和更具体地说,稳定的遗传修饰将是多于 50 代,包括无限期地。

[0035] 本领域技术人员应当理解,遗传改变,包括本文例举的代谢修饰,可参照合适的宿主生物 (例如大肠杆菌) 和它们的相应代谢反应或者所需遗传材料的合适来源生物体 (例如所需代谢途径的基因) 进行描述。然而,考虑到各种生物体的完全基因组测序和基因组学领域中技巧的高水平,本领域技术人员将能够容易地将本文提供的教导和指导应用到基本所有的其它生物体。例如,本文例举的大肠杆菌代谢改变可以容易地应用到其它物种,即通过掺入来自参考物种以外的其它物种的相同或类似编码核酸。这样的遗传改变一般包括例如物种同系物 (species homologs) 的遗传改变,和具体包括例如直系同源

物 (orthologs)、旁系同源物 (paralogs) 或非直系同源基因置换 (nonorthologous gene displacements) 的遗传改变。

[0036] 直系同源物是有垂直亲缘 (vertical descent) 关系的并且在不同生物体中负责基本相同功能或相同功能的一个或多个基因。例如,小鼠环氧化物水解酶和人环氧化物水解酶可以因为环氧化物水解的生物功能而被视为直系同源物。当某些基因共享足够数量的序列相似性以表明它们是同源的,或者是从共同的祖先进化而相关时,它们有垂直亲缘关系。如果某些基因共享三维结构但不一定共享足够数量的序列相似性以表明它们是从共同的祖先进化到一级序列相似性是不可辨识的程度,那么它们也可被视为直系同源物。作为直系同源的基因可以编码具有约 25% 至 100% 氨基酸序列同一性的序列相似性的蛋白质。编码共享氨基酸相似性小于 (less than) 25% 的蛋白质的基因也可被视为因有垂直亲缘关系而提升,如果它们的三维结构也显示很高相似性的话。酶的丝氨酸蛋白酶家族成员,包括组织纤溶酶原激活物和弹性蛋白酶,都被视为因来自共同的祖先而有垂直亲缘关系而提升。

[0037] 直系同源物包括通过例如进化而在结构或总体活性上趋异的基因或它们的编码基因产物。例如,当一个物种编码表现出两种功能的基因产物并且当这样的功能已经在第二个物种中分离到截然不同的基因上时,这三个基因及其相应的产物被视为直系同源物。为了生物化学产品的生产,本领域技术人员应当理解,为了构建非天然存在的微生物,必须对带有代谢活性的要被引入或破坏的直系同源基因进行选择。表现出可分离活性的直系同源物的一个实例是当截然不同的活性已经在两个或更多个物种之间或者在单一物种内分离到截然不同的基因产物上。一个具体的例子是弹性蛋白酶蛋白水解和纤溶酶原蛋白水解 (丝氨酸蛋白酶活性的两种类型) 已经作为纤溶酶原激活物和弹性蛋白酶分离成为两个截然不同的分子。第二个例子是支原体 5' -3' 外切核酸酶和果蝇 (*Drosophila*) DNA 聚合酶 III 活性的分离。来自第一个物种的 DNA 聚合酶可被视为来自第二个物种的外切核酸酶或聚合酶中任一个或二者的直系同源物,反之亦然。

[0038] 相比之下,旁系同源物是因为例如进化趋异后再复制而有亲缘关系的同系物并且具有相似或共同的但不相同的功能。旁系同源物可起源于或来源于例如相同的物种或来自不同的物种。例如,微粒体环氧化物水解酶 (环氧化物水解酶 I) 和可溶性环氧化物水解酶 (环氧化物水解酶 II) 可被视为旁系同源物,因为它们代表了两种独特的酶,从共同的祖先共进化而来,它们催化不同的反应且在相同的物种中具有不同的功能。旁系同源物是来自相同的物种并且彼此具有显著序列相似性的蛋白质,提示它们是同源的,或者通过从共同的祖先共进化而有亲缘关系。旁系同源蛋白质家族的组包括 HipA 同系物、萤光素酶基因、肽酶等。

[0039] 非直系同源基因置换是来自一个物种的非直系同源基因,它可以在不同的物种中取代参考基因功能。取代包括例如,与在不同的物种中的参考功能相比,能够在原始的物种中执行基本上相同或相似的功能。虽然一般而言,非直系同源基因置换将根据与编码参考功能的已知基因在结构上相关性进行鉴别,但是尽管如此,结构上相关性较少但功能上相似的基因及其相应基因产物将仍然落入该术语的含义之内,正如本文中使用的。与编码待取代的功能的基因相比,在非直系同源基因产物的活性位点或结合区中,功能相似性需要例如至少某些结构相似性。因此,非直系同源基因包括例如旁系同源物 (paralog) 或无亲缘关系的基因 (unrelated gene)。

[0040] 因此,在鉴定和构建本发明的具有对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯生物合成能力的非天然存在的微生物中,本领域技术人员应当知道将本文提供的教导和指导应用到某个具体的物种时,代谢修饰的鉴定可包括直系同源物的鉴定和纳入(inclusion)或失活。达到这种程度以致在编码催化相似或基本上相似代谢反应的酶的参考微生物中存在旁系同源物和/或非直系同源基因置换,本领域技术人员也可利用这些进化上相关的基因。

[0041] 直系同源物、旁系同源物和非直系同源基因置换可以通过本领域技术人员熟知的方法进行测定。例如,对于两种多肽,核酸或氨基酸序列的检查将揭示出比较序列之间的序列同一性和相似性。基于这样的相似性,本领域技术人员可以确定相似性是否足够高以表明蛋白质是来自共同的祖先通过进化相关的。本领域技术人员熟知的算法,例如 Align、BLAST、Clustal W 等算法比较并测定原序列的相似性或同一性,并且也测定可比对权重或评分的序列中空位(gaps)的存在或显著性。这样的算法也是本领域已知的并且可同样用于测定核苷酸序列相似性或同一性。具有足够的相似性以确定亲缘关系(relatedness)的参数可基于熟知的方法输入计算机以计算统计相似性,或者在随机多肽中找到相似匹配的机会,并且测定匹配的显著性。两个或更多个序列的计算机比较也可以(如果需要)由本领域技术人员用肉眼最优化。预期有亲缘关系的基因产物或蛋白质可具有高度相似性,例如 25% 至 100% 序列同一性。无亲缘关系的蛋白质可具有与预期偶然发生基本上相同的同一性,如果对足够大小的数据库扫描(约 5%)的话。占 5% 和 24% 之间的序列可以代表或不代表足够的同源性以得出结论是被比较序列是相关的。可以进行另外的统计分析以确定给出数据集大小的这类匹配的显著性,从而确定这些序列的关系。

[0042] 运用 BLAST 算法来测定两个或更多个序列的亲缘关系(relatedness)的示例性参数可以是如下所示的。简而言之,氨基酸序列比对可以使用 BLASTP 2.0.8 版本(1999 年 1 月 5 日)和以下参数进行:矩阵(matrix):0BLOSUM62;空位开放(gap open):11;空位延伸(gap extension):1;x\_dropoff:50;期望值(expect):10.0;字大小(wordsize):3;过滤器(filter):开(on)。核酸序列比对可以使用 BLASTN 2.0.6 版本(1998 年 9 月 16 日)和以下参数进行:匹配(match):1;错配(mismatch):-2;空位开放(gap open):5;空位延伸(gap extension):2;x\_dropoff:50;期望值(expect):10.0;字大小(wordsize):11;过滤器(filter):关(off)。本领域技术人员将会知道对以上参数可以做出什么修改以或者增加或者减少例如比较的严格性和测定两个或更多个序列的亲缘关系。

[0043] 本发明提供一种非天然存在的微生物,包括一种具有(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯途径的微生物,该微生物包含至少一种编码以足以产生(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯的量表达的(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯途径酶的外源核酸,所述(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯途径包含 2-C-甲基-D-赤藓糖醇-4-磷酸脱水酶(参见实施例 I 和图 1,步骤 C)。一种包含(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯途径的非天然存在的微生物还可包含 1-脱氧木酮糖-5-磷酸合酶或 1-脱氧-D-木酮糖-5-磷酸还原异构酶(参见实施例 I 和图 1,步骤 A 和 B)。因此,(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯可包含 52-C-甲基-D-赤藓糖醇-4-磷酸脱水酶、1-脱氧木酮糖-5-磷酸合酶和 1-脱氧-D-木酮糖-5-磷酸还原异构酶。

[0044] 本发明也提供一种非天然存在的微生物,包括一种具有对甲苯甲酸途径的微生物

物,该微生物包含至少一种编码以足够产生对甲苯甲酸的量表达的对甲苯甲酸途径酶的外源核酸,所述对甲苯甲酸途径包含 2-脱氢-3-脱氧磷酸庚糖酸合酶、3-脱氢奎尼酸合酶、3-脱氢奎尼酸脱水酶、莽草酸脱氢酶、莽草酸激酶、3-磷酸莽草酸-2-羧基乙烯基转移酶、分支酸合酶或分支酸裂合酶(参见实施例 II 和图 2,步骤 A-H)。一种具有对甲苯甲酸途径的非天然存在的微生物还可包含(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径(图 1)。(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径可包含例如 2-C-甲基-D-赤藓糖醇-4-磷酸脱水酶、1-脱氧木酮糖-5-磷酸合酶或 1-脱氧-D-木酮糖-5-磷酸还原异构酶(图 1)。

[0045] 本发明另外还提供一种非天然存在的微生物,包括一种具有对苯二甲酸途径的微生物,该微生物包含至少一种编码以足够产生对苯二甲酸的量表达的对苯二甲酸途径酶的外源核酸,所述对苯二甲酸途径包含对甲苯甲酸甲基-单加氧酶还原酶、4-羧基苯甲醇脱氢酶或 4-羧基苯甲醛脱氢酶(参见实施例 III 和图 3)。这样一种含有对苯二甲酸途径的生物体可另外包含对甲苯甲酸途径,其中所述对甲苯甲酸途径包含 2-脱氢-3-脱氧磷酸庚糖酸合酶、3-脱氢奎尼酸合酶、3-脱氢奎尼酸脱水酶、莽草酸脱氢酶、莽草酸激酶、3-磷酸莽草酸-2-羧基乙烯基转移酶、分支酸合酶或分支酸裂合酶(参见实施例 II 和 III 及图 2 和图 3)。这样一种具有对苯二甲酸途径和对甲苯甲酸途径的非天然存在的微生物还可包含(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径(参见实施例 I 和图 1)。(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径可包含例如 2-C-甲基-D-赤藓糖醇-4-磷酸脱水酶、1-脱氧木酮糖-5-磷酸合酶或 1-脱氧-D-木酮糖-5-磷酸还原异构酶(参见实施例 I 和图 1)。

[0046] 在一个另外的实施方案中,本发明提供一种具有对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径的非天然存在的微生物,其中所述非天然存在的微生物包含至少一种编码将底物转变成产物的酶或蛋白质的外源核酸。例如,在(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径中,底物和产物可以选自 3-磷酸甘油醛和丙酮酸与 1-脱氧-D-木酮糖-5-磷酸;1-脱氧-D-木酮糖-5-磷酸与 C-甲基-D-赤藓糖醇-4-磷酸;和 C-甲基-D-赤藓糖醇-4-磷酸与(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯(参见实施例 I 和图 1)。在另一个实施方案中,对甲苯甲酸途径可包含选自以下的底物和产物:(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯与 2,4-二羟基-5-甲基-6-[(磷酰氧基)甲基]噁烷-2-羧酸酯;2,4-二羟基-5-甲基-6-[(磷酰氧基)甲基]噁烷-2-羧酸酯与 1,3-二羟基-4-甲基-5-氧代环己烷-1-羧酸酯;1,3-二羟基-4-甲基-5-氧代环己烷-1-羧酸酯与 5-羟基-4-甲基-3-氧代环己-1-烯-1-羧酸;5-羟基-4-甲基-3-氧代环己-1-烯-1-羧酸与 3,5-二羟基-4-甲基环己-1-烯-1-羧酸酯;3,5-二羟基-4-甲基环己-1-烯-1-羧酸酯与 5-羟基-4-甲基-3-(磷酰氧基)环己-1-烯-1-羧酸酯;5-羟基-4-甲基-3-(磷酰氧基)环己-1-烯-1-羧酸酯与 5-[(1-羧基乙-1-烯-1-基)氧基]-4-甲基-3-(磷酰氧基)环己-1-烯-1-羧酸酯;5-[(1-羧基乙-1-烯-1-基)氧基]-4-甲基-3-(磷酰氧基)环己-1-烯-1-羧酸酯与 3-[(1-羧基乙-1-烯-1-基)氧基]-4-甲基环己-1,5-二烯-1-羧酸酯;和 3-[(1-羧基乙-1-烯-1-基)氧基]-4-甲基环己-1,5-二烯-1-羧酸酯与对甲苯甲酸(参见实施例 II 和图 2)。在又一个实施方案中,对苯二甲酸途径可包含选自以下的底物和产物:对甲苯甲酸与 4-羧基苯甲醇;4-羧基苯甲醇与 4-羧基苯甲醛;和 4-羧基苯甲醛与对苯二甲酸(参见实施例 III 和图 3)。本领域技术人员应当理解,这些都只是示例性的并且本文公开的适合于产生所需产物的和为此

适宜活性可用于将底物转变成产物的任何底物-产物对都可以根据本文公开的内容由本领域技术人员容易地确定。因此,本发明提供一种含有至少一种编码酶或蛋白质的外源核酸的非天然存在的微生物,其中所述酶或蛋白质能转变对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径的底物和产物,如图1-3所示的。

[0047] 尽管在本文中一般描述为含有对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径的微生物,但是应当理解,本发明另外还提供一种非天然存在的微生物,该微生物包含至少一种编码对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径酶的外源核酸,其表达量足以产生对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径的中间体。例如,正如本文所公开的,(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径见图1(参见实施例I)。因此,除了含有产生(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯的(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径的微生物以外,本发明另外还提供一种非天然存在的微生物,该微生物包含至少一种编码(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径酶的外源核酸,其中所述微生物产生(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径中间体,例如1-脱氧-D-木酮糖-5-磷酸或C-甲基-D-赤藓糖醇-4-磷酸。同样地,本发明也提供一种非天然存在的微生物,该微生物含有产生对甲苯甲酸的对甲苯甲酸途径,其中所述非天然存在的微生物包含至少一种编码对甲苯甲酸途径酶的外源核酸,其中所述微生物产生对甲苯甲酸途径中间体,例如2,4-二羟基-5-甲基-6-[(膦酰氧基)甲基]噁烷-2-羧酸酯、1,3-二羟基-4-甲基-5-氧代环己烷-1-羧酸酯、5-羟基-4-甲基-3-氧代环己-1-烯-1-羧酸酯、3,5-二羟基-4-甲基环己-1-烯-1-羧酸酯、5-羟基-4-甲基-3-(膦酰氧基)环己-1-烯-1-羧酸酯或3-[(1-羧基乙-1-烯-1-基)氧基]-4-甲基-3-(膦酰氧基)环己-1-烯-1-羧酸酯或3-[(1-羧基乙-1-烯-1-基)氧基]-4-甲基环己-1,5-二烯-1-羧酸酯。此外,本发明另外还提供一种非天然存在的微生物,该微生物含有对苯二甲酸途径酶,其中所述微生物产生对苯二甲酸途径中间体,例如4-羧基苯甲醇或4-羧基苯甲醛。

[0048] 应当理解,本文公开的任何途径,如在实施例中描述的和在附图中例举的,包括图1-3的途径,都可以用于产生非天然存在的微生物,该微生物按需要产生任何途径中间体或产物。正如本文所公开的,这样一种产中间体微生物可以与表达下游途径酶以产生所需产物的其它微生物联合使用。然而,应当理解,产对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径中间体的非天然存在的微生物可以用于生产作为所需产物的中间体。

[0049] 本文中一般参照代谢反应、其反应物或产物,或具体参照编码与参考代谢反应、反应物或产物相关的酶、或催化参考代谢反应、反应物或产物的酶、或与参考代谢反应、反应物或产物相关的蛋白质的一种或多种核酸或基因,对本发明进行了描述。除非本文另有明确的表述,否则本领域技术人员应当理解,参照反应也构成参照反应的反应物和产物。类似地,除非本文另有明确的表述,否则本领域技术人员应当理解,参照反应物或产物也参照反应,和参照这些代谢中的任何一个也参照编码参与参考反应、反应物或产物的酶或蛋白质的一个或多个基因。同样地,考虑到代谢生物化学、酶学和基因组学的公知领域,本文参照基因或编码核酸也构成参照相应的编码酶和它所催化的反应或与反应相关的蛋白质以及反应的反应物和产物。

