

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

C12N 15/32

C12N 15/55



[12] 发明专利申请公开说明书

[21] 申请号 02827413. X

[43] 公开日 2005 年 5 月 11 日

[11] 公开号 CN 1615363A

[22] 申请日 2002. 11. 18 [21] 申请号 02827413. X

[30] 优先权

[32] 2001. 11. 21 [33] US [31] 60/332,060

[86] 国际申请 PCT/CN2002/000821 2002. 11. 18

[87] 国际公布 WO2003/044199 英 2003. 5. 30

[85] 进入国家阶段日期 2004. 7. 21

[71] 申请人 香港大学

地址 中国香港

[72] 发明人 林文量 叶永健

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 刘 玥 孟凡宏

权利要求书 5 页 说明书 54 页 序列表 11 页
附图 23 页

[54] 发明名称 重组芽孢杆菌植酸酶及其用途

[57] 摘要

本发明克隆并分析了来自于公认安全的微生物：地衣芽孢杆菌和枯草芽孢杆菌 168 的两种植酸酶基因。建立了植酸酶超表达及纯化的方法。这些酶的分子量约为 48kDa(千道尔顿)，具有细胞外植酸酶水解活性。这些重组的酶可用于增强植酸酶在各种商业领域的应用，包括制备动物饲料和成熟生长、开花和结果速度增加的转基因植物。

I S S N 1 0 0 8 - 4 2 7 4

1. 一种分离的核酸分子，其包含编码 phyL 多肽的核苷酸序列，所述多肽具有 SEQ ID NO:2 的氨基酸序列。
- 5 2. 权利要求 1 中的分离的核酸分子，其中所述核苷酸序列为 SEQ ID NO:1。
3. 权利要求 1 中的分离的核酸分子，其中该分子为 DNA。
4. 权利要求 1 中的分离的核酸分子，其中该分子为 RNA。
5. 包含权利要求 3 中核酸分子的载体。
- 10 6. 包含权利要求 5 中载体的宿主细胞。
7. 包含权利要求 3 中核酸分子的宿主细胞，所述核酸分子与异源启动子操纵性连接。
8. 权利要求 7 中的宿主细胞，其中异源启动子为 ϕ 105 原噬菌体启动子。
- 15 9. 权利要求 6、7 或 8 的宿主细胞，它是芽孢杆菌。
10. 权利要求 9 的宿主细胞，它是枯草芽孢杆菌 MU331。
11. 一种生产 phyL 多肽的方法，包括表达由权利要求 6、7 或 8 的宿主细胞的核酸所编码的多肽，以及回收 phyL 多肽。
12. 权利要求 11 的方法，其中宿主细胞为芽孢杆菌。
- 20 13. 权利要求 12 的方法，其中宿主细胞为枯草芽孢杆菌 MU331。
14. 一种制备能表达 phyL 多肽的细胞或其子代的方法，该方法包括用权利要求 6 的载体转染所述细胞。
15. 权利要求 14 的方法，其中细胞为芽孢杆菌。
- 25 16. 权利要求 15 的方法，其中细胞为枯草芽孢杆菌 MU331。
17. 一种分离的核酸分子，其可在严格条件下与权利要求 1 或 2 中核酸分子或其互补分子杂交，其中所述核酸分子编码在中性 pH 下催化肌醇六磷酸盐转化成无机磷酸盐和肌-肌醇磷酸盐的氨基酸序列。
- 30 18. 一种分离的多肽，其由权利要求 17 中的核酸分子所编码。
19. 一种分离的 phyL 多肽，其具有 SEQ ID NO:2 的氨基酸序列

或其能催化肌醇六磷酸盐转化成无机磷酸盐和肌-肌醇磷酸盐的至少20个氨基酸的片段。

20. 包含核苷酸序列的分离的核酸分子, 其中该核苷酸序列编码具有 SEQ ID NO:4 氨基酸序列的 168phyA 多肽。

5 21. 权利要求 20 的分离的核酸分子, 其中所述核苷酸序列为 SEQ ID NO:3。

22. 权利要求 20 的分离的核酸分子, 其中所述分子为 DNA

23. 权利要求 20 的分离的核酸分子, 其中所述分子为 RNA

24. 包含权利要求 22 中核酸分子的载体。

10 25. 包含权利要求 24 中载体的宿主细胞。

26. 包含权利要求 22 中的核酸分子的宿主细胞, 所述核酸分子与异源启动子操纵性连接。

27. 权利要求 26 中的宿主细胞, 其中异源启动子为 $\phi 105$ 原噬菌体启动子。

15 28. 权利要求 25、26 或 27 的宿主细胞, 它是芽孢杆菌。

29. 权利要求 28 的宿主细胞, 它是枯草芽孢杆菌 MU331。

30. 一种生产 168phyA 多肽的方法, 包括表达由权利要求 25、26 或 27 的宿主细胞的核酸所编码的多肽, 以及回收 168phyA 多肽。

31. 权利要求 30 的方法, 其中宿主细胞为芽孢杆菌。

20 32. 权利要求 31 的方法, 其中宿主细胞为枯草芽孢杆菌 MU331。

33. 一种制备能表达 168phyA 多肽的细胞或其子代的方法, 该方法包括用权利要求 24 的载体转染所述细胞。

34. 权利要求 33 的方法, 其中细胞为芽孢杆菌。

25 35. 权利要求 34 的方法, 其中细胞为枯草芽孢杆菌 MU331。

36. 一种分离的核酸分子, 其可在严格条件下与权利要求 20 或 21 中核酸分子或其互补分子杂交, 其中所述核酸分子编码在中性 pH 下催化肌醇六磷酸盐转化成无机磷酸盐和肌-肌醇磷酸盐的氨基酸序列。

30 37. 一种分离的多肽, 其由权利要求 36 中核酸分子所编码。

38. 一种分离的 168phyA 多肽, 其具有 SEQ ID NO:4 的氨基酸

序列或其能催化肌醇六磷酸盐转化成无机磷酸盐和肌-肌醇磷酸盐的至少 20 个氨基酸的片段。

39. 包含权利要求 18 中多肽的动物饲料。
40. 包含权利要求 19 中多肽的动物饲料。
- 5 41. 包含权利要求 31 中多肽的动物饲料。
42. 包含权利要求 32 中多肽的动物饲料。
43. 包含权利要求 6、7、8、25、26 及 27 中任一项的宿主细胞的动物饲料。
44. 包含权利要求 9 中宿主细胞的动物饲料。
- 10 45. 包含权利要求 10 中宿主细胞的动物饲料。
46. 包含权利要求 28 中宿主细胞的动物饲料。
47. 包含权利要求 29 中宿主细胞的动物饲料。
48. 包含编码植酸酶的核苷酸序列的嵌合表达盒，其中所述植酸酶来自于芽孢杆菌菌株，并且所述核苷酸序列与调控性核苷酸序列操纵性连接以使所述调控性核苷酸序列能导致所述核苷酸序列在植物细胞中表达，其中调控性核苷酸序列与所述核苷酸序列异源。
- 15 49. 权利要求 48 的嵌合表达盒，其中所述植酸酶在细胞内表达。
50. 权利要求 49 的嵌合表达盒，其中所述植酸酶具有 SEQ ID NO:2 或 SEQ ID NO:4 的氨基酸序列。
- 20 51. 权利要求 49 的嵌合表达盒，其中所述植酸酶具有 SEQ ID NO:22 或 SEQ ID NO:24 的氨基酸序列。
52. 权利要求 49 的嵌合表达盒，其中所述核苷酸序列为 SEQ ID NO:1 或 SEQ ID NO:3。
53. 权利要求 49 的嵌合表达盒，其中所述核苷酸序列为 SEQ ID NO:21 或 SEQ ID NO:23。
- 25 54. 权利要求 48 的嵌合表达盒，其中所述植酸酶具有 SEQ ID NO:2 或 SEQ ID NO:4 的氨基酸序列，但是 SEQ ID NO:2 的氨基酸残基 1~80 的全部或 N-端的一部分，或 SEQ ID NO:4 的氨基酸残基 1~80 的全部或 N-端的一部分由一段植物信号肽所取代以便使所述植酸酶从植物细胞中分泌。
- 30 55. 权利要求 48 的嵌合表达盒，其中所述植酸酶具有 SEQ ID

NO:2 或 SEQ ID NO:4 的氨基酸序列, 但是 SEQ ID NO:2 中氨基酸残基 1~20 的全部或 N-端的一部分, 或 SEQ ID NO:4 中氨基酸残基 1~26 的全部或 N-端的一部分由一段植物信号肽所取代以便使所述植酸酶从植物细胞中分泌。

5 56. 权利要求 48 的嵌合表达盒, 其中所述核苷酸序列为 SEQ ID NO:1 或 SEQ ID NO:3, 但是 SEQ ID NO:1 中核苷酸 241~480 序列的全部或部分, 或是 SEQ ID NO:3 中核苷酸 100~339 序列的全部或部分由一段植物信号序列所取代以便使该植酸酶从植物细胞中分泌。

10 57. 权利要求 48 的嵌合表达盒, 其中所述核苷酸序列为 SEQ ID NO:1 或 SEQ ID NO:3, 但是 SEQ ID NO:1 中核苷酸 241~300 序列[请注意, 核苷酸 241~300, 与核苷酸 241~323 相对, 相应于 SEQ ID NO:1 的氨基酸残基 1~20]的全部或部分, 或是 SEQ ID NO:3 中核苷酸 100~177 序列的全部或部分由一段植物信号序列所取代以便使所述植酸酶从植物细胞中分泌。

15 58. 一种嵌合表达盒, 其包含编码在中性 pH 下具有催化活性的植酸酶的核苷酸序列, 其中所述核苷酸序列在严格条件下与 SEQ ID NO:1 和/或 SEQ ID NO:3 的核苷酸序列杂交, 并与调控性核苷酸序列操纵性连接, 以使所述调控性核苷酸序列能导致该核苷酸序列在植物细胞中表达, 其中调控性核苷酸序列与所述核苷酸序列异源。

20 59. 包含权利要求 48~58 中任一项的表达盒的表达载体。

60. 包含权利要求 54 中表达载体的转化植物细胞, 其中该植物细胞表达所述植酸酶。

61. 权利要求 60 的转化植物细胞, 其中所述细胞为单子叶物种。

25 62. 权利要求 61 的转化植物细胞, 其中所述单子叶物种选自玉米、高粱、小麦、棕榈及稻。

63. 权利要求 60 的转化植物细胞, 其中所述细胞为双子叶物种。

64. 权利要求 63 的转化植物细胞, 其中所述双子叶物种选自大豆、油菜、加州希蒙得木、中国脂树、烟草、番红花、花生及向日葵。

65. 包含权利要求 61 转化植物细胞的体外培养物。

30 66. 包含权利要求 63 转化植物细胞的体外培养物。

67. 一种转化植物, 其中所述植物的细胞包含权利要求 48-58

中任一项的表达盒，并表达所述植酸酶。

68. 权利要求 67 的转化植物，其中所述植物为稻。

69. 权利要求 67 的转化植物，其中所述植物为油菜。

70. 权利要求 67 的转化植物，其中所述植物为向日葵。

5 71. 权利要求 67 的转化植物，其中所述植物为番红花。

72. 权利要求 67 的转化植物，其中所述植物为花生。

73. 从植物细胞中的植物肌醇六磷酸盐转移无机磷酸盐以改善植物生长、开花和/或结果的方法，该方法包括将包含权利要求 48-58 中任一项的嵌合表达盒的核酸分子导入所述植物细胞以产生转化植物
10 细胞，由此所述转化植物细胞可表达可从植物肌醇六磷酸盐转移无机磷酸盐的所述植酸酶。

74. 权利要求 73 的的方法，还包括从转化植物细胞生成完整植株的步骤，其中所述植物包含表达所述植酸酶的细胞。

75. 权利要求 74 的的方法，还包括有性繁殖或克隆繁殖所述完整植株的步骤，其中所述完整植株的子代包含表达所述植酸酶的细胞。
15

76. 权利要求 73 的的方法，其中所述表达盒用电穿孔法导入所述植物细胞中。

77. 权利要求 73 的的方法，其中所述表达盒用微粒轰击法导入
20 所述植物细胞中。

78. 权利要求 73 的的方法，其中所述表达盒用微注射法导入所述植物细胞中。

79. 从植物肌醇六磷酸盐转移无机磷酸盐以在土壤杆菌敏感性双子叶植物中改善植物生长、开花和/或结果的方法，该方法包括用
25 含包含权利要求 48-58 中任一项的表达盒的土壤杆菌感染所述植物的植物细胞以产生转化植物细胞，由此所述感染植物细胞表达可转移无机磷酸盐的所述植酸酶。

80. 包含权利要求 67 中转化植物的动物饲料。

重组芽孢杆菌植酸酶及其用途

5 本申请对 2001 年 11 月 21 日提交、系列号为 No. 60/332,060 的美国临时申请享有并要求优先权，其整体在此引入作为参考。

1. 引言

10 本发明涉及来自于两种公认安全(GRAS)微生物 - 地衣芽孢杆菌和枯草芽孢杆菌 168 的植酸酶基因，和它们各自所编码的蛋白质产物及其片段、衍生物、类似物和变体。本发明还提供了生产和纯化植酸酶、衍生物、类似物、变体和抗体的方法。本发明还提供了这些植酸酶在动物饲料中的用途。本发明还提供了这两种植酸酶在中性 pH 时具有活性（“中性植酸酶”）的转基因植物及其他中性植酸酶。这些转
15 基因植物的生长、开花及果实的生长增加。

2. 发明背景

肌醇六磷酸盐(phytate)为植酸的盐形式（肌-肌醇 1,2,3,4,5,6-六磷酸二氢盐, (myo-inositol 1,2,3,4,5,6-hexakis dihydrogen phosphate) ），
20 占谷类和豆类中总磷的 80%以上，谷类及豆类与含油种子(oilseed)作物共占据全世界作物收获面积的 90%以上(Reddy N. R., Pierson M. D., Sathe S. k.和 Salunkhe D. K., 1989, 豆类和谷类中的植酸酶, CRC 出版社, Boca Raton, 佛罗里达)。尽管肌醇六磷酸盐是磷的储存形式，但其中的磷并不易为动植物所用，这是因为需要特异性的酶将肌醇六
25 磷酸盐水解成无机磷酸盐。

植酸酶偏向于以肌醇六磷酸盐作为底物，可通过将肌醇六磷酸盐催化生成无机磷酸盐和磷酸肌醇并释放可被动植物利用的磷酸盐，使可利用的磷增加。

30 由于植酸酶在经济和环境方面的重要性，它的超表达(over-expression)是生物技术和酶生产工业中长期的、竞争性的课题。研究者们已找到使用最少的纯化步骤获得最高活性的酶的超表达的途径。

早期对植酸酶表达的研究牵涉到从真菌来源提取和生产该酶。对于动物饲料来说，真菌来源直到现在仍是唯一的已知来源。

早在 20 世纪 80 年代，植酸酶就可在无花果曲霉/黑曲霉的细胞外培养基中表达 (Ullah A.H. 和 Cummins B.J., 1988, 无花果曲霉细胞外 pH6.0 最适酸性磷酸酶: 纯化, N-末端氨基酸序列及生化特征, *Preparative Biochemistry*, 18(1): 37-65)。Ullah 等在同一年对该酶进行了广泛研究。直到现在, 来自于黑曲霉的植酸酶仍然是最重要的商用植酸酶。在 20 世纪 90 年代, 利用新的生物技术, 在有望提高异源蛋白收率的外源菌株中表达重组蛋白, 提高了该酶的产量。据报道, 包括 *fusarium venenatum* (Berka, R.M., Rey M.W, Brown K.M., Byun T 和 Klotz A.V., 1998, 来源于嗜热真菌 *fusarium venenatum* 的植酸酶基因的分子特征及表达, *Applied and environmental microbiology*, 64(11): 4423-4427)、黑曲霉菌及其他曲霉菌 (Pasamontes L, Haiker M, Wyss M, Tessier M 和 Loon A.P.G., 1997, 一种来源于真菌 - 烟曲霉菌的热稳定植酸酶的基因克隆、纯化及鉴定, *Applied and environmental microbiology*, 63(5): 1691-1700; NO. 5,830,733、5,436,156 和 6,153,418 号美国专利)、土生克雷伯菌 (Greiner R., Haller, E., Konietzny U. 和 Jany K.D., 1997, 来源于土生克雷伯菌的植酸酶的纯化和鉴定, *Archives of Biochemistry and Biophysics*, 341(2): 201-206)、*Thermomyces* 菌 (No. 5,866,118 号美国专利) 及 *Schwanniomyces occidentalis* (No. 5,840,561 号美国专利) 在内的真菌菌株可表达大量具有明显活性的异源植酸酶。人们做了许多尝试用酶将肌醇六磷酸盐水解, 中等程度地改善了饲料的营养价值, 减少了动物排泄的磷, 从而保护环境 (Pen J. Verwoerd T.C. 和 Hoekema A., 1993, 含植酸酶的转基因种子作为改善磷利用的全新饲料添加剂, *Bio/Technology*, 11:811-814)。虽然用真菌生产植酸酶仍在继续, 但其他研究团体已将其焦点转到在酵母 (Mayer A. F., Hellmuth, K., Schlieker H., Ulibarri R.L., Oertel S., Dahlems U., Strasser A. W.M. Strasser 和 Loon A.P.G.M., 1998, 一种表达系统成熟: 一种利用多形汉森酵母重组菌株高效、低成本生产植酸酶的方法, *Biotechnology and Bioengineering*, 63(3): 373-381; Han Y., Wilson, D.B. 和 Lei X.G., 1999, 黑曲霉植酸酶基因

(*phyA*) 在酿酒酵母中的表达, *Applied and Environmental microbiology*, 65(5):1915-1918; Han Y 和 Lei X.G., 1999, 糖基化对曲霉菌植酸酶在巴氏毕赤酵母中功能性表达的作用, *Archives of Biochemistry and Biophysics*, 364(1):83-90; Rodriguez E., Mullaney E.J.和 Lei X.G., 2000, 烟曲霉植酸酶基因在巴氏毕赤酵母中的表达及该重组酶的特性, *Biochemical and Biophysical Research Communications*, 268: 373-378)、植物 (Ullah A.H.J., Sethumadhavan K., Mullaney E.J., 1999, 表达于烟叶中的重组真菌植酸酶 (*phyA*) 的特征, *Biochemical and Biophysical Research Communications*, 264: 201-206) 和肠道细菌 - 大肠杆菌 (*E. Coli*) (Dassa J., Marck C.和 Boquet P.L., 1990, 大肠杆菌 *appA* 基因的全核苷酸序列显示 pH2.5 酸性磷酸酶和葡萄糖-1-磷酸酶间具有显著的同源性, *Journal of Bacteriology*, 172(9):5497-5500; Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.M.和 Van Etten R.L., 1992, 大肠杆菌酸性磷酸酶的过度表达、定点诱变及机制, *Journal of Bacteriology*, 267(32): 22830-22836; 以及 Rodriguez E., Han Y.和 Lei X.G., 1999, 从猪结肠分离出来的大肠杆菌酸性磷酸酶/植酸酶基因 (*appA2*) 的克隆、测序及表达, *Biochemical and Biophysical Research Communications*, 257: 117-123) 中表达植酸酶。也研究了其他来源于植物 (Maugenest S., Martinez I 和 Lescure A, 1997, 一种编码玉米种子植酸酶 cDNA 的克隆和鉴定, *Biochemistry Journal*, 322: 511-517) 和哺乳动物 (Craxto A., Caffrey J.J., Burkhart W., Safrany S.T.和 Shears S.B., 1997, 大鼠肝中多个肌醇多聚磷酸酶的分子克隆和表达, *Biochemistry Journal*, 328: 75-81) 的植酸酶。

大肠杆菌和乳酸杆菌中的数种植酸酶基因, 包括 *EcAP*(Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.M.和 Van Etten R.L., 1992, 大肠杆菌酸性磷酸酶的过度表达、定点诱变及机制, *Journal of Bacteriology*, 267(32): 22830-22836)、*appA*(Dassa J., Marck C.和 Boquet P.L., 1990, 大肠杆菌 *appA* 基因的全核苷酸序列显示 pH2.5 酸性磷酸酶和葡萄糖-1-磷酸酶间具有显著的同源性, *Journal of Bacteriology*, 172(9):5497-5500)、*appA2*(Rodriguez E., Han Y.和 Lei X.G., 1999, 从猪结肠分离出来的大肠杆菌酸性磷酸酶/植酸酶基因 (*appA2*) 的克

隆、测序及表达, Biochemical and Biophysical Research Communications, 257: 117-123)及植物乳杆菌 (Zamudio 等, 2001, 植物乳杆菌植酸酶活性归功于其非特异性酸性磷酸酶, Lett. APP. Microbiol. 32: 181-184), 已识别和全部鉴定为酸性磷酸酶, 在 pH 低于 6.0 时具有最佳酶活性。NO. 6,183,740 和 6,190,897 号美国专利公开了其他来源于大肠杆菌的植酸酶。

尽管真菌和大肠杆菌的植酸酶已经可以表达达到具有实际意义的量, 但这些酶的纯化方法复杂, 另外, 这些异源性表达的酶常常不能正确折叠。例如, 已发现大肠杆菌不能表达起源于黑曲霉菌的活性植酸酶, 因为大肠杆菌产生一种非糖基化的大分子量细胞内包涵体蛋白 (Phillippy B.Q.和 Mullaney E.J., 1997, 黑曲霉菌植酸酶(*phyA*)在大肠杆菌中的表达, Journal of Agricultural Food Chemistry, 45: 3337-3342)。而且, 大肠杆菌是一种具有感染动物胃肠道危险性的肠道细菌。

已知有数种芽孢杆菌(*Bacillus*)是 GRAS 细菌菌株。已从枯草芽孢杆菌菌株 VTT E-68013 (*phyC*; Kerovuo J., Laurarus M., Nurminen P., Kalkkinen N.和 Apajalahti J., 1998, 一种来自于枯草芽孢杆菌的全新植酸酶的分离、鉴定、分子基因克隆及测序, Applied and Environmental Microbiology, 64(6):2079-2085, 将其整体在此引入作为参考)和 DS11 (*phyK*; Kim Y.O., Lee J.K., Kim H.K., Yu J.H.和 Oh T.K., 1998, 从芽孢杆菌 DS11 克隆热稳定的植酸酶基因 (*phy*) 及其大肠杆菌中的超表达, FEMS Microbiology Letters, 162:182-191;以及 NO. 6,255,098 号美国专利, 在此将其整体引入作为参考)克隆出编码植酸酶的基因。这些报道表明, 芽孢杆菌植酸酶与真菌、大肠杆菌、植物及哺乳动物植酸酶的特征性区别在于芽孢杆菌植酸酶不含在已知植酸酶中发现的保守 RHGX RXP 结构域序列 (Kerovuo 等, 1998, 同上; Kim 等, 1998, 同上)。此外, 来自于枯草芽孢杆菌的植酸酶的活性和热稳定性特异性地依赖于钙 (Kerovuo 等, 2000, 枯草芽孢杆菌植酸酶的金属依赖性, Biochem. Biophys. Res. Commun. 268: 365-369, 将其整体在此引入作为参考), 这一点在任何其他报道的来源于真菌、大肠杆菌、植物和哺乳动物的植酸酶均未发现。而且, 枯草芽孢杆菌植酸酶活性的

最佳 pH 也不同于来自于真菌和大肠杆菌的植酸酶。许多报道证明，真菌和大肠杆菌的植酸酶是最佳 pH 为 2.5 (Rodriguez 等, 1999, 同上; 及 Dassa 等, 1990, 同上) 到 5.5(Han 等, 1999, 同上) 的酸性磷酸酶。相反, Kerovuo 等 (1998, 同上) 报道的枯草芽孢杆菌植酸酶活性最佳 pH 为 7。因此, 利用公认安全(GRAS)的细菌菌株生产植酸酶具有很大的实用性, 其可以提供一种新而安全的植酸酶来源补充到商用饲料中。

Maugenest 等 (1997, 一种编码玉米种子植酸酶 cDNA 的克隆和鉴定, *Biochemistry Journal*, 322: 511-517) 报道了一种编码玉米种子植酸酶 cDNA 的克隆和鉴定。NO. 6,291,224 号美国专利公开了一种来源于玉米的植酸酶, NO. 6,303,766 号美国专利公开了来源于大豆的植酸酶, 两者均已知为酸性植酸酶。但是, 一般来说植物植酸酶正常情况下不能产生足够的量以适应产业化需求, 而且通常在非发芽种子可检测到非常低的内源活性。细胞外的植酸酶活性明显不足以驱动 (mobilizing) 封闭 (locked up) 在土壤中的肌醇六磷酸盐。

植物可从水中获得碳、氢和氧进行光合作用, 而磷、氮、金属离子、钙及微量元素主要从土壤获得。因此, 土壤中的磷和氮可利用性成为植物生长的限制因子。主要以无机磷酸盐形式存在的磷由根从土壤中吸收, 然后无机磷酸盐转运到植物的其他组织用于各种生物合成过程, 包括 DNA 和 RNA 合成, 等等。但是, 大部分磷在植物中被封闭并以肌醇六磷酸盐的形式储存下来。对于植物来说, 以肌醇六磷酸盐形式封闭的磷像土壤中的肌醇六磷酸盐一样不能为植物所利用。为了供应植物的营养需要, 常在肥料中补充无机磷酸盐来促进植物的生长, 这样又产生了环境的另一个污染源。

在植物中表达植酸酶的尝试还未产生有用的表型。一种来源于真菌 - 黑曲霉菌的酸性植酸酶 (*phyA*) 已成功在转基因烟草中表达 (Ullah 等, 1999, 同上)。从转基因烟草回收的重组植酸酶从催化活性来说与天然的植酸酶没什么差别, 仅仅是最佳 pH 从 pH 5 移到了 4。同一基因在阿布斯 (*Arabidopsis*) 获得超表达 (Richardson 等, 2000, 曲霉菌植酸酶从阿布斯的根分泌至细胞外, 使得植物可从肌醇六磷酸盐获取磷, *Plant Journal* 25(6): 641-649)。NO. 6,022,846 号美国专利公

开了无花果曲霉、黑曲霉菌、泡盛曲霉、*Aspergillus nidulans*、及各种农作物果实、叶和根中酸性植酸酶的表达（也参见 NO. 5,900,525 号美国专利）。这些酸性真菌植酸酶的细胞内表达在转化植物内不会产生任何有意义的表型变化。

- 5 包括猪和鸡在内的许多单胃动物用由大豆粗粉、玉米、小麦、大麦、米糠和芸苔(*canola*)粗粉组成的饲料喂养。既然大部分磷均封闭在肌醇六磷酸盐中，常常使用低最佳 pH、主要来自真菌的外源植酸酶作为饲料添加剂。在动物饲料中引入表达活性植酸酶的转基因植物来代替外源植酸酶也可增强用此类饲料饲养的动物的磷酸盐利用率。
- 10 因此，对通过基因工程可在植物中影响和创造生化途径的方法的需求仍继续存在。

