

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

[12] 发明专利申请公开说明书

C12N 15/82
C12N 15/54 C12N 9/10
C07K 16/40 C12N 5/10
C12Q 1/68 A01H 5/00
A01H 5/10

[21] 申请号 99815641.8

[43] 公开日 2002 年 1 月 30 日

[11] 公开号 CN 1333835A

[22] 申请日 1999.11.16 [21] 申请号 99815641.8

[30] 优先权

[32] 1998.11.17 [33] US [31] 60/108763

[86] 国际申请 PCT/US99/27152 1999.11.16

[87] 国际公布 WO00/29596 英 2000.5.25

[85] 进入国家阶段日期 2001.7.16

[71] 申请人 孟山都技术有限公司

地址 美国密苏里州

[72] 发明人 G·F·贝里

[74] 专利代理机构 中国专利代理(香港)有限公司

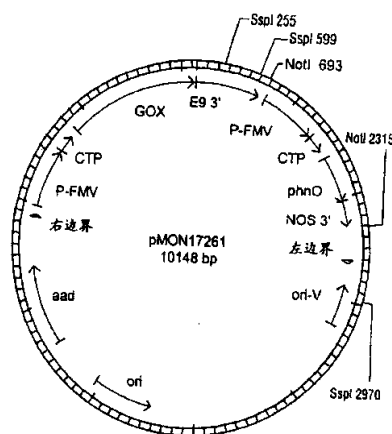
代理人 谭明胜

权利要求书 53 页 说明书 119 页 附图页数 8 页

[54] 发明名称 磷酸代谢植物

[57] 摘要

本发明总的来讲涉及植物的除草剂抗性,更具体地讲涉及一类新的磷酸代谢基因,以及涉及应用这些基因改进植物对磷酸除草剂耐性的方法。



ISSN 1008-4274

权 利 要 求 书

1. 用多核苷酸序列转化的重组植物, 所述多核苷酸序列包含:

- 5 a) 一个植物功能性启动子序列, 其有效连接于;
b) 一个编码酰基转移酶的结构 DNA 序列, 其有效连接于;
c) 一个 3' 序列, 其在植物中起作用, 引起转录终止;

其中所述启动子序列引起所述酶在植物组织中足够表达, 以增强
用所述多核苷酸序列转化的植物的膦酸除草剂(phosphonate
herbicide)耐性, 并且其中所述酶将酰基从酰化供体化合物转
10 移至膦酸除草剂的末端胺。

2. 权利要求 1 的植物, 其中所述酰基转移酶被定位至所述植物的
质体。

3. 权利要求 2 的植物, 其中所述质体包括叶绿体。

4. 权利要求 3 的植物, 其中所述结构 DNA 序列包含一个编码氨
15 基末端叶绿体转运肽的 5' 序列, 所述 5' 序列有效连接于所述结构 DNA
序列的 5', 其中所述酶被定位至所述植物的叶绿体或质体。

5. 权利要求 2 的植物, 其中所述酰化供体是酰基辅酶 A。

6. 权利要求 5 的植物, 其中所述酰基辅酶 A 选自乙酰辅酶 A、
丙酰辅酶 A、丙二酸单酰辅酶 A、琥珀酰辅酶 A 和甲基丙二酸单酰
20 辅酶 A。

7. 权利要求 6 的植物, 其中所述酰基辅酶 A 是乙酰辅酶 A。

8. 权利要求 6 的植物, 选自: 玉米、小麦、棉花、水稻、大豆、
甜菜、canola、亚麻、大麦、油料种子油菜、向日葵、马铃薯、烟草、
番茄、苜蓿、莴苣、苹果、杨树、松树、桉树、金合欢、杨树、美国
25 枫香、radiata pine、火炬松树、云杉、柚木、苜蓿、三叶草和其它饲
料作物、草坪草、油棕、甘蔗、香蕉、咖啡、茶树、可可、苹果、胡
桃、杏仁、葡萄、花生、pulses、矮牵牛、万寿菊、长春花、秋海棠、
天竺葵、三色堇、凤仙花、燕麦、高粱和小米。

9. 权利要求 2 的植物, 其中所述启动子序列衍生自植物 DNA 病毒启动子序列。

10. 权利要求 9 的植物, 其中所述启动子序列选自: CaMV35S、FMV35S、增强的 CaMV35S、增强的 FMV35S、comalina 黄斑驳病毒启动子和甘蔗杆状 DNA 病毒启动子。

11. 权利要求 10 的植物, 其中所述结构 DNA 序列是 SEQ ID NO:3 中叙述的大肠杆菌 *phnO* 基因序列或与其互补。

12. 权利要求 11 的植物, 其中所述结构 DNA 序列编码选自 SEQ ID NO:4 和 SEQ ID NO:8 的肽。

13. 权利要求 12 的植物, 其中所述结构 DNA 序列衍生自微生物, 其中所述序列是选自以下的序列或与能够与选自以下的序列杂交的多核苷酸序列互补: SEQ ID NO:3、SEQ ID NO:7、SEQ ID NO:11、SEQ ID NO:16、SEQ ID NO:17、SEQ ID NO:18 和 SEQ ID NO:19。

14. 权利要求 13 的植物, 其中所述酰基转移酶基本上与大肠杆菌 PhnO 肽相似, 它在植物中起作用, 以将酰基从酰化供体化合物转移至磷酸除草剂的末端胺。

15. 权利要求 14 的植物, 其中所述磷酸除草剂选自草甘膦和 AMPA。

16. 权利要求 4 的植物, 其中编码氨基末端叶绿体转运肽的所述 5' 序列选自: SEQ ID NO:9、SEQ ID NO:11、SEQ ID NO:13 和 SEQ ID NO:14。

17. 按照权利要求 3 的植物, 其表现出对一种或更多种磷酸除草剂的增强的耐性, 所述除草剂选自草甘膦和 AMPA。

18. 由权利要求 17 的植物产生的种子, 其中所述种子包含所述多核苷酸序列。

19. 由权利要求 18 的种子生长的植物。

20. 稳定转化的磷酸除草剂耐性重组植物, 其含有一种多核苷酸序列, 所述多核苷酸序列包含:

- a) 一个植物功能性启动子序列，其有效连接于；
- b) 一个编码酰基转移酶的结构 DNA 序列，其有效连接于；
- c) 一个 3' 序列，其在植物中起作用，引起转录终止；

其中所述启动子序列相对于所述结构 DNA 序列是异源的，并且引起所述酶在植物组织中足够表达，以增强用所述多核苷酸序列转化的植物的所述磷酸除草剂耐性，其中所述酶将酰基从酰化供体转移至磷酸除草剂的末端胺，而且其中所述植物表达编码植物功能性草甘膦氧化还原酶的 GOX 基因。