[0050] 本发明的非天然存在的微生物可以通过导入编码参与一种或多种对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯生物合成途径的酶或蛋白质中的一种或多种的可表达核酸来产生。根据生物合成所选出的宿主微生物,可以表达具体的对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯生物合成途径中的某些或全部的核酸。例如,如果所选宿主在所需生物合成途径的一种或多种酶或蛋白质中是有缺陷的,那么可以将这种有缺陷的酶或蛋白质的可表达核酸导入宿主中用于后续的外源表达。或者,如果所选宿主表现出某些途径基因的内源表达,但是其它的是有缺陷的,那么对于这种有缺陷的酶或蛋白质来说,需要编码核酸才能实现对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯生物合成。因此,本发明的非天然存在的微生物可以通过引入外源酶或蛋白质活性以获得所需生物合成途径来产生,或者所需生物合成途径可以通过引入一种或多种外源酶或蛋白质活性来获得,连同一种或多种内源的酶或蛋白质,产生所需产物,例如对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯。

[0051] 可以从例如细菌、酵母、真菌或可应用于发酵过程中的各种各样的其它微生物中的任何一种中选出宿主微生物,和在它们当中产生非天然存在的微生物。示例性的细菌包括选自以下的菌种:大肠杆菌(*Escherichia coli*)、产酸克雷伯氏菌(*Klebsiella oxytoca*)、产琥珀酸厌氧螺菌(*Anaerobiospirillum succiniciproducens*)、琥珀酸放线杆菌(*Actinobacillus succinogenes*)、产琥珀酸曼氏杆菌(*Mannheimia succiniciproducens*)、豆根瘤菌(*Rhizobium etli*)、枯草芽孢杆菌(*Bacillus subtilis*)、谷氨酸棒杆菌(*Corynebacterium glutamicum*)、氧化葡萄糖杆菌(*Gluconobacter oxydans*)、运动发酵单胞菌(*Zymomonas mobilis*)、乳酸乳球菌(*Lactococcus lactis*)、植物乳杆菌(*Lactobacillus plantarum*)、天蓝色链霉菌(*Streptomyces coelicolor*)、丙酮丁醇梭菌(*Clostridium acetobutylicum*)、荧光假单胞菌(*Pseudomonas fluorescens*)和恶臭假单胞菌(*Pseudomonas putida*)。示例性的酵母或真菌包括选自以下的菌种:酿酒酵母(*Saccharomyces cerevisiae*)、粟酒裂殖酵母(*Schizosaccharomyces pombe*)、乳酸克鲁维酵母(*Kluyveromyces lactis*)、马克斯克鲁维酵母(*Kluyveromyces marxianus*)、土曲霉(*Aspergillus terreus*)、黑曲霉(*Aspergillus niger*)、巴斯德毕赤酵母(*Pichia pastoris*)、少根根霉(*Rhizopus arrhizus*)、米根霉(*Rhizopus oryzae*)等。大肠杆菌是一种特别有用的宿主生物,因为它是一种经过充分表征的微生物,适用于遗传工程。其它特别有用的宿主生物包括酵母,例如酿酒酵母。应当理解,任何合适的微生物宿主生物体都可以用于引入代谢和/或遗传修饰以产生所需产物。

[0052] 根据所选宿主微生物的对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯生物合成途径组分,本发明的非天然存在的微生物将包括一种或多种对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯生物合成途径的至少一种外源表达的对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯途径编码核酸和多达所有的编码核酸。例如,可以在途径酶或蛋白质中有缺陷的宿主中通过相应编码核酸的外源表达来建立对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯生物合成。在对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯途径的所有酶或蛋白质中有缺陷的宿主中,可以包括所有酶或蛋白质在该途径中的外源

表达,虽然应当理解,某一途径的所有酶或蛋白质可以表达,即使该宿主含有途径酶或蛋白质中的至少一个。例如,可以包括所有酶或蛋白质在产生对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯的途径中的外源表达。例如,可以包括对甲苯甲酸途径中的所有酶,例如2-脱氢-3-脱氧磷酸庚糖酸合酶、3-脱氢奎尼酸合酶、3-脱氢奎尼酸脱水酶、莽草酸脱氢酶、莽草酸激酶、3-磷酸莽草酸-2-羧基乙烯基转移酶、分支酸合酶和分支酸裂合酶。另外,可以包括对苯二甲酸途径中的所有酶,例如对甲苯甲酸甲基-单加氧酶还原酶、4-羧基苯甲醇脱氢酶和4-羧基苯甲醛脱氢酶。而且,可以包括(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径中的所有酶,例如2-C-甲基-D-赤藓糖醇-4-磷酸脱水酶、1-脱氧木酮糖-5-磷酸合酶和1-脱氧-D-木酮糖-5-磷酸还原异构酶。

[0053] 考虑到本文提供的教导和指导,本领域技术人员应当理解,以可表达形式引入的编码核酸的数目将至少平行于所选宿主微生物的对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径缺陷。因此,根据所述的具体途径,本发明的非天然存在的微生物可具有一、二、三、四、五、六、七或八种,乃至多达所有编码本文公开的构成对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯生物合成途径的酶或蛋白质的核酸。在有些实施方案中,非天然存在的微生物也可包括其它遗传修饰,该遗传修饰促进或优化对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯生物合成或者赋予宿主微生物其它有用的功能。一种这样的其它功能性可包括例如加强对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径前体中的一种或多种例如3-磷酸甘油醛、丙酮酸、(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯或对甲苯甲酸的合成。此外,正如本文所公开的,在单个生物体中可以包括多个途径,例如产生对甲苯甲酸(图2)、对苯二甲酸(图3)和(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯(图1)的途径(按需要)。

[0054] 一般而言,对宿主微生物进行选择使得它产生对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径的前体,该前体或者作为天然产生的分子或者作为工程产物,以或者提供所需前体的从头(de novo)产生或使由宿主微生物天然产生的前体的产生增加。例如,3-磷酸甘油醛和磷酸烯醇丙酮酸是在宿主生物(例如大肠杆菌)中天然产生的。宿主生物可以经工程改造以增加前体的产生,正如本文所公开的。另外,经过了工程改造以产生所需前体的微生物可以用作宿主生物并且经过进一步工程改造以表达对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径的酶或蛋白质。

[0055] 在有些实施方案中,本发明的非天然存在的微生物从含有合成对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯的酶促能力的宿主中产生。在这个具体的实施方案中,它可用于增加对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径产物的合成或累积以例如驱动对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径反应朝向对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯产生。可以通过例如编码上述对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径酶或蛋白质中的一种或多种的核酸的过量表达来完成增加的合成或累积。例如,通过一个或多个内源基因的外源表达,或者通过一个或多个异源基因的外源表达,可以发生对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯途径的一种或多种酶和/或一种或多种蛋白质的过量表达。因此,天然存在的

生物体可以容易地产生出本发明的非天然存在的微生物,例如产生对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯,即通过一、二、三、四、五种……,乃至多达所有编码对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯生物合成途径酶或蛋白质的核酸的过表达。另外,非天然存在的生物体可以通过内源基因的诱变导致对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯生物合成途径中酶的活性增加来产生。

[0056] 在特别有用的实施方案中,采用了编码核酸的外源表达。外源表达赋予宿主定制表达和/或调节元件的能力并且用于达到由使用者控制的所需表达水平。然而,内源表达也可以用于其它的实施方案中,例如通过去除负调节效应子或者当与诱导型启动子或其它调节元件连接时诱导基因的启动子。因此,具有天然存在的诱导型启动子的内源基因可以通过提供适宜的诱导剂而上调,或者内源基因的调节区可以经工程改造以掺入诱导型调节元件,从而允许内源基因在所需时间的表达增加予以调节。类似地,可以包括诱导型启动子作为调节元件以将外源基因引入到非天然存在的微生物中。

[0057] 应当理解,在本发明的方法中,可以将任何一种或多种外源核酸引入到某种微生物中以产生本发明的非天然存在的微生物。所述核酸可以被引入以例如赋予所述微生物对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯生物合成途径。或者,为了赋予对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯生物合成能力,编码核酸可以被引入以产生具有生物合成能力以催化某些所需反应的中间微生物。例如,具有对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯生物合成途径的非天然存在的微生物可包含至少两种编码所需酶或蛋白质的外源核酸。例如,在(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯途径中,已表达酶的组合可以是2-C-甲基-D-赤藓糖醇-4-磷酸脱水酶和1-脱氧木酮糖-5-磷酸合酶的组合,或者2-C-甲基-D-赤藓糖醇-4-磷酸脱水酶和1-脱氧-D-木酮糖-5-磷酸还原异构酶的组合。在对甲苯甲酸途径中,已表达酶的组合可以是2-脱氢-3-脱氧磷酸庚糖酸合酶和3-脱氢奎尼酸脱水酶的组合;莽草酸激酶和3-磷酸莽草酸-2-羧基乙烯基转移酶的组合;莽草酸激酶和莽草酸脱氢酶的组合;等等。同样地,在对苯二甲酸途径中,已表达酶的组合可以是对甲苯甲酸甲基-单加氧酶还原酶和4-羧基苯甲醇脱氢酶的组合;或4-羧基苯甲醇脱氢酶和4-羧基苯甲醛脱氢酶的组合;等等。因此,应当理解,本发明的非天然存在的微生物中可包括生物合成途径的两种或更多种酶或蛋白质的任何组合。同样地,应当理解,本发明的非天然存在的微生物中根据需要可包括生物合成途径的三种或更多种酶或蛋白质的任何组合,例如,3-脱氢奎尼酸合酶、莽草酸脱氢酶和莽草酸激酶的组合;莽草酸激酶、分支酸合酶和分支酸裂合酶的组合;3-脱氢奎尼酸脱水酶、分支酸合酶和分支酸裂合酶的组合等等,只要所需生物合成途径的酶和/或蛋白质的组合可以导致产生相应的所需产物即可。同样地,本发明的非天然存在的微生物中根据需要可包括生物合成途径的四、五、六、七种或更多种酶或蛋白质的任何组合(取决于本文所公开的途径),只要所需生物合成途径的酶和/或蛋白质的组合可以导致产生相应的所需产物即可。

[0058] 除了本文所述的对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯的生物合成以外,本发明的非天然存在的微生物和方法还可以彼此和与本领域熟知的其它微生物和方法以各种组合使用以达到产物通过其它路径的生物合成。例如,不使用对

甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯生产者来产生对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯的一种替代方法是通过加入另一种能够将甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯途径中间体转变成对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯的微生物。一种这样的程序包括例如产生对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯途径中间体的微生物的发酵。然后,对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯途径中间体可以用作第二种微生物的底物,从而将对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯途径中间体转变成对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯。可以将对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯途径中间体直接加入到第二种生物体的另一种培养物中,或者对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯途径中间体生产者的原始培养物可以通过例如细胞分离来清除掉这些微生物,然后可以向发酵肉汤中后续添加第二种生物体用于生产最终产物而无需中间纯化步骤。

[0059] 在其它实施方案中,本发明的非天然存在的微生物和方法可以按各种各样的亚途径装配以实现例如对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯的生物合成。在这些实施方案中,本发明所需产物的生物合成途径分离在不同的微生物中,并且这些不同的微生物可以共培养以产生最终产物。在这样一种生物合成方案中,一种微生物的产物是第二种微生物的底物直到最终产物被合成。例如,对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯的生物合成可以通过构建含有将一种途径中间体转变成另一种途径中间体或产物的生物合成途径的微生物来完成。或者,对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯也可以通过在同一容器中使用两种生物体的共培养或共发酵由微生物经生物合成生产,其中第一种微生物生产对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯中间体,而第二种微生物将此中间体转变成对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯。

[0060] 考虑到本文提供的教导和指导,本领域技术人员将会理解,存在着本发明的非天然存在的微生物和方法与其它微生物、与具有亚途径的其它非天然存在的微生物的共培养物 and 与本领域熟知的生产对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯的其它化学和/或生物化学方法步骤的组合一起的各种组合和排列。

[0061] 对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)磷酸酯途径酶或蛋白质的编码核酸来源可包括例如其中所编码的基因产物能够催化参考反应的任何物种。这样的物种既包括原核生物又包括真核生物,包括但不限于细菌,包括古细菌(*archaea*)和真细菌(*eubacteria*);和真核生物,包括酵母、植物、昆虫、动物和哺乳动物,包括人。这些来源的示例性物种包括例如大肠杆菌(*Escherichia coli*)、结核分枝杆菌(*Mycobacterium tuberculosis*)、根癌土壤杆菌(*Agrobacterium tumefaciens*)、枯草芽孢杆菌(*Bacillus subtilis*)、集胞藻(*Synechocystis sp.*)、拟南芥(*Arabidopsis thaliana*)、运动发酵单胞菌(*Zymomonas mobilis*)、产酸克雷伯氏菌(*Klebsiella oxytoca*)、鼠伤寒沙门氏菌(*Salmonella typhimurium*)、伤寒沙门氏菌(*Salmonella typhi*)、丘状菌落乳杆菌(*Lactobacillus collinoides*)、肺炎克雷伯氏菌(*Klebsiella pneumoniae*)、巴氏梭菌(*Clostridium pasteurianum*)、弗氏柠檬酸杆菌(*Citrobacter freundii*)、丁酸梭菌

(*Clostridium butyricum*)、*Roseburia inulinivorans*、硫磺矿硫化叶菌 (*Sulfolobus solfataricus*)、粗糙链孢霉 (*Neurospora crassa*)、费氏中华根瘤菌 (*Sinorhizobium fredii*)、幽门螺杆菌 (*Helicobacter pylori*)、激烈热球菌 (*Pyrococcus furiosus*)、流感嗜血菌 (*Haemophilus influenzae*)、菊欧文氏菌 (*Erwinia chrysanthemi*)、金黄色葡萄球菌 (*Staphylococcus aureus*)、盐生杜氏藻 (*Dunaliella salina*)、肺炎链球菌 (*Streptococcus pneumoniae*)、酿酒酵母 (*Saccharomyces cerevisiae*)、构巢曲霉 (*Aspergillus nidulans*)、卡氏肺囊虫 (*Pneumocystis carinii*)、天蓝色链霉菌 (*Streptomyces coelicolor*)，来自伯克氏菌属 (*Burkholderia*)、产碱菌属 (*Alcaligenes*)、假单胞菌属 (*Pseudomonas*)、鞘氨醇单胞菌属 (*Shingomonas*) 和丛毛单胞菌属 (*Comamonas*) (例如睾丸酮丛毛单胞菌 (*Comamonas testosteroni*)) 的菌种，以及本文公开的或可作为相应基因的来源生物体获得的其它示例性物种。然而，关于从现在起超过 550 个种可获得的完全基因组序列 (其中超过一半可在公共数据库例如 NCBI 上获得)，包括 395 种微生物基因组和各种各样的酵母、真菌、植物和哺乳动物基因组，在近缘或远缘物种中的一个或多个基因中，包括例如已知基因的同系物 (homologue)、直系同源物、旁系同源物和非直系同源基因置换，编码所必需的对甲苯甲酸、对苯二甲酸或 (2-羟基-3-甲基-4-氧代丁氧基) 磷酸酯生物合成活性的基因的鉴定，以及生物体间遗传改变的交换在本领域是常规的和众所周知的。因此，对于某一特定的生物体例如大肠杆菌来说，允许本文所述的对甲苯甲酸、对苯二甲酸或 (2-羟基-3-甲基-4-氧代丁氧基) 磷酸酯的生物合成的代谢改变可同样容易地适用于其它微生物，包括原核生物和真核生物。考虑到本文提供的教导和指导，本领域技术人员将会知道，一种生物体中例举的代谢改变可同等地应用于其它生物体。