3. 发明概述

有效利用磷不仅对于植物和动物的生长很重要，对于减少由动物
15 废物及肥料所致的环境污染也很重要，所述动物废物及肥料含有以肌醇六磷酸盐形式存在的不能利用的磷。为了利用各种食物来源中的磷，可将各种来源的植酸酶加入动物饲料中以便单胃动物可有效地利用磷，同时将更少的致污染磷排泄到环境中。如果在中性 pH 具有活性的植酸酶可以在植物中表达，这些转基因植物也可明显提高生长速度，
20 缩短成熟和/或开花时间。因此，需要这样的植酸酶，它在动物饲料和植物中表现出最佳活性，同时对于动物和植物的健康也是安全的。而且，需要生成大量的植酸酶以用于商业应用。

本发明部分地依赖于两种新植酸酶基因的发现（见图 1 和图 2；
SEQ ID NOS: 1, 2, 3, 4）和中性植酸酶的表达可促进植物生长、开花
25 和结果这一观察结果，所述两种基因分别来自于两种微生物：地衣芽孢杆菌和枯草芽孢杆菌 168。相应地，本发明涉及到来自于两种公认安全（GRAS）的微生物的，被分别命名为 *phyL* 和 *168phyA* 的两种植酸酶基因的核苷酸序列（分别为 SEQ ID NOS: 1 和 3；见图 1A 和图 2A），及其所编码的蛋白质的氨基酸序列以及其片段、衍生物、类似物和变体。
30 相应地，本发明提供了分离或重组制备的植酸酶，它们分别源自于地衣芽孢杆菌（*phyL*，具有 SEQ ID NO:2 的氨基酸序列，

见图 1B) 和枯草芽孢杆菌菌株 168 (168phyA, 具有 SEQ ID NO:4 的氨基酸序列, 见图 2B), 及其片段、类似物、衍生物和变体, 如本文所定义的, 将其总称为“本发明的肽”或“本发明的蛋白质”。另外, 本发明提供了编码本发明多肽的核酸分子, 在此将其总称为“本发
5 发明的核酸”, 它包括 cDNA、基因组 DNA 和 RNA。

正如本文所用的那样, 将基因名字排成斜体指的是基因本身, 而基因名字不排成斜体则指的是该基因所编码的蛋白质或多肽产物。例如, “Gene”应该指 *Gene* 基因, 而“Gene”则指 *Gene* 基因所编码的蛋白质或多肽产物。

10 由此, 本发明提供了分离的核酸分子, 该核酸分子包括或由以下核苷酸序列组成, 所述核苷酸序列约有 30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%或 98%与 SEQ ID NO:1 的核苷酸序列或其互补序列、SEQ ID NO:3 或其互补序列相同, 并编码具有 phyL 或 168phyA 活性的蛋白质或多肽。这些活性包
15 括抗原性、免疫原性、催化活性(如植酸酶活性)及其他易于检测的活性。此外, 这些活性还包括可在中性 pH 下发挥功能, 尤其是其酶活性具有宽范围的最适温度, 并在中性 pH、各自的最适温度下具有最高的活性(见图 7B 及 6.4 部分, 见下文)。而且应表现出良好的热稳定性, 尤其是在 Ca^{2+} 存在的条件下。在特定的实施方案中, 这样
20 的核酸分子不包括编码 *phyC* (SEQ ID NO:21)、*phyK* (SEQ ID NO:23) 以及长度分别为至少 5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、90、100、120、140、160、180、200、220、240、260、280、300、320、340、360 或 380 个氨基酸残基的 *phyC* (SEQ ID NO:22)、*phyK* (SEQ ID NO:24) 并具有植酸酶催化活性的片段的
25 核苷酸序列。

本发明进一步提供了分离的核酸分子, 该核酸分子包括或由 SEQ ID NO:1、SEQ ID NO:3 的核苷酸序列或其互补序列中约 25、30、35、40、45、100、150、200、250、300、350、400、450、500、550、600、650、700、750、800、850、900、950、1000、1050、1100、1150、1200、
30 1250、1300、1350 或更多相邻核苷酸组成, 其中所述互补序列编码具有一个或更多 phyL 或 168phyA 活性的蛋白质或多肽。在特定的实

5 实施方案中，这样的核酸分子不包括编码 *phyC* (SEQ ID NO:21)、*phyK* (SEQ ID NO:23) 以及长度分别至少为 5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、90、100、120、140、160、180、200、220、240、260、280、300、320、340、360 或 380 个氨基酸残基的 *phyC* (SEQ ID NO:22)、*phyK* (SEQ ID NO:24) 并具有植酸酶催化活性的片段的核苷酸序列。

10 本发明提供了分离的多肽或蛋白质，它们由一种核酸序列编码，所述核酸序列包含或由以下核苷酸序列组成，所述核苷酸序列至少约有 30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95% 或 98% 与 SEQ ID NO:1 的核苷酸序列或其互补序列、SEQ ID NO:3 的核苷酸序列或其互补序列相同，其中这些蛋白质或多肽也表现出本发明多肽的至少一个结构和/或功能特性。所述的本发明多肽功能特性包括抗原性、免疫原性、催化活性及其他易于检测的活性。在特定的实施方案中，这样的多肽或蛋白质不包括分别由 *phyC* (SEQ ID NO:21)、*phyK* (SEQ ID NO:23) 的核苷酸序列，以及长度至少为 25、30、35、40、45、100、150、200、250、300、350、400、450、500、550、600、650、700、750、800、850、900、950、1000、1050、1100、1150、1200、1250 或 1280 个核苷酸的 *phyC* (SEQ ID NO:21) 片段和长度为至少 25、30、35、40、45、100、150、200、250、300、350、400、450、500、550、600、650、700、750、800、850、900、950、1000、1050、1100、1150、1200、1250、1300、1350、1400、1450、1500、1550、1600、1650 或 1700 个核苷酸的 *phyK* (SEQ ID NO:23) 片段编码的多肽或蛋白质。

25 本发明提供了分离的多肽或蛋白质，它们由一种核酸序列编码，所述核酸分子包含或由以下核苷酸序列组成，所述核苷酸序列包含 SEQ ID NO:1、3 的核苷酸序列或其互补序列中至少约 25、30、35、40、45、100、150、200、250、300、350、400、450、500、550、600、650、700、750、800、850、900、950、1000、1050、1100、1150、1200、1250、1300、1350 或更多相邻核苷酸，其中这些蛋白质或多肽也至少表现出本发明多肽的一个结构和/或功能特性。在具体的实施方案中，这样的多肽或蛋白质不包括分别由 *phyC* (SEQ ID NO:21)、*phyK*

30

(SEQ ID NO:23), 长度至少为 15、30、45、60、90、120、180、240、300、420、540、780、1020、1140、1260 或 1280 个核酸的 *phyC* (SEQ ID NO:21) 片段, 以及长度至少为 15、30、45、60、90、120、180、240、300、420、540、780、1020、1140、1260、1280、1300、1350、
5 1400、1450、1500、1550、1600、1650 或 1700 个核酸的 *phyK* (SEQ ID NO:24) 片段编码的多肽或蛋白质。

本发明也提供了分离的核酸分子, 它们包含编码蛋白质的核苷酸序列, 其中该蛋白质的氨基酸序列至少有约 40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95% 或 98% 与 SEQ ID NO:2
10 或 4 或该蛋白的片段、衍生物、类似物或变体的氨基酸序列相同, 并表现出 *phyC* 和 *phyK* 的抗原性、免疫原性、催化活性及其他易于检测的活性。在特定的实施方案中, 这样的核酸分子不包括编码 *phyC* (SEQ ID NO:21)、*phyK* (SEQ ID NO:23) 以及长度至少为 5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、90、
15 100、120、140、160、180、200、220、240、260、280、300、320、340 或 360 个氨基酸残基的 *phyC* (SEQ ID NO:22) 片段和长度至少为 5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、90、100、120、140、160、180、200、220、240、260、280、300、320、340 或 360 个氨基酸残基 *phyK* (SEQ ID NO:24) 片段的核苷酸
20 序列。

本发明进一步提供了分离的核酸分子, 或其互补序列, 所述核酸分子包含编码蛋白质的核苷酸序列, 其中所述蛋白质的序列包含 SEQ ID NO:2 或 4 或其片段、衍生物、类似物或该蛋白变体中至少约 10、15、20、25、30、50、75、100、125、150、175、200、225、250、275、
25 300、325、350、375 或更多相邻氨基酸, 或由这些相邻氨基酸组成, 并表现出 *phyC* 和 *phyK* 的抗原性、免疫原性、催化活性及其他易于检测的活性。在特定的实施方案中, 这样的核酸分子不包括编码 *phyC* (SEQ ID NO:21)、*phyK* (SEQ ID NO:23), 以及长度至少为 5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、90、
30 100、120、140、160、180、200、220、240、260、280、300、320、340 或 360 个氨基酸残基的 *phyC* (SEQ ID NO:22) 片段和长度至少

为 5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、90、100、120、140、160、180、200、220、240、260、280、300、320、340 或 360 个氨基酸残基 *phyK* (SEQ ID NO:24) 片段的核苷酸序列。

5 此外，本发明还提供了分离的多肽或蛋白质，它们包含至少约 40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95% 或 98% 与 SEQ ID NO:2 或 4 或该蛋白片段、衍生物、类似物或变体的氨基酸序列相同的氨基酸序列，其中这些蛋白质或多肽也表现出本发明多肽的至少一个结构和/或功能特性。在特定的实施方案
10 中，这样的多肽或蛋白质不包括分别由 *phyC* (SEQ ID NO:21)、*phyK* (SEQ ID NO:23) 的核苷酸序列编码的多肽或蛋白质，以及长度分别至少为 5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、90、100、120、140、160、180、200、220、240、260、280、300、320、340 或 360 个氨基酸残基的 *phyC* (SEQ ID NO:22)
15 和 *phyK* (SEQ ID NO:24) 片段。

本发明也提供了分离的多肽或蛋白质，它们包含了 SEQ ID NO:2 或 4 或其片段、衍生物、类似物或该蛋白变体中至少约 10、15、20、25、30、50、75、100、125、150、175、200、225、250、275、300、325、350、375 或更多相邻氨基酸的氨基酸序列，其中这些蛋白质或多
20 肽也表现出本发明多肽的至少一个结构和/或功能特性。在特定的实施方案中，这样的多肽或蛋白质不包括分别由 *phyC* (SEQ ID NO:21)、*phyK* (SEQ ID NO:23) 的核苷酸序列编码的多肽或蛋白质，以及长度分别至少为 5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、90、100、120、140、160、180、200、220、240、
25 260、280、300、320、340 或 360 个氨基酸残基的 *phyC* (SEQ ID NO:22) 和 *phyK* (SEQ ID NO:24) 片段。

在一个实施方案中，本发明提供了分离的核酸分子，该核酸分子在本文定义的严格条件下与具有 SEQ ID NO:1 或 3 序列或其互补序列的核酸杂交，其中该核酸分子编码表现出本发明多肽至少一种结构
30 和/或功能特性的蛋白质或多肽。

此外，本发明也提供了适于用作引物或杂交探针的核酸分子，这

些引物或杂交探针用来检测编码本发明多肽或与本发明多肽类似的其他序列的核酸。

5 本发明的另一方面提供了包含本发明核酸分子的载体，如重组表达载体。而且本发明也提供了包含这样载体或被改造成包含和/或表达本发明核酸分子的宿主细胞，以及包含操纵性连接于异源启动子的本发明核苷酸序列的宿主细胞。在一些实施方案中，宿主细胞为芽孢杆菌，优选枯草芽孢杆菌 MU331。在具体实施方案中，这样的异源启动子为强原噬菌体启动子。

10 本发明还提供了利用重组 DNA 技术制备本发明多肽的方法，该方法包括培养宿主细胞，产生和分离本发明的多肽，其中所述宿主细胞包含编码本发明多肽的重组表达载体，或操纵性连接于异源启动子的编码本发明多肽的核苷酸序列。在一些实施方案中，宿主细胞为芽孢杆菌，优选枯草芽孢杆菌 MU331。在具体实施方案中，本发明提供了一种利用 $\phi 105$ 噬菌体超表达系统快速生产大量本发明多肽的方法。

15 本发明的另一方面是提供包含本发明多肽的动物饲料及其制备方法，其中本发明多肽将磷从肌醇六磷酸盐中释放出来，供动物利用。

20 在另一个实施方案中，本发明提供了包含一种核酸分子的转基因植物，其中所述核酸编码在中性 pH 具有催化活性的植酸酶。在一个特定的实施方案中，本发明提供了包含本发明核酸分子的转基因植物，所述核酸分子表达本发明的植酸酶或其具有功能活性的片段、同系物(homolog)或类似物，或编码来源于芽孢杆菌的植酸酶。在优选实施方案中，植酸酶在细胞内表达。在另一个优选实施方案中，植酸酶在细胞外表达，例如从转基因植物的根部表达。所表达的植酸酶在中性 pH 下具有活性，并使植物将磷从储存于植物或环境（如土壤）中的肌醇六磷酸盐释放出来。本发明也提供了生成此类转基因植物的方法。

25 本发明还提供了与本发明多肽免疫特异性结合的抗体。这样的抗体包括但不局限于可与本发明多肽免疫特异性结合、来源于各种动物的抗体、人源化抗体、嵌合抗体、多克隆抗体、单克隆抗体、双特异性抗体、多特异性抗体、单链抗体、Fab 片段、 $F(ab')_2$ 片段、二硫键

连接的 Fvs, 包含 VL 或 VH 结构域或包含补体决定区 (CDR) 的片段。

5 在一个实施方案中, 本发明提供了检测诸如细胞、培养基等生物材料中本发明多肽或类似多肽存在与否、活性或表达的方法。将生物材料与可直接或间接检测本发明多肽存在与否、活性或表达的试剂接触, 可以测定与对照样品相比样品中多肽的活性或表达水平的升高或降低。在具体实施方案中, 这样的试剂是免疫特异性结合本发明多肽的抗体或其片段。在另一具体实施方案中, 这样的试剂是肌醇六磷酸盐。

10 在另一个实施方案中, 本发明提供了包含生物活性分子及本发明多肽一个或多个结构域或其片段的融合蛋白。特别地, 本发明提供了包含与本发明多肽一个或多个结构域或其片段进行重组融合或化学结合 (包括共价及非共价结合) 的生物活性分子的融合蛋白。

15 3.1 定义

术语“酸性的”或“酸性 pH”在这里指小于 6.0、小于 5.5、小于 5.0 及小于 4.0 的 pH 值。

术语“类似物”在这里指具有与 phyL 或 168phyA、phyL 或 168phyA 片段类似或相同功能的多肽, 但不一定包含与 phyL 或 168phyA、phyL 或 168phyA 片段类似或相同的氨基酸序列或与 phyL 或 168phyA、phyL 或 168phyA 片段类似或相同的结构, 也不一定是抗体或抗体片段。含有类似氨基酸序列的多肽指至少满足下列条件中其中一条的多肽:

20 (i) 多肽含有至少 30%、至少 35%、至少 40%、至少 45%、至少 50%、至少 55%、至少 60%、至少 65%、至少 70%、至少 75%、至少 80%、至少 85%、至少 90%、至少 95% 或至少 99% 与 phyL 或 168phyA、phyL 或 168phyA 片段氨基酸序列相同的氨基酸序列, 而且该多肽不是 phyC, 也不是 phyK, 也不是长度至少为 5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、90、100、120、140、160、180、200、220、240、260、280、300、320、340 或 360 个氨基酸残基的 phyC 或 phyK 片段; (ii) 由至少 30%、至少 35%、至少 40%、至少 45%、至少 50%、至少 55%、至少 60%、至少 65%、至少 70%、

25 30

- 至少 75%、至少 80%、至少 85%、至少 90%、至少 95%或至少 99% 与编码 phyL 或 168phyA、phyL 或 168phyA 片段之核苷酸序列相同的核苷酸序列所编码的多肽，而且该多肽不是 phyC，也不是 phyK，也不是 phyC 或 phyK 的片段；(iii)由在严格条件下与核苷酸序列杂交的核苷酸所编码的多肽，其中所杂交的核苷酸编码 phyL 或 168phyA，phyL 或 168phyA 片段，或其中的至少 10 个氨基酸残基、15 个氨基酸残基、20 个氨基酸残基、25 个氨基酸残基、40 个氨基酸残基、80 个氨基酸残基、90 个氨基酸残基、100 个氨基酸残基、125 个氨基酸残基、150 个氨基酸残基、175 个氨基酸残基、200 个氨基酸残基、225 个氨基酸残基、250 个氨基酸残基、275 个氨基酸残基、300 个氨基酸残基、325 个氨基酸残基、350 个氨基酸残基、375 个氨基酸残基，而且该多肽不是 phyL 或 168phyA，phyL 或 168phyA 片段。表现出与 phyL 或 168phyA，phyL 或 168phyA 片段相似的抗原性、免疫原性、催化活性及其他易于检测的活性，具有相似结构和功能的多肽，是指具有与 phyL 或 168phyA，phyL 或 168phyA 片段类似的二级、三级或四级结构的多肽。多肽的结构可用本领域内熟练技术人员所知的方法来测定，这些方法包括但不限于 X 射线晶体学、核磁共振及晶体成像电子显微镜，多肽的功能可用各种分析测定其生物学活性而确定。
- 20 术语“与 phyL 或 168phyA 免疫性特异性结合的抗体或抗体片段”在这里指与 phyL 或 168phyA，phyL 或 168phyA 片段免疫性特异性结合而不与其他多肽发生非特异性结合的多肽及其片段。优选的与 phyL 或 168phyA，phyL 或 168phyA 片段免疫性特异性结合的抗体及其片段不与其他抗原发生交叉反应。可用诸如免疫测定或其他本领域内熟练技术人员已知的方法来鉴定与 phyL 或 168phyA，phyL 或 168phyA 片段免疫性特异性结合的抗体及其片段。与 phyL 或 168phyA 免疫性特异性结合的抗体及其片段也可分别称为“抗 phyL 抗体”或“抗 168phyA 抗体”。
- 30 术语“衍生物”在这里指对给定肽或蛋白修饰后的肽或蛋白质，如将任意类型分子共价粘附到肽或蛋白上，包括导入非天然生成的氨基酸，优选具生物活性的物质。所得的生物活性保留原有肽或蛋白的

一个或多个生物活性。

术语“片段”在这里指含有相关核酸分子中至少约 25、30、35、40、45、100、150、200、250、300、350、400、450、500、550、600、650、700、750、800、850、900、950、1000、1050、1100、1150、1200、1250、1300、1350 个或更多相邻核苷酸的核酸分子片段，并含有该核酸分子至少一种功能特性（或所编码的蛋白含有该核酸分子所编码蛋白的一种功能特性）；或指含有长度为相关蛋白质或多肽中至少 5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、90、100、120、140、160、180、200、220、240、260、280、300、320、340 或 360 个氨基酸残基的蛋白质片段或多肽，并含有该蛋白或多肽的至少一种功能特性。

术语“公认安全（GRAS）”在这里指某一物质的特性就其计划用途来说被食品和药品监督管理局（FDA）归属于“公认安全”。只要执行“生产质量管理规范”，这样的物质就可以用于食物生产。酶制品的 GRAS 状态可在提交文档的基础上由 FDA 授予。GRAS 状态与那些来自于植物及动物的酶相称，也与那些来自于已长期用作人用酶源而不导致严重健康问题的微生物的酶相称。

“分离的”或“纯化的”肽或蛋白基本上不含细胞材料或来源于该蛋白来源之细胞或组织的其他污染蛋白，或者化学合成时不含化学前体或其他化学物质。“基本上不含细胞材料”这一措辞包括从细胞组分分离的多肽或蛋白制品，其中细胞组分来自所述细胞。因此，基本上不含细胞材料的多肽或蛋白质包括含有少于约 30%、20%、10%、5%、2.5%或 1%（以干重计算）污染蛋白的多肽或蛋白质。当多肽或蛋白为重组生成时，也优选为基本上不含培养基，即培养基的含量约少于蛋白制品的 20%、10%、5%。当多肽或蛋白通过化学合成生成时，优选为基本上不含化学前体或其他化学物质，即它与参加该蛋白合成的化学前体或其他化学物质是分开的。相应地，这样的多肽/蛋白质制品，除了目的多肽/蛋白质片段之外的化学前体或其他化合物含量少于约 30%、20%、10%、5%（以干重计算）。在本发明的一个优选实施方案中，多肽/蛋白质是分离或纯化的。

“分离”的核酸分子是与其他存在于该核酸分中天然来源的核酸

分子分离的核酸分子。此外，“分离”的核酸分子如 cDNA 分子，用重组技术生成时也可基本上不含其他细胞材料或培养基，化学合成时基本上不含化学前体或其他化学物质，但存在于重组 DNA 文库的核酸分子除外。在一个优选实施方案中，编码本发明多肽/蛋白质的核酸分子是分离或纯化的。

“中性 pH”在这里指值介于约 5.5 至 8.5 间的 pH 值，优选的中性 pH 为约 6.0~约 8.0，更优选的中性 pH 为约 6.5~约 7.5，最优选约为 7.0。

术语“操纵性连接”在这里指当转录在“操纵性连接”启动子控制下产生一条具有功能的信使 RNA，其翻译的结果是产生了与该启动子操纵性连接的 DNA 所编码的多肽。

术语“在严格条件下”指彼此之间至少 70%、至少 75%、至少 80%、至少 85%、至少 90%、至少 95% 相同的核苷酸序列相互杂交的杂交和洗脱条件。例如，这样的杂交条件可见：当代分子生物学方案, John Wiley&Sons, N.Y. (1989), 6.3.1-6.3.6.; 分子生物学基础方法, Elsevier Science Publishing Co., Inc., N.Y. (1986), 75-78 页及 84-87 页; 分子克隆, Cold Spring Harbor Laboratory, N.Y. (1982), 387-389 页, 这些均为本领域内熟练技术人员所熟知, 但并不局限于这些。一个优选的、非限制性严格杂交条件的实例是在 6X 氯化钠/枸橼酸钠 (SSC)、0.5%SDS 中于约 68°C 的温度下杂交, 然后在 2X SSC、0.5%SDS 中于室温洗涤一次或多次。另一个优选的、非限制性严格杂交条件的实例是在 6X 氯化钠/枸橼酸钠 (SSC) 中于约 45°C 的温度下杂交, 然后在 0.2X SSC、0.1%SDS 中于 50~65°C 洗涤一次或多次。

术语“变体”在这里指给定肽天然存在的等位基因变体, 或给定肽或蛋白中一个或多个氨基酸残基经过氨基酸替换、插入或缺失等修饰的重组制备变体。

4. 附图简述

下面的图阐明了本发明的实施方案, 并不限制权利要求书包括的本发明范围。

图 1A 和图 1B 分别显示 *phyL*(SEQ ID NO: 1) 的序列和 *phyL*(SEQ

ID NO: 2)的氨基酸序列。

图 2A 和图 2B 分别显示 *168phyA*(SEQ ID NO: 3)的序列和 *168phyA*(SEQ ID NO: 4)的氨基酸序列。

图 3 描绘了 PCR 克隆 *phyL* 的策略。DP 代表用于兼并 PCR 的兼并引物，IP 代表反向 PCR 引物。先进行兼并 PCR 然后进行反向 PCR 来克隆 *phyL* 基因。反向 PCR 的片段与从该基因上游区域（5'至 ATG 翻译起始密码子）至终止密码子之间克隆的全长基因进行对比。

图 4A 和图 4B 显示用于植酸酶过生产的表达载体的构建体。(A) 显示了用于枯草芽孢杆菌 168 植酸酶超表达的表达质粒。构建体携带有 $\phi 105$ 启动子,后面接有一段 Shine-Delgarno(SD)序列、天然 *168phyA* 基因及其天然终止子。(B)显示了用于地衣芽孢杆菌植酸酶超表达的表达质粒。构建体携带有 $\phi 105$ 启动子,后面接有一段 SD 序列、天然 *phyL* 基因及来自于地衣芽孢杆菌的 α -淀粉酶终止子。

图 5A 和 5B 显示了两种植酸酶的表达水平。样品直接取自细菌培养物,离心后装入 10%SDS-聚丙烯酰胺柱。细菌培养物于孵育前及热孵育 0 至 5 小时后收集。可发现热孵育后随着时间的增加酶的产量增加。LRM 为含参考蛋白质大小的低分子量标记 (BIO-RAD, 香港),在梯度的左边标记。(A)显示了 *168phyA* 所编码酶的表达。(B)显示了 *phyL* 所编码酶的表达。

图 6A 和 6B 显示两种植酸酶的酶活性。(A)根据每毫升所收集培养基中的酶单位评估酶活性,(B)根据各单独反应中所用每毫克酶中的酶单位来表示酶活性。

图 7A 和 7B 描绘了本发明中酶的温度和 pH 情况。根据 Engelen 等(1994,植酸酶活性的简便、快速测定,Journal of AOAC International, 77(3): 760-764)所述方法测定植酸酶的活性,只是将检测的量减至 1 毫升。通过测定 405nm 处的光密度来进行比色。孵育时间定为 30 分钟。所有反应均补充 5mM 的 CaCl_2 以确保其酶活性。

图 8 为植物表达载体的构建策略。将去除了天然信号肽序列的 *168phyA* 基因克隆入 pBI221 载体的 BamHI 和 SacI 位点,取代了大肠杆菌 β -D-葡萄糖醛酸甙酶 (GUS) 基因。携带 *168phyA* 基因表达盒的 HindIII/EcoRI 片段从质粒 pBI221 释放出来并亚克隆入

pCAMBIA1300 双链载体的 BamHI 和 SacI 位点, 以生成重组克隆 pCX0168phA。

图 9 是 pCX168phA 表达载体的示意图。168phA: 来自于枯草芽孢杆菌 168 株的植酸酶基因; CaMV: 菜花样花叶病毒的 35S 启动子; 5 CaMV35S 多聚腺苷酸信号: 来自于菜花样花叶病毒 3'UTR 的多聚腺苷酸信号; KanR: 卡那霉素耐药性; NOS: 胭脂碱合酶基因; pBR322 ori: 来自于 pBR322 的复制起点; pVS1-REP: 来自于 pVS1 的复制起点; pVS1-STA: 来自于 pVS1 质粒的 STA 区; T-border(R): T-DNA 重复序列的右侧边界。来自于 pVS1 的 rep 和 sta 区存在 (Ajdukiewicz 10 等, 1994, Plant Molecular Biology,25:989-994) 可增强这些载体在土壤杆菌中的稳定性, 即使其生长于非选择性培养基中。