21. 在重组植物中选择性增强磷酸除草剂耐性的方法，所述方法包括以下步骤：

- a) 用多核苷酸序列转化所述植物，所述多核苷酸序列包含：
 - i) 一个启动子序列，其在植物中起作用，以引起 RNA 序列产生，其有效连接于；
 - ii) 一个结构 DNA 序列，其能够产生编码酰基转移酶的 RNA 序列，其有效连接于；
 - iii) 一个 3' 非翻译序列，其在植物中起作用，引起聚腺苷酸化核苷酸序列添加至所述 RNA 序列的 3' 端；

其中所述启动子序列相对于所述结构 DNA 序列是异源的，并且引起所述酶在植物组织中充分表达，以增强用所述多核苷酸序列转化的植物的磷酸除草剂耐性，其中所述酶将酰基从酰化供体底物转移至磷酸除草剂底物的末端胺，而且其中所述植物表达编码植物功能性草甘膦氧化还原酶的 GOX 基因；和

- b) 在所述植物中表达除草剂耐性有效量的所述酰基转移酶。

22. 按照权利要求 21 的方法，其中所述酰基转移酶由选自以下的 DNA 序列表达： SEQ ID NO:3、SEQ ID NO:7、SEQ ID NO:11 和 SEQ ID NO:19。

23. 产生遗传转化的磷酸除草剂耐性植物的方法，所述方法包括

以下步骤:

a) 将多核苷酸序列插入植物细胞的基因组中, 所述多核苷酸序列包含:

- 5 i) 一个启动子序列, 其在植物细胞中起作用, 以引起 RNA 序列产生, 其有效连接于;
- ii) 一个结构 DNA 序列, 其能够产生编码酰基转移酶的 RNA 序列, 所述酰基转移酶将酰基从酰化供体转移至膦酸除草剂底物的末端胺, 其有效连接于;
- 10 iii) 一个 3'非翻译序列, 其在植物细胞中起作用, 引起聚腺苷酸化核苷酸序列添加至所述 RNA 序列的 3'端;

其中所述启动子序列相对于所述结构 DNA 序列是异源的, 并且引起所述酶在植物细胞中足够表达, 以增强用所述多核苷酸序列转化的植物细胞的膦酸除草剂耐性;

b) 选择转化的植物细胞; 和

15 c) 由所述转化的植物细胞再生表现出膦酸除草剂耐性提高的遗传转化植株。

24. 产生遗传转化的膦酸除草剂耐性植物的方法, 所述方法包括以下步骤:

a) 将多核苷酸序列插入植物细胞的基因组中, 所述多核苷酸序列包含:

- 20 i) 一个启动子序列, 其在植物细胞中起作用, 以引起 RNA 序列产生, 其有效连接于;
- ii) 一个结构 DNA 序列, 其能够产生编码酰基转移酶的 RNA 序列, 所述酰基转移酶将酰基从酰化供体转移至膦酸除草剂底物的末端胺, 其有效连接于;
- 25 iii) 一个 3'非翻译序列, 其在植物细胞中起作用, 引起聚腺苷酸化核苷酸序列添加至所述 RNA 序列的 3'端;

其中所述启动子序列相对于所述结构 DNA 序列是异源的, 并

且引起所述酶在植物细胞中足够表达，以增强用所述多核苷酸序列转化的植物细胞的所述磷酸除草剂耐性，并且其中所述植物细胞表达编码草甘膦氧化还原酶的 GOX 基因；

5 b) 选择转化的植物细胞；和

c) 由所述转化的植物细胞再生表现出磷酸除草剂耐性提高的遗传转化植株。

25. 按照权利要求 24 的方法，其中转化的植物细胞根据在选择剂存在下生长的能力来选择，其中所述选择剂选自草甘膦和 AMPA。

10 26. 包含酰基转移酶的肽，所述酰基转移酶催化将酰基从酰化供体转移至磷酸除草剂的末端胺。

27. 按照权利要求 26 的肽，其中所述酰基转移酶由选自 SEQ ID NO:4、SEQ ID NO:8、SEQ ID NO:12 和 SEQ ID NO:20 的序列组成。

15 28. 按照权利要求 26 的肽，其中所述磷酸除草剂选自草甘膦和 AMPA。

29. 按照权利要求 26 的肽，其中所述酰基转移酶在细胞中由是选自以下的多核苷酸序列或与选自以下的多核苷酸序列杂交的 DNA 序列表达：SEQ ID NO:3、SEQ ID NO:7、SEQ ID NO:11 和 SEQ ID
20 NO:19。

30. 按照权利要求 27 的肽，其中所述 DNA 序列衍生自微生物，所述微生物是选自以下属的成员：肠杆菌科(Enterobacteriaceae)、链霉菌属、芽孢杆菌属、放线杆菌属、Ascomycota 和 Basidiomycota。

25 31. 选择用载体转化的一种或多种细胞的方法，所述载体含有编码起作用以将磷酸除草剂化合物 N-乙酰化的酶的酰基转移酶基因，所述方法包括以下步骤：

a) 用所述载体转化细胞群体；

b) 在抑制量的磷酸除草剂化合物存在下温育所述转化的细

胞;

c) 鉴定在所述抑制量的所述化合物存在下生长的一种或多种细胞; 和

d) 分离并纯化在所述抑制量的所述化合物存在下生长的所述一种或多种细胞。

32. 包含酰基转移酶基因的按照权利要求 31 的载体, 所述酰基转移酶基因编码在所述一种或多种细胞中起作用以将酰基从酰化供体底物转移至膦酸除草剂化合物的末端胺的酶。

33. 按照权利要求 31 的载体, 其中所述基因在宿主细胞中表达。

34. 按照权利要求 33 的载体, 其中所述宿主细胞不受抑制量的膦酸除草剂存在的抑制, 所述抑制量是对于缺乏编码所述酶的功能性酰基转移酶基因的宿主细胞有抑制作用的量。

35. 按照权利要求 34 的宿主细胞, 选自细菌细胞、真菌细胞、动物细胞和植物细胞。

36. 按照权利要求 35 的宿主细胞, 其中所述细菌细胞选自包括杆菌科、分枝杆菌科 (Mycobacteriaceae)、土壤杆菌科 (Agrobacteriaceae)、放线杆菌科 (Actinobacteriaceae)、链霉菌属和芽孢杆菌属。

37. 按照权利要求 35 的宿主细胞, 其中所述真菌细胞选自包括 Ascomycota、Basidiomycota 和 Deuteromycota 的真菌菌种。

38. 按照权利要求 35 的宿主细胞, 其中所述植物细胞选自包括大豆 (Glycine max)、玉蜀黍 (Zea mays)、烟草 (Nicotiana tabacum)、陆地棉 (Gossypium gossypia)、普通小麦 (Triticum aestivum) 和欧洲油菜的植物物种。

39. 与酰基转移酶蛋白序列结合的抗体, 其中所述酰基转移酶选自 SEQ ID NO:4、SEQ ID NO:8、SEQ ID NO:12 和 SEQ ID NO:20。