[0062] 在有些情况下，例如当无亲缘关系的物种中存在对甲苯甲酸、对苯二甲酸或 (2-羟基-3-甲基-4-氧代丁氧基) 磷酸酯生物合成替代途径时，通过例如来自催化相似、但非相同的代谢反应以代替参考反应的无亲缘关系的物种的一种或多种旁系同源物的外源表达，可以在宿主物种上赋予对甲苯甲酸、对苯二甲酸或 (2-羟基-3-甲基-4-氧代丁氧基) 磷酸酯生物合成。因为不同的生物体之间在代谢网络中存在着某些差异，所以，本领域技术人员应当理解，在不同的生物体之间实际基因使用可能有差异。然而，考虑到本文提供的教导和指导，本领域技术人员也应当理解，采用本文例举的那些的关联 (cognate) 代谢改变以在目标物种中构建将合成对甲苯甲酸、对苯二甲酸或 (2-羟基-3-甲基-4-氧代丁氧基) 磷酸酯的微生物，本发明的教导和方法可适用于所有的微生物。

[0063] 用于构建非天然存在的产对甲苯甲酸、对苯二甲酸或 (2-羟基-3-甲基-4-氧代丁氧基) 磷酸酯宿主并测试其表达水平的方法可以通过例如本领域众所周知的重组和检测方法来实现。这样的方法可以在例如以下文献中找到：Sambrook 等人，*Molecular Cloning :A Laboratory Manual*，第三版，Cold Spring Harbor Laboratory, New York (2001)；和 Ausubel 等人，*Current Protocols in Molecular Biology*, John Wiley and Sons, Baltimore, MD (1999)。

[0064] 采用本领域众所周知的技术，包括但不限于接合、电穿孔、化学转化、转导、转染和超声转化，可以将参与产生对甲苯甲酸、对苯二甲酸或 (2-羟基-3-甲基-4-氧代丁氧基) 磷酸酯的途径的外源核酸序列稳定地或瞬时地导入宿主细胞中。对于在大肠杆菌或其它原核细胞中的外源表达，真核生物核酸的基因或 cDNA 中的某些核酸序列可以编码靶向信号，

例如 N- 端线粒体靶向信号或其它靶向信号, 这些靶向信号可以在转化到原核宿主细胞之前去除掉 ( 如果需要的话)。例如, 线粒体前导序列的去除导致在大肠杆菌中的表达增加 (Hoffmeister 等人, *J. Biol. Chem.* 280:4329-4338 (2005))。对于在酵母或其它真核细胞中的外源表达, 基因可以在胞质溶胶中表达而无需添加前导序列, 或者可以靶向线粒体或其它细胞器, 或者通过添加合适的靶向序列例如适合宿主细胞的线粒体靶向信号或分泌信号而靶向分泌。因此, 应当理解, 为了去除或包括靶向序列而对核酸序列的适宜修饰可以被掺入到外源核酸序列中以带来所需要的特性。此外, 基因可以用本领域众所周知的技术进行密码子优化以达到蛋白质的最佳表达。

[0065] 正如本文例举的, 可以构建一种或多种表达载体以包括一种或多种对甲苯甲酸、对苯二甲酸或 (2- 羟基 -3- 甲基 -4- 氧代丁氧基) 磷酸酯生物合成途径编码核酸, 该编码核酸与宿主生物中有功能的表达调控序列操作性连接。适用于本发明的微生物宿主生物的表达载体包括例如质粒、噬菌体载体、病毒载体、附加体 (episomes) 和人工染色体, 包括能够稳定地整合到宿主染色体中的载体和选择序列或标记。另外, 表达载体可包括一个或多个选择标记基因和适宜的表达调控序列。也可包括例如提供对抗生素或毒素的抗性、代偿营养缺陷或者供给培养基中没有的关键营养物的选择标记基因。表达调控序列可包括组成型和诱导型启动子、转录增强子、转录终止子等, 这些都是本领域众所周知的。当两种或更多种外源编码核酸必须共表达时, 可以将两种核酸例如插入到一个表达载体中或者插入到独立的表达载体中。对于单个载体表达, 编码核酸可以与一个共同的表达调控序列操作性连接或者与不同的表达调控序列 (例如一个诱导型启动子和一个组成型启动子) 操作性连接。参与代谢或合成途径的外源核酸序列的转化可以使用本领域众所周知的方法来证实。这样的方法包括例如核酸分析法, 例如 mRNA 的 RNA 印迹法或聚合酶链式反应 (PCR) 扩增法, 或用于基因产物表达的免疫印迹法, 或测试引入的核酸序列或其相应基因产物表达的其它合适分析方法。本领域技术人员应当理解, 外源核酸以足以产生所需产物的量进行表达, 本领域技术人员还应当理解, 可以采用本领域众所周知的和本文公开的方法对表达水平进行优化以获得足够的表达。

[0066] 本发明另外还提供一种用于生产 (2- 羟基 -3- 甲基 -4- 氧代丁氧基) 磷酸酯的方法, 包括将含有 (2- 羟基 -3- 甲基 -4- 氧代丁氧基) 磷酸酯途径的非天然存在的微生物在足以产生 (2- 羟基 -3- 甲基 -4- 氧代丁氧基) 磷酸酯的条件下培养足够的时间周期。这样一种微生物可具有 (2- 羟基 -3- 甲基 -4- 氧代丁氧基) 磷酸酯途径, 该途径包含至少一种编码以足以产生 (2- 羟基 -3- 甲基 -4- 氧代丁氧基) 磷酸酯的量表达的 (2- 羟基 -3- 甲基 -4- 氧代丁氧基) 磷酸酯途径酶的外源核酸, 所述 (2- 羟基 -3- 甲基 -4- 氧代丁氧基) 磷酸酯途径包含 2-C- 甲基 -D- 赤藓糖醇 -4- 磷酸脱水酶 (参见实施例 I 和图 1, 步骤 C)。(2- 羟基 -3- 甲基 -4- 氧代丁氧基) 磷酸酯途径可任选还包含 1- 脱氧木酮糖 -5- 磷酸合酶和 / 或 1- 脱氧 -D- 木酮糖 -5- 磷酸还原异构酶 (参见实施例 I 和图 1, 步骤 A 和 B)。

[0067] 在另一个实施方案中, 本发明提供一种用于生产对甲苯甲酸的方法, 包括将包含对甲苯甲酸途径的非天然存在的微生物在足以产生对甲苯甲酸的条件下培养足够的时间周期。对甲苯甲酸途径可包含至少一种编码以足以产生对甲苯甲酸的量表达的对甲苯甲酸途径酶的外源核酸, 所述对甲苯甲酸途径包含 2- 脱氢 -3- 脱氧磷酸庚糖酸合酶、3- 脱氢奎尼酸合酶、3- 脱氢奎尼酸脱水酶、莽草酸脱氢酶、莽草酸激酶、3- 磷酸莽草酸 -2- 羧基乙烯

基转移酶、分支酸合酶和 / 或分支酸裂合酶 (参见实施例 II 和图 2, 步骤 A-H)。在另一个实施方案中, 本发明的方法可以使用一种还包含 (2-羟基-3-甲基-4-氧代丁氧基) 磷酸酯途径的非天然存在的微生物 (参见实施例 I 和图 1)。这样一种 (2-羟基-3-甲基-4-氧代丁氧基) 磷酸酯途径可包含 2-C-甲基-D-赤藓糖醇-4-磷酸脱水酶、1-脱氧木酮糖-5-磷酸合酶和 / 或 1-脱氧-D-木酮糖-5-磷酸还原异构酶 (参见实施例 I 和图 1)。

[0068] 本发明还提供一种用于生产对苯二甲酸的方法, 包括将含有对苯二甲酸途径的非天然存在的微生物在足以产生对苯二甲酸的条件培养足够的时间周期。这样一种对苯二甲酸途径可包含至少一种编码以足以产生对苯二甲酸的量表达的对苯二甲酸途径酶的外源核酸, 所述对苯二甲酸途径包含对甲苯甲酸甲基-单加氧酶还原酶、4-羧基苯甲醇脱氢酶和 / 或 4-羧基苯甲醛脱氢酶。这样一种微生物还可包含对甲苯甲酸途径, 其中所述对甲苯甲酸途径包含 2-脱氢-3-脱氧磷酸庚糖酸合酶、3-脱氢奎尼酸合酶、3-脱氢奎尼酸脱水酶、莽草酸脱氢酶、莽草酸激酶、3-磷酸莽草酸-2-羧基乙烯基转移酶、分支酸合酶和 / 或分支酸裂合酶 (参见实施例 2 和实施例 3 及图 2 和图 3)。在另一个实施方案中, 所述非天然存在的微生物还可包含 (2-羟基-3-甲基-4-氧代丁氧基) 磷酸酯途径 (参见实施例 I 和图 1)。因此, 在一个具体的实施方案中, 本发明提供一种非天然存在的微生物及其使用方法, 其中所述微生物含有对甲苯甲酸途径、对苯二甲酸途径和 (2-羟基-3-甲基-4-氧代丁氧基) 磷酸酯途径。

[0069] 可以采用众所周知的方法进行合适的纯化和 / 或测定以测试对甲苯甲酸、对苯二甲酸或 (2-羟基-3-甲基-4-氧代丁氧基) 磷酸酯的产生。对于每个待测的工程菌株, 可以让合适的重复例如一式三份的培养物生长。例如, 可以监测工程生产宿主中产物和副产物的形成。通过诸如 HPLC (高效液相层析)、GC-MS (气相层析-质谱法)、LC-MS (液相层析-质谱法) 和 UV-可见光谱法的方法或者采用本领域众所周知的常规程序的其它合适分析方法, 可以对最终的产物和中间体、和其它有机化合物进行分析。发酵肉汤中产物的释放也可以用培养物上清液进行测试。采用例如折射率检测器 (对于葡萄糖和醇) 和 UV 检测器 (对于有机酸) (Lin 等人, *Biotechnol. Bioeng.* 90:775-779 (2005)) 或本领域众所周知的其它合适的测定和检测方法, 副产物和残留葡萄糖可以通过 HPLC 进行定量。来自外源 DNA 序列的个体酶或蛋白质活性也可以用本领域众所周知的方法进行测定。例如, 对甲苯甲酸甲基-单加氧酶活性可以通过以下方法进行测定: 将纯化的酶与 NADH、FeSO<sub>4</sub> 和对甲苯甲酸底物一起在水浴中孵育, 通过蛋白质的沉淀使反应终止, 和通过 HPLC 分析上清液中的产物 (Locher 等人, *J. Bacteriol.* 173:3741-3748 (1991))。

[0070] 采用本领域众所周知的各种方法, 可以将对甲苯甲酸、对苯二甲酸或 (2-羟基-3-甲基-4-氧代丁氧基) 磷酸酯与培养物中的其它组分分离开来。这样的分离方法包括例如萃取程序以及方法, 包括连续的液液萃取、全蒸发、膜过滤、膜分离、反渗透、电渗析、蒸馏、结晶、离心、萃取过滤、离子交换层析、大小排阻层析、吸附层析和超滤。所有的上述方法都是本领域众所周知的。

[0071] 本文所述的任何非天然存在的微生物可以进行培养以产生和 / 或分泌本发明的生物合成产物。例如, 对甲苯甲酸、对苯二甲酸或 (2-羟基-3-甲基-4-氧代丁氧基) 磷酸酯生产者可以进行培养以经生物合成生产对甲苯甲酸、对苯二甲酸或 (2-羟基-3-甲基-4-氧代丁氧基) 磷酸酯。

[0072] 为了生产对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯,将重组菌株在含碳源和其它必需营养物的培养基中进行培养。有时较为理想的是在发酵罐中维持厌氧条件以降低总过程的成本。这样的条件可以用下述方法获得:例如,首先在培养基中充入氮气,然后将带有隔片和皇冠盖的烧瓶密封。对于其中在厌氧条件下未观察到生长的菌株来说,可以采用微好氧条件,即在隔片上钻小洞供有限通气。示例性的厌氧条件先前已有描述并且是本领域众所周知的。示例性的好氧和厌氧条件描述于例如于2007年8月10日申请的美国专利公布2009/0047719。发酵可按本文公开的分批方式、补料-分批方式或连续方式进行。

[0073] 如有需要,可以将培养基的pH维持在所需要的pH,特别是中性pH,例如pH约为7,即通过按需要加入碱(例如NaOH或其它碱)或酸以将培养基维持在所需要的pH。生长速率可以通过使用分光光度计(600nm)测量光密度来测定,葡萄糖摄取速率可以通过监测碳源随时间的消耗量来测定。

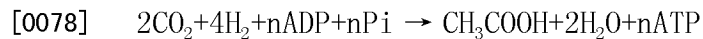
[0074] 生长培养基可包括例如可以给所述非天然存在的微生物供应碳源的任何碳水化合物来源。这样的来源包括例如糖,例如葡萄糖、木糖、阿拉伯糖、半乳糖、甘露糖、果糖、蔗糖和淀粉。碳水化合物的其它来源包括例如可再生贮存物(renewable feedstocks)和生物质。在本发明的方法中可用作贮存物的示例性生物质类型包括纤维素类生物质、半纤维素类生物质和木质素贮存物或贮存物部分。这样的生物质类贮存物含有例如可用作碳源的碳水化合物底物例如葡萄糖、木糖、阿拉伯糖、半乳糖、甘露糖、果糖和淀粉。考虑到本文提供的教导和指导,本领域技术人员将会理解,除了以上例举的那些以外的可再生贮存物和生物质也可以用来培养本发明的微生物以生产对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯。

[0075] 除了可再生贮存物例如以上例举的那些以外,本发明的对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯微生物也可以进行修饰以在作为其碳源的合成气上进行生长。在这个具体的实施方案中,使一种或多种蛋白质或酶在产对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基)膦酸酯的生物体中进行表达以提供利用合成气或其它气体碳源的代谢途径。

[0076] 合成气体(synthesis gas),也称为合成气(syngas)或发生炉煤气(producer gas),是煤的气化和含碳材料例如生物质材料(包括农作物和残留物)的气化的主要产物。合成气是主要为 $H_2$ 和CO的混合物并且可以得自任何有机贮存物,包括但不限于煤、煤油、天然气、生物质和废有机物质的气化。气化一般在高燃料/氧比率下进行。尽管主要是 $H_2$ 和CO,但是合成气也可包括少量的 $CO_2$ 和其它气体。因此,合成气体提供了一种成本有效的的气体碳源例如CO,另外还有 $CO_2$ 。

[0077] Wood-Ljungdahl途径催化CO和 $H_2$ 转变成乙酰-CoA和其它产物(例如乙酸盐(酯)(acetate))。能够利用CO和合成气的生物体一般也具有通过Wood-Ljungdahl途径所包括的一组相同的基础酶和转化作用来利用 $CO_2$ 和 $CO_2/H_2$ 混合物的能力。由微生物将 $CO_2$ 经 $H_2$ 依赖性转变成乙酸酯在揭示出CO也可以被相同的生物体所利用并且涉及相同的途径之前不久就被认识到。许多产乙酸菌(acetogen)已显示能在 $CO_2$ 存在下生长并产生诸如乙酸盐的化合物,只要有氢存在以供给必需的还原当量即可(参见例如,Drake, Acetogenesis, pp. 3-60 Chapman and Hall, New York, (1994))。这可以用以下反应

式概述：



[0079] 因此，具备 Wood-Ljungdahl 途径的非天然存在的微生物也可以利用  $\text{CO}_2$  和  $\text{H}_2$  混合物来产生乙酰 -CoA 和其它所需产物。

[0080] Wood-Ljungdahl 途径是本领域众所周知的并且由 12 个反应组成，这 12 个反应可分成两支：(1) 甲基支和 (2) 羧基支。甲基支将合成气转变成甲基 - 四氢叶酸（甲基 -THF），而羧基支将甲基 -THF 转变成乙酰 -CoA。甲基支中的反应由以下酶或蛋白质按顺序催化：铁氧还蛋白氧化还原酶、甲酸脱氢酶、甲酰基四氢叶酸合成酶、次甲基四氢叶酸环化脱水酶、亚甲基四氢叶酸脱氢酶和亚甲基四氢叶酸还原酶。羧基支中的反应由以下酶或蛋白质按顺序催化：甲基四氢叶酸：类咕啉（corrino）蛋白甲基转移酶（例如 AcsE）、类咕啉铁 - 硫蛋白、镍蛋白装配蛋白（例如 AcsF）、铁氧还蛋白、乙酰 -CoA 合酶、一氧化碳脱氢酶和镍蛋白装配蛋白（例如 CooC）。根据本文提供的用于引入足够数目的编码核酸以产生对甲苯甲酸途径、对苯二甲酸途径或（2- 羟基 -3- 甲基 -4- 氧代丁氧基）磷酸酯途径的教导和指导，本领域技术人员应当理解，就至少导入在宿主生物中不存在的编码 Wood-Ljungdahl 酶或蛋白质的核酸而论，也可以进行相同的工程设计。因此，将一种或多种编码核酸导入到本发明的微生物中使得经修饰的生物体含有完全 Wood-Ljungdahl 途径将会赋予合成气利用能力。