图 10A 和 10B 分别显示潮霉素基因 (A) 和 *168phA* 基因 (B) 在转基因烟草中的筛选。(A) 泳道 1: 1KB 加(plus) DNA 梯度; 泳道 2: 作为阴性对照的未转化烟草; 泳道 3-4: 仅用 pCAMBIA 1300 15 载体转化的转基因烟草(对照); 泳道 5-8: 004 的转基因烟草系(0041、0042、0043、0044); 泳道 9-12: 013 的转基因烟草系(0131、0132、0133、0134)。(B) 泳道 1: 1KB 标记物; 泳道 2-3: 作为阳性对照的 pCX-*168phA* 对照质粒; 泳道 4: 作为阴性对照的未转化烟草; 泳道 5-6: 仅用 pCAMBIA 1300 载体转化的转基因烟草(对照); 泳道 7-10: 004 的转基因烟草系(0041、0042、0043、0044); 泳道 11-14: 20 013 的转基因烟草系(0131、0132、0133、0134)。

图 11 显示了 F0 转基因烟草的 RNA 分析。从植物提取的总 RNA 取 20 μg 进行 1% (w/v) 琼脂糖电泳。用 DIG-PCR 试剂盒 (Roche 25 Diagnostics, 香港) 标记的 *168phA* cDNA 作为探针。在转基因系中 (0042、0043 和 0134) 检测到了 mRNA 信号, 但在对照系中未检测到。

图 12 显示了转基因烟草的蛋白质分析。每孔加入 30 μL 从烟草叶中分离的可溶性蛋白。在转基因烟草样品 0042、0043 和 0134 中检测到了植酸酶, 但在对照植物中未检测到。

30 图 13 显示了转基因烟草植物 (图 13b-13d) 和对照植物 (仅用载体转化的转基因植物, 图 13a) 开花茎的数目。

图 14 显示了转基因烟草植物 (图 14a-14c) 和对照植物 (图 14d) 主茎的数目。

图 15 显示了 F1 转基因烟草的 DNA 印迹分析。各泳道加入 10 μg 来自于各种 F1 系基因组 DNA 的 HindIII 限制性消化产物。用 PCR 试剂盒进行放射性标记的 168phyA cDNA 作为探针。在转基因系 (0042 和 0134) 中检测到了特异性的条带, 但在对照系中未检测到。来自于 0042 的 F1 系 (42-1 和 42-2) 包含单拷贝的基因, 而来自于 0134 的 F1 系 (134-1) 则含有两个拷贝。

图 16 显示了 F1 转基因烟草的 RNA 分析。从植物提取的总 RNA 取 20 μg 进行 1% (w/v) 琼脂糖电泳。用 DIG-PCR 试剂盒 (Roche Diagnostics, 香港) 标记的 168phA cDNA 作为探针。在转基因系中 (0042 和 0134) 检测到了 mRNA 信号, 但在对照系中未检测到。

图 17 显示了 F1 转基因烟草的蛋白质分析。每孔加入 10 μg 从烟草叶中分离的可溶性蛋白。在来自于 0042 的 F1 样品中检测到了植酸酶, 但在对照系的 F1 样品中未检测到。

图 18 显示了在磷酸盐缺乏条件下 F2 转基因烟草幼苗的生长情况。F2 幼苗在 MS 培养基 (含 1.25 mM 磷酸盐和 30 g/L 蔗糖) 中生长 15 天, 然后转移到修正 MS 培养基 (不含磷酸盐, 蔗糖浓度降至 5 g/L) 继续生长 17 天。发现转基因系的生物量多于对照系。

图 19 显示了低磷酸盐条件下转基因烟草幼苗在琼脂中的生长情况。幼苗先在含 10^{-3}M 或 10^{-5}M 磷酸盐的 MS 琼脂培养基生长 20 天, 然后在含 10^{-3}M 、 10^{-4}M 或 10^{-5}M 磷酸盐的 MS 琼脂培养基再生长 30 天。将植物干燥, 各自称重。各柱为 18 棵个体植物的平均值 ($N=18$)。

图 20 显示了低磷酸盐条件下转基因烟草幼苗在流体培养基中的生长情况。幼苗先在含 1.25mM 磷酸盐的 MS0 培养基生长 10 天, 然后在含 0.01 mM 或 0.1 mM 磷酸盐的 MS 流体培养基中再生长 20 天。将植物干燥, 各自称重。各柱为 25 棵个体植物的平均值 ($N=25$)。在低磷酸盐条件下, 转基因系 (0042 和 0134) 所得干重比对照系要重。

图 21 显示转基因植物中的内源植酸酶活性升高。将提取的叶蛋白 (200 μg) 与外源 IP6 (400 μg) 一起 37 $^{\circ}\text{C}$ 孵育 4、6 和 8 小时。

然后用阴离子交换层析纯化肌醇磷酸盐 (IP6、IP5、IP4 和 IP3)，用 HPLC 分析并用屈光指数检测器测定各自的 IP6 及 IP5 峰。如图所示，42 系的植物提取物 (N=4) 所得 IP6/IP5 比值比对照植物的比值 (N=4) 低，说明转基因植物具有比对照植物更高的内源性植酸酶活性。

5. 发明详述

5.1 PhyL 和 168phyA

现已在枯草芽孢杆菌 168 的基因组中发现与枯草芽孢杆菌中两种已公开植酸酶具有高度序列同源性的开放阅读框 (ORF)。如第 6.3 部分所述，表达的克隆 *168phyA* 表现为分子量 (MW) 为 44 kDa 的成熟植酸酶 168phyA，分子量由 SDS-PAGE 测定 (见图 5A)。利用基于 168phyA、phyK (Kim Y.O.等, 1998, 从 *Baillus sp. DS11* 从克隆热稳定的植酸酶基因 (phy) 及其在大肠杆菌中的超表达, FEMS Microbiology Letters:162:182-191) 和 phyC (Kerovuo J.等, 1998, 一种来自于枯草芽孢杆菌的全新植酸酶的分离、鉴定、分子基因克隆及测序, Applied and Environmental Microbiology, 64(6):2079-2085) 之间保守氨基酸序列的兼并寡核苷酸, 通过兼并 PCR 从地衣芽孢杆菌克隆了 *phyL*。从核苷酸序列推断出的氨基酸序列表明它是一种 381 个氨基酸残基的蛋白, 并且象 168phyA 和其他枯草芽孢杆菌植酸酶一样没有真菌及大肠杆菌植酸酶中发现的高度保守 RHGXRXP 序列基序。SDS-PAGE 测定的 *phyL* 分子量约为 47 kDa (见图 5B)。

本发明两种植酸酶的活性用 Engelen A.J.等所述检测方法 (1994, 植酸酶活性的简便、快速测定, Journal of AOAC International, 77(3): 760-764) 测定。结果表明, 对于酶活性来说, 168phyA 和 *phyL* 均有着宽的最佳温度范围, *phyL* 在 65°C、168phyA 在 55°C 达到峰值 (见图 7B 和下文的第 6.4 部分)。而且, 本发明的两种酶均展现出良好的热稳定性, 尤其是在有 Ca²⁺ 存在时 (见第 6.4 部分)。本发明多肽的这些特征, 即宽最适温度范围、良好热稳定性及中性 pH 下具有最佳酶活性, 提示这些多肽如 5.10 部分讨论的那样具有很大的商业用途。

因此, 本发明提供了分别具有 SEQ ID NOS: 1 和 3、*phyL* 和 *168phyA*

序列的核酸分子，及其所编码分别含有 SEQ ID NOS: 2 和 4、phyL 和 168phyA 的多肽。

5.2 phyL 和 168phyA 的类似物、衍生物及变体

5 除了上述核酸分子及多肽之外，本发明的核酸分子和多肽还包括那些具有与本发明上述核酸分子及多肽相同生物活性、类似或相同结构域和/或足够多核苷酸序列或氨基酸相同（类似物）的核酸分子和多肽。

10 本发明多肽的共同生物活性包括抗原性、免疫原性、催化活性尤其是中性 pH 下的催化活性及其他熟练技术人员易于检测的活性。

含有类似氨基酸序列的多肽指至少满足下列条件中的一条的多肽：(i) 具有与 phyL (SEQ ID NO:2) 或 168phyA (SEQ ID NO:4)、phyL 或 168phyA 片段氨基酸序列有至少 30%、至少 35%、至少 40%、至少 45%、至少 50%、至少 55%、至少 60%、至少 65%、至少 70%、至少 75%、至少 80%、至少 85%、至少 90%、至少 95%或至少 99% 15 相同的氨基酸序列的多肽，并且具有本发明多肽的至少一个功能特性，而且该多肽不是 phyC，也不是 phyK，也不是 phyC 或 phyK 的片段；(ii) 由与编码 *phyL* (SEQ ID NO:1) 或 *168phyA* (SEQ ID NO:3)、phyL 或 168phyA 片段的核苷酸序列有至少 30%、至少 35%、至少 40%、至少 45%、至少 50%、至少 55%、至少 60%、至少 65%、至少 70%、至少 75%、至少 80%、至少 85%、至少 90%、至少 95%或至少 99% 20 相同的核苷酸序列所编码的多肽，并且具有本发明多肽的至少一个功能特性，而且该多肽不是 phyC，也不是 phyK，也不是 phyC 或 phyK 的片段；(iii) 由本文所定义严格条件下与核苷酸序列杂交的核苷酸所编码的多肽，其中所杂交的核苷酸编码 *phyL* (SEQ ID NO:1) 或 *168phyA* (SEQ ID NO:3)，或编码具有本发明多肽至少一个功能特性的 phyL 或 168phyA 片段，或编码其中的至少 10 个氨基酸残基、15 个氨基酸残基、20 个氨基酸残基、25 个氨基酸残基、40 个氨基酸残基、80 个氨基酸残基、90 个氨基酸残基、100 个氨基酸残基、125 个氨基酸残基、150 个氨基酸残基、175 个氨基酸残基、200 个氨基酸残基、225 个氨基酸残基、250 个氨基酸残基、275 个氨基酸残基、300

个氨基酸残基、325 个氨基酸残基、350 个氨基酸残基、375 个氨基酸残基，而且该多肽既不是 phyL 或 168phyA 也不是 phyL 或 168phyA 片段。具有与 phyL 或 168phyA，phyL 或 168phyA 片段类似结构的多肽，是指具有与 phyL 或 168phyA，phyL 或 168phyA 片段类似的
5 二级、三级或四级结构的多肽。多肽的结构可用本领域内熟练技术人员所知的方法来测定，这些方法包括但不限于 X 射线晶体学、核磁共振及晶体成像电子显微镜。在一优选实施方案中，本发明的多肽来源于芽孢杆菌细胞的 GRAS 菌株。

本发明也包括本发明多肽的衍生物。例如，衍生物可包括经过修
10 饰的肽或蛋白质，如糖基化、乙酰化、聚乙二醇化、磷酸化、酰胺化，使用已知的保护/封闭基团、蛋白裂解、连接上细胞配基或其他蛋白等来使其衍生化，但并不局限于此。可利用已知的技术实施大量化学修饰，这些技术包括但不限于特异性化学切割、乙酰化、甲酰基化，等等。此外，衍生物可包含一个或多个非经典氨基酸。

15 另一方面，本发明分离的核酸分子编码本发明多肽的变体，其中氨基酸序列经基因工程修饰，使该肽的活性升高或降低或其局部结构发生改变而不明显改变其生物活性。一方面，这些变体可作为激动剂或拮抗剂。激动剂可基本保持与本发明多肽相同或其部分生物活性，而拮抗剂可抑制本发明多肽的一个或多个活性。这样的修饰包括氨基
20 酸替换、删除和/或插入。氨基酸修饰可用本领域内任何已知方法完成，各方法对于那些本领域内熟练技术人员来说便于使用并常规。

例如，可根据本领域内任何已知的技术完成诱变，包括但不限于合成指定需修饰多肽中一个或多个修饰的寡核苷酸。定点诱变可用特异的寡核苷酸序列完成，其中寡核苷酸序列除了多肽中足够数量邻近
25 核酸之外还编码包含所需要突变的核苷酸序列。这样的寡核苷酸序列可以充当能在横断(traversed)的缺失连接处两端均形成稳定双链结构的引物。典型地，优选长约为 17 至 75 个核苷酸或更长的引物，其中序列的连接处两端约有 10 至 25 个或更多残基发生改变。在一个或多个位置引入大量不同突变的许多引物都可用于生成突变体文库。

30 正如各种出版物所述，定点诱变的技术在本领域内众所周知（如 Keunkel 等，Methods Enzymol., 154: 367-82, 1987，在此将其整体引

入作为参考)。一般地, 定点诱变的完成首先要获得单链载体或双链载体中两条链的解旋部分, 其中载体的序列包括编码所需肽的 DNA 序列。通常通过合成来制备含所需突变序列的寡核苷酸引物。将引物与单链一起退火, 然后与 DNA 聚合酶如 T7 DNA 聚合酶接触, 以完成含突变链的合成。这样就形成异源双链, 其中一条链编码原始的非突变序列, 第二条链含有所需的突变。然后将些异源双链用于转化或转染适当的细胞如大肠杆菌, 并挑选包括重组载体的克隆, 其中该重组载体带有突变的序列排列。正如可理解的那样, 该技术一般采用既存在单链形态也存在双链形态的噬菌体载体。可在定点诱变中应用的典型载体包括诸如 M13 噬菌体之类的载体。这些噬菌体易于从商业途径获得, 其用法通常也为本领域内熟练技术人员所知。取消了将目的基因从质粒转移到噬菌体这一步骤的双链质粒也可常规用于定点诱变。

另外, 利用商用热稳定酶如 Taq DNA 聚合酶的 PCR 方法, 将致突变的寡核苷酸引物引入可克隆入合适克隆载体或表达载体的扩增片段。例如, 关于 PCR 介导的突变程序, 可参见 Tomic 等, *Nucleic Acids Res.*, 18(6):1656,1987; 及 Upender 等, *Biotechniques*, 18 (1): 29-30, 32, 1995, 将其整体在此引入作为参考。还通过利用热稳定连接酶和热稳定聚合酶的 PCR, 将磷酸化的突变寡核苷酸引入可克隆入合适克隆载体或表达载体的扩增片段 (例如可见 Michael, *Biotechniques*, 16(3):410-2, 1994, 在此将其整体引入作为参考)。还可以使用其他本领域内熟练技术人员已知的、生成给定多肽或其片段变体的方法。例如, 编码所述多肽或其片段氨基酸序列的重组载体可以用突变试剂如羟胺加以处理, 获得序列变体。

待修饰的氨基酸残基优选为暴露于表面的残基。另外, 在进行氨基酸替换时优选的待替换氨基酸为保守氨基酸替换, 例如, 用极性残基替换极性残基, 用亲水残基替换亲水残基, 用疏水残基替换疏水残基, 用带阳性电荷的残基替换带阳性电荷的残基, 或用带阴性电荷的残基替换带阴性电荷的残基。而且, 待修饰的氨基酸残基优选非高度或完全保守的交叉(across)株或种, 和/或对于维持该蛋白的生物学活性来说是非决定性的。

相应的，编码本发明多肽的核酸分子属于本发明的范围，其中该多肽含有对于其生物活性来说非决定性的氨基酸修饰。

5.3 利用噬菌体 ϕ 105超表达系统生产酶

5 已报道的植酸酶超表达诱导系统包括 IPTG 诱导来自于枯草芽孢杆菌 DS11 的 *phyK* 基因在大肠杆菌中表达 (Kim 等, 1998, 同上)、甲醇诱导曲霉菌 *phyA* 基因在酵母菌巴氏毕赤酵母表达 (Han & Lei, 1999, 同上) 及以底物肌醇六磷酸为诱导剂在大肠杆菌中生成来自于 *Klebsiella terrigena* 的 *phytase* (Greiner 等, 1997) 和 *phyC* 编码的植酸酶。用肌醇六磷酸作为诱导剂是以底物特异性的理论为基础。

10 在以前建立的枯草芽孢杆菌 ϕ 105 系统中 (Thornwell, S.J., Ease A.K., Errington J., 1993, 一种基于枯草芽孢杆菌 ϕ 105 的高效表达及分泌系统及其在蜡样芽孢杆菌 β -内酰胺酶 I 生产中的用途, *Gene*, 133: 47-53, 将其整体在此引入作为参考), 将一种有缺陷的原噬菌体载体 ϕ 105 MU331 加以衍生, 用于枯草芽孢杆菌的高水平蛋白超表达 (Leung Y.C.和 Errington J., 1995, 阻断宿主枯草芽孢杆菌裂解并提供异源基因强表达的噬菌体 ϕ 105 基因中插入的鉴定, *Gene*, 154: 1-6, 将其整体在此引入作为参考)。在此衍生系统中, 将 *lacZ* 报告基因 (即来源于 pSG23 质粒的 *lacZ-cat* 盒; Errington, J., 1986, 一种将大肠杆菌 *lacZ* 基因与枯草芽孢杆菌中染色体基因融合的通用方法. *J. Gen. Microbiol.* 132: 2953-2966) 插入与各种噬菌体如 λ 噬菌体裂解表达盒相似的区域。该系统不仅提供了基因的可高效诱导 (通过加热) 转录产物, 而且还防止该系统的宿主细胞裂解。因此, 培养基中产生的酶易于分离, 而不破坏细胞, 因而纯化步骤可以大大简化。另外, 与大肠杆菌不同, 芽孢杆菌是 GRAS 细菌, 其蛋白产物对于包括人在内的动物也是公认安全的。

25 相应地, 将本发明的核酸分子插入表达载体 pSG 以构建用于表达 *phyL* 的 pSGt-pL 构建体和用于表达 168*phyA* 的 pSGt-pA 构建体。用自 *phyL* 基因起始密码子到终止密码子之间编码区旁侧的引物, PCR 扩增编码成熟 *phyL* 的基因片段, 然后亚克隆入 pSGt 表达载体, 其中该载体由地衣芽孢杆菌 α -淀粉酶终止子插入 pSG 表达载体而构建

(见图 4B 和 6.2 部分, 同上)。因此, pSGt-pL 构建体中 phyL 基因处于原噬菌体 ϕ 105 启动子的控制下。将用 *168phyA* 开放阅读框(ORF) 旁侧引物所得 PCR 产物亚克隆入 pSG 载体, 构建 PSG-pA 构建体。在该构建体中, *168phyA* 基因的两侧是 ϕ 105 启动子和 *168phyA* 基因天然启动子(图 4A 和 6.3 部分)。将这些质粒导入用于抗生素抗性基因扩增和选择的大肠杆菌 JM109 株, 然后导入宿主菌株如枯草芽孢杆菌 MU331, 用于酶的生产。相应地, 本发明还包括植酸酶的载体、宿主细胞和重组生产方法(详见 6.2 部分和 6.3 部分)。在一些实施方案中, 宿主细胞为芽孢杆菌, 优选枯草芽孢杆菌 MU331。

10

5.4 融合蛋白

本发明进一步包括了融合蛋白, 其中本发明多肽或其片段与异源多肽以重组或化学的方式融合(即一段不相关的多肽或其部分, 优选该多肽的至少 10 个、20 个、30 个、40 个、50 个、60 个、70 个、80 个、90 个或 100 个氨基酸), 生成融合蛋白。融合可以是直接的, 但也可通过接头序列融合。

一方面, 融合蛋白包括在 N 端与异源信号序列融合的本发明多肽。例如, 本发明多肽中存在的天然信号序列可由异源的信号肽所取代。各种信号序列可从商业途径获得。例如, *phoA* 分泌信号(Sambrook 等, 同上; 当代分子生物学方案, Ausubel 等编, John Wiley & Sons, 1992) 和蛋白 A 分泌信号(Pharmacia Biotech; Piscataway, NJ) 可作为原核异源信号序列。

在另一个实施方案中, 本发明的多肽可与其他多肽中的标记序列如 pQE 载体(QIAGEN, Inc., 9259 Avenue, Chatsworth, Ca, 91311) 中提供的标记如六组氨酸多肽融合, 其中许多可从商业途径获得。例如, 如 Gentsz 等, 1989, *proc. Natl. Acad. Sci. USA*, 86: 821-824, 六组氨酸为融合蛋白的纯化提供了方便。其他肽标记的例子包括与来源于流感血细胞凝集素蛋白表位相对应的血细胞凝集素“HA”标记(Wilson 等, 1984, *Cell*, 37: 767) 和“旗标(flag)”标记(Knappik 等, 1994, *Biotechniques*, 17(4): 754-761)。这些标记对于重组所产生本发明多肽蛋白的纯化非常有用。

融合蛋白可用标准的重组 DNA 技术或蛋白质合成技术，如使用肽合成仪生成。例如，编码融合蛋白的核酸分子可用包括自动化 DNA 合成仪在内的常规技术合成。此外，可利用锚定引物进行基因片段的 PCR 扩增，其中该锚定引物可在两个连续的基因片段间产生互补突出部分，随后退火，重新扩增并产生嵌合基因序列（例如参见：当代分子生物学方案，Ausubel 等编，John Wiley & Sons, 1992）。

编码融合蛋白的核酸分子可引入合适的表达载体中，该载体包含插入编码蛋白序列的转录和翻译所必需的元件。

在一个具体的实施方案中，融合蛋白的表达载体由一个诱导性启动子调控。

5.5 重组蛋白的纯化

一旦本发明的多肽由上述方法生成，即可利用重组蛋白纯化领域内任何已知的方法加以纯化，例如层析（如离子交换层析，亲和层析尤其是特异性抗体的亲和层析，凝胶过滤层析）、离心、溶解度差异等，但并不限于此，还可用纯化蛋白的其他标准技术。而且，本发明的多肽或其片段可以和此处或别处所述的异源多肽序列融合，以便于纯化。

在一个具体的实施方案中，从细菌培养上清中纯化枯草芽孢杆菌表达的 phyL 或 168phyA 的方法包括：用乙醇沉淀，然后离心，将沉淀重悬进行凝胶过滤柱层析（见第 6.2 和 6.3 部分）。

5.6 动物饲料的制备

如上文第 5.1-5.5 部分中所述制备的本发明多肽，在中性 pH 下具有植酸酶活性，可用于提供动物饲料，此饲料中的磷可被以此饲料饲养的动物高效利用。因此，本发明的另一方面是提供含本发明多肽的动物饲料，其中本发明多肽可使磷从肌醇六磷酸盐中释放出来，供动物利用。例如，这样的动物饲料制备方法可以是：饲料与本发明的植酸酶粉末混合并制成小丸，其中酶粉末在约每吨饲料 1 公斤粉末的重量比的条件下酶活性为 200,000-400,000 EU/kg。为混合均匀，酶粉末可先与小量饲料如 10 公斤混合，然后与剩余的饲料混合。酶在饲料

中的用量至少为 200 酶单位(EU)/kg 饲料, 优选至少 250 EU/kg 饲料, 最优选 300 EU/kg 饲料。一个酶单位 (EU) 能够在 37°C 和 pH 7.0 条件下, 于 1 分钟内使 5.1mM 肌醇六磷酸钠释放出 1 μ mol 正-磷酸盐。饲料可以是玉米、谷物、小麦、大麦、米糠、大豆粗粉及芸苔粗粉或其他用于动物饲料的任意普通材料。

5.7 转基因植物的制备

植物的生长需要碳、氢、氧、磷、氮、金属离子和微量元素等元素。尽管植物可从水和光合作用获得, 而磷、氮、金属离子、钙及微量元素主要从土壤获得。因此, 土壤中的磷和氮可利用性成为植物生长的限制因子。

即使两个领域的植酸酶基因和蛋白序列及结构存在具大的差异, 来自于另一实验领域中生物体的植酸酶可以在一植物中利用基本的生物信号组件有效发挥功能, 本发明即基于这一发现。因此, 本发明涉及在植物内产生新的生化信号途径, 使磷从不能利用的肌醇六磷酸转化成可利用的无机磷酸盐, 从而增强植物的生长能力, 例如以侧芽数目增加来衡量的植物生长能力。因为磷酸盐也是开花和结果所必需的, 本发明也提供了开花 (如开花更早, 花/芽的数目增加) 和结果 (如果实的数目增加) 改善了的开花植物。

相应地, 本发明提供了转基因植物, 其中该植物含有编码和表达植酸酶的核酸分子, 该植酸酶在中性 pH 下具有最佳的催化活性。本发明的转基因植物与相当的未改造植物即同一种属 (株) 相比具有改良的生长、开花和结果。在一个具体实施方案中, 此种植酸酶来自在中性 pH 时具有最佳催化活性的芽孢杆菌种属。在一个优选实施方案中, 本发明的转基因植物包含本发明的核酸分子, 并表达在中性 pH 具有活性并具有广谱温度范围的 phyL (SEQ ID NO:2) 或 168phyA (SEQ ID NO:4), phyL 的温度范围为约 37°C 至约 70°C, 168phyA 的温度范围为约 37°C 至约 65°C。在一个优选实施方案中, 植酸酶的分泌量检测不到或检测不到有意义的量 (即不超过总植酸酶的 1%、2%、5%或 10%)。SEQ ID NOS: 2 和 4 具有天然信号肽, 但蛋白并无一点分泌。在另一个优选实施方案中, 植酸酶表达至细胞外, 例如

从转基因植物的根分泌出来。植物中这种中性植酸酶细胞外表达可用异源核苷酸序列融合到或取代编码植酸酶基因信号肽的核苷酸序列的 N-末端（即对于 phyL (SEQ ID NO:2) 来说是 1~80 个氨基酸残基的全部或一部分，优选全序列或 1~20 氨基酸残基部分；对于 168phyA (SEQ ID NO:4) 来说是说是 1~80 个氨基酸残基的全部或一部分，优选全序列或 1~26 氨基酸残基部分），其中此异源序列编码可使该植酸酶从特定植物中细胞有效分泌出来的植物信号肽。植物信号肽的例子包括但不局限于来自于延伸多肽或延伸多肽相关的信号肽 (Richardson 等, 2001, Plant Journal 25: 641-649)、酸性磷酸酶 (Haran, -S; Logendra, -S; Seskar, -M; Bratanova, -M; Raskin, -I., oct., 2000, Arabidopsis 酸性磷酸酶启动子的鉴定及酸性磷酸酶表达的调控, Plant-Physiol. 125 (2): 615-626)、内质网信号肽 (Borisjuk, -N-V; Borisjuk, -L-G; Logendra, -S; Petersen, -F; Gleba, -Y; Raskin, -I., May, 1999, 重组蛋白在植物根分泌液中的生成, Nat-Biotechnol. 17(5): 466-9)、 α -淀粉酶 (Park CS, Chang CC, Kim JY, Ogrydziak DM, Ryu DD., 1997, 高丰度 α -淀粉酶在酵母 *Yarrowia lipolytica* 中的表达、分泌和加工, J Biol Chem 272: 6876-6881) 和 PVR3 (Choi, -D-W; Song, -J-Y; Oh, -M-H; Lee, -J-S; Moon, -J; Suh, -S-W; Kim, -S-G., Mar., 1996, 一种编码来自于豆幼苗根 ns-LTP 样蛋白的选择性 cDNA 的分离, Plant-Mol-Bil. 30(5):1059-66)。相应地，在另一个优选实施方案中，本发明的转基因植物包含本发明的核酸分子并表达 phyL (SEQ ID NO:2) 或 168phyA (SEQ ID NO:4)，只是 SEQ ID NO:2 的 1~80 个氨基酸残基的全部或一部分尤其是 N 端部分，优选全序列或 1~20 氨基酸残基部分，或 SEQ ID NO:2 的 1~80 个氨基酸残基的全部或一部分尤其是 N 端部分，优选全序列或 1~26 氨基酸残基部分，由异源植物信号肽通过基因工程加以取代。在这样的转基因植物中，中性植酸酶分泌入土壤中并将土壤中的肌醇六磷酸转化成植物可以摄取的无机磷酸盐。在另一个优选实施方案中，本发明的转基因植物至少包含本发明的两个核酸分子，其中一个核酸分子编码 phyL (SEQ ID NO:2)，另一个编码 SEQ ID NO:2 1~80 个氨基酸残基的全部或一部分尤其是 N 端部分，优选全序列或 1~20 氨基酸残基部分由异源植物信号肽取