40. 鉴定样品中重组酰基转移酶基因的方法, 所述方法包括:

a) 提供能够与所述基因杂交的一种或多种不同的多核苷酸序

列;

b) 提供包含互补于所述不同的多核苷酸序列的一种或多种多核苷酸序列的参比样品; 和

5 c) 提供用于将所述不同的序列、所述参比样品和样品中所述重组酰基转移酶基因混合的说明, 和

d) 检测所述样品中的所述重组基因。

41. 用于检测样品中重组酰基转移酶基因存在的试剂盒, 包括:

a) 提供能够与所述基因杂交的一种或多种不同的多核苷酸序列;

10 b) 提供包含互补于所述不同的多核苷酸序列的一种或多种多核苷酸序列的参比样品; 和

c) 提供用于将所述不同的序列、所述参比样品和样品中所述重组酰基转移酶基因混合的说明,

将以上物质和说明一起包装在试剂盒中。

15 42. 包含一种多核苷酸序列的植物, 所述多核苷酸序列含有一种编码乙酰基转移酶蛋白的基因, 其中所述基因在所述植物中的表达刺激所述植物的生长。

说明书

磷酸代谢植物

5

关于在先申请

该申请要求于 1998 年 11 月 17 日申请的美国临时申请序号 60/108,763 的优先权。

发明领域

10

本发明总的来讲涉及植物的除草剂抗性, 更具体地讲涉及一类新的磷酸(phosphonate)代谢基因, 以及涉及应用这些基因改进植物对磷酸除草剂(phosphonate herbicide)耐性的方法。

现有技术的描述

15

20

25

含磷有机分子可以是天然存在的或合成衍生的。含磷-碳(C-P)键有机分子也在自然界发现或作为合成化合物, 它如果多少被天然酶途径降解, 也通常不是被快速降解。含有直接碳-磷(C-P)键取代更通常已知的磷酸酯的碳-氧-磷键的合成的有机磷酸化合物和次磷酸化合物(Metcalf 等, Gene 129:27-32, 1993)因此已经广泛地用作杀虫剂、抗生素和除草剂(Chen 等, J. Biol. Chem. 265:4461-4471, 1990; Hilderbrand 等, *The role of phosphonates in living systems*, Hilderbrand, R.L. 编辑, 第 5-29 页, CRC Press. Inc., Boca Raton. FL, 1983)。磷酸在自然界中普遍存在, 磷酸被单独发现和多种生物中的多种多样的大分子结构中发现(Jiang 等, J. Bacteriol. 177:6411-6421, 1995)。磷酸分子的降解通过多种已知途径 - C-P 裂解酶途径、磷酸酶途径和 C-N 水解途径进行(Wanner, *Biodegradation* 5:175-184, 1994; Barry 等, 美国专利第 5,463,175 号, 1995)。已经鉴定了能够进行这些步骤的细菌分离物(Shinabarger 等, J. Bacteriol. 168:702-707, 1986; Kishore 等, J. Biol Chem. 262:12,164-

12,168, 1987; Pipke 等, Appl. Environ. Microbiol. 54:1293-1296, 1987; Jacob 等, Appl. Environ. Microbiol. 54:2953-2958, 1988; Lee 等, J. Bacteriol. 174:2501-2510, 1992; Dumora 等, Biochim. Biophys. Acta 997:193-198, 1989; Lacoste 等, J. Gen. Microbiol. 138:1283-1287, 1992)。然而, 除磷酸酶和草甘膦氧化酶(GOX)外, 尚未鉴定能够进行这些反应的其它酶。

几项研究一直集中于鉴定磷酸的 C-P 裂解酶降解所需的基因。Wackett 等(J. Bacteriol. 169:710-717, 1987)公开了放射形土壤杆菌 (*Agrobacterium radiobacter*)对磷酸降解的广泛的底物特异性以及具体利用草甘膦作为唯一磷酸源。Shinabarger 等和 Kishore 等公开了假单胞菌属物种将磷酸除草剂草甘膦通过肌氨酸中间体经 C-P 裂解酶降解为甘氨酸和无机磷酸。

先前已经表明, 大肠杆菌 B 菌株能够利用磷酸(Chen 等), 而大肠杆菌 K-12 菌株不能降解磷酸。然而, 随后表明, K-12 菌株含有一组完整的能够利用磷酸的隐蔽性基因(*psiD* 或 *phn*)(Makino 等), 因为根据在含甲基磷酸或乙基磷酸作为唯一磷源的低磷酸培养基上的生长, 容易选择出突变体。随后表明, 适应于在甲基磷酸或乙基磷酸上生长的这样的 K-12 菌株能够利用其它磷酸作为唯一磷源(Wackett 等, J. Bacteriol. 169:1753-1756, 1987)。

Avila 等(J. Am. Chem. Soc. 109:6758-6764, 1987)关心与氨甲基磷酸相关的生物降解和解毒过程的机制评估, 包括阐明中间体、产物和降解性去磷酸化过程的机制。Avila 等研究了由先前适应于在乙基磷酸上生长的大肠杆菌 K-12 培养物中的多种氨基磷酸底物生成去磷酸化生物降解产物。此外, Avila 等在某些研究中利用 N-乙酰-AMPA (N-乙酰-氨基-甲基-磷酸)作为唯一磷酸源, 以便表明乙酰化 AMPA 对于 C-P 键裂解不是抑制性的。另外, Avila 等注意到, N-乙酰-AMPA 在大肠杆菌 K-12 生长期间能够充当唯一磷酸源, 然而, 当将 AMPA 用作唯一磷酸源时, 他们没有观察到 N-乙酰-AMPA 的生成。他们的结

果表明, AMPA 在大肠杆菌中不是乙酰化的底物。

Chen 等通过互补克隆到磷酸利用缺陷的大肠杆菌 K-12 菌株中, 这使得所述 K-12 菌株能够利用磷酸作为唯一磷酸源, 从大肠杆菌 B 中鉴定出功能性 *psiD* 基因座(J. Biol. Chem. 265:4461-4471, 1990)。