[0081] 还原性三羧酸循环结合一氧化碳脱氢酶和 / 或加氢酶活性也可以允许  $\text{CO}$ 、 $\text{CO}_2$  和 / 或  $\text{H}_2$  转变成乙酰 -CoA 和其它产物（例如乙酸盐（酯））。通过还原性 TCA 途径能够固定碳的生物体可以利用下列酶中的一种或多种：ATP 柠檬酸 - 裂合酶、柠檬酸裂合酶、顺乌头酸酶、异柠檬酸脱氢酶、 $\alpha$  - 酮戊二酸：铁氧还蛋白氧化还原酶、琥珀酰 -CoA 合成酶、琥珀酰 -CoA 转移酶、延胡索酸还原酶、延胡索酸酶、苹果酸脱氢酶、NAD(P)H：铁氧还蛋白氧化还原酶、一氧化碳脱氢酶和加氢酶。具体地讲，通过一氧化碳脱氢酶和加氢酶从  $\text{CO}$  和 / 或  $\text{H}_2$  萃取的还原性当量（reducing equivalents）用于通过还原性 TCA 循环将  $\text{CO}_2$  固定成为乙酰 -CoA 或乙酸盐（酯）。乙酸盐（酯）可通过例如乙酰 -CoA 转移酶、乙酸激酶 / 磷酸转乙酰酶和乙酰 -CoA 合成酶等酶转变成乙酰 -CoA。乙酰 -CoA 可通过丙酮酸：铁氧还蛋白氧化还原酶和糖异生作用（gluconeogenesis）的酶转变成对甲苯甲酸、对苯二甲酸或（2- 羟基 -3- 甲基 -4- 氧代丁氧基）磷酸酯前体、3- 磷酸甘油醛、磷酸烯醇丙酮酸和丙酮酸。根据本文提供的用于引入足够数目的编码核酸以产生对甲苯甲酸、对苯二甲酸或（2- 羟基 -3- 甲基 -4- 氧代丁氧基）磷酸酯途径的教导和指导，本领域技术人员应当理解，就至少导入在宿主生物中不存在的编码还原性 TCA 途径酶或蛋白质的核酸而论，也可以进行相类似的工程设计。因此，将一种或多种编码核酸导入到本发明的微生物中使得经修饰的生物体含有完全还原性 TCA 途径将会赋予合成气利用能力。

[0082] 因此，考虑到本文提供的教导和指导，本领域技术人员应当理解，非天然存在的微生物当在碳源（例如碳水化合物）上生长时就可以产生并分泌本发明的生物合成化合物。这样的化合物包括例如对甲苯甲酸、对苯二甲酸或（2- 羟基 -3- 甲基 -4- 氧代丁氧基）磷酸酯和任何在对甲苯甲酸、对苯二甲酸或（2- 羟基 -3- 甲基 -4- 氧代丁氧基）磷酸酯途径中的中间代谢物。在所需酶或蛋白质活性中的一个或多个中，全都需要进行工程改造以达到所需化合物或中间体，包括例如对甲苯甲酸、对苯二甲酸或（2- 羟基 -3- 甲基 -4- 氧代丁

氧基) 磷酸酯生物合成途径中的某些或全部的生物合成。因此, 本发明提供一种非天然存在的微生物, 当在碳水化合物或其它碳源上生长时能产生和 / 或分泌对甲苯甲酸、对苯二甲酸或 (2- 羟基 -3- 甲基 -4- 氧代丁氧基) 磷酸酯并且当在碳水化合物或其它碳源上生长时能产生和 / 或分泌对甲苯甲酸、对苯二甲酸或 (2- 羟基 -3- 甲基 -4- 氧代丁氧基) 磷酸酯途径中所示的中间代谢物的任何一个。本发明的产对甲苯甲酸、对苯二甲酸或 (2- 羟基 -3- 甲基 -4- 氧代丁氧基) 磷酸酯微生物可以从中间体开始合成。例如, (2- 羟基 -3- 甲基 -4- 氧代丁氧基) 磷酸酯途径中间体可以是 1- 脱氧 -D- 木酮糖 -5- 磷酸或 C- 甲基 -D- 赤藓糖醇 -4- 磷酸 (参见实施例 I 和图 1)。对甲苯甲酸途径中间体可以是例如 2, 4- 二羟基 -5- 甲基 -6- [( 磷酰氧基 ) 甲基 ] 噁烷 -2- 羧酸酯、1, 3- 二羟基 -4- 甲基 -5- 氧代环己烷 -1- 羧酸酯、5- 羟基 -4- 甲基 -3- 氧代环己 -1- 烯 -1- 羧酸酯、3, 5- 二羟基 -4- 甲基环己 -1- 烯 -1- 羧酸酯、5- 羟基 -4- 甲基 -3- ( 磷酰氧基 ) 环己 -1- 烯 -1- 羧酸酯、5- [(1- 羧基乙 -1- 烯 -1- 基) 氧基] -4- 甲基 -3- ( 磷酰氧基 ) 环己 -1- 烯 -1- 羧酸酯或 3- [(1- 羧基乙 -1- 烯 -1- 基) 氧基] -4- 甲基环己 -1, 5- 二烯 -1- 羧酸酯 (参见实施例 II 和图 2)。对苯二甲酸中间体可以是例如 4- 羧基苯甲醇或 4- 羧基苯甲醛 (参见实施例 III 和图 3)。

[0083] 使用本领域众所周知的方法 (如本文例举的方法) 构建本发明的非天然存在的微生物以外源性表达至少一种编码对甲苯甲酸、对苯二甲酸或 (2- 羟基 -3- 甲基 -4- 氧代丁氧基) 磷酸酯途径酶或蛋白质的核酸, 其表达量足以产生对甲苯甲酸、对苯二甲酸或 (2- 羟基 -3- 甲基 -4- 氧代丁氧基) 磷酸酯。应当理解, 本发明的微生物在足以产生对甲苯甲酸、对苯二甲酸或 (2- 羟基 -3- 甲基 -4- 氧代丁氧基) 磷酸酯的条件下进行培养。根据本文提供的教导和指导, 本发明的非天然存在的微生物可以实现生物合成对甲苯甲酸、对苯二甲酸或 (2- 羟基 -3- 甲基 -4- 氧代丁氧基) 磷酸酯, 导致胞内浓度在约 0.1-200mM 之间或更高。一般而言, 对甲苯甲酸、对苯二甲酸或 (2- 羟基 -3- 甲基 -4- 氧代丁氧基) 磷酸酯的胞内浓度是在约 3-150mM 之间, 特别是在约 5-125mM 之间, 和更特别是在约 8-100mM 之间, 包括约 10mM、20mM、50mM、80mM 或更高。在这些示例性范围中的每一个范围内或者高于它们的胞内浓度也可以由本发明的非天然存在的微生物实现。

[0084] 在有些实施方案中, 培养条件包括厌氧或基本上厌氧的生长或维持条件。示例性的厌氧条件先前已有描述并且是本领域众所周知的。用于发酵过程的示例性厌氧条件在本文中有描述并且描述于 2007 年 8 月 10 日申请的美国专利公布 2009/0047719。这些条件中的任何一个都可以与所述非天然存在的微生物以及本领域众所周知的其它厌氧条件一起使用。对甲苯甲酸、对苯二甲酸或 (2- 羟基 -3- 甲基 -4- 氧代丁氧基) 磷酸酯生产者可以在基本上厌氧的条件下合成对甲苯甲酸、对苯二甲酸或 (2- 羟基 -3- 甲基 -4- 氧代丁氧基) 磷酸酯, 达到胞内浓度为 5-10mM 或更高以及本文例举的所有其它浓度。应当理解, 即使上述描述指的是胞内浓度, 但是产对甲苯甲酸、对苯二甲酸或 (2- 羟基 -3- 甲基 -4- 氧代丁氧基) 磷酸酯微生物可以在细胞内产生对甲苯甲酸、对苯二甲酸或 (2- 羟基 -3- 甲基 -4- 氧代丁氧基) 磷酸酯和 / 或将该产物分泌到培养基中。

[0085] 除了本文公开的培养和发酵条件以外, 用于实现对甲苯甲酸、对苯二甲酸或 (2- 羟基 -3- 甲基 -4- 氧代丁氧基) 磷酸酯生物合成的生长条件还可包括往培养条件中添加渗透保护剂。在某些实施方案中, 本发明的非天然存在的微生物可以在渗透保护剂存在下按本文所述方法进行保持、培养或发酵。简而言之, 渗透保护剂是指起到渗透质的作用的

并且有助于本文所述微生物在渗透胁迫下存活的化合物。渗透保护剂包括但不限于甜菜碱、氨基酸和糖（如海藻糖）。这类渗透保护剂的非限制性实例有甘氨酸甜菜碱、果仁糖甜菜碱、二甲基噻亭、二甲基磺酸基（sulfonio）丙酸盐、3-二甲基磺酸基（sulfonio）-2-甲基丙酸盐、哌可酸、二甲基磺酸基乙酸盐、胆碱、L-肉毒碱和（S）-2-甲基-1,4,5,6-四氢甲基嘧啶-4-羧酸（ectoine）。在一个方面，所述渗透保护剂是甘氨酸甜菜碱。本领域普通技术人员应当理解，适合保护本文所述微生物免于渗透胁迫的渗透保护剂的含量和类型将取决于所使用的微生物。渗透保护剂在培养条件下的含量可以例如为不超过约0.1mM、不超过约0.5mM、不超过约1.0mM、不超过约1.5mM、不超过约2.0mM、不超过约2.5mM、不超过约3.0mM、不超过约5.0mM、不超过约7.0mM、不超过约10mM、不超过约50mM、不超过约100mM或不超过约500mM。

[0086] 培养条件可包括例如液体培养程序以及发酵和其它大规模培养程序。如本文所述的，本发明的生物合成产物的特别有用的产能可以在厌氧或基本上厌氧培养条件下获得。

[0087] 如本文所述的，用于实现生物合成对甲苯甲酸、对苯二甲酸或（2-羟基-3-甲基-4-氧代丁氧基）膦酸酯的一种示例性生长条件包括厌氧培养或发酵条件。在某些实施方案中，本发明的非天然存在的微生物可以在厌氧或基本上厌氧的条件下进行保持、培养或发酵。简而言之，厌氧条件是指无氧的环境。基本上厌氧的条件包括例如培养、分批发酵或连续发酵使得培养基中的溶解氧浓度保持在0和10%饱和之间。基本上厌氧的条件也包括密封室里面的液体培养基中或固体琼脂上的生长细胞或静止细胞维持在小于1%氧的气氛下。氧的百分比可以通过例如在培养物中通入N<sub>2</sub>/CO<sub>2</sub>混合物或其它合适的一种或多种非氧气体来维持。

[0088] 为了制备对甲苯甲酸、对苯二甲酸或（2-羟基-3-甲基-4-氧代丁氧基）膦酸酯，可以将本文所述的培养条件连续放大和生长。示例性生长程序包括例如补料-分批发酵和分批分离；补料-分批发酵和连续分离；或连续发酵和连续分离。所有的这些过程都是本领域众所周知的。发酵程序特别用于经生物合成生产商业数量的对甲苯甲酸、对苯二甲酸或（2-羟基-3-甲基-4-氧代丁氧基）膦酸酯。一般而言，并根据非连续培养程序，对甲苯甲酸、对苯二甲酸或（2-羟基-3-甲基-4-氧代丁氧基）膦酸酯的连续和/或接近连续生产包括将本发明的非天然存在的产对甲苯甲酸、对苯二甲酸或（2-羟基-3-甲基-4-氧代丁氧基）膦酸酯生物体在足够的营养物和培养基中进行培养以保持和/或几乎保持其以指数期生长。在这样的条件下连续培养可以包括例如生长1天、2、3、4、5、6或7天或更长时间。另外，连续培养可包括1周、2、3、4或5周或更多周，甚至达到数月的较长时间周期。或者，本发明的生物体可以培养数小时，如果适合具体应用的话。应当理解，连续和/或接近连续培养条件也可包括在这些示例性周期之间以内的所有时间间隔。还应当理解，本发明微生物的培养时间是产生足够量的产物用于所需目的的足够时间周期。

[0089] 发酵程序是本领域众所周知的。简而言之，用于生物合成性生产对甲苯甲酸、对苯二甲酸或（2-羟基-3-甲基-4-氧代丁氧基）膦酸酯的发酵可用于例如补料-分批发酵和分批分离；补料-分批发酵和连续分离；或连续发酵和连续分离。分批和连续发酵程序的实例是本领域众所周知的。

[0090] 除了使用本发明的对甲苯甲酸、对苯二甲酸或（2-羟基-3-甲基-4-氧代丁氧基）膦酸酯生产者来连续生产相当大量的对甲苯甲酸、对苯二甲酸或（2-羟基-3-甲基-4-氧

代丁氧基) 膦酸酯的上述发酵程序以外, 所述对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基) 膦酸酯生产者也可以例如同时经历化学合成程序以将产物转变成其它化合物, 或者产物可以从发酵培养物中分离出来后再经历化学转变以将产物转变成其它化合物, 如果需要的话。

[0091] 为了产生更好的生产者, 可以采用代谢建模来使生长条件最优化。建模也可用来设计另外最优化途径利用的基因敲除 (gene knockout) (参见例如, 美国专利公布 US 2002/0012939、US 2003/0224363、US 2004/0029149、US 2004/0072723、US 2003/0059792、US 2002/0168654 和 US 2004/0009466 和美国专利第 7, 127, 379 号)。建模分析允许可靠地预测对代谢朝着更有效产生对甲苯甲酸、对苯二甲酸或(2-羟基-3-甲基-4-氧代丁氧基) 膦酸酯转换的细胞生长的影响。

[0092] 用于鉴定和设计代谢改变有利于生物合成所需产物的一种计算方法是 OptKnock 计算框架 (framework) (Burgard 等人, *Biotechnol. Bioeng.* 84:647-657 (2003))。OptKnock 是一种代谢建模和模拟程序, 它给出基因缺失或破坏策略的建议, 导致产生遗传上稳定的微生物, 从而过量产生靶产物。具体地讲, 该框架检查微生物的完全代谢和 / 或生物化学网络, 以便给出遗传操作建议, 从而迫使所需生物化学品变成细胞生长的专性 (obligatory) 副产物。通过生物化学生产与细胞生长通过策略性置入的基因缺失或其它功能性基因破坏相结合, 在生物反应器中长时间周期之后施加在工程菌株上的生长选择压力导致生产性能的改进, 这种改进是强制生长偶联的生物化学生产所致的结果。最后, 当构建基因缺失时, 经设计的菌株回复成它们的野生型状态的可能性可忽略不计, 因为通过 OptKnock 选出的基因已从基因组中完全取出。因此, 这种计算方法可以用于或者鉴定导致生物合成所需产物的替代途径或者结合所述非天然存在的微生物使用以进一步最优化所需产物的生物合成。

[0093] 简而言之, OptKnock 是在本文中使用的的一个术语, 指的是用于细胞代谢建模的一种计算方法和系统。OptKnock 程序与将特殊约束 (constraint) 掺入到流量平衡分析 (FBA) 模型的模型和方法的框架相关。这些约束包括例如定性动态信息、定性调节信息和 / 或 DNA 微阵列实验数据。OptKnock 也计算各种代谢问题的解决方案, 例如通过收紧通过流量平衡模型获得的流量边界, 并随后探测在基因添加或缺失存在下代谢网络的性能限制。OptKnock 计算框架允许构建模型格式, 模型格式允许有效查询代谢网络的性能限制并提供解决所得混合整数线性编程问题的方法。本文称为 OptKnock 的代谢建模和模拟方法描述于例如美国专利公布 2002/0168654 (申请日为 2002 年 1 月 10 日)、国际专利申请第 PCT/US02/00660 号 (申请日为 2002 年 1 月 10 日) 和美国专利公布 2009/0047719 (申请日为 2007 年 8 月 10 日)。