代的 phyL。在另一优选实施方案中，本发明的转基因植物至少包含本发明的两个核酸分子，其中一个核酸分子编码 168phyA (SEQ ID NO:4)，另一个编码 SEQ ID NO:4 1~80 个氨基酸残基的全部或部分尤其是 N 端部分，优选全序列或 1~26 氨基酸残基部分由异源植物信号肽取代的 168phyA。这样的转基因植物可同时在细胞内和细胞外表达植酸酶。在另一个优选实施方案中，本发明的转基因植物包含本发明的核酸分子，并表达与本发明多肽功能或结构特性至少一项相同的类似物、衍生物及或其片段。在另一个优选实施方案中，本发明的转基因植物包含本发明的核酸分子，该核酸分子可在本文定义的严格条件下与具有 SEQ ID NO:1 或 3 或其互补序列的核酸分子杂交，并编码表现出至少与本发明多肽相同的结构和/或功能特性的蛋白或多肽。特定地，本发明提供了可生成中性植酸酶的转基因烟草和稻谷的生产，该植酸酶有助于改善植物生理如植物生长速度和特征，例如改善开花反应。而且，当这些生成中性植酸酶的植物喂给动物时，植酸酶可作用于动物食物中其他来源的肌醇六磷酸盐，水解肌醇六磷酸盐从而为动物消化释放无机磷酸盐。这样就减少或排除了将植酸酶或无机磷酸盐补充到动物饲料中的必要性，并减少了因动物排泄磷而导致的环境污染问题。相应地，本发明还提供了包含本发明转基因植物(特别是来自于这些转基因植物的种子或果实)的动物饲料。

相应地，本发明也提供了植物基因修饰的嵌合基因构建体，以便通过提高磷的利用率来提高其生长速度、缩短开花所需要的时间。嵌合基因构建体包含一段基本上全部编码植酸酶的序列，其中该植酸酶在中性 pH 时可催化肌醇六磷酸盐水解。优选植酸酶来源于芽孢杆菌细菌。在一个特定的实施方案中，嵌合基因构建体包含具有 SEQ ID NO:1 和/或 SEQ ID NO:3 序列的核酸分子。在另一个实施方案中，嵌合基因构建体包含具有 SEQ ID NO:1 和/或 SEQ ID NO:3 序列的核酸分子，SEQ ID NO:1 中核苷酸 241~480 的全部或部分，尤其是 241~300 核苷酸的全部或部分由编码植物信号肽的异源核酸分子所取代，和/或 SEQ ID NO:3 中核苷酸 100~339 的全部或部分，尤其是核苷酸 100~177 的全部或部分由编码植物信号肽的异源核酸分子所取代。在另一个优选实施方案中，嵌合基因构建体包含编码本发明

多肽的类似物或其片段的核酸，其中，类似物或其片段与本发明多肽具有相同的至少一个结构或功能特性。在另一个特定的实施方案中，嵌合基因构建体包含一段在本文定义严格条件下与具有 SEQ ID NO:1 或 3 序列之核酸杂交的序列，其中该序列编码具有与本发明多肽相同的至少一个结构或功能特性的蛋白或多肽。而且，嵌合基因构建体中包含的核酸分子其所编码的植酸酶可以是其他任意植酸酶，只要在中性 pH 具有最佳催化活性，具有与本发明植酸酶类似的结构特征，如多个钙结合点。这样的植酸酶包括但不限于下列多肽：来自于芽孢杆菌的植酸酶（登记号为：AAC38573、AAC31775、7767024）、来自于枯草芽孢杆菌的植酸酶（登记号为：AAC31775、AAG17903、AAB72078、AAA87722）、来自于液化淀粉芽孢杆菌的植酸酶（登记号为：7246002、7245653）、来自于 *Caulobacter crescentus* 的植酸酶（登记号为：AAK23276）和来自于蓝色链霉菌的植酸酶（登记号为：CAC17528）。

编码植酸酶的序列的上游和下游操作性地连接调控组件，优选与该植酸酶异源的组件；例如在植物细胞中使该基因表达（酶生成）的 CMV 35S 启动子（见 6.5.1-6.5.4 部分）。当用常规转化方法，如微粒轰击、土壤杆菌感染、微注射等将含有本发明植酸酶基因的构建体导入植物细胞中时，该基因在调控序列控制下在细胞中表达。表达的植酸酶成功地与天然存在于植物中的生物合成器官相互作用，催化无机磷酸盐在中性 pH 下从肌酸六磷酸盐释放出来。本发明也可通过提高无机磷酸盐的利用率来促进植物的生长，从而改善开花和结果。这样，植物成熟和开花所需的时间就缩短了。相应地，本发明也提供了肌酸六磷酸盐水平降低的植物细胞及完整植株，植物细胞含本发明的嵌合基因构建体。也提供了提高植物细胞及完整植株中无机磷酸盐利用率的方法，包括将与本发明一致的嵌合基因构建体导入该植物细胞或该完整植株的细胞中这一步聚。

在一个特定的实施方案中，采用稻谷（见 6.5.3 部分）及烟草的植物（见 6.5.4 部分）作为两种模型系统。将两个编码植酸酶嵌合构建体导入这两种植物中。

在本发明的一个优选实施方案中，使用来自于枯草芽孢杆菌的植

酸酶。它含有用于分泌的信号肽，因而植酸酶从细胞中分泌出来。此酶能使无机磷酸盐在中性 pH 下从肌酸六磷酸盐释放出来，并具有高温稳定性。至此已发现即使两个领域的植酸酶基因和蛋白序列及结构存在巨大的差异，来自于另一实验领域中生物体的植酸酶可以在一植物中利用基本的生物信号组件有效发挥功能。因此，涉及在植物中生成可使肌酸六磷酸中的磷转化成无机磷酸盐形式的新生化信号途径。本发明所获得的结果表明该新生化途径可提高植物的生长速度（见 6.5.9 部分及图 12 和图 13）。

还观察到开花也需要磷。因此，本发明也提供了转基因开花植物，其开花所需的时间因中性植酸酶基因表达使磷利用率上升而得以缩短。尽管任何植物均可用本发明的表达盒和方法加以修饰，优选的是来自于下列各类的物种（括号内为代表性物种）但并不局限于此：

单子叶植物：天冬属（芦笋）、雀麦属(cheatgrass)、萱草属（萱草）、大麦属（大麦）、黑麦草属（黑麦草）、稻属（稻）、黍属（毛线稷草）、狼尾草属（狼尾草）、蜀黍属、trigonella（fenu grass）、小麦属（小麦）、玉蜀黍属（玉米），和：

双子叶植物：金鱼草属（flower sp.）、Arabidopsis（拟南芥）、花生属（花生）、茄属（颠茄）、芸苔属（油菜(rapeseed)）、长管弯头花、辣椒属（胡椒）、红花属（番红花）、菊苣属（菊苣）、橙属（桔、柠檬）、菊花、香瓜属（黄瓜）、曼陀罗属（曼陀罗）、萝卜属（胡萝卜）、毛地黄属（毛地黄）、草莓属（草莓）、老鹳草属（flower sp.）、大豆属(大豆)、向日葵属（向日葵）、hyscyamus、番薯属（牵牛花）、莴苣属（莴苣）、百脉根属（flower sp.）、番茄属（番茄）、majorana、锦葵属（棉）、木薯属、苜蓿属（苜蓿）、nemesia、烟草属（烟草）、onobrychis、天竺葵属（驱蚊香草）、牵牛属（flower sp.）、毛茛属（flower sp.）、莱服属（萝卜）、salpiglossis、千里光属（flower sp.）、芥属（白芥子）、茄属（土豆）、红花草属（红花）、豇豆属（绿豆、蚕豆）、葡萄属（葡萄）。

可通过数种途径对植物进行基因改造。最常用的方法是土壤杆菌介导的转化。在此方法中，在天然感染植物中根瘤土壤杆菌的植物基因组中插入致癌基因，使之发生改变。在实验条件下，将选择的基因

加入根土壤杆菌 Ti(肿瘤诱导的)细菌质粒的 T-DNA 中,使得 T-DNA 通过细胞自身内部转移机制转移到植物中时该基因整合到植物的染色体中。T-DNA 的唯一必要部分是其两小段(25 个碱基对)边界重复序列,其中至少有一段是植物转化所必须。编码促进植物生长之植物激素的细菌基因从 T-DNA 切除并用一段 DNA 序列取代,一般地,该序列包含:选择标记(如抗生素抗性基因,通常是卡那霉素抗性)、一个限制性酶切位点——含有限制性酶可切割 DNA 的特异核苷酸序列的位点,以及需要导入植物的目标基因(B. Tinland, 1996. T-DNA 整合入植物基因组中, Trends in Plant Science 1, 178-184; D. Grierson (ed.) 1991. 植物基因工程, Blackie, Glasgow)。土壤杆菌可加到培养的植物原生质体(去掉细胞壁的植物细胞)中,使之可以重新生成细胞壁,此时未转化细胞被抗生素杀死,而转化的细胞则产生抗性基因。然后利用标准的植物组织培养技术将存活的转化细胞再生成植物。在另一种替代技术中,将无菌皿或植物利于生长部分的片段与土壤杆菌一起置于流体培养基中,然后用激素诱导根再生从而使长在选

5 择培养基的苗再生。用于传送基因的第三种技术可能可以用于诸如拟南芥的一些植物中,该方法中土壤杆菌或“裸露”的 DNA 可融入种皮促使其转化(Clough SJ 和 Bent AF, 1998, 植物的浸润:一种用于土壤杆菌介导阿拉伯芥转化的简便方法. Plant J 16: 735-43)

10 20 近来一种用于植物基因工程的基因枪方法有了进一步的发展,开始广泛应用。此方法中,利用电子脉冲、气压或火药撞击将非常小生物活性 DNA 包被的钨或金颗粒(微发射体)高速推进植物细胞中。当颗粒穿过细胞时, DNA 溶解并整合入该细胞及其子代的基因组中。已证明此方法可以生成稳定的转化体(Christou, P.等, 1988, 用 DNA 包被的金颗粒稳定转化大豆愈伤组织, Plant Physiology 87: 671-674)。

25 此方法可应用于完整植株,尤其是在分生组织特别有效。它也能将 DNA 输送到核或线粒体(Johnston, S.A.等, 1988, 用微发射体通过轰击的方法将酵母中的线粒体转化, Science 240, 1538-41)和叶绿体(Svab, Z.等, 1990, 高等植物中质体的稳定转化, Proc Natl Acad Sci. USA 87, 8526-8530)。

30

植物基因工程的电穿孔方法获得的成功就少一些。这种方法中,

当用膜活性剂或高压直流电快速脉冲的电穿孔处理时，培养的原生质体摄取纯 DNA。一旦 DNA 进入原生质体即可整合入细胞的基因组中。然后用标准的组织培养技术再生转基因植物。

5 植物基因工程的微注射方法可能是最困难的。在此方法中，用非常细的玻璃针以在动物中所用的类似方法将 DNA 微注射入靶植物细胞中。这种技术费力低效，而且对于大量转基因植物来说不实用。

选择基因改造植物的方法常常取决于靶植物的种属和哪种方法已证明是有效的。

10 5.8 抗体的制备

可使用特异性识别本发明多肽或其片段的抗体来检测、筛选和分离本发明多肽或其片段，或编码来自于其他生物体类似酶的类似序列。例如，在一特定的实施方案中，能与 phyL 或 168phyA 或其片段特异性结合的抗体可用于各种体外检测分析，包括酶联免疫测定
15 (ELISA)、放射免疫测定、蛋白质印迹等，来检测生物材料如细胞、细胞培养基（如细菌培养基、哺乳动物细胞培养基、昆虫细胞培养基、酵母细胞培养基等）以及血液、血浆、血清、组织等样品中的本发明多肽或其片段、衍生物、类似物或变体，或具有与本发明多肽类似酶活性的类似分子。

20 可用本领域内任何合适的已知方法生产本发明多肽的特异性抗体。针对目标抗原的多克隆抗体可用本领域内已知的各种方法产生。例如，可将来源于本发明多肽的抗原施用于包括但不限于兔、小鼠、大鼠等在内各种宿主动物，以诱导生成含该抗原特异性多克隆抗体的抗血清。根据宿主种属的不同，可用各种佐剂来增强免疫应答，
25 包括但不限于弗氏（完全及不完全）佐剂、矿物凝胶如氢氧化铝、表面活性剂如溶血卵磷脂、pluronic polyols、聚阴离子、肽、油性乳剂、钥孔噉血蓝蛋白、二硝基苯酚以及用于人类的潜在性有用佐剂如 BCG（卡介苗）和小棒状杆菌。此类佐剂在本领域内亦是众所周知。

30 可利用本领域内已知的各种技术制备单克隆抗体，包括杂交瘤、重组及噬菌体显示技术或其组合。例如，单克隆抗体可用杂交瘤技术产生，这些杂交瘤技术包括本领域已知的技术及 Harlow 等《抗体：

实验手册》(Cold Spring harbor laboratory press,第二版, 1988) 及 Hammerling 等在《单克隆抗体与 T 细胞杂交瘤》563-681 页 (Elsevier, N.Y., 1981, 在此将两者整体均引入作为参考) 所述技术。这里所用术语“单克隆抗体”并不限于由杂交技术所产生的抗体。术语“单克隆抗体”指来源于单克隆的抗体, 包括任何真核、原核或噬菌体克隆, 而非指产生这些抗体的方法。

利用杂交瘤技术生成和筛选特异性抗体的方法是常规方法, 在本领域内众所周知。在一非限制性的实施例中, 可用目标抗原或表达该抗原的细胞免疫小鼠。一旦检测到免疫应答如在小鼠血清中检测到该抗原的特异性抗体, 就收集小鼠脾并将脾细胞分离出来。然后用熟知的方法将脾细胞与适当的骨髓瘤细胞融合。用限制性稀释法筛选和克隆杂交瘤。然后用本领域内已知方法分析可产生与该抗原结合的杂交瘤克隆。用阳性杂交瘤克隆在小鼠腹腔内接种, 产生腹水, 它通常含有高水平的抗体。

识别特异表位的抗体片段可用已知的技术生成。例如, Fab 和 $F(ab')_2$ 片段可利用诸如木瓜蛋白酶 (产生 Fab 片段) 或胃蛋白酶 (产生 $F(ab')_2$ 片段) 的酶通过免疫球蛋白分子的蛋白裂解产生。 $F(ab')_2$ 片段包含完整的轻链及重链的 CH1 区和铰链区。

本发明的抗体或其片段也可用本领域内任何已知的抗体合成方法生成, 尤其是用化学合成或优选重组表达技术生成。

编码抗体的核苷酸序列可由本领域内熟练技术人员可利用的任何信息所获得 (即从 Genbank、文献或常规克隆获得)。如果一个含编码特定抗体或其结合表位片段的核酸分子的克隆, 但特定抗体或其结合表位片段的序列是已知的, 则编码该免疫球蛋白的核酸分子可以化学合成, 或从适当的来源用可与 3' 及 5' 端杂交的合成引物通过 PCR 扩增获得 (如抗体 cDNA 文库或表达该抗体的任意细胞或组织如选择性表达抗体的杂交瘤细胞所产生的 cDNA 文库, 或是从这些细胞或组织分离的核酸分子, 优选含 poly A 的 RNA), 或利用需鉴定基因如来自于 cDNA 文库、编码该抗体的 cDNA 克隆的特异性寡核苷酸探针进行克隆获得。PCR 扩增产生的扩增核酸分子可利用本领域内任何已知的方法进一步克隆入可复制克隆载体中。

一旦抗体的核苷酸序列得以测定，即可用本领域内控制核苷酸序列的熟知方法如重组 DNA 技术、定点诱变、PCR 等（例如，可见 Sambrook 等，同上；及 Ausebel 等编，1998，当代分子生物学方案，John Wiley & Sons, NY 所描述的技术，在此将两者的整体引入作为参考），对抗体的核苷酸序列加以调节，以产生含有不同氨基酸序列的抗体，例如通过在抗体的表位结合区域或可增强或减弱抗体生物活性的任何抗体部分引入氨基酸替代、删除和/或插入。抗体的重组表达需要构建含编码抗体的核苷酸序列的表达载体。得知编码抗体或其轻链或重链或其中一部分的核苷酸序列之后，可用如前面章节所述、本领域内众所周知的方法通过重组 DNA 技术生成产生该抗体分子的载体。可使用本领域内熟练技术人员所熟知的方法，构建含编码抗体序列及适当转录和翻译控制信号的表达载体。例如，这些方法包括体外 DNA 重组技术、合成技术以及体内基因重组。然后将编码抗体重链可变区、轻链可变区以及同时编码重链可变区和轻链可变区、重链和/或轻链可变区中表位结合片段或抗体中的一个或多个抗原决定簇（CDRs）的核苷酸序列导入用于表达该抗体的适当宿主细胞。相应的，本发明包括了含聚核苷酸的宿主细胞，其中该聚核苷酸编码本发明多肽或其片段的特异性抗体。在一些实施方案中，宿主细胞为芽孢杆菌，优选枯草芽孢杆菌 MU331。

宿主细胞可与本发明的两个表达载体共转染，其中第一个载体编码重链衍生多肽，第二个载体编码轻链衍生多肽。这两个载体可含有相同的选择标记使重链及轻链多肽的表达水平相等，也可含有不同的选择标记以确保两质粒的维持。此外，可使用同时编码并表达重链及轻链多肽的单个载体。在这种情况下，轻链应置于重链之前以避免具有毒性的游离重链过量生成（Proudfoot, 1986, Nature, 322:52; 和 Kohler, 1980, Proc. Natl. Acad. Sic. USA, 77:2197）。重链及轻链的编码序列可包含 cDNA 或基因组 DNA。

在另一实施方案中，抗体也可用本领域内已知的各种噬菌体显示方法生成。在噬菌体显示方法中，抗体功能结构域在噬菌体颗粒的表面上显示出来，其中该颗粒携带有编码抗体的聚核苷酸序列。在一个具体的实施方案中，此类噬菌体可用于显示从所有组成部分或组合抗

体文库（如人或鼠的文库）表达的抗原结合域，如 Fab 及 Fv 或二硫键稳定的 Fvs。然后对表达与目标抗原结合之结构域的噬菌体加以筛选或鉴定，例如可使用标记抗原或固定或捕获于固体表面或珠子的抗原。在这些方法中所用的噬菌体通常为丝状噬菌体，包括 fd 和 M13。

5 抗原结合域以与噬菌体基因 III 或基因 VIII 蛋白的重组融合蛋白方式表达。可用于制造本发明免疫球蛋白或其片段的噬菌体显示方法实例包括 Brinkman 等，1995, *J. Immunol. Methods* 182: 41-50; Ames 等，1995, *J. Immunol. Methods* 184: 177-186; Kettleborough 等，1994, *Eur. J. Immunol.*, 24:952-958; Persic 等，1997, *Gene*, 187: 9-18; Burton

10 等，1994, *Advances in Immunology* 57: 191-280; NO. PCT/GF91/01134 号 PCT 申请; WO 90/02809、WO 91/10737、WO 92/01047、WO 92/18619、WO 93/11236、WO 95/15982、WO 95/20401 号 PCT 公告; 及 5,698,426、5,223,409、5,403,484、5,580,717、5,427,908、5,750,753、5,821,047、5,571,698、5,427,908、5,516,637、5,780,225、5,658,727、

15 5,733,743 和 5,969,108 号美国专利等公开的方法，并将其各自的整体在此引入作为参考。

正如上述文献所述，来自于噬菌体的抗体编码区可加以分离，并用于产生包括人类抗体的完整抗体或其他任何所需的片段，并可在任意需要的宿主细胞如哺乳动物细胞、昆虫细胞、植物细胞、酵母、细菌中表达，详情如下所述。例如，重组生成 Fab、Fab'和(Fab')₂ 片段的技术也可用本领域内已知的方法如 WO 92/22324 号 PCT 公告、Mullinax 等，1992, *BioTechniques* 12(6): 864-869 和 Swai 等，1995, *AJRI* 34: 26-34 及 Better 等，*Science*, 240: 1041-1043, 1988 (将其各自的整体在此引入作为参考) 所公开的方法来实施。可用于生成单

20 链 Fvs 和抗体的技术包括 4,946,778 和 5,258,498 号美国专利; Huston 等，1991, *Methods in Enzymology* 203:46-88; Shu 等，1993, *PNAS* 90: 7995-7999 和 Skerra 等，1988, *Science* 240:1038-1040 所公开的方法。

一旦本发明的抗体分子已用上述的任意方法生成，即可用本领域内任何已知的免疫球蛋白分子纯化方法加以纯化，例如层析（如离子交换层析，亲和层析尤其是 A 蛋白或 G 蛋白纯化后特异性抗原的亲

30 和层析，凝胶过滤层析）、离心、溶解度差异等，或纯化蛋白的其他

任意标准技术。而且，本发明的抗体或其片段可与本文所述或本领域内其他已知异源多肽序列融合，便于纯化。

就某些用途包括在人体内使用抗体及体外检测分析来说，优选使用嵌合的人源化或人类抗体。嵌合抗体是分子中不同部分来源于不同物种的分子，如含有来源于鼠单克隆抗体的可变区及来源于人免疫球蛋白的恒定区的抗体。生成嵌合抗体的方法在本领域内是已知的。例如参见 Morrison, *Science*, 229: 1202, 1985; Oi 等, *Biotechniques*, 4: 214, 1986; gillies 等, *J. Immunol. Methods*, 125: 191-202, 1989; 5,807,715、4,816,567 和 4,816,397 号美国专利，将其各自整体在此引入作为参考。人源化抗体是来自于非人类的抗体分子，它与含有一个或多个非人类来源抗原决定簇、人免疫球蛋白框架区及恒定区的抗原结合。人框架区中的框架残基常常用于来源于供体 CDR 的抗体中相应的残基所取代，以改变优选改善抗原结合。可用本领域内的熟知方法对框架替换进行鉴定，如模拟 CDR 和框架残基的相互作用来鉴别抗原结合中的重要框架残基，进行序列比较以鉴别特定位置中的不常见框架残基。例如，可参见 Queen 等的 5,585,089 号美国专利; Riechmann 等, 1988, *Nature*, 332:323, 将各自的整体在此引作作为参考。可用本领域内各种已知技术将抗体人源化，包括 CDR-移植 (EP 239,400; WO 91/09967 号 PCT 公告; 5,225,539、5,530,101 和 5,585,089 号美国专利) 和链替换 (5,565,332 号美国专利)，在此将各自的整体引入作为参考。

完整的人类抗体对于人类患者的治疗性处置来说特别理想。人类抗体可用本领域内各种已知的方法制造，包括前面所述、使用来源于人类免疫球蛋白序列的抗体文库的噬菌体显示方法。参见 4,444,887 和 4,716,111 号美国专利及 WO 98/46645、WO 98/50433、WO 98/24893、WO 98/16654、WO 96/34096、WO 9633735、WO 91/10741 号 PCT 公告，在此将各自的整体引入作为参考。

人类抗体也可用不能表达内源性功能免疫球蛋白、但其可表达人类免疫球蛋白基因的转基因小鼠生成。用于生成人类抗体的此类技术概貌参见 Longerg 和 Huszar, 1995, *Int. Rev. Immunol.* 13: 65-93。用于生成人类抗体、人类单克隆抗体及生成这些抗体的方案的此类技

术详细讨论可见 WO 98/24893、WO 92/01047、WO 96/34096、WO 96/33735 号 PCT 公告，0 598 877 号欧洲专利，5,413,923、5,625,126、5,633,425、5,569,825、5,661,016、5,545,806、5,814,318、5,885,793、5,916,771 和 5,939,598 号美国专利，在此将各自的整体引入作为参考。

5 此外，诸如 Abgenix, Inc.(Fremont, CA)、Medarex (NJ) 和 Genpharm (San Jose, CA) 之类的公司可负责用与上述方法类似的技术提供直接针对选定抗原的人类抗体。

识别选定表位的完整人类抗体可用一种称为“导向选择”的技术生成。在此途径中，选定的非人类单克隆抗体如小鼠抗体用于指导识
10 别同样表位的完整人类抗体选择 (Jespers 等, 1988, Bio/Technology 12: 899-903)。

与异源多肽融合或结合的抗体可在本领域内众所周知的体外免疫测定及纯化方法 (如亲和层析) 中使用。可参见 WO 03/21232 号 PCT 公告; EP 439,095 欧洲专利; Naramura 等, 1994, Immunol. Lett. 39:
15 91-99; 5,474,981 号美国专利; Gillies 等, 1992, PNAS 89: 1428-1432 和 Fell 等, 1991, J. Immunol. 146:2445-2452; 在此将各自的整体引入作为参考。

抗体也可粘附到固体支持物上，这对本发明多肽或其片段、衍生物、变体或具有本发明多肽类似酶学活性的类似分子进行免疫测定或
20 纯化时特别有用。此类固体支持物包括但不限于玻璃、纤维素、聚丙烯酰胺、尼龙、聚氯乙烯或聚丙烯。

5.9 检测分析

检测生物样品中本发明多肽或核酸分子的示例性方法涉及：从各
25 种来源获得生物样品，并使样品与一种能检测本发明多肽或核酸 (如 mRNA、基因组 DNA) 的化合物或试剂接触，以检测样品中是否存在本发明多肽或核酸。用于检测编码本发明多肽的 mRNA 或基因组 DNA 的优选试剂为标记的、可与编码本发明多肽之 mRNA 或基因组 DNA 杂交的核酸探针。例如，核酸探针可以是全长 cDNA 如 SEQ ID
30 NO:1 或 3 的核酸序列，也可以是其中一部分如长度上至少含有 15、20、25、30、50、100、250、500 或更多个连续核苷酸、可在严格条