5 Chen 等因此公开了在 15.5 kb *Bam*HI 片段上鉴定出的 *psiD* 互补性基因座的 DNA 序列, 含有 17 个可读框, 命名为 *phnA-phnQ*, 包含大肠杆菌 B 的 *phn* 操纵子。随后发现大肠杆菌 K-12 的隐蔽性 *phn (psiD)* 操纵子在 *phnE* 中含有一个 8 个碱基对的插入。导致 *phnE* 中移码, 这不仅产生缺陷型 *phnE* 基因产物, 而且明显引起对该操纵子中下游基因表达的极性效应, 这防止磷酸利用(Makino 等, J. Bacteriol. 173:2665-2672, 10 1991)。Makino 等的研究已经更精确地描述该操纵子含有基因 *phnC-phnP*。进一步的研究已经涉及对该操纵子中每个基因的功能性质的了解(Chen 等, J. Biol. Chem. 265:4461-4471, 1990; Makino 等, J. Bacteriol. 173:2665-2672, 1991; Wanner 等, FEMS Microbiol. Lett. 100:133-140, 15 1992; Metcalf 等, Gene 129:27-32, 1993; Ohtaki 等, Actinomyceteol. 8:66-68, 1994)。在所有这些研究中, 已经根据与含结构螺旋-回折-螺旋基序的其它核苷酸结合蛋白的相似性, 暗示 *phnO* 基因作为调节蛋白。此外, *phn* 操纵子中基因的诱变证明至少对于所测试的磷酸, *phnO* 不是磷酸利用所需的, 这进一步支持了所提出的该基因的调节功能(Metcalf 等, J. Bacteriol. 173:587-600, 1991)。已经从其它细菌中鉴定出同源 *phn* 序列, 包括与大肠杆菌 *phnO* 基本上相似的基因, 采用由大肠杆菌 *phnO* 基因的核苷酸序列推导出的核苷酸序列, 从 *S. griseus* 中分离出该基因(Jiang 等, J. Bacteriol. 177:6411-6421, (1995); McGrath 等, Eur, J. Biochem. 234:225-230 (1995); Ohtaki 等, Actinomyceteol. 8:66-69 (1994))。然而, 关于 *phnO*, 除作为调节因子外没有提出其它功能。20 在最近的综述中, 再次引述了 CP 裂解酶操纵子中 *phnO* 的调节作用(Berlyn, Microbiol. Molec. Biol. Rev. 62:814-984, 1998)。

分子生物学的进展, 特别是植物学结合重组 DNA 技术的进展,

已经使得人们能够构建含有农学上重要的非天然基因的重组植物。另外，当将这类基因掺入植物并在植物中表达时，这类基因理想地付给所述重组植物以某些有益性状或特征。一种这样的性状是除草剂抗性。能够在除草剂存在下生长的重组植物具有优于除草剂敏感性物种的非常大的优点。另外，除草剂抗性植物通过减少耕地除草和控制自生植物的需要，为农作物生产提供更加成本有效的方法。

数十年来，一直在使用化学除草剂来抑制植物的代谢，特别是用于农学目的，作为作物大田中除草或控制自生植物的方法。已经证明对于这些目的特别有用的一类除草剂被称为膦酸除草剂。也许在农学上最为成功的膦酸除草剂是草甘膦(N-膦酰基-甲基-甘氨酸)。

已经构建了耐受膦酸除草剂草甘膦的重组植物。当应用于植物时，草甘膦被吸收到植物组织中，抑制芳族氨基酸的形成，这是通过抑制质体定位的 5-烯醇丙酮酰(pyruvyl)-3-磷酸莽草酸合酶的活性而介导的，5-烯醇丙酮酰(pyruvyl)-3-磷酸莽草酸合酶也称为 EPSP 合酶或 EPSPS，该酶被普遍认为是植物、细菌和真菌特有的。已经用对草甘膦抑制的敏感性低得多的细菌 EPSPS 酶转化了重组植物。因此，表达该细菌 EPSPS 的植物对草甘膦的敏感性低得多，通常被鉴定为草甘膦抗性。因此，可以将更大量的草甘膦应用于这样的重组植物，确保对该除草剂感病或敏感的植物死亡。然而，已经鉴定出其它基因，当将其转化到植物基因组中时，它们编码也提供草甘膦抗性的酶。一种这样的酶被描述为 GOX 或草甘膦氧化还原酶。GOX 的作用通过催化草甘膦降解为氨甲基膦酸(AMPA)和乙醛酸，提供保护植物抵抗膦酸除草剂草甘膦。由于草甘膦降解产生的 AMPA 可以引起褪色和矮化或植物生长减低以及其它不希望有的特性。许多植物种也对外源应用的 AMPA 敏感，以及对由于 GOX 介导的草甘膦除草剂降解产生的内源 AMPA 敏感。没有描述公开了保护植物抵抗膦酸除草剂(例如 AMPA)应用的方法。

Barry 等(美国专利第 5,633,435 号)公开了编码 EPSP 合酶的基因，

它们可用于产生转化的细菌和耐受作为除草剂的草甘膦的植物，以及公开了应用这样的基因作为在种植的转基因作物大田中选择性除草的方法。 Barry 等(美国专利第 5,463,175 号)公开了编码草甘膦氧化还原酶(GOX)酶的基因，所述基因可用于产生降解草甘膦除草剂的转化细菌和植物以及耐受作为除草剂的草甘膦的作物。 Barry 等(美国专利第 5,463,175 号)公开了 AMPA 作为 GOX 介导的草甘膦代谢产物的形成。已经报道了 AMPA 对于大多数植物种(Franz, 1985)而不是对于所有植物种(Maier, 1983; Tanaka 等, 1986)的植物毒性比草甘膦低得多。其编码的蛋白能够中和或代谢草甘膦降解产生的 AMPA 的基因的共表达将提供比单独应用 GOX 大得多的改进。因此，克服对由于草甘膦降解而形成 AMPA 的敏感性的方法、或当 AMPA 用作除草剂或在植物转化方法中用作选择剂时抗 AMPA 的方法将可用于在转基因植物和对这类化合物敏感的其他生物中提供增强的或改进的除草剂耐性。

已经描述了应用草甘膦作为化学杀配子剂(美国专利第 4,735,649)。其中，公开了草甘膦可以在最适条件下杀伤约 95% 的雄配子，而留下约 40-60% 的能够受精的雌配子。另外，在所公开的应用水平上通常观察到矮化效应，表现为植株大小减小并且有少量褪色。因此，同大多数杀配子剂的情况一样，采用草甘膦作为杀配子剂的主要缺点是由于对雄配子缺乏足够的选择性而产生的植物毒性副作用。可以通过用草甘膦处理后在表达 GOX 的转基因植物中产生 AMPA，实现这些植物毒性表现。因此，提供在表达 GOX 的转基因植物中防止作为应用草甘膦作为杀配子剂的副作用的矮化效应和褪色的方法可能是有利的。此外，一个更有效的方法可能最理想杀伤 95% 以上的雄配子或防止雄配子成熟的方法，并且将留下 60% 以上的基本上不受影响的雌配子。认为与编码能够将 AMPA N-酰基化的酶的转酰基酶基因一起进行 GOX 的组织特异性共表达，将达到这一目标。

现在已经发现，大肠杆菌 *phnO* 基因编码具有转酰基酶、酰基转移酶或酰基-CoA 转酰基酶活性的酶，其中优选的底物是显示出末端胺

的磷酸,特别是氨甲基磷酸(AMPA)。从酰基-CoA 将酰基转移至 AMPA 的游离末端胺上,导致形成 N-酰化 AMPA。不知道植物在多大程度上将 AMPA 酰化,并且已经表明某些植物对 AMPA 敏感,而对酰基-AMPA 不敏感。因此,在植物中表达 *phnO* 将在增强磷酸除草剂耐性方面是有用的,特别是当在植物转化中将 AMPA 用作除草剂或选择剂时,更尤其是当在结合表达 GOX 基因的重组植物,将草甘膦用作除草剂时是有用的。