[0094] 用于鉴定和设计代谢改变有利于生物合成性生产产物的另一种计算方法是一种称为 SimPheny<sup>®</sup> 的代谢建模和模拟系统。这种计算方法和系统描述于例如美国专利公布 2003/0233218 (申请日为 2002 年 6 月 14 日) 和国际专利申请第 PCT/US03/18838 号 (申请日为 2002 年 6 月 13 日)。SimPheny<sup>®</sup> 是一种计算系统, 其可用于产生计算机模拟 (in silico) 网络模型并且模拟质量、能量或电荷通过生物系统的化学反应的流通量 (flux), 以定义在该系统中含有化学反应的任何和所有可能功能性的解空间 (solution space), 从而确定生物系统的允许活性范围。这种方法被称为基于约束的建模, 因为解空间由约束例

如所包括反应的已知化学计量以及与通过反应的最大流通量相关的反应热力学和容量约束限定。由这些约束限定的空间可以进行询问以确定生物系统或其生化组分的表型能力和行为。

[0095] 这些计算方法与生物的真实性相一致,因为生物系统是灵活多变的并且能够以许多不同的方式达到相同的结果。通过进化机制设计出生物系统一直都受到所有有生命的系统必须面对的基础约束的限制。因此,基于约束的建模策略包括这些普遍 (general) 真实性。此外,通过约束收紧而将进一步限制连续施加到网络模型上的能力导致解空间大小的缩小,从而提高可预测生理性能或表型的精确度。

[0096] 考虑到本文提供的教导和指导,本领域技术人员将能够应用各种计算框架进行代谢建模和模拟以设计并实施所需化合物在宿主微生物中的生物合成。这样的代谢建模和模拟方法包括例如以上例举的计算系统,如 SimPheny<sup>®</sup>和 OptKnock。为了举例说明本发明,本文参照 OptKnock 的建模和模拟计算框架描述了一些方法。本领域技术人员将会知道如何使用 OptKnock 将代谢改变的鉴定、设计和实施应用到本领域众所周知的任何这类其它代谢建模和模拟计算框架和方法中。

[0097] 上述方法将提供一组代谢反应予以破坏。该组代谢反应或代谢修饰内各反应的消除可导致产生所需产物,该所需产物在生物体的生长阶段期间作为专性产物。因为反应都是已知的,所以解决 bilevel OptKnock 问题的方案也将提供编码催化该组反应内各反应的一种或多种酶的一个或多个相关基因。一组反应及其编码参与各反应的酶的相应基因的鉴定一般都是自动化过程,通过找出反应与具有在酶和编码基因之间的关系的反应数据库的联系来完成。

[0098] 一旦被鉴定出来,就在靶细胞或靶生物体中,通过功能性破坏至少一个编码该组内各代谢反应的基因来进行已被破坏的该组反应,以达到产生所需产物。一种特别有用的达到功能性破坏反应组的手段是通过剔除各编码基因。然而,在有些情况下,通过其它遗传畸变(包括例如调节区(例如启动子)或调节因子的顺式结合位点的突变、缺失),或者通过在任何数目的位置上编码序列的截短,来破坏反应可能是有益的。导致基因组小于总缺失的这些后来遗传畸变,都可用于例如当需要产物偶联的快速评价时,或者当遗传回复(reversion)几乎不可能发生时。

[0099] 为了确认上述 bilevel OptKnock 问题的额外生产方案,可以执行一种称为 integer cut 的最优化方法,上述 bilevel OptKnock 问题可导致更多组反应要破坏或代谢修饰从而生物合成(包括生长偶联的生物合成)所需产物。该方法向前进行以迭代解决以上例举的 OptKnock 问题并在各次迭代上掺入称为 integer cut 的额外约束。integer cut 约束有效地阻止求解过程(solution procedure),避免挑选出在将产物生物合成与生长强制偶联的任何之前迭代中鉴定的切实相同的反应组。例如,如果先前鉴定的生长偶联的代谢修饰对于破坏具体到反应 1、2 和 3,那么以下约束阻止相同反应以避免在后续求解中同时考虑。integer cut 方法是本领域众所周知的并且可以在例如 Burgard 等人, Biotechnol. Prog. 17:791-797(2001) 中找到。关于本文所述的所有方法及其结合代谢建模和模拟的 OptKnock 计算框架的应用,在迭代计算分析中降低冗余度的 integer cut 方法也可以结合本领域熟知的其它计算框架(包括例如 SimPheny<sup>®</sup>)使用。

[0100] 本文例举的方法允许构建经生物合成生产所需产物的细胞和生物体,包括将靶生

物化学产物的生产与经工程改造以包含已鉴定的遗传改变的细胞或生物体的生长专性相偶联。因此,本文所述的计算方法允许鉴定和实施通过选自 OptKnock 或 SimPheny<sup>®</sup> 的计算机模拟方法鉴定的代谢修饰。该组代谢修饰可包括例如添加一种或多种生物合成途径酶和 / 或功能性破坏一个或多个代谢反应,包括例如通过基因缺失来破坏。

[0101] 正如上所论述的, OptKnock 方法在突变型微生物网络当经历长时期的生长选择时可以朝向其计算预测的最大生长表型进化的前提下开发出来。换句话说,该方法促使生物体在选择压力下自身最优化的能力发生改变。OptKnock 框架允许基因缺失组合的穷尽性列举,迫使在生物化学生产和基于网络化学计量的细胞生长之间有偶联。最佳基因 / 反应敲除的鉴定需要 bilevel 最优化问题的方案,从而挑选出主动反应组使得所得网络的最佳生长方案过量生产目标生物化学产物 (Burgard 等人, *Biotechnol. Bioeng.* 84:647-657(2003))。

[0102] 大肠杆菌代谢的计算机模拟化学计量模型可用于鉴定代谢途径中的必需基因,如先前列举的和描述于例如美国专利公布 US 2002/0012939、US 2003/0224363、US 2004/0029149、US 2004/0072723、US 2003/0059792、US 2002/0168654 和 US 2004/0009466 ;和美国专利第 7,127,379 号。正如本文所公开的, OptKnock 数学框架可用于准确地指出导致所需产物的生长偶联生产的基因缺失。此外, bilevel OptKnock 问题的方案提供了唯一的一组缺失。为了列举出所有有意义的方案,也就是说,所有导致生长偶联生产形成的敲除组,可以执行一项称为 integer cut 的最优化技术。这项技术需要在各次迭代上掺入称为 integer cut 的额外约束来迭代解决所述 OptKnock 问题,正如以上所讨论的。

[0103] 正如本文所公开的,可以将编码对甲苯甲酸、对苯二甲酸或 (2-羟基-3-甲基-4-氧代丁氧基) 膦酸酯途径所需活性的核酸导入宿主生物中。在有些情况下,可能理想的是改变对甲苯甲酸、对苯二甲酸或 (2-羟基-3-甲基-4-氧代丁氧基) 膦酸酯途径酶或蛋白质的活性以增加对甲苯甲酸、对苯二甲酸或 (2-羟基-3-甲基-4-氧代丁氧基) 膦酸酯的产量。例如,可以将增加蛋白质或酶活性的已知突变引入到编码核酸分子中。另外,最优化方法可以用来增加酶或蛋白质的活性和 / 或减少抑制活性,例如减少负调节物的活性。

[0104] 一种这样的最优化方法是定向进化 (directed evolution)。定向进化是一种十分有效的方法,包括引入靶向特定基因的突变以便改进和 / 或改变酶的特性。经改进和 / 或改变的酶可以通过灵敏的高通量筛选测定的开发和实施来进行鉴定,该测定允许许多酶变异体 (例如,  $>10^4$ ) 的自动化筛选。通常要进行多轮重复的诱变和筛选才能获得具有最优化特性的酶。能够有助于鉴定诱变用基因的区域计算算法也已经开发出来并且可以明显地减少需要产生和筛选的酶变异体数目。已开发出多项定向进化技术 (关于综述参见 Hibbert 等人, *Biomol. Eng.* 22:11-19(2005); Huisman 和 Lalonde, *Biocatalysis in the pharmaceutical and Biotechnology industries* (制药和生物化学工业中的生物催化), 第 717-742 页 (2007), Patel (编著), CRC Press; Otten 和 Quax, *Biomol. Eng.* 22:1-9(2005); 和 Sen 等人, *Appl Biochem. Biotechnol.* 143:212-223(2007)) 以在产生多样性变异体文库上有效,并且这些方法已经成功地应用于横跨许多酶类别的多个特性的改进。已通过定向进化技术改进和 / 或改变的酶特征包括例如:选择性 / 专一性,用于非

天然底物的转变；温度稳定性，用于严酷高温加工处理；pH 稳定性，用于在较低或较高 pH 条件下的生物加工过程；底物或产物耐受性，使得可以达到高的产物效价；结合 ( $K_m$ )，包括拓宽底物结合以包括非天然底物；抑制 ( $K_i$ )，以去除被产物、底物或关键中间体抑制；活性 (kcat)，以提高酶促反应速率从而达到所需流通量；表达水平，以增加蛋白质收率和总体途径流通量；氧稳定性，用于在好氧条件下空气敏感酶的操作；和厌氧活性，用于在氧不存在时需氧酶的操作。

[0105] 许多示例性的方法已经开发出来用于靶向特定酶所需特性的基因的诱变和多样化。这样的方法都是本领域技术人员众所周知的。这些方法中的任何一个都可以用于改变和 / 或最优化对甲苯甲酸、对苯二甲酸或 (2- 羟基 -3- 甲基 -4- 氧代丁氧基) 磷酸酯途径酶或蛋白质的活性。这样的方法包括但不限于 EpPCR, EpPCR 通过在 PCR 反应中降低 DNA 聚合酶的保真度而引入随机点突变 (Pritchard 等人, *J. Theor. Biol.* 234:497-509(2005))；易错滚环扩增 (Error-prone Rolling Circle Amplification, epRCA), epRCA 类似于 epPCR, 只是用完全环状质粒作为模板, 并且用在最后 2 个核苷酸上具有外切核酸酶抗性硫代磷酸酯键的随机六聚体来扩增质粒, 然后将其转化到细胞中, 在转化的细胞中, 质粒在串联重复上被重新环化 (Fujii 等人, *Nucleic Acids Res.* 32:e145(2004))；和 Fujii 等人, *Nat. Protoc.* 1:2493-2497(2006))；DNA 或家族改组 (Family Shuffling), 其通常包括用核酸酶 (例如 Dnase I 或 EndoV) 消化两种或更多种变异基因以产生随机片段库, 该随机片段可在 DNA 聚合酶存在下通过退火和延伸的多次循环进行重新装配以产生嵌合基因文库 (Stemmer, *Proc Natl Acad Sci USA* 91:10747-10751(1994))；和 Stemmer, *Nature* 370:389-391(1994))；交错延伸 (Staggered Extension, StEP), 其需要模板引导, 然后经过 2 步 PCR 的重复循环, 并且经变性和非常短持续时间的退火 / 延伸 (短至 5 秒) (Zhao 等人, *Nat. Biotechnol.* 16:258-261(1998))；随机引导重组 (Random Priming Recombination, RPR), 其中用随机序列引物来产生大量互补于模板不同区段 (segment) 的短 DNA 片段 (Shao 等人, *Nucleic Acids Res.* 26:681-683(1998))。

[0106] 另外的方法包括异源双链体重组, 其中用线性化质粒 DNA 来形成被错配修复所修复的异源双链体 (Volkov 等人, *Nucleic Acids Res.* 27:e18(1999))；和 Volkov 等人, *Methods Enzymol.* 328:456-463(2000))；过渡模板随机嵌合生长 (Random Chimeragenesis on Transient Templates, RACHITT), RACHITT 使用单链 DNA (ssDNA) 的 Dnase I 片段化和大小分级分离 (Coco 等人, *Nat. Biotechnol.* 19:354-359(2001))；截短模板重组延伸 (Recombined Extension on Truncated templates, RETT), RETT 需要在用作模板库的单向 ssDNA 片段存在下模板转换来自引物的单向生长链 (Lee 等人, *J. Molec. Catalysis* 26:119-129(2003))；简并寡核苷酸基因改组 (Degenerate Oligonucleotide Gene Shuffling, DOGS), 其中用简并引物来控制分子间的重组 (Bergquist 和 Gibbs, *Methods Mol. Biol.* 352:191-204(2007))；Bergquist 等人, *Biomol. Eng.* 22:63-72(2005)；Gibbs 等人, *Gene* 271:13-20(2001))；用渐增截短法产生杂合酶 (Incremental Truncation for the Creation of Hybrid Enzymes, ITCHY), ITCHY 是用目标基因或基因片段的 1 个碱基对缺失来产生组合文库 (Ostermeier 等人, *Proc. Natl. Acad. Sci. USA* 96:3562-3567(1999))；和 Ostermeier 等人, *Nat. Biotechnol.* 17:1205-1209(1999))；用硫代渐增截短法产生杂合酶 (Thio-Incremental

Truncation for the Creation of Hybrid Enzymes, THIO-ITCHY), THIO-ITCHY 类似于 ITCHY, 只是用硫代磷酸 dNTPs 产生截短物 (Lutz 等人, *Nucleic Acids Res.* 29:E16(2001)); SCRATCHY, 它将 ITCHY 和 DNA 改组这两种重组基因的方法结合起来 (Lutz 等人, *Proc. Natl. Acad. Sci. USA* 98:11248-11253(2001)); 随机漂移诱变 (Random Drift Mutagenesis, RNDM), 其中通过 epPCR 产生的突变再进行筛选 / 选择并将有可利用活性的那些保留下来 (Bergquist 等人, *Biomol. Eng.* 22:63-72(2005)); 序列饱和诱变 (Sequence Saturation Mutagenesis, SeSaM), 一种使用随机掺入硫代磷酸核苷酸并裂解来产生随机长度片段库的随机诱变方法, 其用作模板以在“通用”碱基 (例如肌苷) 存在下延伸, 并且含肌苷互补序列的复制得到随机碱基掺入, 因此诱变 (Wong 等人, *Biotechnol. J.* 3:74-82(2008); Wong 等人, *Nucleic Acids Res.* 32:e26(2004); 和 Wong 等人, *Anal. Biochem.* 341:187-189(2005)); 合成改组 (Synthetic Shuffling), 其使用设计的重叠寡核苷酸来编码“靶标中的所有遗传多样性”并且允许经改组的子代的非常高多样性 (Ness 等人, *Nat. Biotechnol.* 20:1251-1255(2002)); 核苷酸交换和切除技术 (Nucleotide Exchange and Excision Technology, Next), 它利用 dUTP 掺入后再用尿嘧啶 DNA 糖基化酶处理后用哌啶处理以进行终点 DNA 片段化的组合 (Muller 等人, *Nucleic Acids Res.* 33:e117(2005))。

[0107] 更多的方法包括不依赖序列同源性的蛋白质重组 (Sequence Homology-Independent Protein Recombination, SHIPREC), 其中使用接头 (linker) 来促进两个远缘基因或无亲缘关系的基因之间的融合, 并且在这两种基因之间产生许多嵌合体, 导致产生单交换杂合体文库 (Sieber 等人, *Nat. Biotechnol.* 19:456-460(2001)); 基因位点饱和诱变™ (Gene Site Saturation Mutagenesis™, GSSM™), 其中起始材料包括含有插入片段的超螺旋双链 DNA (dsDNA) 质粒和在突变的所需位点上发生简并的两种引物 (Kretz 等人, *Methods Enzymol.* 388:3-11(2004)); 组合盒式诱变 (Combinatorial Cassette Mutagenesis, CCM), 其包括使用短寡核苷酸盒以替代具有大量可能氨基酸序列改变的限制区 (Reidhaar-Olson 等人, *Methods Enzymol.* 208:564-586(1991)); 和 Reidhaar-Olson 等人, *Science* 241:53-57(1988)); 组合多重盒式诱变 (Combinatorial Multiple Cassette Mutagenesis, CMCM), CMCM 基本上类似于 CCM 并以高突变率使用 epPCR 以鉴定热点和热区, 然后通过 CMCM 延伸以覆盖蛋白质序列空间的限定区 (Reetz 等人, *Angew. Chem. Int. Ed Engl.* 40:3589-3591(2001)); 致突变菌株技术 (Mutator Strains technique), 该技术中使用条件 ts 致突变质粒 (利用 mutD5 基因, 编码 DNA 聚合酶 III 的突变亚单位), 以在选择期间使随机和天然突变频率增加 20-4000 倍, 而当不需要选择时阻断缺失突变的累积 (Selifonova 等人, *Appl. Environ. Microbiol.* 67:3645-3649(2001)); Low 等人, *J. Mol. Biol.* 260:359-3680(1996))。