件下充分与编码本发明多肽之 mRNA 或基因组 DNA 杂交的寡核苷酸。

用于检测本发明多肽的优选试剂是能与本发明多肽结合的抗体，尤其是带有可检测标签的抗体。抗体可以是多克隆抗体，优选单克隆抗体。可使用完整抗体或其片段（即 Fab 或 F(ab')₂）。关于抗体的详细描述也可参见 5.5 部分。

就探针或抗体来说，术语“标记的”包括可检测物质与该探针或抗体偶合（即物理连接）的探针或抗体直接标记，以及该探针或抗体通过反应与其他直接标记的试剂相连接的探针或抗体间接标记。间接标记的例子包括用荧光标记的第二抗体检测第一抗体和用生物素标记 DNA 探针的末端以便其可用荧光标记的链霉蛋白抗生素检测。本发明的检测方法可用于在体外和体内检测样品中的 mRNA、蛋白或基因组。例如，mRNA 的体外检测技术包括 RNA 杂交及原位杂交。本发明多肽的体外检测技术包括酶联免疫测定(ELISA)、蛋白质 Western 印迹、免疫沉淀和免疫荧光法。体外检测基因组 DNA 的技术包括 DNA 杂交。而且，在体内检测本发明多肽的技术包括将针对该多肽的标记抗体引入受试生物体中。例如，抗体可用放射性标记物标记，其在受试生物体中的存在与否及存在的位置可用包括放射自显影在内的标准成像技术检测。

在一个特定的实施方案中，这些方法还包括从对照受试者获取对照样品，使对照样品与能检测本发明多肽或编码本发明多肽的 mRNA 或基因组 DNA 的化合物或试剂接触，检测样品中多肽及编码该多肽的 mRNA 或基因组 DNA 存在与否，并将对照样品中本发明多肽或编码本发明多肽的 mRNA 或基因组 DNA 的存在度与待测样品中本发明多肽或编码本发明多肽的 mRNA 或基因组 DNA 的存在度相比较。

本发明还包括检测待测样品中本发明多肽或核酸的试剂盒。

例如试剂盒可包含能检测样品中本发明多肽或编码该多肽的 mRNA 的标记化合物及测定样品中多肽或 mRNA 含量的工具（例如与该多肽或寡核苷酸探针结合的抗体，其中该探针与编码该多肽的 DNA 或 mRNA 结合）。试剂盒也包括使用指南。

对于基于抗体的试剂盒来说，它可以包含：（1）第一抗体（如

粘附于固定支持物的抗体)，与本发明多肽结合，有时还可包括(2)不同的第二抗体，与多肽或第一抗体结合，并与可检测试剂相连。

对于基于寡核苷酸的试剂盒来说，它可以包含：(1)可与编码本发明多肽的核酸序列杂交的寡核苷酸，如可检测的标记寡核苷酸，或(2)一对用于扩增编码本发明多肽之核酸的引物。试剂盒还包括缓冲试剂、防腐剂或蛋白稳定试剂。试剂盒也可包括检测该可检测试剂的组分(如酶或底物)。试剂盒也可包括一种对照样品或一系列的对照样品，它们可以加以测定并与待测样品相比较。该试剂盒的每一组分通常密封于单独的容器里，这些各种各样的容器与使用指南置于单一的包装中。

5.10 植酸酶及转基因植物的商业应用

正如前面所陈述的那样，肌醇六磷酸盐作为动物膳食的主要成分大量存在于食物来源中。但是，包括家禽动物和鱼在内的单胃动物不能利用此磷源，而且当它排入外界环境中时肌醇六磷酸盐导致大量对生态系统的污染问题。环境的改变可能不是马上就能发现，因为它们主要发生于食物链的最下游，但随着污染的继续，这些改变的效应将累积起来并向整个生态系统渗透，从而导致整个生态系统的永久损伤。因此，本发明的多肽对单胃动物没有毒性，利用本发明多肽的超表达系统大量生产本发明的多肽并制备含有本发明多肽的动物饲料，具有极大的商业价值。本发明的多肽作为动物饲料为单胃动物所用将减少未利用磷排泄入环境中，从而使环境污染降至最低。

而且，为了促进植物的生长，肥料中常常加入磷酸盐，从而进一步加重环境污染。尽管土壤中确实存在磷来源，但它以肌醇六磷酸形态被封闭住并不能为植物所用。因此，带有与本发明一致的嵌合基因构建体的转基因植物，在细胞内和/或细胞外表达植酸酶，在磷的利用方面具有很大的优势，否则这些磷既不能为植物本身所用也不能被动物利用。也就是说，转基因植物的高效利用磷不仅有助于减少磷带来的环境污染，而且可增强植物的生长，包括开花及结果行为，这意味着在农业和园艺业有着重要的应用。而且，在动物饲料中引入带有本发明自身表达的细胞内植酸酶的转基因植物，有助于动物利用磷，

使其排泄物给环境带来的污染减少。

6. 实施例

下面的实施例讲解了所述植酸酶及抗体的克隆、生成、分离和鉴定。这些实施例不具有限制性。

6.1 phyL 基因的分子克隆

来自于地衣芽孢杆菌的 phyL 基因克隆策略如图 3 所示。地衣芽孢杆菌细胞从商业途径获得 (ATCC 号: 10716)。细菌细胞于 37°C 在营养琼脂板 (2.5% [w/v] 营养肉汤粉末, 1.5% [w/v] 细菌学琼脂) 上生长, 并作为兼并 PCR 反应的模板。兼并寡核苷酸 (SEQ ID NO:5 和 SEQ ID NO:6) 根据 PhyK (Kim Y.O., Lee J.K., Kim H.K., Yu J.H. 和 Oh T.K., 1998, 从芽孢杆菌 DS11 克隆热稳定的植酸酶基因 (*phy*) 及其大肠杆菌中的超表达, FEMS Microbiology Letters, 162:182-191)、PhyC (Kerovuo J., Laurarus M., Nurminen P., Kalkkinen N. 和 Apajalahti J., 1998, 一种来自于枯草芽孢杆菌的全新植酸酶的分

10 离、鉴定、分子基因克隆及测序, Applied and Environmental Microbiology, 64(6):2079-2085) 和 168phyA (SEQ ID NO:4) 间的保守氨基酸序列设计, 并作为 PCR 反应的引物。扩增在 PCR 机器 (Robocycler gradient 40, Stratagene, USA) 中以 {94°C 45 秒, 50°C 45 秒, 72°C 2 分 30 秒} 的条件进行 30 个循环。将目标 PCR 产物从 2% (w/v) 琼脂糖凝胶切下来, 并用 GeneClean III 试剂盒 (Qbiogene, Inc, CA) 纯化。将纯化产物克隆入 pBSSK, 用 X-gal/IPTG 诱导, 进行氨苄青霉素抗性选择。在补充有 100µg/mL 氨苄青霉素的 LB 肉汤中培养的阳性克隆质粒用 Quantum mini-prep 试剂盒 (Bio-Rad, 香港) 提取, 并进行测序 (MWG Biotech AG, 德国)。

15 20 25

用基因组纯化试剂盒 (Promega, 香港) 提取地衣芽孢杆菌的基因组 DNA, 并用紫外分光光度计于 260nm 测定 DNA 浓度。两组基因组 DNA (各 20µg) 分别用 10 单位 Hae III (Boehringer Mannheim, 香港) 和 10 单位 Sau3 AI (Boehringer Mannheim, 香港) 部分限制性酶消化 1 小时。消化的 DNA 纯化并稀释到 1µg/mL, 然后用 T4 DNA 连接酶 (Life Technologies, 香港) 环化。环化 DNA 用酚-氯仿提取

30

和乙醇沉淀进行纯化。

正义和反义寡核苷酸 (SEQ ID NOS: 7-12) 从兼并 PCR 所产生序列和 5'和 3'末端旁侧设计。反向 PCR 以部分消化的基因组 DNA 为模板, 在 {94°C 45 秒, 55°C 45 秒, 72°C 两分钟} 条件下进行 30 个循环。阳性 PCR 产物用乙醇沉淀并用相应的限制性酶 (Boehringer Mannheim, 香港) 消化, 然后亚克隆入 pBSSK 的 EcoRI 和 Bam HI 位点。筛选出阳性克隆并如上所述提取出来。对克隆进行测序, 将序列数据汇总并用包括 MAC DNASIS (Hitachi, 日本) 和 DNA Strider (Christian Marck, Service de Biochimie, Department de Biologie, Institut de Recherche Fondamentale, CEA, 法国) 在内的 DNA 处理软件分析。用 GeneWorks forMac (Intelligenetics, Mountain View, CA) 完成种类分析。

phyL 的 DNA 序列及推导出的氨基酸序列分别如 SEQ ID NOS: 1 和 2 所示。在起始密码子 ATG 上游 12bp 处发现一个带有共有序列 GGAGG 的假想核糖体结合位点。从核苷酸序列所推导的氨基酸序列显示为一个 381 个氨基酸残基的蛋白, 比其他三种枯草芽孢杆菌植酸酶要短。用 BLAST 检索将 DNA 序列及其推导出的氨基酸序列与 NCBI 数据库相比。发现在蛋白水平上 *phyL* 有 64% 与 *phyK* 相同, 有 65% 与 *phyC* 相同, 有 69% 与 168*phyA* 相同, 有 64% 与 *phyK* 相同, 而在 DNA 水平上有 79% 与 *phyK* 相同, 有 79% 与 *phyC* 相同, 有 90% 与 168*phyA* 相同。与这三种枯草芽孢杆菌植酸酶类似, *phyL* 编码的植酸酶不具备高度保守的 RHGXRXP 序列基序, 而该基序在所有的真菌及大肠杆菌植酸酶中均可鉴定出来。

25 6.2 *phyL* 所编码植酸酶的超表达

PCR 引物 (SEQ ID NOS: 13 和 14) 设计从 *phyL* 基因翻译起始密码子 ATG 和终止密码子之间的编码区旁侧设计。编码成熟酶的基因片段用 *Pfu* 聚合酶 (Promega, WI) 扩增并亚克隆入 pSGt 表达载体中, 其中 pSGt 通过将地衣芽孢杆菌 α -淀粉酶基因的终止子亚克隆入 pSG 载体而得以构建。这样, *phyL* 基因处于 ϕ 105 原噬菌体载体记启动子的调控之下。此载体命名为 pSGt-pL (图 4B)。

通过电穿孔将质粒导入大肠杆菌 JM109 株中。用氨苄青霉素抗性筛选菌落，并用 Quatum Mini-prep 试剂盒 (BIO-RAD, 香港) 将阳性克隆加以提取、纯化。

- 5 为了将重组质粒转化入用于酶生产的宿主株，用 Osborne 等人所述方法 (Osborne M. S., Craig R.J 和 Rothstein D.M., 1985, 可利用温和噬菌体 $\phi 105$ 调控元件的枯草芽孢杆菌热诱导转录系统, J. of Bacteriology 16: 1101-1108) 制备枯草芽孢杆菌 MU331 活性细胞。将转化体分开种在补充有氯霉素和红霉素的琼脂上，筛选转化体。对两种抗生素均有抗性的克隆进一步用 $\phi 105$ 特异性引物和 *phyL* 特异性引物 (SEQ ID NO:14) 通过 PCR 进行筛选。以此方式建造用于酶特征研究中使用的重组株 pL-01，在 30% (v/v) 甘油中于 -80°C 冻存。

细菌培养和酶生产的培养基如下：

脑心浸渍肉汤由下列成分组成：

小牛脑浸渍固体	12.5%
牛心浸渍固体	5%
蛋白酶胨	10%
葡萄糖	2%
氯化钠	5%
磷酸氢二钠	2.5%
酵母中性 pH 提取物	2.5%

- 15 将 pL-01 株以划线方式接种到补充有 5 $\mu\text{g}/\text{mL}$ 氯霉素的 LB 琼脂板上。第二天，挑出单个克隆并转移到补充有 5 $\mu\text{g}/\text{mL}$ 氯霉素的细菌培养基中。细胞以 280 rpm 转速摇动培养直到 OD_{600} 读数达到 7.0。取 1 mL 培养物转移到 15 mL 不含抗生素的细菌培养基中。让细胞长到 OD_{600} 读数为 4.5，然后在 50°C 水浴中轻轻振摇 5 分钟进行热诱导。
- 20 诱导后于不同时间点采样。

如无其他说明，所有的酶纯化步骤均在 0°C-4°C 进行。收集生长于细菌培养基中的细菌，于 3000rpm 离心 30 分钟。将收集上清与三倍体积的冰乙醇 (-20°C) 混合，于 4°C 搅拌沉淀过夜。6000rpm 离心 30 分钟收集沉淀，自然干燥，重悬于 100 mM 的 pH7、含 5 mM CaCl_2

的 Tris-HCl 中。重悬的酶过 NAP-10 葡聚糖凝胶过滤柱 (Marsha Pharmacia, 香港) 进行缓冲液交换。在酶分析之前将用预设分析缓冲液洗脱的酶于 -20°C 保存。

phyL 所编码成熟植酸酶用 SDS-PAGE 测定的分子量约为 47 kDa (图 5B)。热诱导后 5 小时的引集液所产生的 *phyL* 所编码植酸酶的产量可达 175 mg/L。酶活性达到每 mL 培养基 4.1 单位和每毫克所用酶 23.6 单位, 其中 1 个单位酶活性定义为在指定条件下释放 1 μmol 无机磷酸盐所需要的酶用量 (图 6)。与以前研究中所测定的芽孢杆菌植酸酶活性 (Powar 和 Jagannathan, 1982, 来自于枯草芽孢杆菌的肌醇六磷酸特异性磷酸酶的纯化和性质, J. of Bacteriology 151 (3): 1102-1108) 相比, 本发明的新 *phyL* 酶在培养时间缩短 14 倍的同时其活性提高了 17 倍。该酶用双向 SDS-PAGE 测定的等电点约为 5.1。

6.3 168*phyA* 所编码植酸酶的生产及其活性

通过检索枯草芽孢杆菌基因组的序列同源性发现枯草芽孢杆菌 168 的基因组中有一个与两种已公开的枯草芽孢杆菌植酸酶具有高度序列同源的开放阅读框 (ORF)。设计 PCR 引物 (SEQ ID NOS: 15 和 16) 扩增此 ORF 旁侧的基因片段, 将 PCR 产物亚克隆入 pSG 表达载体, 构建 pSG-pA。在此构建体中, *168phyA* 基因的旁侧为 $\phi 105$ 启动子和 *168phyA* 基因的天然终止子 (图 4A)。按前文第 6.2 部分所述将 pSG-pA 质粒转化入枯草芽孢杆菌 MU331 株组件中, 构建重组株 pA-01。阳性克隆用 $\phi 105$ 特异性引物和 *168phyA* 特异性引物 (SEQ ID NO:16) 通过 PCR 进行筛选。然后将其用于前文第 6.2 部分的酶生产。

168phyA 所编码成熟植酸酶用 SDS-PAGE 测定的分子量约为 44 kDa (图 5A), 进一步确认了从氨基酸序列 (SEQ ID NO:4) 计算的分了量。热诱导后 4 小时的引集液所产生的 *168phyA* 所编码植酸酶的产量可达 246.2 mg/L。酶活性达到每 mL 培养基 5.3 单位和每毫克所用酶 36.8 单位 (图 6)。与以前研究中所测定的芽孢杆菌植酸酶活性 (Powar 和 Jagannathan, 1982, 同上) 相比, 本发明的新 168*phyA* 酶在培养时间缩短 18 倍的同时其活性提高了 22 倍。该酶用双向

SDS-PAGE 测定的等电点约为 5.0。

为了提高酶生产的收率，用 pA-01 进行 2 升 (2-L) 规模的原料分批发酵。在此发酵过程中，用 pH-状态法控制碳源 (葡萄糖) 和氮源 (胰脲) 的加入。诱导后 6 小时的酶活性达到 28 EU/mL 培养物，
5 与产述简单振摇所产生的酶活性相比提高了 5 倍。

6.4 植酸酶活性的测定

在不同的 pH 和温度下测定指定缓冲液中的酶活性。用于 pH 试验的缓冲液包括 100 mM 柠檬酸-HCl, pH 为 3.5 和 6.5; 100 mM 醋酸-HCl, pH 为 4.5~6.0; 100 mM Tris-HCl, pH 为 7~8.5, 100 mM 甘氨酸-NaOH, pH 为 9、9.5 和 10.5。上述所列缓冲液均添加 5 mM CaCl_2 。
10 用标准 Bradford 蛋白测定法 (BIO-RAD, 香港) 测定酶的浓度。纯化后的酶用测定缓冲液稀释, 按 Engelen 等所述方法 (1994, 同上) 进行比色测定, 只是将测定降到 1 mL。简单地说, 将酶用各种指定的测定缓冲液稀释到 200 μL 总体积。将 4 mL 用蒸馏水稀释到 10 mM 的肌醇六磷酸钠加到该酶中, 将 *168phyA* 及 *phyL* 所编码植酸酶的混合物分别于 55°C 和 65°C 孵育 30 分钟。在反应液中加入 0.4 mL 新鲜配制的终止液以猝灭酶活性。5 分钟后, 取 200 μL 已猝灭的混合物转到 96 孔 ELISA 板 (Nunc, 丹麦) 于 405 nm 处进行光密度测定。
15

20 中性 pH 下进行的温度试验 (图 7A) 表明 *168phyA* 及 *phyL* 所编码的植酸酶均表现出宽的最佳温度范围, *phyL* 所编码的植酸酶在 65°C、*168phyA* 所编码的植酸酶在 55°C 达到峰值。图 7B 显示了各个最佳温度下 pH 对植酸酶在指定测定缓冲液 (如上所述) 中活性的影响。两种植酸酶均在中性 pH 时表现出最高活性。

25 至于热稳定性试验, 将稀释的酶等分, 在 70°C -90°C 之间的不同高温下孵育 10 分钟, 室温冷却 1 小时使蛋白重新折叠, 然后进行活性实验。

30 结果发现, *phyL* 所编码的植酸酶即使在低 Ca^{2+} 浓度 (1mM) 的条件下高温变性后可恢复 60-70% 的正常活性。它甚至可以经受高到 95°C 的变性, 此时仍保持超过 50% 的原始活性。

枯草芽孢杆菌 *168phyA* 所编码的植酸酶即使在高 Ca^{2+} 浓度 (5

mM)的条件下高温变性后可恢复 60%-60%的原始活性。它甚至可以经受高到 95°C 的变性,此时仍保持超过 46.7%的原始活性。但是, 168phyA 在低 Ca²⁺浓度 (1 mM) 所保持的活性比高 Ca²⁺浓度 (5mM) 少 20%。

5

6.5 转基因植物的产生

稻是全世界尤其是亚洲的重要农作物。在中国,稻在总农作物谷类产量中占 42%,种植面积占了 29%。稻是单子叶植物,并依气候及生长条件的不同,一些热带变种可在一年内完成三次生命周期。当在 24°C 以上温度,光周期长于 14 小时生长时,从秧苗到开花需要 60 天。从开花到收割种子还需要 30 天。一棵稻在完成一个生命周期时一般可产 500 粒种子。

烟草是植物转化中很好的模型系统,因为它的转化率高,在组织培养中易于繁殖。烟草因其阔叶被认为是双子叶植物,其阔叶具有很高的商业价值。烟草为一年生植物,可在 120 天内完成其生命周期。当在 22°C 以上温度,光周期长于 14 小时生长时,从秧苗到第一次开花需要 96~100 天。从一棵烟草植物可获得 20-30 粒果实,每粒果实重约 0.3~0.4 克,含 1000 粒以上的种子。一般地,一棵烟草植物在第一朵花盛开 30 天后产生能发芽的种子。

在植物细胞如稻和烟草细胞中导入含编码本发明植酸酶的基因的嵌合构建体,由于提高了储存于植物和/或土壤中的无机磷酸盐的利用率,可以提高植物的生长速度,从而缩短成熟及开花所需的时间。

6.5.1 植物表达载体的构建

构建植物表达载体的策略如图 8 所示。用一对该基因旁侧的引物 (SEQ ID NOS: 17 和 18) 通过 PCR 扩增 168phyA 基因。用 168phyA 基因在 BamHI 及 SacI 限制位点取代 pBI221 载体的大肠杆菌β-D-葡萄糖醛酸酶 (GUS) 基因,以获得 phyA-221 中间载体。用 HindIII 和 EcoRI 消化含有用于筛选、由 CaMV 35S 启动子驱动的潮霉素抗生基因的 pCAMBIA1300 双链载体 (GeneBank 登记号为 AF234296), 与经 HindIII/EcoRI 线性化的 phyA-221 中间载体连接,生成新的表达

载体 pCX-168phyA (图 9)

6.5.2 土壤杆菌培养和转化

用 Hofgen 和 Willmitzer 所述方法 (1988, 用于土壤杆菌转化的
5 活性细胞储存, Nucleic Acids Res. 16: 9877) 将两个单独的 pCX-
168phyA 克隆 (克隆 04 和 13) 转化入 EHA 105 根瘤土壤杆菌。将一个
集落接种到含 50 $\mu\text{g}/\text{mL}$ 卡那霉素和 25 $\mu\text{g}/\text{mL}$ 氯霉素的 20 mL LB
液体培养基中, 于 28 $^{\circ}\text{C}$ 快速振荡培养 2 天, 直到培养悬液的 OD600
10 约为 0.8~0.1。培养物于 4000 rpm 离心 10 分钟, 沉淀重悬于 20 mL
的 AMM 培养基 (见表 1, 见下文) 中用于植物转化。

6.5.3 转基因稻的产生

用日本稻 *Oryza sativa* L. 的驯化品系之一中华 11 号实施试验。将
成熟的种子消毒, 在 N6D 培养基上发芽两周。将鳞部的愈伤组织在
15 N6D 培养基上再传代培养一周。三周龄的愈伤组织在细胞悬液中浸
泡 20 分子, 用无菌滤纸将过量的细菌吸去。与细胞粘附的愈伤组织
转移到置于 N6DC 培养基上的滤纸上, 于 25 $^{\circ}\text{C}$ 避光共培养 3 天。共
培养后, 用含 500 mg/mL AAD 羧苄西林的培养基 (见表 1, 下文)
洗涤受感染的愈伤组织 3 次, 用无菌滤纸干燥, 然后转移到 N6DS1
20 培养基上 (见表 1, 下文)。

愈伤组织在 N6DS1 培养基上培养 2 周, 然后转移到 N6DS2 培养
基上 (见表 1, 下文) 进行 3~4 周的进一步选择。耐受的愈伤组织转
移到 HIGROW 培养基上 (见表 1, 下文) 进行 10 天避光预分化, 然
后各自转移到 MSRS 培养基 (见表 1, 下文) 中在生长小室中于
25 24 $^{\circ}\text{C}$ ~26 $^{\circ}\text{C}$ 进行发芽再生, 期间用来自荧光灯管、光通量密度为 120
 $\mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ 的光线照射 16 小时。再生植物转移到 MSCN 培养基上 (见
表 1, 下文) 进一步生长。当耐受植物约 10 cm 高时, 将其转移到温
室的土壤中, 直到生长成熟。

6.5.4 转基因烟草的产生

烟草驯化品种 “GeXin NO 1” (红花烟草) 的种子用 30% (v/v)

的 Clorox 消毒 15 分钟，用无菌水洗 5 次，在 Murashige 和 Skoog 基础培养基（MS 培养基，Sigma M-9274, St. Louis, MO）发芽。秧苗在相同的培养基上于 22°C 进行 16 小时光/8 小时暗期培养，光线为来自荧光灯管、光通量密度为 $50 \mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ 的光线。

- 5 将一个含目标基因的 EHA 105 根瘤土壤杆菌集落接种到含 50 $\mu\text{g}/\text{mL}$ 卡那霉素、25 $\mu\text{g}/\text{mL}$ 氯霉素和 50 $\mu\text{g}/\text{mL}$ 利复霉素的 20 mL LB 液体培养基中，于 28°C 快速振荡孵育 2 天。将烟草的叶切成约 1 平方厘米的小片，在 20 mL 细胞悬液中浸泡 2~3 分钟。用无菌滤纸将过量的细菌吸去，将外植体转到补充有 2 mg/L 6-BA（MSB 培养基）
- 10 的 MS 培养基上（见表 1，下文）于避光于 25~26°C 培养 2 天。两天的共培养后，外植体转到含 30 mg/L 潮霉素及 500 mg/L 羧苄西林的 MSB 培养基在标准光照条件下于 26°C 发芽再生 3~4 周。当长到约 1 cm 长时将耐受的芽切下来，转移到含 25 mg/L 潮霉素及 500 mg/L 羧苄西林的 MS 培养基中让其生根。当耐受植物约 8cm 高时，转移到温室的土壤中，直到生长成熟。每个 pCX-168phyA 克隆（004 和 013）
- 15 可生成 4 棵植物，分别命名为 0041、0042、0043、0044 和 0131、0132、0133、0134。

表 1 用于组织培养和植物转化的培养基

培养基	成分
N6D	N6 ¹ , 500 mg/L 酪蛋白, 30 g/L 蔗糖, 2.5 mg/L 2,4-D, 2.5g/L phytigel, pH 5.7
N6DC	N6D 培养基加 10 g/L 葡萄糖、100 $\mu\text{mol}/\text{L}$ 乙酰丁香酮, pH 5.2
AAM	AA ² , 500 mg/L 酪蛋白, 68.5g/L 蔗糖, 36g/L 葡萄糖, 200 $\mu\text{mol}/\text{L}$ 乙酰丁香酮, pH 5.2
AAD	AA, 30 g/L 蔗糖, 2,4-双氯苯氧乙酸 (2,4-D), pH 5.7
N6DS1	N6D 培养基加 500 mg/L 头孢噻肟及 50 mg/L 潮霉素
N6DS2	N6D 培养基加 300 mg/L 头孢噻肟及 50 mg/L 潮霉素
HIGROW	培养基 (Gibco BRL 10924-017) 加 2.5 g/L phytigel
MSRS	MS 培养基 ³ (sigma M-9274) 加 2 mg/L 6-苄氨基嘌呤

	(6-BA), 0.2 mg/L α -萘酸 (NAA), 0.5mg/L 玉米素 (ZT), 200 mg/L 头孢噻肟及 50 mg/L 潮霉素, pH 5.8
MSCN	MS 培养基加 0.2 mg/L NAA 及 0.5 mg/L 矮壮素 (氯化氯胆碱(CCC))

¹ Zhu Z-Q, Wang J-J, Sun J-S, Xu Z, Yin G-C, Zhu -Y, Bi F-Y, 1975, 通过氮来源比较试验建立稻的花药培养的高效培养基, Sci Sin 18: 659-668 (英文摘要)。

² Hiei Y, Ohta S, Komari T, Kumashiro T., 1994, 土壤杆菌介导的高效稻 (*Oryza sativa* L.) 转化及 T-DNA 的边界序列分析, Plant J 6: 271-282。