发明概述

因此简而言之,本发明涉及包含一类编码能够将磷酸化合物 N-酰化的蛋白的新基因的物质组合物,涉及应用这些基因和所编码的蛋白以改进植物对磷酸除草剂耐性的方法。本发明也涉及用编码能够将磷酸化合物 N-酰化的蛋白的基因转化的重组植物和微生物的选择方法,涉及能够将化合物 N-氨基-甲基-磷酸(N-AMPA)和其它相关磷酸化合物 N-酰化的肽。另外,本发明也涉及应用用转酰基酶基因转化植物以防止自花受精的方法或涉及增强植物中异花受精的方法。

因此在发现由本发明达到的几个优点中,可以注意到提供了产生稳定转化的除草剂耐性重组植物,它们已经在其基因组中插入了编码所需基因产物、最好是 N-酰基-转移酶的多核苷酸序列。所述多核苷酸序列最好由含启动子序列的盒组成,所述启动子在植物中起作用并有效连接于结构 DNA 序列的 5',所述结构 DNA 序列当转录出 RNA 序列时,编码 N-酰基转移酶肽。所述启动子序列可以相对于所述结构 DNA 序列是异源的,并且引起在植物组织中所述转移酶的足够表达,以为用所述多核苷酸序列转化的植物提供除草剂耐性。所述结构序列最好有效连接于在植物中起作用的 3'非翻译聚腺苷酸化序列的 3',并且当所述 3'非翻译聚腺苷酸化序列与所述结构序列一起转录为 RNA 时,使得聚腺苷酸核苷酸序列添加至所转录的 RNA 的 3'端。所述结构 DNA 序列的表达在所述植物组织中产生足够水平的所述转酰基

酶，以增强所述转化植物的除草剂耐性。

5 作为再一实施方案，所述结构 DNA 序列也可以含有一个额外的编码氨基末端肽序列的 5' 序列，该序列在植物中起作用，以将所述结构序列转录产生的肽导向胞内细胞器。这一额外的编码序列最好与编码所述转酰基酶的结构序列符合读框地连接。所述氨基末端肽序列可以或者是信号肽或者是转运肽。所述胞内细胞器可以是叶绿体、线粒体、液泡、内质网或其它这样的结构。所述结构 DNA 序列也可以连接于可以用来增强所述所需基因产物表达的 5' 序列，例如非翻译前导序列(UTL's)、内含子序列或这些序列的组合等。也可以将内含子序列引入编码所述转酰基酶的结构 DNA 序列中。另一方面，叶绿体转化或质体转化可以导致转酰基酶编码序列和酶定位至叶绿体或质体，消除了核基因组转化、从所述核基因组表达和随后将所述基因产物导向亚细胞器的需要。

15 最好是，所述重组植物表达编码催化 AMPA 形成的酶的基因。AMPA 形成可以产生于天然存在下前体的代谢、例如提供给所述植物的草甘膦的前体，或可以产生于通过某些分解代谢途径形成 AMPA。GOX 与 AMPA 转酰基酶表达的共表达，提供了意想不到的对某些膦酸除草剂更具抗性的植物。然而，允许仅用 N-转酰基酶转化的植物在 AMPA 或者相似或相关化合物存在下生长的一个实施方案，将提供用于鉴定遗传转化的植物、愈伤组织或胚胎发生组织的有用的选择方法。

25 按照本发明的另一方面，提供用于选择性地增强或改进重组植物的除草剂耐性的方法，所述重组植物在其核基因组、叶绿体基因组、质体基因组或线粒体基因组中已经插入了一种包含编码 N-酰基转移酶的多核苷酸序列的盒。

再一实施方案包括对在表达 GOX 基因的转化植物中选择性地增强除草剂耐性的方法的改进，所述 GOX 基因编码在产生转酰基酶的不同植物中表达的草甘膦氧化还原酶。

按照本发明的另一方面，提供通过将包含编码 N-酰基转移酶的多核苷酸序列的盒插入植物细胞基因组中，产生遗传转化的除草剂耐性植物的方法。

5 再一实施方案包括产生对来自表达 GOX 基因的植物细胞的遗传转化除草剂耐性植物的方法的改进，所述 GOX 基因编码在产生转酰基酶的不同植物细胞中表达的草甘膦氧化还原酶。

在任一上述实施方案中，所述除草剂耐性植物或植物细胞可以选自：玉米、小麦、棉花、水稻、大豆、甜菜、canola、亚麻、大麦、油料种子油菜、向日葵、马铃薯、烟草、番茄、苜蓿、莴苣、苹果、10 杨树、松树、桉树、金合欢、杨树、美国枫香、radiata pine、火炬松、云杉、柚木、苜蓿、三叶草和其它饲料作物、草坪草、油棕、甘蔗、香蕉、咖啡、茶树、可可、苹果、胡桃、杏仁、葡萄、花生、pulses、矮牵牛、万寿菊、长春花、秋海棠、天竺葵、三色堇、凤仙花、燕麦、高粱和小米。

15 按照本发明的另一方面，提供能够将化合物 N-氨甲基磷酸(N-AMPA 或 AMPA) N-酰化的肽或当应用于、引入植物代谢或由植物代谢产生时能够引起植物毒性效应的其它这样的化合物。一种这样的肽是得自大肠杆菌 *phnO* 结构基因序列表达的 N-氨甲基磷酸转酰基酶 (AAT)。也设想了结构和功能与大肠杆菌 *phnO* 基因产物相似的其它20 肽。

本发明的另一方面提供用含转酰基酶基因的载体转化的细胞的选择方法，所述转酰基酶基因表达能够将 AMPA 和相似的化合物 N-酰化的酶。所述方法包括以下步骤：用所述载体转化一个细胞群体，以及在通过在足以抑制任何非转化细胞生长或存活的量的 AMPA 存在25 下温育来选择所述转化细胞后，从所述群体中的非转化细胞分离和纯化所述转化细胞。所述转化细胞可以是细菌细胞、植物细胞或真菌细胞。细菌细胞可以是其中肠杆菌科(*Enterobacteraceae*)、分枝杆菌科(*Mycobacteraceae*)、土壤杆菌科(*Agrobacteraceae*)和放线细菌科

(*Actinobacteraceae*) 包括的任一科的成员。真菌细胞可以是 *Ascomycota*、*Basidiomycota* 等的成员。植物细胞可以得自 *Plantae* 科的任何成员。

5 本发明的再一实施方案提供从来源于用转酰基酶基因转化的植物的组织、细胞或植物其它部分产生植物的方法，所述转酰基酶基因为 *phnO* 基因、*gox* 基因、其中 GOX 和转酰基酶肽由翻译融合体或转录融合体、或编码 GOX 和转酰基酶肽的多顺反子基因的产生。