[0108] 另外的示例性方法包括精细诱变 (Look-through Mutagenesis, LTM), 它是一种多维诱变方法, 用于评价和最优化所选氨基酸的组合突变 (Rajpal 等人, *Proc. Natl. Acad. Sci. USA* 102:8466-8471(2005)); 基因重新装配 (Gene Reassembly), 它是一种 DNA 改组方法, 可一度应用到多基因或用来产生单个基因的嵌合体 (多点突变) 的大文库 (由 Verenum Corporation 公司供应的 TGR™ (Tunable Gene Reassembly™) 技术); 计算机模拟蛋白质设计自动化 (in Silico Protein Design Automation, PDA), 它是一

种最优化算法,可锚定具有特殊折叠的结构确定的蛋白质骨架,并检索可以稳定折叠和总体蛋白质动力学用于氨基酸取代的序列空间,而且一般来说对具有已知三维结构的蛋白质最有效地工作 (Hayes 等人, Proc. Natl. Acad. Sci. USA 99:15926-15931 (2002)); 和迭代饱和诱变 (Iterative Saturation Mutagenesis, ISM), ISM 包括利用结构/功能知识来挑选用于酶改进的最可能位点,使用诱变方法例如 Stratagene QuikChange (Stratagene; San Diego CA) 在所选位点上进行饱和诱变,再根据所需特性进行筛选/选择,和使用改进的克隆,在另一位点上重新开始并继续重复直到达到所需要的活性 (Reetz 等人, Nat. Protoc. 2:891-903 (2007); 和 Reetz 等人, Angew. Chem. Int. Ed Engl. 45:7745-7751 (2006))。

[0109] 任何上述用于诱变的方法可以单独或者以任何组合使用。另外,定向进化方法的任何一个或组合可以结合本文所述的适应性进化技术使用。

[0110] 应当理解,基本上不影响本发明各种实施方案的活性的修改也在本文提供的本发明定义内提供。因此,以下实施例意在说明而不是限制本发明。

[0111] 实施例 I

[0112] 用于生产 (2-羟基-3-甲基-4-氧代丁氧基) 磷酸酯的示例性途径

[0113] 本实施例描述了一种用于生产对苯二甲酸 (PTA) 前体 (2-羟基-3-甲基-4-氧代丁氧基) 磷酸酯 (2H3M4OP) 的示例性途径。

[0114] 对甲苯甲酸和 PTA 途径的前体是 2H3M4OP。这种化学物质可衍生自中心代谢物 3-磷酸甘油醛 (G3P) 和丙酮酸 (在三个酶促步骤中), 如图 1 所示。前两个步骤对于大肠杆菌和利用甲基赤藓糖醇磷酸 (非甲羟戊酸) 途径进行类异戊二烯 (isoprenoid) 生物合成的其它生物体来说是天然的。丙酮酸和 G3P 首先通过 DXP 合酶缩合生成 1-脱氧-D-木酮糖 5-磷酸 (DXP)。接下来,碳骨架的还原和重排由 DXP 还原异构酶催化。最后,一种新的二醇脱水酶将 2-C-甲基-D-赤藓糖醇-4-磷酸转化成对甲苯甲酸前体 2H3M4OP。

[0115] A. 1-脱氧木酮糖-5-磷酸 (DXP) 合酶。丙酮酸和 G3P 通过 DXP 合酶 (EC2. 2. 1. 7) 缩合生成 DXP。该酶催化类异戊二烯生物合成的非甲羟戊酸途径中的第一步。该酶需要二磷酸硫胺素作为辅因子,并且还需要还原型 FAD, 尽管没有净的氧化还原变化。大肠杆菌酶的晶体结构是可获得的 (Xiang 等人, J. Biol. Chem. 282:2676-2682 (2007) (doi:M610235200, pii;10.1074/jbc.M610235200doi)。其它的酶已在结核分枝杆菌 (*M. tuberculosis*) (Bailey 等人, Glycobiology 12:813-820 (2002) 和根癌土壤杆菌 (*Agrobacterium tumefaciens*) (Lee 等人, J. Biotechnol. 128:555-566 (2007) (doi:S0168-1656(06)00966-7, pii;10.1016/j.jbiotec.2006.11.009, doi) 中进行了克隆和表征。来自枯草芽孢杆菌 (*B. subtilis*) 和集胞藻 (*Synechocystis* sp.) PCC6803 的 DXP 合酶已被克隆到大肠杆菌中 (Harker 和 Bramley, FEBS Lett. 448:115-119 (1999) (doi:S0014-5793(99)00360-9, pii)。

[0116]

基因	GenBank 登录号	GI No.	生物体
dxs	AAC73523.1	1786622	大肠杆菌

dxs	POA554.1	61222979	结核分枝杆菌
dxs11	AAP56243.1	37903541	根癌土壤杆菌
dxs	P54523.1	1731052	枯草芽孢杆菌
s111945	BAA17089.1	1652165	集胞藻 PCC 6803

[0117]

[0118] B.1-脱氧-D-木酮糖-5-磷酸还原异构酶 (EC 1.1.1.267)。在类异戊二烯生物合成的非甲羟戊酸途径的第二步中,1-脱氧-D-木酮糖-5-磷酸(DXP)经依赖NAD(P)H的还原和重排为2-C-甲基-D-赤藓糖醇-4-磷酸由DXP还原异构酶(DXR, EC 1.1.1.267)催化。依赖NADPH的大肠杆菌酶由dxr编码(Takahashi等人, Proc. Natl. Acad. Sci. USA 95:9879-9884(1998))。来自拟南芥(*Arabidopsis thaliana*)的重组酶在大肠杆菌中功能性表达(Carretero-Paulet等人, Plant Physiol. 129:1581-1591(2002) (doi:10.1104/pp.003798(doi))。来自运动发酵单胞菌和结核分枝杆菌的DXR酶已被表征并且其晶体结构是可获得的(Grolle等人, FEMS Microbiol. Lett. 191:131-137(2000) (doi:S0378-1097(00)00382-7, pii); Henriksson等人, Acta Crystallogr. D. Biol. Crystallogr. 62:807-813(2006) (doi:S0907444906019196, pii;10.1107/S0907444906019196, doi))。最充分表征的DXR酶是严格NADPH依赖性的,但是来自拟南芥(*A. thaliana*)和结核分枝杆菌的酶与NADH以降低的速率起反应(Argyrou和Blanchard, Biochemistry 43:4375-4384(2004) (doi:10.1021/bi049974k, doi); Rohdich等人, FEBS J. 273:4446-4458(2006) (doi:EJB5446, pii;10.1111/j.1742-4658.2006.05446.x, doi))。

[0119]

基因	GenBank 登录号	GI No.	生物体
dxr	AAC73284.1	1786369	大肠杆菌
dxr	AAF73140.1	8131928	拟南芥 ( <i>Arabidopsis thaliana</i> )
dxr	CAB60758.1	6434139	运动发酵单胞菌
dxr	NP_217386.2	57117032	结核分枝杆菌

[0120] C.2-C-甲基-D-赤藓糖醇-4-磷酸脱水酶。需要二醇脱水酶来将2-C-甲基-D-赤藓糖醇-4-磷酸转变成对甲苯甲酸前体(Altmiller和Wagner, Arch. Biochem. Biophys. 138:160-170(1970))。虽然该转化一直没有实验证明,但是有几种酶能催化类似的转化,包括二羟基酸脱水酶(EC 4.2.1.9)、丙二醇脱水酶(EC 4.2.1.28)、甘油脱水酶(EC 4.2.1.30)和肌型-肌糖脱水酶(EC4.2.1.44)。

[0121] 能够将第二个二醇即2,3-丁二醇转变成2-丁酮的二醇脱水酶或丙二醇脱水酶(EC 4.2.1.28)是用于该转化的优异候选物。依赖腺苷酰钴胺素的二醇脱水酶含

有  $\alpha$ 、 $\beta$ 、 $\gamma$  亚基,它们都是酶功能所需要的。示例性的候选基因能在肺炎克雷伯氏菌 (*Klebsiella pneumoniae*) (Tobimatsu 等人, *Biosci. Biotechnol. Biochem.* 62:1774-1777 (1998); Toraya 等人, *Biochem. Biophys. Res. Commun.* 69:475-480 (1976))、鼠伤寒沙门氏菌 (*Salmonella typhimurium*) (Bobik 等人, *J. Bacteriol.* 179:6633-6639 (1997))、产酸克雷伯氏菌 (*Klebsiella oxytoca*) (Tobimatsu 等人, *J. Biol. Chem.* 270:7142-7148 (1995)) 和丘状菌落乳杆菌 (*Lactobacillus collinoides*) (Sauvageot 等人, *FEMS Microbiol. Lett.* 209:69-74 (2002)) 中发现。用于在其它生物体中分离出候选二醇脱水酶基因的方法是本领域众所周知的 (参见例如美国专利第 5,686,276 号)。

[0122]

基因	GenBank 登录号	GI No.	生物体
pddA	BAA08099.1	868006	产酸克雷伯氏菌
pddB	BAA08100.1	868007	产酸克雷伯氏菌
pddC	BAA08101.1	868008	产酸克雷伯氏菌
pduC	AAB84102.1	2587029	鼠伤寒沙门氏菌
pduD	AAB84103.1	2587030	鼠伤寒沙门氏菌
pduE	AAB84104.1	2587031	鼠伤寒沙门氏菌
pduC	CAC82541.1	18857678	丘状菌落乳杆菌
pduD	CAC82542.1	18857679	丘状菌落乳杆菌
pduE	CAD01091.1	18857680	丘状菌落乳杆菌
pddA	AAC98384.1	4063702	肺炎克雷伯氏菌
pddB	AAC98385.1	4063703	肺炎克雷伯氏菌
pddC	AAC98386.1	4063704	肺炎克雷伯氏菌

[0123] 甘油脱水酶家族中的酶 (EC 4.2.1.30) 也可用于催化 2-C-甲基-D-赤藓糖醇-4-磷酸脱水。示例性的候选基因在肺炎克雷伯氏菌 (*Klebsiella pneumoniae*) 中由 *gldABC* 和 *dhaB123* 编码 (WO 2008/137403) 和 (Toraya 等人, *Biochem. Biophys. Res. Commun.* 69:475-480 (1976))、在巴氏梭菌 (*Clostridium pasteurianum*) 中由 *dhaBCE* 编码 (Macis 等人, *FEMS Microbiol. Lett.* 164:21-28 (1998)) 和在弗氏柠檬酸杆菌 (*Citrobacter freundii*) 中由 *dhaBCE* 编码 (Seyfried 等人, *J. Bacteriol.* 178:5793-5796 (1996))。最近,来自肺炎克雷伯氏菌 (*K. pneumoniae*) 并具有活性增强 80-336 倍的依赖 B12 的二醇脱水酶的变异体通过在  $\beta$  亚基的两个残基中引入突变而进行了工程改造 (Qi 等人, *J. Biotechnol.* 144:43-50 (2009) (doi:S0168-1656(09)00258-2, pii;10.1016/

j. jbiotec. 2009. 06. 015, doi)。杜邦公司 (DuPont) 用易错 PCR (error-prone PCR) 开发出具有失活动力学降低的二醇脱水酶 (WO 2004/056963)。

[0124]

基因	GenBank 登录号	GI No.	生物体
gldA	AAB96343. 1	1778022	肺炎克雷伯氏菌
gldB	AAB96344. 1	1778023	肺炎克雷伯氏菌
gldC	AAB96345. 1	1778024	肺炎克雷伯氏菌
dhaB1	ABR78884. 1	150956854	肺炎克雷伯氏菌
dhaB2	ABR78883. 1	150956853	肺炎克雷伯氏菌
dhaB3	ABR78882. 1	150956852	肺炎克雷伯氏菌
dhaB	AAC27922. 1	3360389	巴氏梭菌
dhaC	AAC27923. 1	3360390	巴氏梭菌
dhaE	AAC27924. 1	3360391	巴氏梭菌
dhaB	P45514. 1	1169287	弗氏柠檬酸杆菌
dhaC	AAB48851. 1	1229154	弗氏柠檬酸杆菌
dhaE	AAB48852. 1	1229155	弗氏柠檬酸杆菌

[0125] 如果利用依赖 B12 的二醇脱水酶,那么建议相应再激活因子的异源表达。依赖 B12 的二醇脱水酶经受被底物和一些下游产物的基于机制的自杀活化。在依赖 ATP 的过程中,由与无活性钴胺素紧密缔合引起的失活,可以被二醇脱水酶再激活因子部分地克服。B12 辅因子的再生需要额外的 ATP。二醇脱水酶再生因子是双亚基蛋白质。示例性的候选物能在产酸克雷伯氏菌 (Mori 等人, J. Biol. Chem. 272:32034-32041(1997))、鼠伤寒沙门氏菌 (Bobik 等人, J. Bacteriol. 179:6633-6639(1997);Chen 等人, J. Bacteriol. 176:5474-5482(1994))、丘状菌落乳杆菌 (Sauvageot 等人, FEMS Microbiol. Lett. 209:69-74(2002)) 和肺炎克雷伯氏菌 (WO 2008/137403) 中发现。

[0126]

基因	GenBank 登录号	GI No.	生物体
ddrA	AAC15871. 1	3115376	产酸克雷伯氏菌
ddrB	AAC15872. 1	3115377	产酸克雷伯氏菌

pduG	AAL20947.1	16420573	鼠伤寒沙门氏菌
pduH	AAL20948.1	16420574	鼠伤寒沙门氏菌
pduG	YP_002236779	206579698	肺炎克雷伯氏菌
pduH	YP_002236778	206579863	肺炎克雷伯氏菌
pduG	CAD01092	29335724	丘状菌落乳杆菌
pduH	CAD01093	29335725	丘状菌落乳杆菌

[0127] 不依赖 B12 的二醇脱水酶利用 S-腺苷酰甲硫氨酸 (SAM) 作为辅因子, 在严格厌氧条件下起作用, 并且需要被特异性激活酶活化 (Frey 等人, Chem. Rev. 103:2129-2148(2003))。丁酸梭菌中由 dhaB1 和 dhaB2 编码的甘油脱氢酶及其相应的激活因子已被充分表征 (O' Brien 等人, Biochemistry 43:4635-4645(2004); Raynaud 等人, Proc. Natl. Acad. Sci. USA 100:5010-5015(2003))。最近, 该酶已应用于大肠杆菌的 1,3-丙二醇过量产生菌株并且能够达到非常高的产物效价 (Tang 等人, Appl. Environ. Microbiol. 75:1628-1634(2009) (doi:AEM.02376-08, pii;10.1128/AEM.02376-08, doi))。来自 *Roseburia inulinivorans* 的另外不依赖 B12 的二醇脱水酶和激活因子显示能催化将 2,3-丁二醇转变成 2-丁酮 (美国专利公布 2009/09155870)。

[0128]

基因	GenBank 登录号	GI No.	生物体
dhaB1	AAM54728.1	27461255	丁酸梭菌
dhaB2	AAM54729.1	27461256	丁酸梭菌
rdhtA	ABC25539.1	83596382	<i>Roseburia inulinivorans</i>
rdhtB	ABC25540.1	83596383	<i>Roseburia inulinivorans</i>