³ Murashige T.和 Skoog F., 1962, 一种用于烟草组织培养快速速生长及生物测定的修正培养基, Plant Pysiol. 15: 473-479。

10 这些转基因烟草的克隆如下部分所述进行鉴定和描述。

6.5.5 DNA 制备及 PCR 分析

用 PCR 检测特异的潮霉素 B 抗性基因 DNA 序列。用下面的方法从已转化和未转化 (对照) 的植物制备基因组 DNA: 植物叶称重, 匀浆前于液氮中冷冻。50 mg 植物组织中加入 600 μ L 提取缓冲液 (100 mM Tris-HCl, pH 8.0, 50 mM EDTA, 500 mM NaCl 和 10 mM β -巯基乙醇), 混合物煮沸 10 分钟。混合物置冰上冷却, 以 16000 g 离心 15 分钟。然后将上清中的植物基因组 DNA 用 0.1 倍体积的 10 M 醋酸铵和 2.0 倍体积的无水乙醇于 -20°C 沉淀 2 小时。以 16000 g 离心 30 分钟将基因组 DNA 沉淀下来。DNA 沉淀用 75% (v/v) 洗涤, 重新溶于水中。一般地, 0.1 g 植物组织可得 50 μ g 植物基因组 DNA。

PCR 反应由 {94°C 变性 30 秒, 56°C 退火 40 秒, 72°C 延伸 60 秒} 的条件循环 30 次构成。潮霉素 B 抗性基因的正义和反义引物分别为: 5'-CTACAAA-GATCGTTATGTTTATCGGCA-3' (SEQ ID NO:19) 和 5'-AGACCAATGCGGAG-CATATACG-3' (SEQ ID NO:20), 并扩增出一个潮霉素 B 抗性基因 (E00287) 序列 162~803 之间的 641 bp 片段。正义及反义引物 (SEQ ID NOS: 17 和 18) 用于扩增 *168phyA* 基因。转基因烟草的 PCR 筛选结果如图 10 所示。如图 10A 所示, 所

有转基因烟草均产生了潮霉素 B 抗性基因的 PCR 产物，包括用 pCAMBIA 1300 转化的植物。相反，用 pCX-168phyA 而不是用 pCAMBIA 1300 转化的植物产生了 *168phyA* 基因 PCR 产物。这些结果表明，pCX-168phyA 和 pCAMBIA 1300 成功地用上述方法导入了
5 相应烟草的基因组中。

6.5.6 RNA 印迹(Northern blotting)

用从植物叶提取的总 RNA 作为 RNA 印迹检测的模板。不同转基因植物个体的 RNA 上样到 1%琼脂糖凝胶适当地分离开来，通过毛
10 细管样作用过夜转移到尼龙膜上。以不含细菌信号肽、DIG 标记的 *168phyA* 基因 cDNA 的探针进行 RNA 杂交。所有涉及试剂均购自 Roche Diagnostics (香港)，所有程序均按厂商的手册操作。转基因烟草系 F0 及 F1 的 RNA 印迹结果分别如图 11 和 16 所示。F0 系 0042、0043 及 0134 的蛋白提取物中检测到了植酸酶。如图 16 所示，mRNA 表
15 达遗传到了 0042 和 0134 的 F1 系。

6.5.7 蛋白质印迹(Western blotting)

用烟草叶提取的总蛋白进行蛋白质印迹试验。用 SDS-PAGE 在
20 10%丙烯酰胺凝胶上适当分离各个蛋白质样品，于 4°C 以 100V 电压转膜 1 小时，转到硝化纤维素膜上。用枯草芽孢杆菌中超表达的 *168phyA* 基因所编码植酸酶免疫兔产生多克隆抗-168phyA 抗体。多克隆抗体用野生型烟草吸附，然后作为探针加到样品蛋白。用 NBT/BCIP 底物根据厂商的方案进行信号检测。

F0 和 F1 转基因烟草的蛋白质印迹结果分别如图 12 和 17 所示。
25 从 F0 系 0042、0043 和 0134 的蛋白提取物中检测到了植酸酶。如图 17 所示，植酸酶的表达遗传到了 0042 的 F1 系。一般地，转基因蛋白在转基因植物提取物中检测不到，因为其表达水平太低。例如，烟草叶中表达的重组真菌植酸酶 (*phyA*) 经过层析纯化 (Ullah 等, 1999, 同上) 后仅能达到蛋白质印迹的信号。因此，图 12 和图 17 所示蛋白
30 质印迹中植酸酶显影表明转基因植物中的植酸酶表达水平很高。

6.5.8 DNA 印迹(Southern blotting)

使用植物叶提取的基因组 DNA 进行 DNA 印迹。在 RNase 活性存在的条件下用 HindIII 限制性内切酶切割基因组 DNA。不同转基因植物个体的消化 DNA 上样到 0.7%琼脂糖凝胶，进行适当分离，通过毛细管样作用过夜转移到尼龙膜上。以不含细菌信号肽、DIG 标记的 *168phyA* 基因 cDNA 的探针进行 RNA 杂交。所有涉及试剂均购自 Roche Diagnostics (香港)，所有程序均按厂商的手册操作。

如图 15 所示，在 F1 系 (0042 和 0134) 中检测到了特异性条带，但在对照系中未检测到。0042 的 F1 系 (42-1 和 42-2) 含单个拷贝的基因，而 0134 (134-1) 的 F1 系则有两个基因拷贝。

6.5.9 对照烟草及转基因烟草的表型

烟草是植物转化中很好的模型系统，因为它的转化率高，在组织培养中易于繁殖。烟草因其阔叶被认为是双子叶植物，其阔叶具有很高的商业价值。烟草为一年生植物，可在 120 天内完成其生命周期。当在 22°C 以上温度，光周期长于 14 小时生长时，从秧苗到第一次开花需要 96~100 天。从一棵烟草植物可获得 20-30 粒果实，每粒果实重约 0.3~0.4 克，含 1000 粒以上的种子。一般地，一棵烟草植物在第一朵花盛开 30 天后产生能发芽的种子。

转化后转基因植物的表型也继承下来了，其表型如表 2 所示。一般地，用植酸酶基因转化的植物在第一朵花盛开时高 101~130cm，比用载体转化的植物第一朵花盛开时的高度 (138~158cm) 矮。开花后，用植酸酶基因转化的植物其高度 (142~168cm) 仍比用载体转化的植物 (182-206) 矮。即使用植酸酶基因转化的植物一般要矮一些，其通常比用载体转化的植物 (6 枝) 具有更多的花枝 (每棵植物 8~10 枝) (见表 2 及图 13)。从形态学上看，植酸酶转化的植物中有 4 种发育成不止一个侧枝 (图 14)。至于花蕾的数目，用植酸酶基因转化的植物其花蕾数比仅用载体转化的植物多 (见表 2)。至于花期，用植酸酶基因转化的植物其花期 (50 到超过 88 天) 比仅用载体转化的植物 (35~37 天) 长 (见表 2)

6.5.10 转基因烟草幼苗在磷酸盐缺乏条件下生长

对照及转基因（42-1 系）F2 烟草种子用 33%（v/v）的 Clorox 消毒 15 分钟，用无菌水洗 5 次，接种到 MS 培养基（包括 1.25mM 磷酸盐和 30g/L 蔗糖）中。15 天后，幼苗转移到改进 MS 培养基（不含磷酸盐，蔗糖浓度降到 5g/L）中再培养 17 天。发现在磷酸盐饥饿试验中，与对照幼苗相比，转基因系的生物量多于对照系（图 18）。

6.5.11 转基因烟草幼苗在低磷酸盐条件下生长

表面消毒的对照种子及转基因烟草种子种到含 20 mL 改进 MS 琼脂培养基（标准 MS 培养基，另外加入 10 g/L 蔗糖、 10^{-3} M 或 10^{-5} M 磷酸盐）的 petri 皿（每个 9cm 皿 60 粒种子）中。20 天后，从高磷酸盐（ 10^{-3} M）或低磷酸盐（ 10^{-5} M）板中的 9 棵幼苗转移到含不同浓度磷酸盐（ 10^{-3} M、 10^{-4} M 或 10^{-5} M）的组织培养盒（7cmX7cm, 50mL 改进 MS 琼脂）中。在收获、测定干重前幼苗再生长 30 天。每棵植物单独称重。每组称 18 棵，其平均值如图 19 所示。转基因系 0042 比对照系长得粗，尤其是磷的利用受到限制时，两者的差别具有统计学意义。另外，也完成了幼苗在液体培养基中生长的试验。简单地说，烟草种子种到 MS0 基质中（30g 蔗糖，1.25mM 磷）培养 10 天。每个系的 25 棵植物分成 5 组，每组测定干重。如图 20 所示，在低磷酸盐条件下转基因系（0042 和 0134）的干重比对照系重。

6.5.12 用 HPLC 分析植酸酶的体外活性

5 g 嫩叶组织浸入 10 mL 预冷的提取缓冲液（0.1 M Tris-HCl, pH7.0, 1 mM 苯甲磺酰氟及 0.1 mM CaCl_2 ）。以 12000 g 离心 20 分钟收集水相中的可溶蛋白，用 Bradford 蛋白质测定法（Bio-Rad）定量蛋白浓度。为了评价植物提取物中的植酸酶活性，取对照系及转基因植物的 200 μg 蛋白与 400 μg IP6（Sigma, P8810）于 37°C 孵育。在 4、6 及 8 小时时，加入一倍体积的 0.05 M HCl 终止酶反应。为了比较植物提取物的植酸酶活性，用阴离子交换层析（Sandberg 和 Ahderne, 1986, 测定食物及小肠内容物中肌醇三磷酸盐、肌醇四磷酸盐、肌醇五磷酸盐和肌醇六磷酸盐的 HPLC 方法, Journal of Food Science 51

(3): 547-550) 纯化肌醇磷酸盐 (IP6、IP5、IP4 和 IP3)。简单地说, 0.5 mL 酶混合物上样到 2 mL AG-1X8 阴离子交换柱 (Bio-Rad), 用 10 倍柱体积的 0.025 M HCl 洗去杂质。随后, 用 3 M HCl 洗脱肌醇磷酸盐。洗脱样品冻干, 重悬于 100 μ L 移动相[50% (v/v) 甲醇, 0.1% (v/v) 甲酸], 1.5% (v/v) 盐酸四丁铵及 0.05 M EDTA 进行 HPLC (waters 600) 分析 (Sandberg 和 Ahderne, 1986, 同上)。在 C18 柱 (Alltech Alltima C18) 中注入 20 μ L 进行肌醇磷酸盐测定。用屈光指数检测器 (Shimadzu R1D-10A, Shimadzu Corporation, 日本) 衡量 IP6 和 IP5 各自的峰值, 并计算 IP6/IP5 比值。与植物提取物刚混合时, 肌醇六磷酸底物的 IP6/IP5 比值为 3.61 ± 0.14 (n=4)。在孵育过程中, IP6 逐渐降解成低磷酸化的磷酸盐 (IP5、IP4 和 IP3), 因而随着时间的推移 IP6/IP5 比值也下降。如图 21 所示, 42 系 (N=4) 的植物提取特获得的 IP6/IP5 比值比从对照植物 (N=4) 获得的比值低。

15 总而言之, 用植酸酶基因转化的烟草植物具有下列表型: (1) 开花茎增多; (2) 主干茎增多; (3) 蕾数增多, 及 (4) 花期延长 (见表 2)。可以预料的是植酸酶基因转化的烟草植物将比对照植物具有更多的果实, 因为前者的花蕾数增加了。

表 2
转基因烟草植物的表型

植物系	转移至土壤的时间	第一朵花开的时间	第一朵花开时的植物高度(cm)	开花后的植物高度(cm)	主要花茎数	主茎数*	花蕾数*	开花天数
0042	6月8日	8月25日	102	142	10	3	74	>88+
0043	6月8日	9月8日	101	142	9#	3	31#	58
0131	6月8日	9月12日	125	160	8	1	33	3
0133	6月8日	9月10日	130	168	8#	3	47#	44
0134	6月8日	9月4日	120	150	9#	2	36#	50^
载体对照 1	6月8日	9月4日	135	182	6	1	30	36
载体对照 2	6月15日	9月16日	158	206	6	1	37	34
非转基因对照	6月8日	9月22日	150	185	6	1	32	35

*数据于 2001 年 10 月 4 日采集

#仅计算第一主茎上的花蕾。侧枝上的花蕾还未成熟到用于计算。

+至 2001 年 11 月 20 日

^新花于 2001 年 11 月 8 日开。

7. 等效物

本领域内的熟练技术人员将认识到，或仅使用常规试验能确定本文所述本发明特定实施方案的许多等效物。此类等效物也属于后面权利要求所包含的范围。

- 5 本说明书中提到的所有出版物、专利及专利申请均被引入说明书作为参考，如同每一单独出版物、专利或专利申请具体地、单独地指出将其引入作为参考一样。

在本文中引用或讨论的参考文献不应解释为承认它是本发明的现有技术。

<110> 香港大学

<120> 重组芽孢杆菌植酸酶及其用途

<130> 9661-020-888

<140> 60/332,060

<141> 2001-11-21

<160> 24

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 1390

<212> DNA

<213> *Bacillus licheniformis*

<220>

<221> CDS

<222> (241)...(1390)

<400> 1

```

ttttaccoga tggatgggga cttaaacgaa ottgogttg agatatacat tccgattcat 60
tgagagatag cgatgttaa ggcagccccc ggaaaaaatt ccgggggttt tctttggggt 120
togtactota gagtatoggc ggtctttttt agccatocact ttaacaaaa gtttacatac 180
cctcaaatga taattttcat tggtttgcta ggataaatgt tatgaaaagg aggttaatat 240
atgaactttt acaaaacgot cgctttatca acactcgcag catccttatg gtotccctca 300
tggagcagtc tccccataa cgaagctgcg gctcacaag aattcacggg gactgoogat 360
gcagagacag agooggtgga taccctgac gacgcggcag atgaccggc gatttggggt 420
catccgaagc agcctgaaaa aagcaggctc atcaccacaa acaaaaagtc gggcttaate 480
gtotatgat tgaagggaaa acagcttgcg gcctatccgt ttggcaaat aaacaatgtc 540
gaoctgogot acaattttoc gctcogatgc aaaaaattg atattgoogg ggoccaaac 600
cgtcagacg gcaaaaacac ggttgaatt tacgccttg acggcgaaa aagcaagctg 660
aagaacatcg tcaatcctca aaaaacctt caaacggata tccaggagg atatggcttc 720
agcctgtatc acagccagaa aacgggcaag ttctacgcca tggtagcgg aaagaacgga 780
gaattcgagc aatatgaact gtttgacaac ggaaaaggac aagtcgaggg caaaaaggtc 840
cgctcattca aaatgagctc tcaaacagaa gggcttgcg cagatgatga atacggcmaa 900
atgtacatcg cgaagaaga cgttgcgatt tggcttttca gcgccgagc ggacggcgga 960
gataaaggaa aaatcgctga tcgtgccgac ggaccgcac taacttctga tattgaagg 1020
ctgacgattt actacggaga agacggagaa gggatattga tcgcgtocag tcagggcgat 1080
aacogctatg ccatctatga cggcgcggg aaaaacgact acgtcactgc ttttcaatt 1140
gaggacgcca aagaaatcga cgggacaagc gataccgatg gaatcgact catcgcttc 1200
ggcctcggca aaacatatcc atacggcaco tttgtogccc aagacggoga aaatacggaa 1260
aatggacaac cggccaatca gaacttcaa attgtctcot gggaaaaaat cgcgcagcgc 1320
ctggacgaca aacctgatat cgatgatcag gtcgatccoc gnaaactgaa aaaccgagcc 1380
aaataaggac

```

<210> 2

<211> 381

<212> PRT

<213> *Bacillus licheniformis*

<400> 2

```

Met Asn Phe Tyr Lys Thr Leu Ala Leu Ser Thr Leu Ala Ala Ser Leu
1                               5                               10                               15
Trp Ser Pro Ser Trp Ser Ser Leu Pro His Asn Glu Ala Ala Ala His
20                               25                               30

```

```

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

```

```

<210> 3
<211> 1380
<212> DNA
<213> Bacillus subtilis

```

```

<220>
<221> CDS
<222> (100)...(1248)

```

```

<400> 3
gagtcagaaa ccotataaaa aaagggttcat tttcctcagc gtaatcacct gtatatatt 60
tacaatagta gtgttagtga taaaagagga gggtagcaaa atg aag gtt cca aaa 114
Met Lys Val Pro Lys
      1      5

```

```

aca atg ctg cta agc act ggc gcg ggt tta ttg ctt agc ctg aca gca 162

```

Thr Met Leu Leu Ser Thr Ala Ala Gly Leu Leu Leu Ser Leu Thr Ala	
10 15 20	
acc tcg gtg tcg gct cat tat gtg aat gag gaa cat cat ttc aaa gtg	210
Thr Ser Val Ser Ala His Tyr Val Asn Glu Glu His His Phe Lys Val	
25 30 35	
act gca cac acg gag aca gat ccg gtc gca tct ggc gat gat gca gca	258
Thr Ala His Thr Glu Thr Asp Pro Val Ala Ser Gly Asp Asp Ala Ala	
40 45 50	
gat gac ccg gcc att tgg gtt cat gaa aaa cac ccg gaa aaa agc aag	306
Asp Asp Pro Ala Ile Trp Val His Glu Lys His Pro Glu Lys Ser Lys	
55 60 65	
ttg att aca aca aat aag aag tca ggg ctc gtt gtg tat gat tta gac	354
Leu Ile Thr Thr Asn Lys Lys Ser Gly Leu Val Val Tyr Asp Leu Asp	
70 75 80 85	
gga aaa cag ctt cat tct tat gag ttt ggc aag ctc aat aat gtc gat	402
Gly Lys Gln Leu His Ser Tyr Glu Phe Gly Lys Leu Asn Asn Val Asp	
90 95 100	
ctg cgc tat gat ttt cca ttg aac ggc gaa aaa att gat att gct gcc	450
Leu Arg Tyr Asp Phe Pro Leu Asn Gly Glu Lys Ile Asp Ile Ala Ala	
105 110 115	
gca tcc aac cgg tcc gaa gga aaa aat aca att gaa gta tat gca ata	498
Ala Ser Asn Arg Ser Glu Gly Lys Asn Thr Ile Glu Val Tyr Ala Ile	
120 125 130	
gac ggg gat aaa gga aaa ttg aaa agc att aca gat ccg aac cat cct	546
Asp Gly Asp Lys Gly Lys Leu Lys Ser Ile Thr Asp Pro Asn His Pro	
135 140 145	
att tcc acc aat att tct gag gtt tat gga ttc agc ttg tat cac agc	594
Ile Ser Thr Asn Ile Ser Glu Val Tyr Gly Phe Ser Leu Tyr His Ser	
150 155 160 165	
cag aaa aca gga gca ttt tac gca tta gtg aca ggc aaa caa ggg gaa	642
Gln Lys Thr Gly Ala Phe Tyr Ala Leu Val Thr Gly Lys Gln Gly Glu	
170 175 180	
ttt gag cag tat gaa att gtt gat ggt gga aag ggt tat gta aca ggg	690
Phe Glu Gln Tyr Glu Ile Val Asp Gly Gly Lys Gly Tyr Val Thr Gly	
185 190 195	
aaa aag gtg cgt gaa ttt aag ttg aat tct cag acc gaa ggc ctt gtt	738
Lys Lys Val Arg Glu Phe Lys Leu Asn Ser Gln Thr Glu Gly Leu Val	
200 205 210	
gcg gat gat gag tac gga aac cta tac ata goa gag gaa gat gag gcc	786
Ala Asp Asp Glu Tyr Gly Asn Leu Tyr Ile Ala Glu Glu Asp Glu Ala	
215 220 225	
atc tgg aaa ttt aac gct gag ccc ggc gga ggg tca aag ggg cag gtt	834
Ile Trp Lys Phe Asn Ala Glu Pro Gly Gly Gly Ser Lys Gly Gln Val	
230 235 240 245	
gtt gac cgt gcg aca gga gat cat ttg aca gct gat att gaa gga ctg	882
Val Asp Arg Ala Thr Gly Asp His Leu Thr Ala Asp Ile Glu Gly Leu	
250 255 260	

aca atc tat. tat gca cca aat ggc aaa gga tat ctc atg gct tca agt 930
 Thr Ile Tyr Tyr Ala Pro Asn Gly Lys Gly Tyr Leu Met Ala Ser Ser
 265 270 275

caa gga aat aac agc tat gca atg tat gaa cgg cag ggg aaa aat cgc 978
 Gln Gly Asn Asn Ser Tyr Ala Met Tyr Glu Arg Gln Gly Lys Asn Arg
 280 285 290

tat gta gcc aac ttt gag att aca gat ggc gag aag ata gac ggt act 1026
 Tyr Val Ala Asn Phe Glu Ile Thr Asp Gly Glu Lys Ile Asp Gly Thr
 295 300 305

agt gac acg gat ggt att gat gtt ctc ggt ttc gga ctt ggc cca aaa 1074
 Ser Asp Thr Asp Gly Ile Asp Val Leu Gly Phe Gly Leu Gly Pro Lys
 310 315 320 325

tat ccg tac ggg att ttt gtg gcg cag gac ggc gaa aat att gat aac 1122
 Tyr Pro Tyr Gly Ile Phe Val Ala Gln Asp Gly Glu Asn Ile Asp Asn
 330 335 340

gga caa gcc gtc aat caa aat ttc aaa att gta tcg tgg gaa caa att 1170
 Gly Gln Ala Val Asn Gln Asn Phe Lys Ile Val Ser Trp Glu Gln Ile
 345 350 355

gca cag cat ctc ggc gaa atg cct gat ctt cat aaa cag gta aat ccg 1218
 Ala Gln His Leu Gly Glu Met Pro Asp Leu His Lys Gln Val Asn Pro
 360 365 370

agg aag ctg aaa gac cgt tct gac ggc tag aatagaagc agcttgtgca 1268
 Arg Lys Leu Lys Asp Arg Ser Asp Gly *
 375 380

gctgcttttt totatgaata aaaaaatcgt tcatagcaat gaagatttt tcaagaaagc 1328
 gccagatgaa togcttttag tttgcagga agctcatcaa acgtaaagc gg 1380

<210> 4

<211> 382

<212> PRT

<213> *Bacillus subtilis*

<400> 4

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


```

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

```

```

<210> 5
<211> 20
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> 引物

```

```

<220>
<221> misc_feature
<222> 6, 9, 18
<223> n = .deoxyinosine

```

```

<400> 5
gaygcngcng aygayocngc 20

```

```

<210> 6
<211> 20
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> 引物

```

```

<220>
<221> misc_feature
<222> 18
<223> n = deoxyinosine

```

```

<400> 6
tortaytgyt craaytcncc 20

```

<210> 7
 <211> 28
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> 引物

 <400> 7
 atgaattcgt cgggottaat cgtctatg 28

 <210> 8
 <211> 27
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> 引物

 <400> 8
 atggatcctc aggctgcttc ggatgaa 27

 <210> 9
 <211> 27
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> 引物

 <400> 9
 atgaattcat ggcttcagcc gtatcac 27

 <210> 10
 <211> 27
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> 引物

 <400> 10
 atggatccgt gtttttgccg tctgacc 27

 <210> 11
 <211> 27
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> 引物

 <400> 11
 gcgaattcga taccgatgga atogacg 27

 <210> 12
 <211> 27
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> 引物

<400> 12
 atggatcctc atagatggca tagcggg 27

 <210> 13
 <211> 30
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> 引物

 <400> 13
 atttaacata tgaactttta caaaacgctc 30

 <210> 14
 <211> 26
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> 引物

 <400> 14
 atggatccgt ccttatttgg ctcggt 26

 <210> 15
 <211> 33
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> 引物

 <400> 15
 aggaattcca tatgaagggt ccaaaaacaa tgc 33

 <210> 16
 <211> 27
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> 引物

 <400> 16
 taggatcctc atctggcgct ttcttgt 27

 <210> 17
 <211> 35
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> 引物

 <400> 17
 atggatccat ggctcattat gtgaatgagg aacat 35

 <210> 18
 <211> 30
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> 引物

 <400> 18
 attagagctc ctagccgtca gaacggtcctt 30

 <210> 19
 <211> 27
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> 引物

 <400> 19
 ctacaaagat cgttatgttt atcggca 27

 <210> 20
 <211> 22
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> 引物

 <400> 20
 agaccaatgc ggagcatata cg 22

 <210> 21
 <211> 1290
 <212> DNA
 <213> *Bacillus subtilis*

 <400> 21
 cacatttgac aattttcaca aaaacttaac actgacaatc atgtatata gttacaattg 60
 aagtgcacgt tcataaaagg aggaagtaaa atgaatcatt caaaaacact tttgttaacc 120
 gggggccg gactgatgct cacatgctgt ggggtgtctt cccaggcaaa gcataagctg 180
 tccgatcctt atcattttac cgtgaatgca gcggcggaaa cggaacccgt tgatacggcc 240
 ggtgacgctg ctgatgatcc tgcgatttgg ctggacccca agactcctca gaacagcaaa 300
 ttgattacga ccaataaaaa atcaggttta gtcgtttaca gccttgatgg taagatgctt 360
 cattcctata ataocgggaa gctgaacaat gtcgatatcc gttatgattt tccggtgac 420
 ggcaaaaaag tcgatatogc ggoagcatcc aatcggctctg aaggaaaaaa taccattgag 480
 atttaogota ttgatggaaa aaaocggcaca ttacaaagca tgacagatcc agaccatccg 540
 attgcaacag caattaatga ggtatacggg tttaccttat accacagtca aaaaacagga 600
 aaatattacg cgatggtagc aggaaaagag ggtgaatttg aacaatacga attaaaggcg 660
 gacaaaaatg gatacatatc cggcaaaaag gtacgggctt ttaaaatgaa ttcccagacg 720
 gaagggatgg cagcagacga tgaataoggc aggtttata tcgoagaaga agatgaggcc 780
 atttggaagt tcagoccca gcoggaoggc ggcagtaacg gaacggttat cgaacgtgcc 840
 gacggcaggg atttaactcg tgnattgaa ggattgacga tttactacgc tgctgacggg 900
 aaaggctatc tgatggcacc aagocaggga aacagcagct acgocattta tgacagacaa 960
 gyaagaaca aatatgttgc ggattttcgc ataacagacg gtcctgaaac agacgggaca 1020
 agcgatacag aoggaattga ogttotgggt ttoggaotgg ggoctgaata toogttoggt 1080
 atttttgtcg cacaggacgg tgaaaatata gatcacggcc aaaaggccaa tcaaaatfff 1140
 aaaatcgtgc catgggaaag aattgctgat caaatcgggt tccgcccgtc ggcaaatgaa 1200
 caggttgacc cgagaaaact gaccgacaga agocgaaaat aaacatgcaa aaagcagctt 1260
 atacaagctg ctttttgcac gtgaagaacg 1290

 <210> 22
 <211> 383
 <212> PRT
 <213> *Bacillus subtilis*

<400> 22
Met Asn His Ser Lys Thr Leu Leu Leu Thr Ala Ala Ala Gly Leu Met
1 5 10 15
Leu Thr Cys Gly Ala Val Ser Ser Gln Ala Lys His Lys Leu Ser Asp
20 25 30
Pro Tyr His Phe Thr Val Asn Ala Ala Ala Glu Thr Glu Pro Val Asp
35 40 45
Thr Ala Gly Asp Ala Ala Asp Asp Pro Ala Ile Trp Leu Asp Pro Lys
50 55 60
Thr Pro Gln Asn Ser Lys Leu Ile Thr Thr Asn Lys Lys Ser Gly Leu
65 70 75 80
Val Val Tyr Ser Leu Asp Gly Lys Met Leu His Ser Tyr Asn Thr Gly
85 90 95
Lys Leu Asn Asn Val Asp Ile Arg Tyr Asp Phe Pro Leu Asn Gly Lys
100 105 110
Lys Val Asp Ile Ala Ala Ala Ser Asn Arg Ser Glu Gly Lys Asn Thr
115 120 125
Ile Glu Ile Tyr Ala Ile Asp Gly Lys Asn Gly Thr Leu Gln Ser Met
130 135 140
Thr Asp Pro Asp His Pro Ile Ala Thr Ala Ile Asn Glu Val Tyr Gly
145 150 155 160
Phe Thr Leu Tyr His Ser Gln Lys Thr Gly Lys Tyr Tyr Ala Met Val
165 170 175
Thr Gly Lys Glu Gly Glu Phe Glu Gln Tyr Glu Leu Lys Ala Asp Lys
180 185 190
Asn Gly Tyr Ile Ser Gly Lys Lys Val Arg Ala Phe Lys Met Asn Ser
195 200 205
Gln Thr Glu Gly Met Ala Ala Asp Asp Glu Tyr Gly Arg Leu Tyr Ile
210 215 220
Ala Glu Glu Asp Glu Ala Ile Trp Lys Phe Ser Ala Glu Pro Asp Gly
225 230 235 240
Gly Ser Asn Gly Thr Val Ile Asp Arg Ala Asp Gly Arg His Leu Thr
245 250 255
Arg Asp Ile Glu Gly Leu Thr Ile Tyr Tyr Ala Ala Asp Gly Lys Gly
260 265 270
Tyr Leu Met Ala Ser Ser Gln Gly Asn Ser Ser Tyr Ala Ile Tyr Asp
275 280 285
Arg Gln Gly Lys Asn Lys Tyr Val Ala Asp Phe Arg Ile Thr Asp Gly
290 295 300
Pro Glu Thr Asp Gly Thr Ser Asp Thr Asp Gly Ile Asp Val Leu Gly
305 310 315 320
Phe Gly Leu Gly Pro Glu Tyr Pro Phe Gly Ile Phe Val Ala Gln Asp
325 330 335
Gly Glu Asn Ile Asp His Gly Gln Lys Ala Asn Gln Asn Phe Lys Ile
340 345 350
Val Pro Trp Glu Arg Ile Ala Asp Gln Ile Gly Phe Arg Pro Leu Ala
355 360 365
Asn Glu Gln Val Asp Pro Arg Lys Leu Thr Asp Arg Ser Gly Lys
370 375 380

<210> 23
<211> 1724
<212> DNA
<213> Bacillus sp.