10 本发明的再一实施方案提供用于产生以组织特异性方式表达作为反义基因的全部或部分 *phnO* 基因或相似转酰基酶基因或 GOX 基因的植物的方法。

其它方面也包括试剂，例如针对 AMPA 转酰基酶的抗体，和用于鉴定转酰基酶基因序列的多核苷酸。这些试剂可以包括在试剂盒中，所述试剂盒含有 AMPA 转酰基酶、为 AMPA 转酰基酶基因序列或与其互补的多核苷酸、用于热扩增 AMPA 转酰基酶基因序列的多核苷酸、针对 AMPA 转酰基酶、用于在实验室中或大田中检测 AMPA 转酰基酶的抗体、以及用于试剂盒形式以及用于本文设想的其它测定中的任何其它必需的试剂。

20 本发明的再一目的是提供应用磷酸除草剂作为化学杂交剂的方法。该方法提供选择性杀配子效应和产生雄性不育植物。可以将这样的植物工程改造，使得 *gox* 或 *phnO*、或 *gox* 和 *phnO* 不能在繁殖所需的植物组织中表达，引起对所应用的抑制成熟配子结构形成的植物毒性化合物的敏感性。

附图简述

25 图 1 说明了 [¹⁴C] 同位素检测 HPLC 色谱图，表示仅含 [¹⁴C] 草甘膦 (11.3 分钟, 98.8%) 和微量 [¹⁴C] AMPA (5.8 分钟, 0.16%) 和未鉴定的 [¹⁴C] 物质 (10.2 分钟, 1%) 的施药溶液样品。

图 2 说明了所观察的放射性代谢物 [¹⁴C] AMPA、[¹⁴C] 草甘膦和 N-

乙酰-[¹⁴C]-AMPA 以及鉴定为 N-乙酰-N-甲基-[¹⁴C]-AMPA 的杂质的标准的混合物的 HPLC 图形。

图 3 说明了来自用 GOX 和 AMPA 转酰基酶转化、并用 [¹⁴C] 草甘膦处理的玉米愈伤组织的提取物的代表性 HPLC 图形。所述峰表示 [¹⁴C] 草甘膦(10.8 分钟, 所观察的总 [¹⁴C] 的 92.5%)、[¹⁴C] AMPA 主要由 GOX 介导的草甘膦降解产生(5.98 分钟, 所观察的总 [¹⁴C] 的 1.71%)和由于在愈伤组织中表达的重组 AMPA 转酰基酶介导的 [¹⁴C] AMPA 酰化产生的 N-乙酰-[¹⁴C] AMPA(13.29 分钟, 所观察的总 [¹⁴C] 的 4.54%)。

图 4 说明质粒 pMON17261。

图 5 说明质粒 pMON32571。

图 6 说明质粒 pMON32936。

图 7 说明质粒 pMON32946。

图 8 说明质粒 pMON32948。

发明详述

提供本发明的以下详细描述, 以有助于本领域技术人员实施本发明。虽然如此, 以下详细描述不应该解释为过度限制本发明, 因为在不偏离本发明性发现的精神或范围的情况下, 本领域技术人员可以对本文讨论的实施方案进行各种修改和改变。

许多词语和短语是分子生物学、微生物学、蛋白质化学和植物学领域众所周知的, 一般具有其普遍和通常理解的含义, 除非另有说明 (otherwise to be taken in context)。然而, 本文所用的以下词语和短语具有以下一般叙述的含义。

AMPA 转酰基酶。 本文所用的 AMPA 转酰基酶是指在从酰基载体化合物(例如辅酶 A, 它是众所周知的, 并且在生物学和化学领域中缩写为 CoA)转移酰基化学基团中起作用的酶。特别是, AMPA 转酰基酶将酰基化学基团从酰基载体转移至氨甲基磷酸的游离氨基, 众所周知氨甲基磷酸是草甘膦氧化还原酶介导的草甘膦代谢的副产物。

AMPA 转酰基酶(AAT)在本文中也可以称为 AMPA 乙酰基转移酶、AMPA 转酰基酶或乙酰-AMPA 合酶(AAS), 在本文中已经表明它具有乙酰转移酶活性、丙酰转移酶活性、丙二酸单酰(malonyl)转移酶活性和琥珀酰转移酶活性。因此, 用作能够以 AMPA 转酰基酶起作用的酰基载体形式的底物的这些化合物的任何生物功能等同物(乙酰、丙酰、丙二酸单酰或琥珀酰), 均在本发明范围内。已经鉴定一种 AMPA 转酰基酶, 并且该酶在本文中通过实施例表明按照本文包括的描述起作用, 该酶先前在本领域中已经被成为 PhnO, 这是一种由大肠杆菌 *phn* 操纵子中的 *phnO* 基因编码的蛋白。

10 **生物功能等同物:** 在本发明 AMPA 转酰基酶蛋白方面本文使用的这种等同物是这样的肽、多肽和蛋白, 它们所含的序列或部分表现出与本发明的新肽(例如 PhnO)有序列相似性, 并且表现出与本文公开的多肽相同或相似的功能特性, 包括转酰基酶活性。生物等同物也包括与针对 PhnO 产生的抗体反应即与其特异性结合、并表现出相同或相似转酰基酶活性的肽、多肽和蛋白, 所述抗体包括单克隆抗体和多克隆抗体。

15 在编码转酰基酶的基因方面本文使用的生物功能等同物是这样的多核苷酸, 它们与本文设想并描述的多核苷酸序列反应, 即它们能够与编码在 AMPA 酰基转移中起作用的转酰基酶、或编码本文设想并描述的大致相似的转酰基酶蛋白的多核苷酸或与其互补的多核苷酸的多核苷酸序列杂交。与本文描述的蛋白大致相似的蛋白是生物功能等同物, 并表现出与本文公开的所述多肽相同或相似的功能特性, 包括改进的除草剂耐性或改进的除草剂抗性。生物等同肽含有例如表现出与本发明的新肽(例如 PhnO)有序列相似性的一个或更多个活性位点的序列或部分。生物等同物也包括与针对 PhnO 序列和 PhnO 样肽序列产生的抗体反应即与其特异性结合、并表现出相同或相似除草剂耐性或除草剂抗性方面改进的肽、多肽和蛋白, 所述抗体包括单克隆抗体和多克隆抗体。

如本文所用的叶绿体或质体定位的是指或者为多核苷酸或者为多肽的生物分子，它们定位于叶绿体或质体中，使得所述分子与细胞质环境分离，并在叶绿体或质体胞质中起作用以提供本发明中宣布的效应。生物分子定位至叶绿体或质体，在多核苷酸方面，可以通过人工机械方法，例如电穿孔、机械微注射或通过多核苷酸包被的微粒轰击而发生，或在多肽方面，可以通过分泌或输入方法而发生，其中使用用来将所连接的多肽导向、插入、辅助或定位到叶绿体或质体中的天然存在的、非天然存在的或异源质体或叶绿体引导肽序列。