[0129] 二羟基酸脱水酶 (DHAD, EC 4.2.1.9) 是一种参与支链氨基酸生物合成的不依赖 B12 的酶。在其原始作用中, 它将 2,3-二羟基-3-甲基戊酸转变成 2-酮-3-甲基-戊酸 (异亮氨酸的前体)。在缬氨酸生物合成中, 该酶催化将 2,3-二羟基-异戊酸脱水生成 2-酮异戊酸。来自硫磺矿硫化叶菌 (*Sulfolobus solfataricus*) 的 DHAD 的底物范围拓宽, 并且在大肠杆菌中表达的重组酶活性在多种醛糖酸上得到证实 (Kim 和 Lee, J. Biochem. 139:591-596(2006) (doi:139/3/591, pii;10.1093/jb/mvj057, doi))。同许多二醇脱水酶不一样, 硫磺矿硫化叶菌 (*S. solfataricus*) 酶能耐受氧。由 *ilvD* 编码的大肠杆菌酶对氧敏感, 氧能使其铁-硫聚簇失活 (Flint 等人, J. Biol. Chem. 268:14732-14742(1993))。同样的酶已在粗糙链孢霉 (Altmeppen 和 Wagner, Arch. Biochem. Biophys. 138:160-170(1970)) 和鼠伤寒沙门氏菌 (Armstrong 等人, Biochim. Biophys. Acta 498:282-293(1977)) 中进行了表征。

[0130]

基因	GenBank 登录号	GI No.	生物体
ilvD	NP_344419.1	15899814	硫磺矿硫化叶菌
ilvD	AAT48208.1	48994964	大肠杆菌
ilvD	NP_462795.1	16767180	鼠伤寒沙门氏菌
ilvD	XP_958280.1	85090149	粗糙链孢霉

[0131] 二醇脱水酶肌型-肌糖 (myo-inosose)-2-脱水酶 (EC 4.2.1.44) 是另一个示例性的候选酶。肌型-肌糖是一种含有相邻醇基团的六元环。编码肌型-肌糖-2-脱水酶功能性的纯化酶在产气克雷伯氏菌 (*Klebsiella aerogenes*) 中在肌型-肌醇降解方面进行了研究 (Berman 和 Magasanik, *J. Biol. Chem.* 241:800-806 (1966)), 但是迄今为止还没有与某个基因联系在一起。费氏中华根瘤菌 (*Sinorhizobium fredii*) 的肌型-肌糖-2-脱水酶已在大肠杆菌中进行了克隆和功能表达 (Yoshida 等人, *Biosci. Biotechnol. Biochem.* 70:2957-2964 (2006) (doi:JST. JSTAGE/bbb/60362, pii)。同样, 也研究了来自枯草芽孢杆菌并由 *ioIE* 编码的酶 (Yoshida 等人, *Microbiology* 150:571-580 (2004))。

[0132]

基因	GenBank 登录号	GI No.	生物体
<i>ioIE</i>	P42416.1	1176989	枯草芽孢杆菌
<i>ioIE</i>	AAX24114.1	60549621	费氏中华根瘤菌

[0133] 实施例 II

[0134] 通过莽草酸途径酶从 (2-羟基-3-甲基-4-氧代丁氧基) 磷酸酯合成对甲苯甲酸的示例性途径

[0135] 本实施例描述了使用莽草酸途径酶合成对甲苯甲酸的示例性途径。

[0136] 对甲苯甲酸的化学结构非常类似于对羟基苯甲酸酯 (电子载体泛醌的前体)。4-羟基苯甲酸酯通过在细菌、植物和真菌中存在的莽草酸途径中的酶从中心代谢前体合成。莽草酸途径由将 D-赤藓糖-4-磷酸 (E4P) 和磷酸烯醇丙酮酸 (PEP) 转化成分支酸的 7 个酶促步骤组成。该途径的酶包括 2-脱氢-3-脱氧磷酸庚糖酸 (DAHP) 合酶、脱氢奎尼酸 (DHQ) 合酶、DHQ 脱水酶、莽草酸脱氢酶、莽草酸激酶、5-烯醇丙酮酰基莽草酸-3-磷酸 (EPSP) 合酶和分支酸合酶。在该途径的第一步中, D-赤藓糖-4-磷酸和磷酸烯醇丙酮酸因加入 DAHP 合酶而生成 7-磷酸-3-脱氧-D-阿拉伯庚糖醛酸。然后, 该化合物脱磷酸、脱水并还原生成莽草酸。莽草酸通过莽草酸激酶、3-磷酸莽草酸-2-羧基乙烯基转移酶和分支酸合酶等三种酶的作用转变成分支酸。接下来, 将分支酸转变成 4-羟基苯甲酸酯由分支酸裂合酶催化。

[0137] 对甲苯甲酸的合成按类似的方式进行, 如图 2 所示。该途径源自于 PEP 和 2H3M4OP, 一种类似于 E4P 的化合物, 其中甲基替代 E4P 的 3-羟基。E4P 的羟基没有直接参与莽草酸

途径反应的化学过程,所以预期甲基取代的 2H3M40P 前体能作为替代底物起反应。定向进化或适应性进化可以用于改善对 2H3M40P 和下游衍生物作为底物的偏爱。这样的方法都是本领域众所周知的。

[0138] 用于提高通过莽草酸途径酶的流通效率的菌株工程策略在此处也是适用的。通过改变葡萄糖转运系统可以增加途径前体 PEP 的利用度 (Yi 等人, *Biotechnol. Prog.* 19:1450-1459 (2003) (doi:10.1021/bp0340584, doi)。过量产生 4-羟基苯甲酸酯的菌株经工程改造以借助 7-磷酸-3-脱氧-D-阿拉伯庚糖醛酸合酶的反馈不敏感同工酶 (isozyme) 的过量表达来提高通过莽草酸途径的流通量 (Barker 和 Frost, *Biotechnol. Bioeng.* 76:376-390 (2001) (doi:10.1002/bit.10160, pii)。另外,莽草酸途径酶和分支酸裂合酶的表达水平也提高了。同样的策略可以应用于过量产生对甲苯甲酸的菌株。

[0139] A. 2-脱氢-3-脱氧磷酸庚糖酸合酶 (EC 2.5.1.54)。D-赤藓糖-4-磷酸和磷酸烯醇丙酮酸的缩合由 2-脱氢-3-脱氧磷酸庚糖酸 (DAHP) 合酶 (EC 2.5.1.54) 催化。该酶的三种同工酶在大肠杆菌基因组中分别由 *aroG*、*aroF* 和 *aroH* 编码并分别被苯丙氨酸、酪氨酸和色氨酸反馈抑制。在基本培养基上生长的野生型细胞中, *aroG*、*aroF* 和 *aroH* 基因产物分别贡献 DAHP 合酶活性的 80%、20% 和 1% (Hudson 和 Davidson, *J. Mol. Biol.* 180:1023-1051 (1984) (doi:0022-2836 (84)90269-9, pii)。发现 AroG 的两个残基能减轻被苯丙氨酸抑制 (Kikuchi 等人, *Appl. Environ. Microbiol.* 63:761-762 (1997))。通过单碱基对变化去除 AroF 被酪氨酸的反馈抑制 (Weaver 和 Herrmann, *J. Bacteriol.* 172:6581-6584 (1990))。酪氨酸不敏感的 DAHP 合酶在大肠杆菌的 4-羟基苯甲酸酯过量产生菌株中过量表达 (Barker 和 Frost, *Biotechnol. Bioeng.* 76:376-390 (2001) (doi:10.1002/bit.10160, pii)。*aroG* 基因产物显示接受多种多样的替代 4 和 5 个碳长度的底物 (Sheflyan 等人, *J. Am. Chem. Soc.* 120(43):11027-11032 (1998); Williamson 等人, *Bioorg. Med. Chem. Lett.* 15:2339-2342 (2005) (doi:S0960-894X(05)00273-8, pii; 10.1016/j.bmcl.2005.02.080, doi)。该酶与 (3S)-2-脱氧赤藓糖-4-磷酸 (一种类似于 D-赤藓糖-4-磷酸但在 2 位上缺少羟基的底物) 有效地起反应 (Williamson 等人, (同上) 2005)。来自幽门螺杆菌 (*Helicobacter pylori*) 和激烈热球菌 (*pyrococcus furiosus*) 的酶也接受这种替代底物 (Schofield 等人, *Biochemistry* 44:11950-11962 (2005) (doi:10.1021/bi050577z, doi; Webby 等人, *Biochem. J.* 390:223-230 (2005) (doi:BJ20050259, pii; 10.1042/BJ20050259, doi) 并且已经在大肠杆菌中表达。DAHP 合酶的一种进化变体,它与野生型大肠杆菌 AroG 酶有 7 个氨基酸不相同,显示在  $K_{cat}/K_M$  中表现出有 60 倍改进 (Ran 和 Frost, *J. Am. Chem. Soc.* 129:6130-6139 (2007) (doi:10.1021/ja067330p, doi)。

[0140]

基因	GenBank 登录号	GI No.	生物体
<i>aroG</i>	AAC73841.1	1786969	大肠杆菌
<i>aroF</i>	AAC75650.1	1788953	大肠杆菌

aroH	AAC74774.1	1787996	大肠杆菌
aroF	Q9ZMU5	81555637	幽门螺杆菌
PF1690	NP_579419.1	18978062	激烈热球菌

[0141]

[0142] B. 3-脱氢奎尼酸合酶 (EC 4.2.3.4)。图 2 中所示的底物 (2) (2,4-二羟基-5-甲基-6-[(磷酰氧基)甲基]噁烷-2-羧酸酯) 脱磷酸生成底物 (3) (1,3-二羟基-4-甲基环己-1-烯-1-羧酸酯) 类似于 7-磷酸-3-脱氧-阿拉伯糖基-庚酮糖酸被 3-脱氢奎尼酸合酶脱磷酸。该酶已在大肠杆菌 (Mehdi 等人, *Methods Enzymol.* 142:306-314(1987)、枯草芽孢杆菌 (Hasan 和 Nester, *J. Biol. Chem.* 253:4999-5004(1978)) 和结核分枝杆菌 H37Rv (de Mendonca 等人, *J. Bacteriol.* 189:6246-6252(2007) (doi:JB.00425-07, pii;10.1128/JB.00425-07, doi) 中进行了表征。该大肠杆菌酶经受住被 L-酪氨酸抑制 (Barker 和 Frost, *Biotechnol. Bioeng.* 76:376-390 2001) (doi:10.1002/bit.10160, pii)。

[0143]

基因	GenBank 登录号	GI No.	生物体
aroB	AAC76414.1	1789791	大肠杆菌
aroB	NP_390151.1	16079327	枯草芽孢杆菌
aroB	CAB06200.1	1781064	结核分枝杆菌

[0144] C. 3-脱氢奎尼酸脱水酶 (EC 4.2.1.10)。3-脱氢奎尼酸脱水酶, 亦称 3-脱氢奎尼酸酶 (quinase) (DHQase), 天然催化将 3-脱氢奎尼酸脱水生成 3-脱氢莽草酸, 类似于图 2 的对甲苯甲酸途径中的步骤 C。DHQase 酶可以基于机制、立体化学和序列同源性分成两大类 (Gourley 等人, *Nat. Struct. Biol.* 6:521-525. (1999) (doi:10.1038/9287, doi)。一般来说, 1 型酶参与生物合成, 而 2 型酶以反 (降解) 方向发挥作用。来自大肠杆菌 (Kinghorn 等人, *Gene* 14:73-80. 1981) (doi:0378-1119(81)90149-9, pii)、伤寒沙门氏菌 (Kinghorn 等人, (同上) 1981; Servos 等人, *J. Gen. Microbiol.* 137:147-152(1991)) 和枯草芽孢杆菌 (Warburg 等人, *Gene* 32:57-66(1984) (doi:0378-1119(84)90032-5, pii) 的 1 型酶已被克隆和表征。示例性的 II 型 3-脱氢奎尼酸脱水酶在结核分枝杆菌、天蓝色链霉菌 (*Streptomyces coelicolor*) (Evans 等人, *FEBS Lett.* 530:24-30(2002)) 和幽门螺杆菌 (Lee 等人, *Proteins* 51:616-7(2003)) 中被发现。

[0145]

基因	GenBank 登录号	GI No.	生物体
aroD	AAC74763.1	1787984	大肠杆菌
aroD	P24670.2	17433709	伤寒沙门氏菌

aroC	NP_390189.1	16079365	枯草芽孢杆菌
aroD	P0A4Z6.2	61219243	结核分枝杆菌
aroQ	P15474.3	8039781	天蓝色链霉菌
aroQ	Q48255.2	2492957	幽门螺杆菌

[0146] D. 莽草酸脱氢酶 (EC 1.1.1.25)。莽草酸脱氢酶催化 3-脱氢莽草酸的 NAD(P)H 依赖性还原生成莽草酸,类似于图 2 的步骤 D。大肠杆菌基因组编码具有不同辅因子特异性的两种莽草酸脱氢酶旁系同源物 (paralogs)。由 aroE 编码的酶是 NADPH 特异性的,而 ydiB 基因产物是奎尼酸 / 莽草酸脱氢酶,它可利用 NADH(优先的)或 NADPH 作为辅因子 (Michel 等人, J. Biol. Chem. 278:19463-19472(2003) (doi:10.1074/jbc.M300794200, doi;M300794200, pii)。来自结核分枝杆菌 (Zhang 等人, J. Biochem. Mol. Biol. 38:624-631(2005))、流感嗜血菌 (Ye 等人, J. Bacteriol. 185:4144-4151(2003)) 和幽门螺杆菌 (Han 等人, FEBS J. 273:4682-4692(2006) (doi:EJB5469, pii;10.1111/j.1742-4658.2006.05469.x, doi) 的 NADPH 依赖性酶已在大肠杆菌中进行了功能性表达。

[0147]

基因	GenBank 登录号	GI No.	生物体
aroE	AAC76306.1	1789675	大肠杆菌
ydiB	AAC74762.1	1787983	大肠杆菌
aroE	NP_217068.1	15609689	结核分枝杆菌
aroE	P43876.1	1168510	流感嗜血菌
aroE	AAW22052.1	56684731	幽门螺杆菌

[0148] E. 莽草酸激酶 (EC 2.7.1.71)。莽草酸激酶催化莽草酸 3-羟基的 ATP 依赖性磷酸化,类似于图 2 的步骤 E。两种莽草酸激酶在大肠杆菌中由 aroK(SK1) 和 aroL(SK2) 编码 (DeFeyter 和 Pittard, J. Bacteriol. 165:331-333(1986); Lobner-Olesen 和 Marinus, J. Bacteriol. 174:525-529(1992))。由 aroL 编码的 SK2 的 Km 比 SK1 的 Km 低 100 倍,表明该酶负责芳族化合物的生物合成 (DeFeyter 等人, (同上)1986)。来自结核分枝杆菌 (Gu 等人, J. Mol. Biol. 319:779-789(2002) (doi:10.1016/S0022-2836(02)00339-X, doi;S0022-2836(02)00339-X, pii);oliveira 等人, Protein Expr. Purif. 22:430-435(2001) (doi:10.1006/prev.2001.1457, doi;S1046-5928(01)91457-3, pii)、幽门螺杆菌 (Cheng 等人, J. Bacteriol. 187:8156-8163(2005) (doi:187/23/8156, pii;10.1128/JB.187.23.8156-8163.2005, doi) 和菊欧文氏菌 (Krell 等人, Protein Sci. 10:1137-1149(2001) (doi:10.1110/ps.52501, doi) 的其它莽草酸激酶已在大肠杆菌中进行了克隆。

[0149]

基因	GenBank 登录号	GI No.	生物体
aroK	YP_026215.2	90111581	大肠杆菌
aroL	NP_414922.1	16128373	大肠杆菌
aroK	CAB06199.1	1781063	结核分枝杆菌
aroK	NP_206956.1	15644786	幽门螺杆菌
SK	CAA32883.1	42966	菊欧文氏菌

[0150] F. 3-磷酸莽草酸-2-羧基乙烯基转移酶 (EC 2.5.1.19)。3-磷酸莽草酸-2-羧基乙烯基转移酶,亦称 5-烯醇丙酮酰基莽草酸-3-磷酸合酶 (EPSPS),催化将磷酸烯醇丙酮酸的烯醇丙酮酰基部分转移到莽草酸-3-磷酸的 5-羟基上。该酶在大肠杆菌中由 aroA 编码 (Anderson 等人, *Biochemistry* 27:1604-1610(1988))。来自结核分枝杆菌 (Oliveira 等人, *Protein Expr. Purif.* 22:430-435(2001) (doi:10.1006/prev.2001.1457, doi;S1046-5928(01)91457-3, pii)、盐生杜氏藻 (*Dunaliella salina*) (Yi 等人, *J. Microbiol.* 45:153-157(2007) (doi:2519, pii) 和金黄色葡萄球菌 (Priestman 等人, *FEBS Lett.* 579:728-732(2005) (doi:S0014-5793(05)00012-8, pii;10.1016/j.febslet.2004.12.057, doi) 的 EPSPS 酶已在大肠杆菌中进行了克隆和功能性表达。