<400> 23
catatggtga acaatttcag cgagttaatg aaagaacca ataatcaaa aattagagaa 60
aaacattaat ctgatgcgct ttcatatcgc gttaccgat taatagaata gaaattacaa 120
ataaaacatt gtactaata ttcattttaa atatttgctc acgtcaattt tttctcttca 180
taaatctca cattcggaca atcttcacaa aaottaaca ctgaacttcc tgtatgtatt 240
ttacaattaa agtgcacggt cataaaagga ggatggaaaa tgaatcattc aaaaacactt 300

```

ttgtaaccg cggcagcgg attgatgctc acatgcggtg cggtttcttc tcaggccaaa 360
cataagctgt ctgacacctt tcatTTTacc gtgaatgcgg cggcggaaac ggagccggtt 420
gatacagccg gtgatgcagc tgatgatcct gcgatttggc tggaccccaa gaatootcag 480
aacagcaaat tgatcacaac caataaaaaa tcaggcttag ccgtgtacag cctagagggg 540
aagatgcttc attcctatca taccgggaag ctgaacaatg ttgatatccg atatgatttt 600
ccgttgaacg gaaaaaaagt cgatattgcg gcggcatcca atcgggtctga aggaaagaat 660
accattgaga tttacgccat tgacgggaaa aacggcacat tacaagcat tacggatcca 720
aacccgccga ttgcatcagc aattgatgaa gtatacggtt tcagcttgta ccacagtcaa 780
aaaacaggaa aatattacgc gatggtgaca ggaaaagaag gcgaatttga acaatacгаа 840
ttaaattgagg ataaaaatgg atacatatcc ggcaaaaagg taagggcggt taaaatgaat 900
tctcagacag aagggatggc agcagacgat gaatacggca gtctttatat cgcagaagaa 960
gatgaggcca tctggaagtt cagcgtgtag ccggacggcg gcagtaacgg aacggttatc 1020
gatcgtgccg atggcaggca ttaaccctct gatattgaag gactgacgat ttactacgct 1080
gctgacggga aaggctatct gottgectca agccagggtg acagcageta tgcgatttat 1140
gaaagacagc gacagaacaa atagtgtgag gactttcaga taacagacgg gcctgaaaca 1200
gacggcacia gogatacaga cggcaattgac gttotgggtt tggggctggg goctgaaat 1260
ccgttcggtc tttttgtcgc acaggacgga gagaatatag atcagggcca aaaggccaat 1320
caaaatttta aaatgtgccc atgggaaaaga atcgtgata aaatcggctt tcaaccgcag 1380
gtcaataaac aggtcgaccc gagaaaaatg accgacagaa ggggaaaata aacatgaaaa 1440
aagcagctta tccaagctgc tttttgatgt gaagagcgtt tcatgagaaa gtcttggaac 1500
ggatagccgt aagcacagcc ggcagccggt catacgtgta cgcgggtact gtctcttgat 1560
aattaagcgc cgcgatttgt ttacgctcac ccgggttgt catataaaaa tggatcttat 1620
ccggaaaaac cgcaaacccg ctgtaagaaa caaatgttga aaacgggggc gcgggagaaa 1680
ggtctgtcat ctgaaaaggcc tgacaagccg caatgtctaa gctt 1724

```

```

<210> 24
<211> 383
<212> PRT
<213> Bacillus sp.

```

```

<400> 24
Met Asn His Ser Lys Thr Leu Leu Leu Thr Ala Ala Ala Gly Leu Met
1 5 10 15
Leu Thr Cys Gly Ala Val Ser Ser Gln Ala Lys His Lys Leu Ser Asp
20 25 30
Pro Tyr His Phe Thr Val Asn Ala Ala Ala Glu Thr Glu Pro Val Asp
35 40 45
Thr Ala Gly Asp Ala Ala Asp Asp Pro Ala Ile Trp Leu Asp Pro Lys
50 55 60
Asn Pro Gln Asn Ser Lys Leu Ile Thr Thr Asn Lys Lys Ser Gly Leu
65 70 75 80
Ala Val Tyr Ser Leu Glu Gly Lys Met Leu His Ser Tyr His Thr Gly
85 90 95
Lys Leu Asn Asn Val Asp Ile Arg Tyr Asp Phe Pro Leu Asn Gly Lys
100 105 110
Lys Val Asp Ile Ala Ala Ala Ser Asn Arg Ser Glu Gly Lys Asn Thr
115 120 125
Ile Glu Ile Tyr Ala Ile Asp Gly Lys Asn Gly Thr Leu Gln Ser Ile
130 135 140
Thr Asp Pro Asn Arg Pro Ile Ala Ser Ala Ile Asp Glu Val Tyr Gly
145 150 155 160
Phe Ser Leu Tyr His Ser Gln Lys Thr Gly Lys Tyr Tyr Ala Met Val
165 170 175
Thr Gly Lys Glu Gly Glu Phe Glu Gln Tyr Glu Leu Asn Ala Asp Lys
180 185 190
Asn Gly Tyr Ile Ser Gly Lys Lys Val Arg Ala Phe Lys Met Asn Ser
195 200 205
Gln Thr Glu Gly Met Ala Ala Asp Asp Glu Tyr Gly Ser Leu Tyr Ile
210 215 220
Ala Glu Glu Asp Glu Ala Ile Trp Lys Phe Ser Ala Glu Pro Asp Gly
225 230 235 240
Gly Ser Asn Gly Thr Val Ile Asp Arg Ala Asp Gly Arg His Leu Thr

```

```

                245                250                255
Pro Asp Ile Glu Gly Leu Thr Ile Tyr Tyr Ala Ala Asp Gly Lys Gly
                260                265                270
Tyr Leu Leu Ala Ser Ser Gln Gly Asn Ser Ser Tyr Ala Ile Tyr Glu
                275                280                285
Arg Gln Gly Gln Asn Lys Tyr Val Ala Asp Phe Gln Ile Thr Asp Gly
                290                295                300
Pro Glu Thr Asp Gly Thr Ser Asp Thr Asp Gly Ile Asp Val Leu Gly
305                310                315
Phe Gly Leu Gly Pro Glu Tyr Pro Phe Gly Leu Phe Val Ala Gln Asp
                325                330                335
Gly Glu Asn Ile Asp His Gly Gln Lys Ala Asn Gln Asn Phe Lys Met
                340                345                350
Val Pro Trp Glu Arg Ile Ala Asp Lys Ile Gly Phe His Pro Gln Val
                355                360                365
Asn Lys Gln Val Asp Pro Arg Lys Met Thr Asp Arg Ser Gly Lys
                370                375                380

```

```

ttttaccga tggatgggga cttaaacgaa cttgctgttg agatatacat tccgattcat tgagagatag cgatgttaaa
ggcagcccc ggaataaatt ccgggggttt tctttgggtt tctgtactta gagtatcggc ggtctttttt agccatcaact
ttaaacaata gtttacatac cctcaaatga taattttcat tggtttgcta ggataaatgt tatgaaaagg aggttaatat
ATGAACCTTT ACAAAACGCT CGCTTATCA ACACTCGCAG RAITCACGGT GACTCCCGAT AACCCACAA ACZAAAAGTC GGGCTTAATC
CGAAGCTGCG GCTCACAAAG GATTCGAAGC CATCCGAAAG AGCTGAAA AAAGAGGCTC ATCACCACAA ACCCCCTGAC GACCGGGCAG
ATGACCCGGC GATTTGGT TGAAGGAAA ACAGCTTGG GCCTATCCGT TTGGCAAAT AAACAATGTC GACCTGCGCT ACAATTTTCC
GCTCGATGGC AAAAAAATG ATATTGCCGG GGCTCAAAAC CGGTCAGACG GCAAAAACAC GGTTGAAATTT TACGCCTTTG
ACGGCGAAA AACCAAGCTG AAGAACATCG TCAATCCTCA AAAACCTATT CAAACCGATA TCCAGGAGGT ATATGGCTTC
AGCCTGTATC ACAGCCAGAA AACCGGCAAG TTCTACGCCA TGGTGACCGG AAAGAACGTA GAATTCGAGC AATATGAACT
GTTTGACAAAC GHAAGGAC AAGTCGAGGG CAAAAGGTC CGCTCATTCA AATGAGCTC TCAAAACAGAA GGGCTTGCGG
CAGATGATGA ATACGGCAA ATGTACATCG CCGAAGAAGA CGTTGGGATT TGGTCTTTCA GCGCCGAGCC GGACCGGGGA
GATAAAGGAA AATCGTGA TCGTGCGGAC GGACCGCATC TAACTTCTGA TATGGAAGGG CTGACGATTTT ACTACGGAGA
AGACGGAGAA GGGTATTGA TCGCGTCCAG TCAGGGCGAT AACCGCTATG CCAATCTATGA CCGGCGCGGG AAAAACGACT
ACGTCACTGC TTTTTCATTT GAGGACGGCA AAGRAATCGA CCGGACAAGC GATACCGGATG GAATCGACGT CATCGGCTTC
GGCCTCGGCA AAACATATCC ATACGGCATC TTTGTGCGCC AAGACGGGGA AAATACGGAA AATGGACAAC CGGCCAATCA
GNACTTCAA AATTGCTCCT GGGAAAAAT CGCCGACGGG CTGGACGACA AACCTGTAT ATCGATGATCAG GTCGATCCCC
GAAAACTGAA AAACCGAGCC AATAAAGGAC

```

图 1A

MNFYKTLALS TLAASLWSPS WSSLPHNEAA AHKEFTVTAD AETEPVDTPD DAADDPAINW HPKQPEKSRL ITTNKKSGLI
 VYDLKGGKQLA AYPFGKLNIV DLRYNFFLDG KKIDIAGASN RSDGKNTVEI YAFDGEKSKL KNIVNPQKPI QTDIQEVYGF
 SLYHSQKTGK FYAMVTGKNG EFEQYELFDN GKQVEGKKV RSFKMSSQTE GLAADDEYVK MYIAEEDVAI WSFSAEPDGG
 DKGKIVDRAD GPHLTSIDIEG LTIYYGEDGE GYLIASSQGD NRYAIYDRRG KNDYVTAFSI EDGKEIDGTS DTDGIDVIGF
 GLGKTYPYGI FVAQDGENTE NGQPANQNFK IVSWEKILADA LDDKPDIDDQ VDPRKLNRA K

图 1B

gagtcagaaa ccctataaaa aaaggttcat tttcctcacg gtaatcacct gtatatatt tacaatagta gtgttagtga
 taaaagagga gggtaacaaa tgaagggttcc aaaaacaatg ctgctaagca ctgcccggg tttatttgctt agcctgacag
 caacctcggg gtcggctcat tatgtgatg aggaacatca tttcaaaagtg actgcacaca cggagacaga tccgggtcgc
 tctggcgatg atgcagcaga tgacccggcc atttgggttc atgaaaaaca cccgaaaaa agcraagtga ttacaacaaa
 taagaagtca ggsctcgttg tgtatgattt agacgaaaa cagcttcatt ctatgagtt tggcraagctc aataatgtcg
 atctcggcta tgattttcca ttgaaacggc aaaaattga tattgctgcc gcattccaacc gctccgaag aaaaatata
 attgaagtat atgcaataga cggggataaa ggaarattga aaagcaattac agatccgaac catcctattt ccaccaatat
 ttctgaggtt tatggatfca gcttgtatca cagccagaaa acaggagcat tttacgcat agtgacagcc aaacaagggg
 aatttgagca gtatgaaatt gttgatggtg garagggtta tctaacacagg aaaaaggtgc gtgaaattaa gttgaaattct
 cagaccgaag gccttctgic ggatgatgag tacggaaacc tatacatagc agaggaagat gaggccatct ggraatttaa
 cgctgagccc ggcggagggf caaagggca ggttctgac cctgcgacag gagatcattt gacagctgat attgaaaggac
 tgacaatcta ttatgcacca aatggcaca gatatctcat gcttccaagt caagaaata acagctatgc aatgtatgaa
 cggcagggga aaatcgcta ttagcccaac tttgagatta cagatggcga gaagatagc ggtactagtg acacggatgg
 tattgattt ctcggtttcg gacttggccc aaatatccg taacggatfca cagatggcga ttgtggcga gcacggcga
 acggacagc cctcaatcaa aatttcaaaa ttgtatcgtg gaaacnaatt gcacagcacc tcggcgaaa gcctgatctt
 cataaacagg taatccgag gaagctgaaa gaccgtctg accgctagaa tagaaagcag ctgtgcagc tgccttttc
 tatgaataaa aaaatcgttc atagcaatga acgatttttc aagaaagcgc cagatgaatc gcttttagtt ttgcaggaag
 ctcatcaaac gtaaatgcgg

图

2A

MKVPKTMLLS TAAGLLLSLT ATSVSAHYVN EHHFKVYFAH TETDPVASGD DAADDPALIW HEKHEPEKSKL ITTNKKSGLV
VYDLGKQLH SYEFGKLNIV DLRVDFPLNG EKIDLAAASN RSEKNTIEV YAIDGDKGKL KSIITDPNHPI STNISEVYGF
SLYHSQKIGA FYALVTGKQG EFEQYEIVDG GKGYVTGKKV REFKLSQTE GLVADDEYGN LYIAEEDAI WKFNAPGPGG
SKGQVVDRAI GDHL-TADIEG LTIYYAPNGK GYLMASSQGN NSYAMYERQG KNRYVANFEI TDGEKIDGTS DTDGIDVLGF
GLGPKYPIGI FVAQDGENID NGQAVNQNFK IVSWEQLAQH LGEMPDLLHKQ VNPRLKDRS DG

图 2B

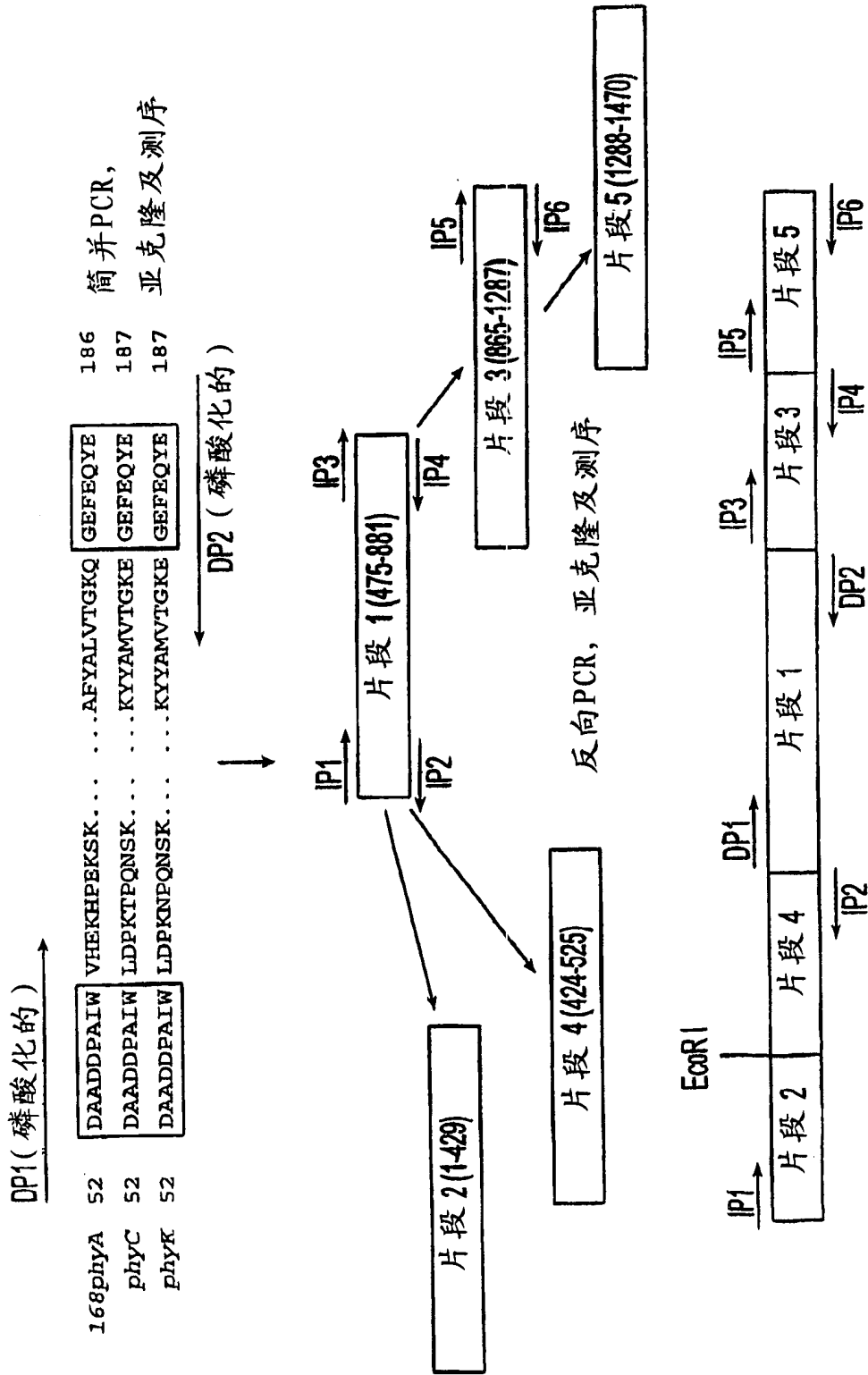
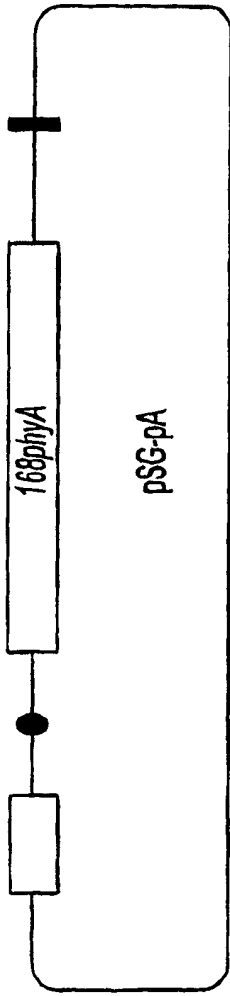