事件是指由于外源 DNA 插入核 DNA、线粒体 DNA、质体 DNA 或叶绿体 DNA 中一个或多个特有位点而得到的转基因植物或植物组织。

表达：编码 DNA 分子(例如结构基因)所经历的产生基因产物的胞内过程的组合，所述胞内过程包括转录、翻译以及其它胞内蛋白和 RNA 的加工和稳定功能。

非天然存在的基因：本发明的非天然存在的酰基转移酶基因含有编码植物功能性 RNA 序列的遗传信息，但最好是编码酰基转移酶蛋白的基因，所述酰基转移酶蛋白无论是天然存在的还是天然存在的蛋白的变体，其制备方式涉及任何类别的遗传分离或操作。这包括从其天然存在的状态分离所述基因；通过密码子修饰对所述基因进行操作；定点诱变、截短、引入或除去限制性内切核酸酶酶切位点；通过体外方法(例如亚磷酰胺化学合成法等)、热扩增方法(例如聚合酶链式反应、连接酶链式反应、反相聚合酶反应等)和任何其它操作性或分离性方法合成或再合成编码本发明的转酰基酶的天然存在的序列。

有效连接的(operably linked)：符合读框地连接、使得一个区段的特性影响另一个区段表达的核酸区段。例如，具有聚合酶加载、结合并起始转录功能的特性的启动子序列影响与所述启动子连接的序列的表达。

植物可表达的编码区：因为含有促进目的基因表达的典型的植物

调节元件而可在植物中表达(即可以在植物中被转录和/或翻译)的编码区。

5 质体转运肽: 可用于将所连接的氨基酸(例如蛋白融合体)导向或定位至亚细胞区室或亚细胞器(例如质体或叶绿体)的任何氨基酸序列。促进进入线粒体的氨基酸序列完全不同于质体转运肽或与其不相似, 也被描述为转运肽, 但它不能起作用以将肽序列导向质体或叶绿体细胞器。

10 转基因植物的后代包括至少含有一种异源基因或转基因的所述转基因植物的任何后代, 或从在其谱系中具有所述转基因的转基因植物得到的任何后来的植物。后代不限于一个世代, 而是包括所述转基因植入的许多后代, 只要它们含有或表达所述转基因。含有转基因胚的种子以及得自所述转基因植物及其后代的种子, 也是本发明的重要部分。含有所需转基因的转基因细胞、组织、种子或植物是原始转基因细胞、组织或植物的后代。

15 启动子: DNA 序列或一组 DNA 序列上的识别位点, 它为结构基因提供表达控制元件, 并且 RNA 聚合酶与其特异性结合并启动该基因的 RNA 合成(转录)。

R_0 是得自培养的植物组织或细胞转化的原代再生植株。得自所述 R_0 的后来的后代或世代称为 R_1 (第一代)、 R_2 (第二代)等。

20 再生: 通过由植物细胞或植物组织(例如植物原生质体或外植体)培育植株而产生全株的过程。

结构编码序列是指编码肽、多肽或蛋白的 DNA 序列, 所述肽、多肽或蛋白是从所述结构编码序列转录出信使 RNA (mRNA)、然后由 mRNA 翻译产生所需肽、多肽或蛋白产物后而产生的。

25 结构基因: 被表达而产生多肽的基因。

大致同源性: 当本文使用该术语时, 大致同源性是指约 40%至约 65%、约 66%至约 75%、约 76%至约 86%、约 87%至约 90%、91%至约 95%、以及约 96%至约 99%的核酸序列与参比多核苷酸序列(例如大

肠杆菌 *phnO* 基因序列)同源。与第二多核苷酸分子大致同源的第一多核苷酸分子是所述第二多核苷酸或与其互补,使得所述第一多核苷酸分子与所述第二多核苷酸分子或其互补序列在严格杂交条件下杂交,严格性定义为导致第一多核苷酸与第二多核苷酸杂交所需的最适盐浓度和温度。改变严格性的方法是本领域众所周知的,但可能在
5 Sambrook 等(编辑), *Molecular Cloning: A Laboratory Manual*, 第二版, 1989, Cold Spring Harbor Press; 或 Ausubel 等(编辑), *Short Protocols in Molecular Biology*, 第三版, 1995, John Wiley and Sons, Inc. 中给出的参考文献。认为在本发明范围内的多肽是与参比多肽序列、最好与大肠杆菌 *PhnO* 肽序列约 40%至约 65%、约 66%至约 75%、约 76%至约 86%、约 87%至约 90%、91%至约 95%、以及约 96%至约 99%相似的那些多肽。
10

终止子: 如本文在计划用于在植物中表达的植物特异性序列方面使用的,为 3'端转录终止和聚腺苷酸化序列。

转化: 将外源多核苷酸序列(例如质粒或病毒载体或重组多核苷酸分子)引入细胞、原生质体、质体或叶绿体、或线粒体中的方法,其中所述外源多核苷酸序列或者掺入包含于所述细胞中的内源多核苷酸序列中,或者能够自主复制。转化的细胞是通过将一种或更多种外源多核苷酸分子引入该细胞中已经改变的细胞。稳定转化的细胞是全部或部分所述外源多核苷酸已经掺入所述细胞的核基因组物质、线粒体基因组物质或质体或叶绿体基因组物质中、使得所述外源多核苷酸赋予所述细胞以及所述转化细胞的后代某个或某些基因型或表型性状,所述性状可通过以下方法测定:检测外源引入的多核苷酸、所述外源多核苷酸的 mRNA 或蛋白产物、通常由所述细胞在缺乏所述外源多核苷酸时不产生的、或在所述细胞中没有发现的代谢物、或目测检查得自
15
20
25 所述转化细胞的细胞、植物组织或植物。

转基因: 转基因是一种多核苷酸序列,所述多核苷酸序列已经被转移至细胞并包含表达盒,所述表达盒含有编码所需多肽的结构基因

序列。所述转基因能够在受体转化细胞、组织或生物中表达。这可以包括完整的质粒或其它载体，或可以仅仅包括所述转移的多核苷酸的植物功能性编码序列。转基因细胞是得自或再生自转化细胞(包括原始转化细胞)的任何细胞。典型的转基因细胞包括得自转化植物细胞和得自稳定转化的转基因植物的特定细胞(例如叶片、根、茎、分生组织和其它组织体细胞或生殖细胞或种系和绒毡层细胞)的植物愈伤组织。转基因事件是将至少一种外源多核苷酸插入植物细胞或原生质体的核基因组、质体基因组或线粒体基因组而得到的植物或其后代。转基因植物是已经被遗传修饰以含有异源多核苷酸序列并将其表达为蛋白或 RNA 或不是先前所述植物组合物中一部分的 DNA 分子的植物或其后代。如本文具体例举的，例如遗传修饰转基因棉花植株，使其含有并表达有效连接于在植物细胞或组织中或在全株中起作用的转录和翻译控制序列并在其调控之下的至少一种异源 DNA 序列。转基因植物也可以称为转化植物。转基因植物也指原始转基因植物的后代，其中那些后代含有所述异源编码序列并能够在本文所述植物可表达转录和翻译控制序列的调节控制下表达所述异源编码序列。转基因植物可以产生转基因花、种子、球茎、根、块茎、果实和花粉等，并且可以通过常规育种方法与植物亲和系杂交，产生杂种转基因植物。