[0151]

基因	GenBank 登录号	GI No.	生物体
aroA	AAC73994.1	1787137	大肠杆菌
aroA	AAA25356.1	149928	结核分枝杆菌
aroA	AAA71897.1	152956	金黄色葡萄球菌
aroA	ABM68632.1	122937807	盐生杜氏藻

[0152] G. 分支酸合酶 (EC 4.2.3.5)。分支酸合酶是莽草酸途径中的第七个酶,催化将 5-烯醇丙酮酰基莽草酸-3-磷酸转化成分支酸。该酶需要还原型黄素单核苷酸 (FMN) 作为辅因子,尽管该酶的净反应没有包括氧化还原变化。与植物和细菌中发现的酶相反,真菌中的分支酸合酶也能够在损失 NADPH 的情况下将 FMN 还原 (Macheroux 等人, *Planta* 207:325-334(1999))。代表性的单功能酶由大肠杆菌 (White 等人, *Biochem. J.* 251:313-322(1988)) 和肺炎链球菌 (Maclean 和 Ali, *Structure* 11:1499-1511(2003) (doi:S0969212603002648, pii) 的 aroC 编码。双功能的真菌酶可在粗糙链孢霉 (Kitzing 等人, *J. Biol. Chem.* 276:42658-42666(2001) (doi:10.1074/jbc.M107249200, doi;M107249200, pii) 和酿酒酵母 (Jones 等人, *Mol. Microbiol.* 5:2143-2152(1991)) 中发现。

[0153]

基因	GenBank 登录号	GI No.	生物体
aroC	NP_416832.1	16130264	大肠杆菌
aroC	ACH47980.1	197205483	肺炎链球菌
U25818.1:19..1317	AAC49056.1	976375	粗糙链孢霉
AR02	CAA42745.1	3387	酿酒酵母

[0154] H. 分支酸裂合酶 (EC 4.1.3.40)。分支酸裂合酶催化泛醌生物合成中的第一个关键步骤：从分支酸中去除丙酮酸生成 4-羟基苯甲酸酯。该酶促反应因 4-羟基苯甲酸酯产物的缓慢释放而被限速 (Gallagher 等人, *Proteins* 44:304-311(2001) (doi:10.1002/prot.1095, pii), 据此认为它在将 4-羟基苯甲酸酯递送到下游膜结合酶中起作用。大肠杆菌的分支酸裂合酶已被克隆和表征并且该酶已经结晶出来 (Gallagher 等人, (同上)2001; Siebert 等人, *FEBS Lett.* 307:347-350(1992) (doi:0014-5793(92)80710-X, pii)。结构研究表明 G90 残基能促进产物抑制 (Smith 等人, *Arch. Biochem. Biophys.* 445:72-80(2006) (doi:S0003-9861(05)00446-7, pii; 10.1016/j.abb.2005.10.026, doi)。两种表面活性胱氨酸残基的修饰能降低蛋白质聚集 (Holden 等人, *Biochim. Biophys. Acta* 1594:160-167(2002) (doi:S0167483801003028, pii)。结核分枝杆菌分支酸裂合酶的一种重组形式已在大肠杆菌中进行了克隆和表征 (Stadthagen 等人, *J. Biol. Chem.* 280:40699-40706(2005) (doi:M508332200, pii; 10.1074/jbc.M508332200, doi)。

[0155]

基因	GenBank 登录号	GI No.	生物体
ubiC	AAC77009.2	87082361	大肠杆菌
Rv2949c	NP_217465.1	15610086	结核分枝杆菌

[0156] B-F. 多功能的 AROM 蛋白质。在大多数细菌中,莽草酸途径的酶都由独立的多肽编码。在真核微生物中,五个酶促功能由五功能超基因所编码的多功能蛋白质催化 (Campbell 等人, *Int. J. Parasitol.* 34:5-13(2004) (doi:S0020751903003102, pii)。多功能 AROM 蛋白质复合体催化的反应类似于图 2 的反应 B-F。AROM 蛋白质复合体已在真菌包括构巢曲霉 (*Aspergillus nidulans*)、粗糙链孢霉 (*Neurospora crassa*)、酿酒酵母 (*Saccharomyces cerevisiae*) 和卡氏肺囊虫 (*Pneumocystis carinii*) 中进行了表征 (Banerji 等人, *J. Gen. Microbiol.* 139:2901-2914(1993); Charles 等人, *Nucleic Acids Res.* 14:2201-2213(1986); Coggins 等人, *Methods Enzymol.* 142:325-341(1987); Duncan, K., *Biochem. J.* 246:375-386(1987))。AROM 的若干组分已显示作为各个多肽独立地起作用。例如,脱氢奎尼酸合酶 (DHQS) 构成 AROM 的氨基端结构域,并且当克隆到大肠杆菌时可以独立地起作用 (Moore 等人, *Biochem. J.* 301(Pt 1):297-304(1994))。来自构巢曲霉的 AROM 组分的若干晶体结构提供了对催化机制的深刻认识 (Carpenter 等人, *Nature*

394:299-302(1998) (doi:10.1038/28431, doi)。

[0157]

基因	GenBank 登录号	GI No.	生物体
AROM	P07547.3	238054389	构巢曲霉
AROM	P08566.1	114166	酿酒酵母
AROM	P07547.3	238054389	构巢曲霉
AROM	Q12659.1	2492977	卡氏肺囊虫

[0158]

[0159] 实施例 III

[0160] 将对甲苯甲酸酶促转化成对苯二甲酸的示例性途径

[0161] 本实施例描述了将对甲苯甲酸转变成对苯二甲酸 (PTA) 的示例性途径。

[0162] 对甲苯甲酸可进一步转化成 PTA, 即通过在三个酶促步骤中将甲基氧化成酸, 如图 3 所示。该途径由对甲苯甲酸甲基-单加氧酶还原酶、4-羧基苯甲醇脱氢酶和 4-羧基苯甲醛脱氢酶组成。在第一个步骤中, 在 O<sub>2</sub> 存在下, 对甲苯甲酸甲基-单加氧酶将对甲苯甲酸氧化成 4-羧基苯甲醇。睾丸酮丛毛单胞菌 (*Comamonas testosteroni*) 酶 (tsaBM) (其也与作为底物的 4-甲苯磺酸酯反应) 已被纯化和表征 (Locher 等人, *J. Bacteriol.* 173:3741-3748(1991))。随后, 4-羧基苯甲醇被 4-羧基苯甲醇脱氢酶 (tsaC) 转变成醛。醛至酸的转化是由 4-羧基苯甲醛脱氢酶 (tsaD) 催化的。催化这些反应的酶可在睾丸酮丛毛单胞菌 T-2 (一种能够利用对甲苯甲酸作为唯一的碳源和能源的生物体) 中发现 (Junker 等人, *J. Bacteriol.* 179:919-927(1997))。将对甲苯甲酸转化成 PTA 的其它基因可以通过序列同源性找到, 特别是对于伯克氏菌属 (*Burkholderia*)、产碱菌属 (*Alcaligenes*)、假单胞菌属 (*Pseudomonas*)、鞘氨醇单胞菌属 (*Shingomonas*) 和丛毛单胞菌属 (*Comamonas*) 等菌属中的蛋白细菌 (proteobacteria) 尤其如此 (美国专利第 6, 187, 569 号和美国专利公布 2003/0170836)。下表列举了与睾丸酮丛毛单胞菌酶相关的 Genbank 标志物 (Genbank identifiers)。

[0163]

基因	GenBank 登录号	GI No.	生物体
tsaB	AAC44805.1	1790868	睾丸酮丛毛单胞菌
tsaM	AAC44804.1	1790867	睾丸酮丛毛单胞菌
tsaC	AAC44807.1	1790870	睾丸酮丛毛单胞菌
tsaD	AAC44808.1	1790871	睾丸酮丛毛单胞菌

[0164] 在本申请全文中, 引用了各种出版物。这些出版物在其实体中的公开内容, 包括包括发表的 GenBank 和 GI No., 特此通过引用结合到本申请中, 以便更全面地描述本发明所

属领域的现有状态。尽管本发明参照以上提供的实施例作了说明,但是应当理解,在不偏离本发明的精神的情况下,可以进行各种修改。

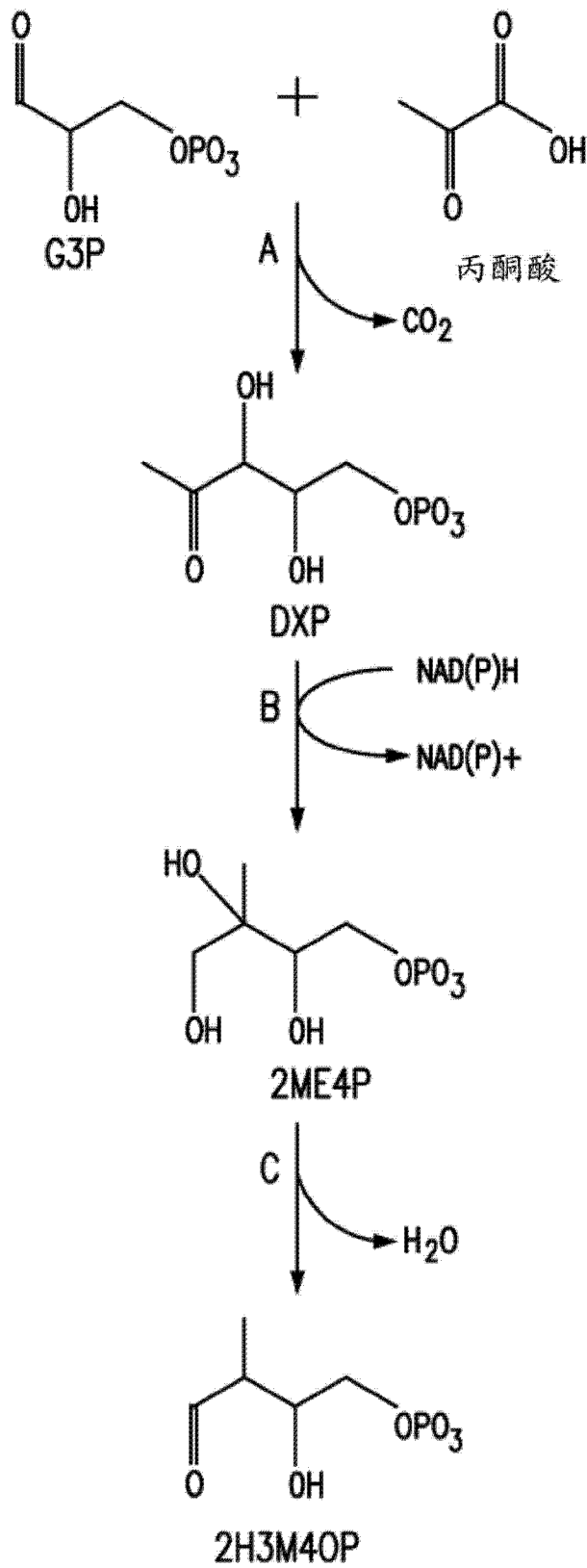


图 1

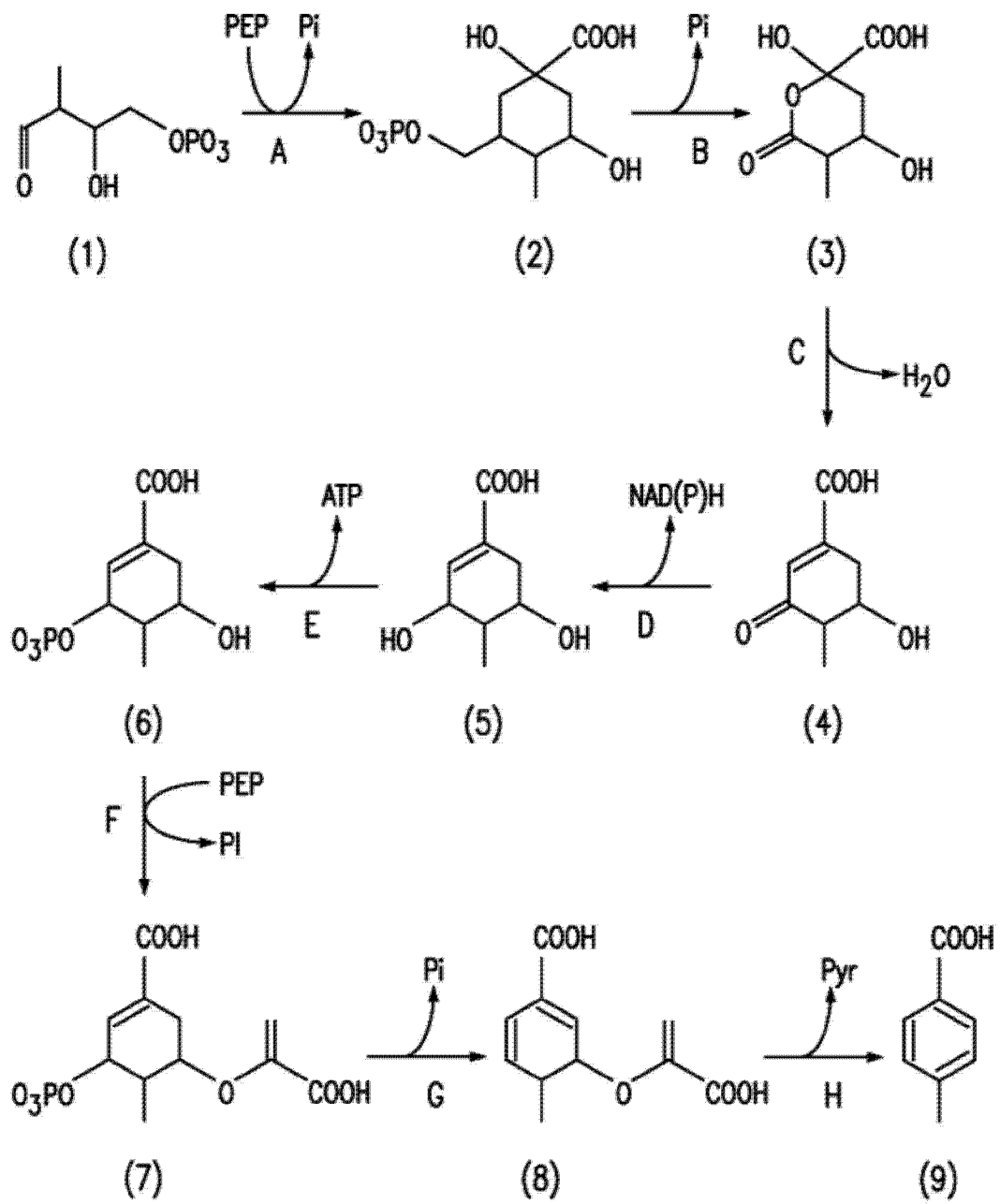


图 2

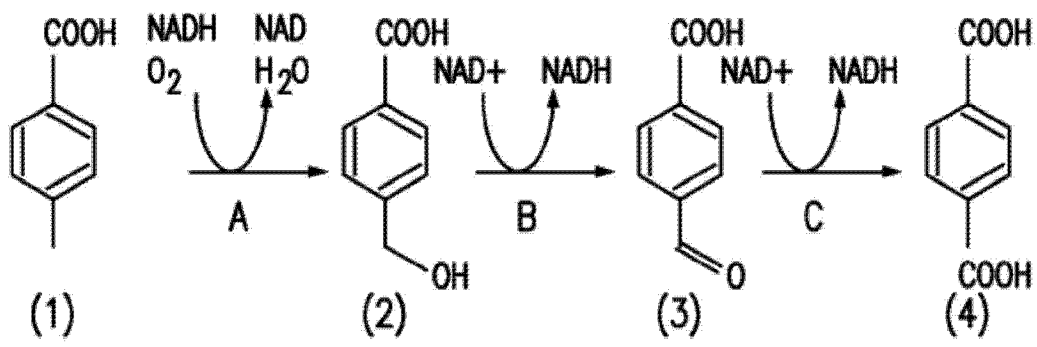


图 3