图 3

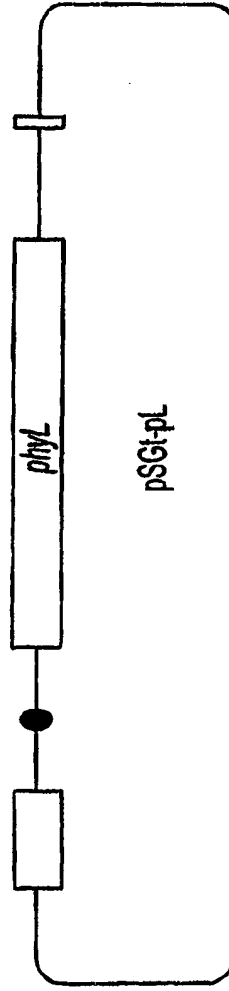


105 启动子

● S.D. 序列

■ 枯草芽孢杆菌 168phyA基因的天然终止子

图 4A



105 启动子

● S.D. 序列

■ 地衣芽孢杆菌 淀粉酶基因终止子

图 4B

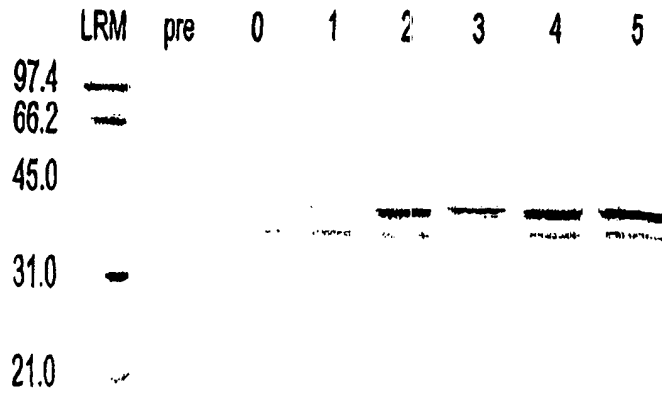


图 5A

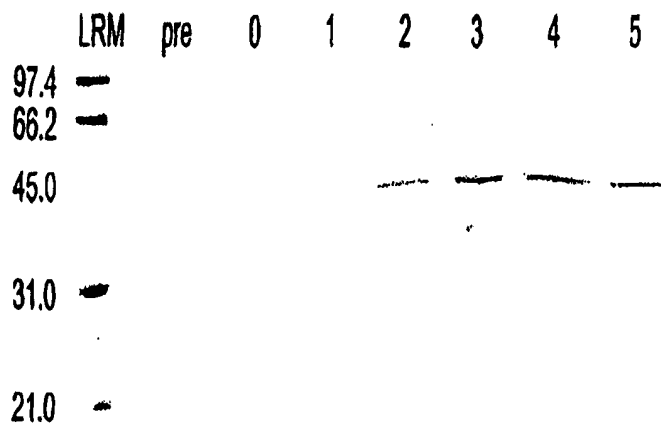


图 5B

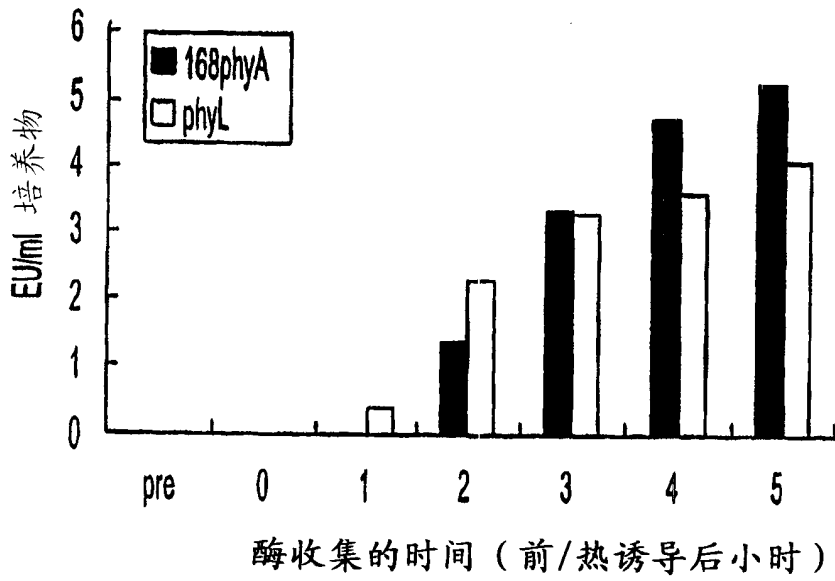


图 6A

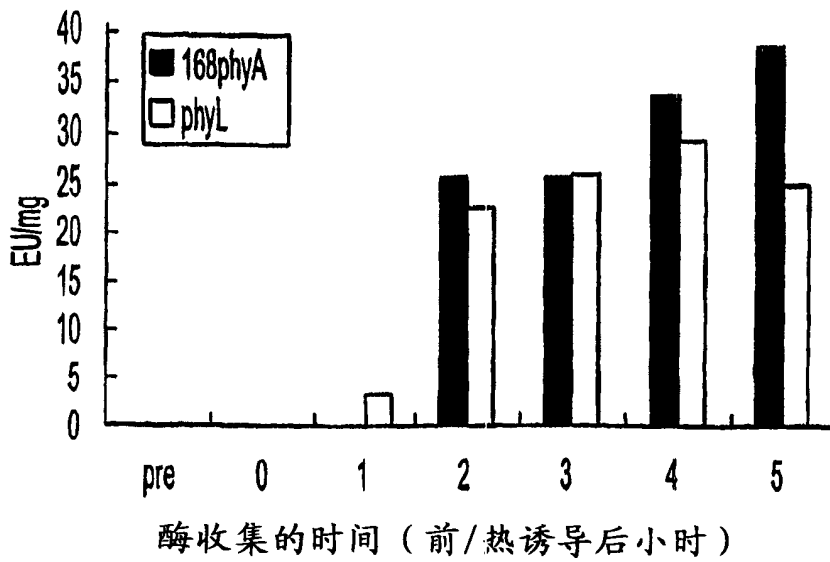


图 6B

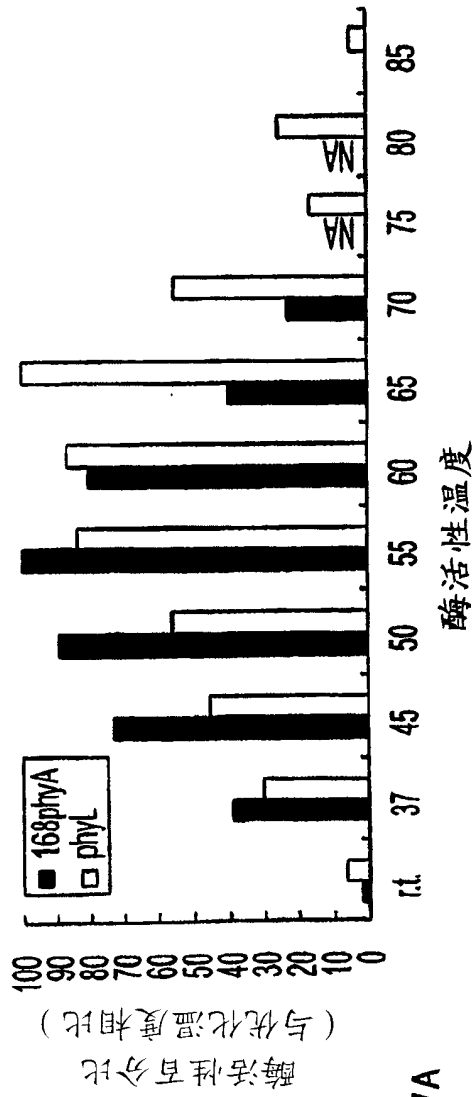


图 7A

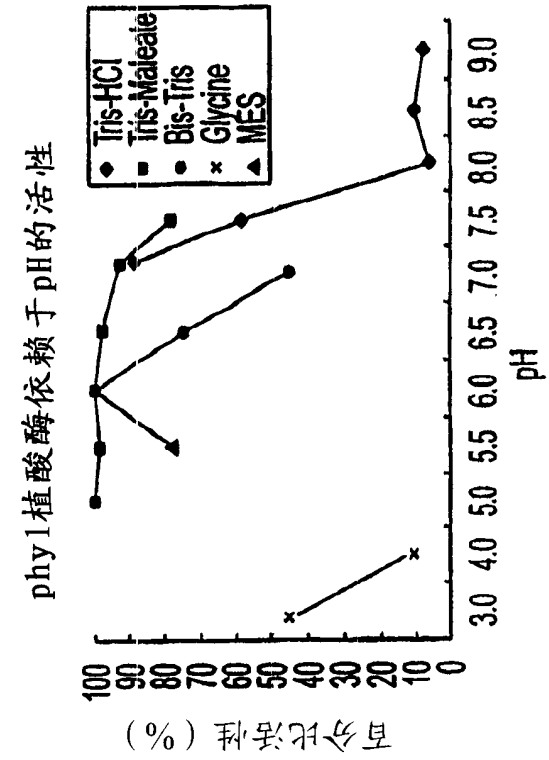
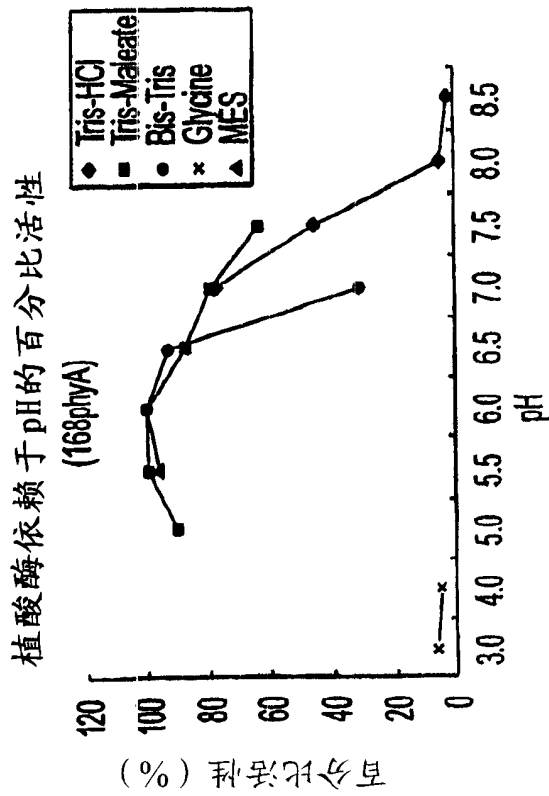


图 7B

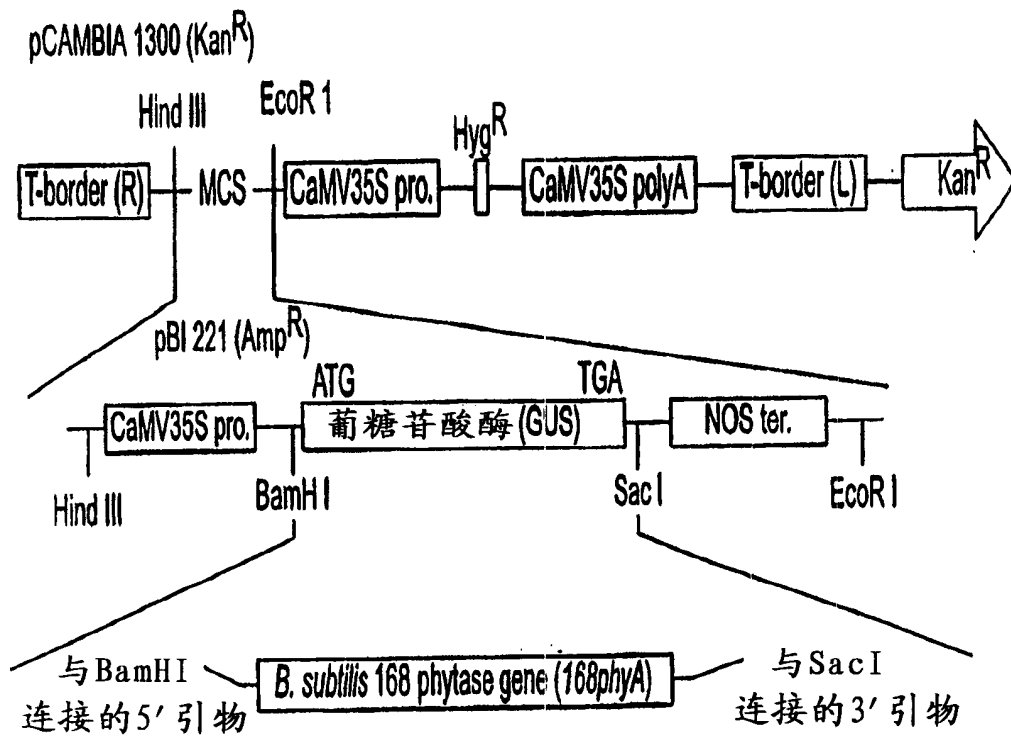


图 8

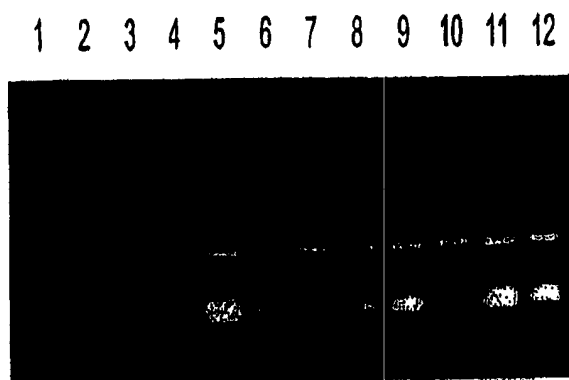


图 10A

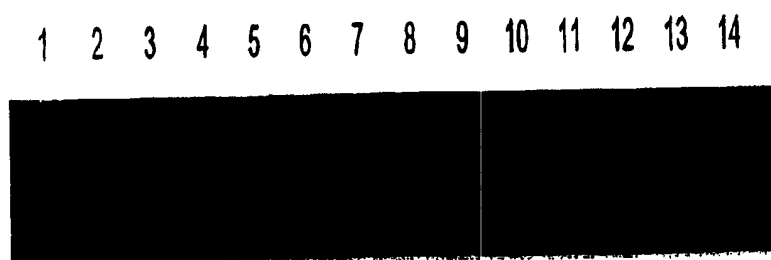


图 10B

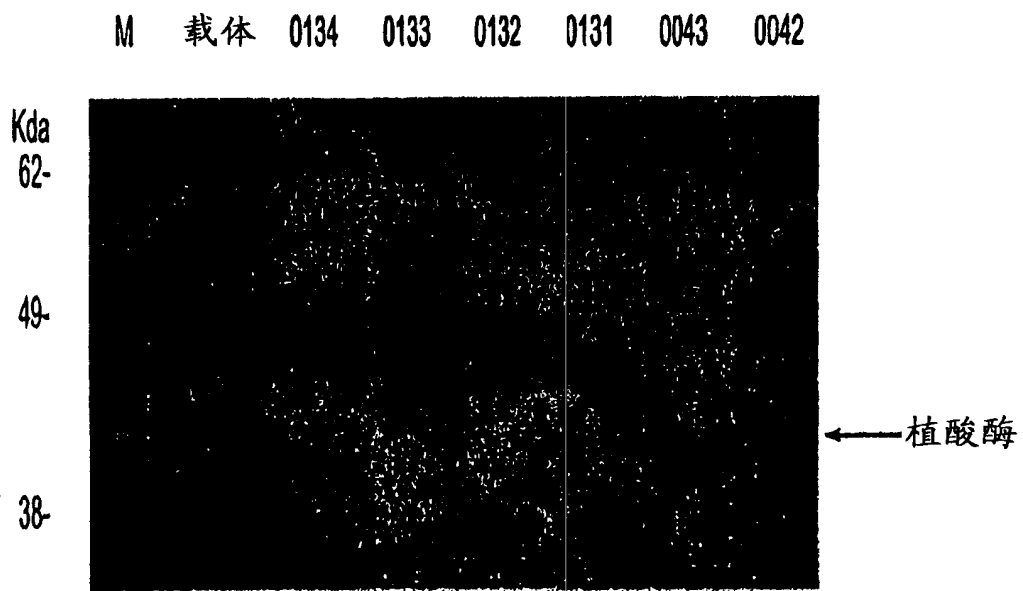


图 11

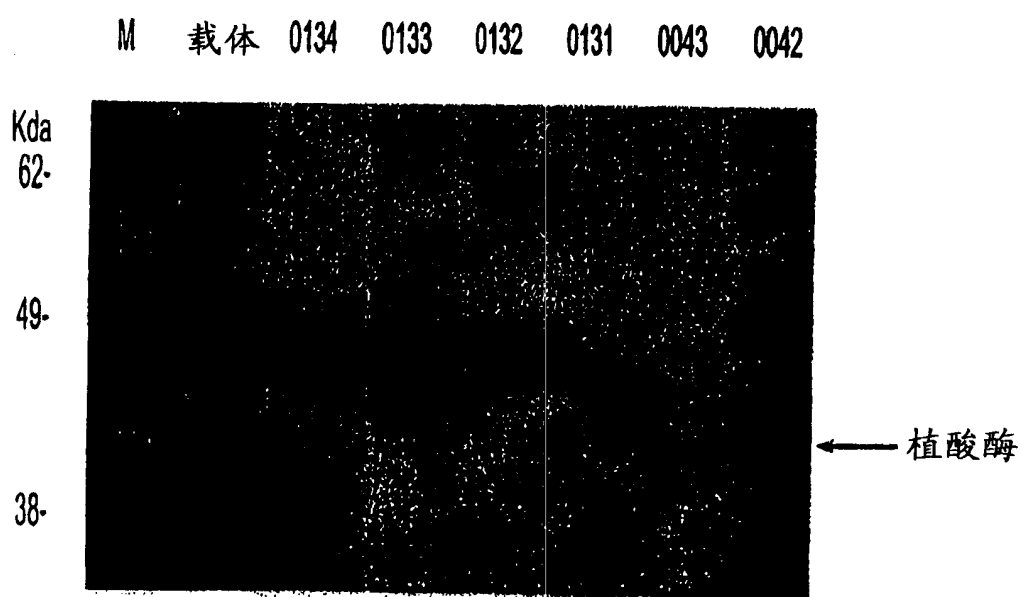


图 12



载体

图 13a



0134

图 13b



0042

图 13c



0043

图 13d



0042

图 14a



0043

图 14b



0134

图 14c



载体

图 14d

Con 1 42-1 42-2 134-1 Con2 Con 3 Con4

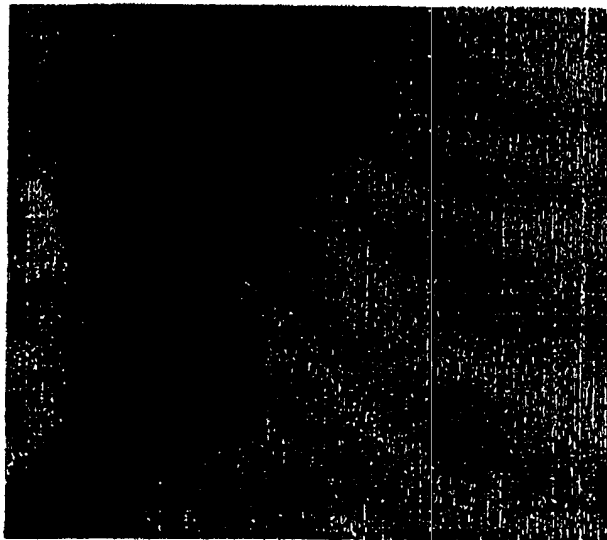


图 15

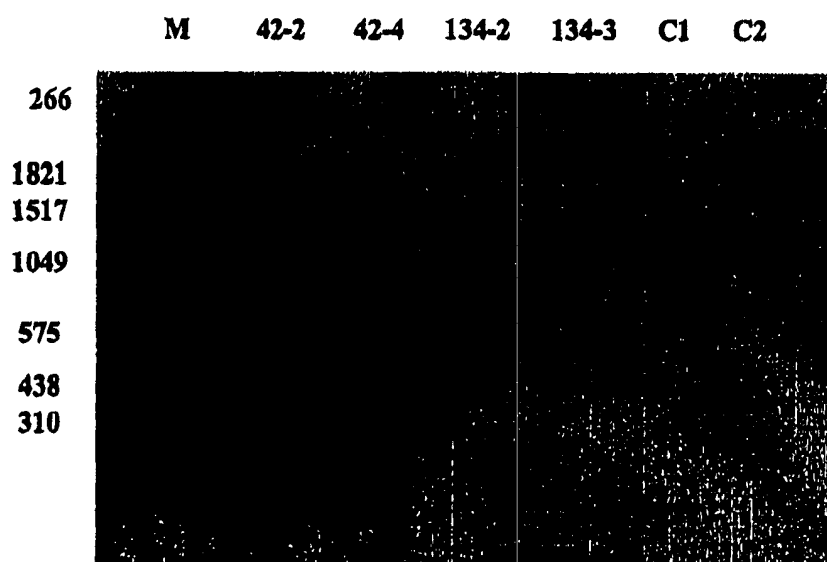


图 16

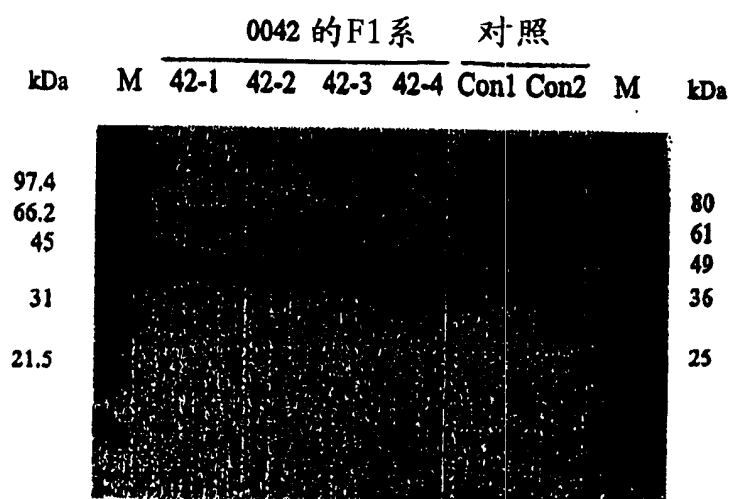


图 17

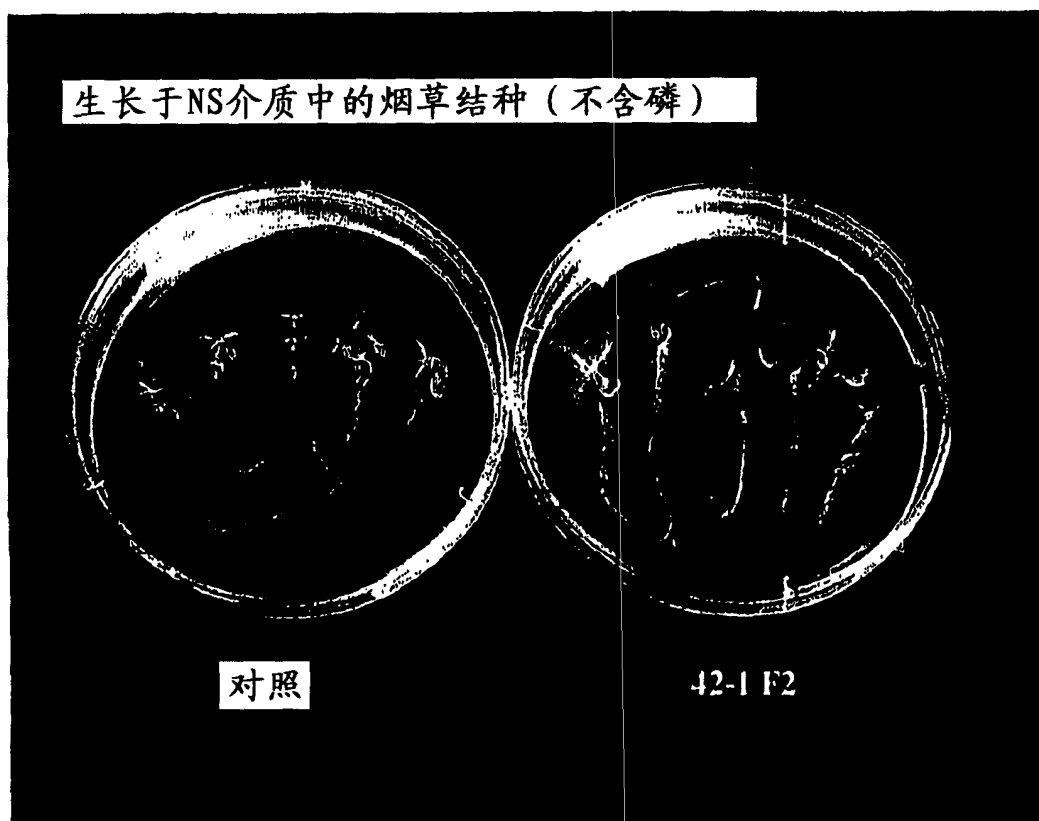


图 18

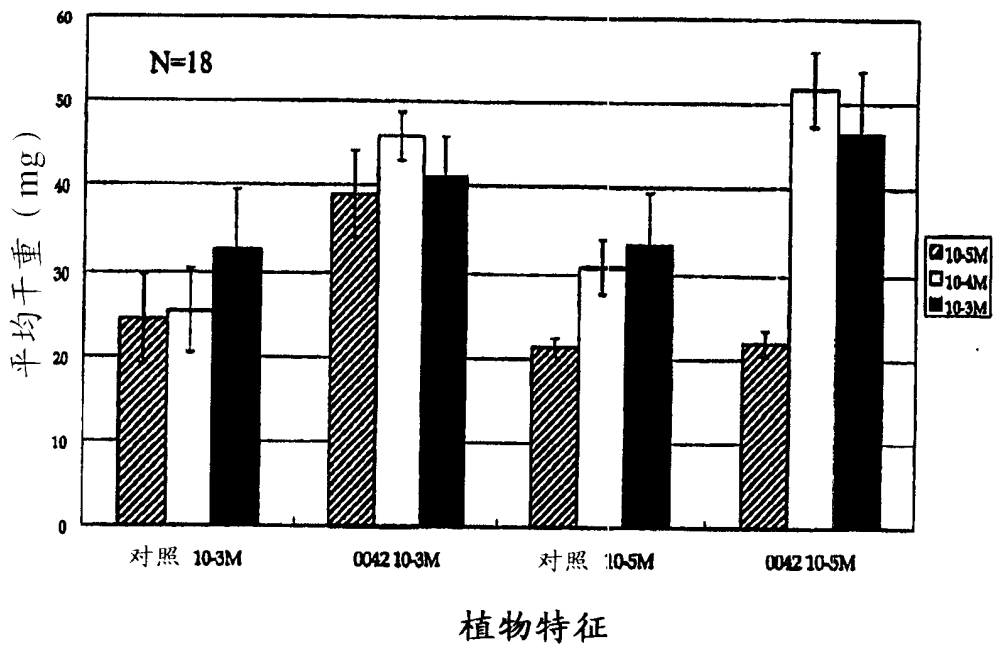


图 19

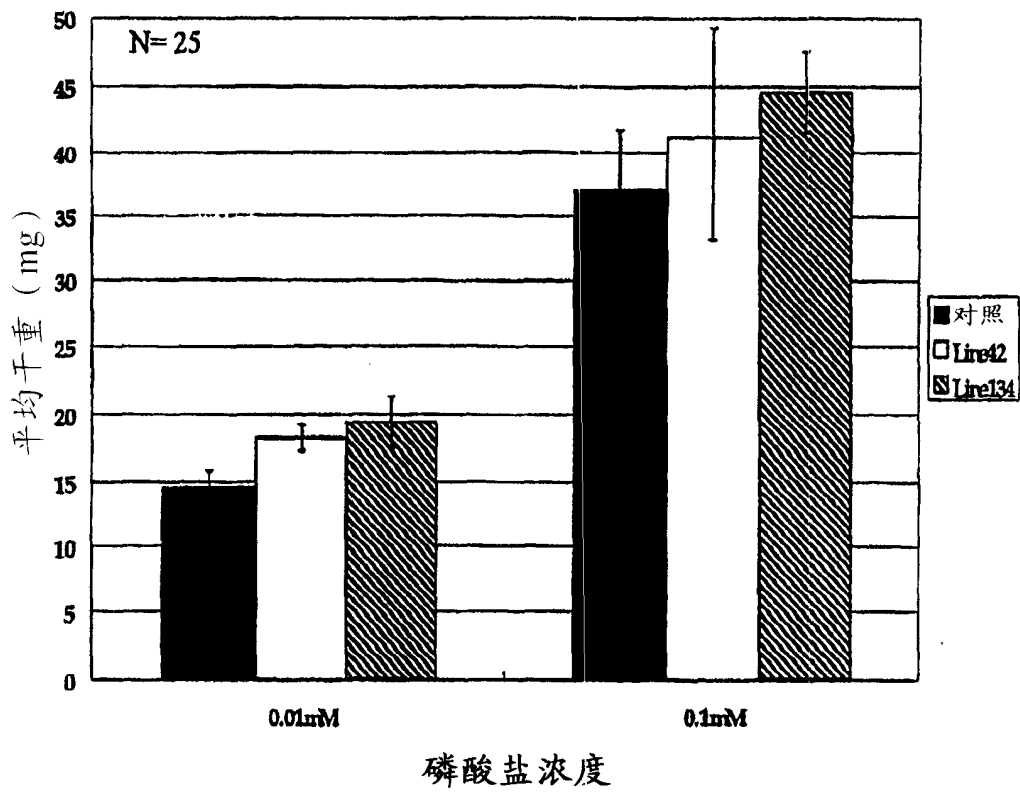


图 20

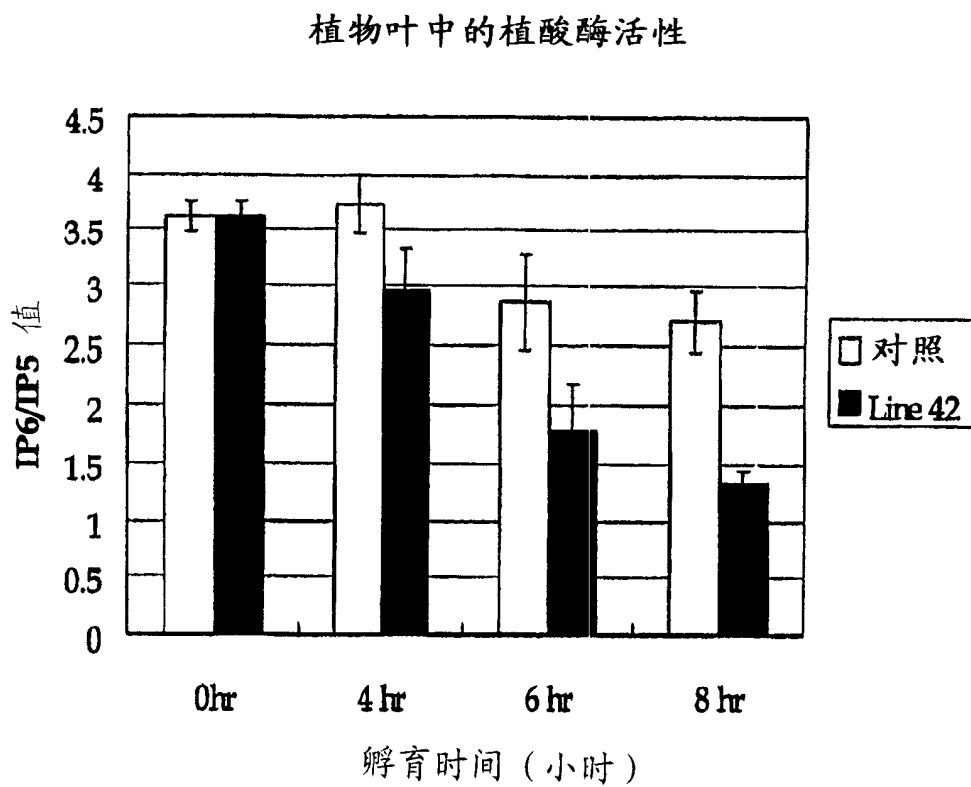


图 21