载体：能够在宿主细胞中复制和/或可以有效连接另一 DNA 或其它多核苷酸序列以导致所连接的序列复制的 DNA 分子或其它多核苷酸分子。质粒是一种典型的载体。

按照本发明，已经发现当用编码能够降解草甘膦的酶的某些基因转化植物时，植物可以产生植物毒性化合物。特别是，草甘膦氧化还原酶(GOX)介导的草甘膦代谢产生鉴定为 N-氨甲基-磷酸(AMPA)的植物毒性化合物。其它研究已经表明，AMPA 的 N-酰化衍生物 - N-酰基-氨甲基-磷酸(N-酰基-AMPA 或酰基-AMPA)对大多数植物种的植物毒性低得多。已经鉴定了能够通过酰化机制共价修饰 AMPA、导致 N-酰基-AMPA 形成的酶。特别是一种酶引起外源供应的 AMPA 被 N-乙

酰化。在表达该酶与 GOX 的植物中，没有观察到 AMPA 植物毒性效应。

在此设想的本发明利用包含用于调节基因表达的元件的重组多核苷酸盒，在所述多核苷酸盒可以插入序列，例如编码有用蛋白的结构基因。最好采用本领域熟知的限制性内切核酸酶，完成将这样的序列插入表达盒中，然而其它插入方法也是已知的。例如定点重组法有效地将所需序列插入这样的表达盒中。表达盒至少含有一种植物有效的启动子，以用于起始信使 RNA 分子的产生，从所述信使 RNA 分子翻译所述有用的蛋白。盒也含有鉴定为 3' 序列的植物有效序列，其作用是终止转录，并提供被 3' 聚腺苷酸化的非翻译序列。因此，计划用于植物的表达盒应该至少含有一种启动子序列，并且在其 3' 端与转录终止和聚腺苷酸化序列的 3' 连接。最好是，存在含有一个或多个特有限制性内切核酸酶酶切位点的多克隆序列或接头序列，将所述启动子和便于插入结构基因序列和其它元件的 3' 序列桥接。计划用于植物的表达盒也最好含有一段插入所述启动子和所述 3' 序列之间的 5' 非翻译序列。已经表明，5' 非翻译序列 (UTL) 增强植物中的基因表达。也设想了内含子作为可以存在于本发明的这种表达盒中的序列。也已经表明，植物有效内含子的存在增强某些植物种 (特别是玉米) 中的基因表达。内含子可以存在于表达盒中沿所述盒序列的任何编号的位置。这可以包括所述启动子和所述 3' 终止序列之间和/或结构基因内的位置。在表达盒中可以存在一个以上的内含子，然而对于本文设想的发明的目的，最好在表达盒用于单子叶植物和植物组织中时存在内含子。增强子序列也是本领域熟知的，并且可以存在，尽管不必作为表达盒的一部分，因为已知增强子序列在存在于驱动目的基因表达的启动子的上游或下游或甚至距启动子很远时起作用。

定位至植物核基因组并且以双链 DNA 形式存在的基因的表达，涉及由 RNA 聚合酶从所述 DNA 的一条链转录以产生初级信使 RNA 转录物 (mRNA)，并且随后在细胞核内加工所述 mRNA 初级转录物。

这种加工涉及将聚腺苷酸核苷酸添加至所述 RNA 的 3'末端的 3'非翻译多核苷酸序列。DNA 转录为 mRNA 由通常称为“启动子”的 DNA 序列来调节。所述启动子包含一端碱基序列，所述碱基序列发送信号使 RNA 聚合酶与该 DNA 结合并启动采用模板 DNA 链的 mRNA 转录，以制备相应的 RNA 互补链。

本领域技术人员会认识到，有许多在植物细胞中有活性的启动子，并且已经描述于文献中。这样的启动子可以得自植物、植物病毒或植物共生体、腐生、共生或致病微生物，包括但不限于胭脂碱合酶(NOS)和章鱼碱合酶(OCS)启动子(它们在根癌土壤杆菌(*Agrobacterium tumefaciens*)的肿瘤诱导质粒上携带)、花椰菜花叶病毒(CaMV) 19S 和 35S 启动子、来自核酮糖 1,5-二磷酸羧化酶小亚基(ssRUBISCO，一种非常丰富的植物多肽)的光诱导型启动子、水稻 *Act1* 启动子、玄参花叶病毒(FMV) 35S 启动子、甘蔗杆状 DNA 病毒启动子、遍在蛋白启动子、花生褪绿线死病病毒启动子、鸭趾草黄色病毒(*comalina yellow virus*)启动子、叶绿素 a/b 结合蛋白启动子和分生组织增强启动子 *Act2*、*Act8*、*Act11* 和 *EF1a* 等。所有这些启动子均已经用来产生各种类型的 DNA 构建体，所述 DNA 构建体已经在植物中表达(参见例如 McElroy 等, 1990; Barry 和 Kishore, USP 5,463,175)并且在本发明范围内。也设想了叶绿体或质体特异性启动子、叶绿体或质体功能性启动子以及叶绿体和质体有效启动子。最好是，所选定的特定启动子应该能够引起在植物中的足够表达，以导致产生有效量的转酰基酶，以赋予植物基本上耐受膦酸除草剂和膦酸除草剂代谢产物。提供所需耐性所需的转酰基酶的量可以根据植物种而变化。

一组优选的启动子是组成型启动子，例如 CaMV35S 或 FMV35S 启动子，它们在大多数植物器官中产生高水平的表达。增强形式或重复形式的 CaMV35S 和 FMV35S 启动子在本发明的实施中特别有用(Kay 等, 1987; Rogers, USP 5,378,619)。另外，也可能优选导致在所述植物的特定组织(例如叶片、茎、根、块茎、种子、果实等)中表达所

述转酰基酶基因，并且所选定的启动子应该具有所需的组织和发育特异性。因此，应该通过选择具有所需组织表达能力和接近的启动子强度的启动子，并且选择在所述靶组织中产生所需除草剂耐性的转化体，来优化启动子功能。这种从转化体库中的选择方法常规用于植物的异源结构基因表达，因为在含有相同的异源基因的转化体之间由于在所述植物基因组中基因插入位点而有变异。（通常称为“位置效应”）。除已知在植物细胞中引起DNA转录(组成型或组织特异性)启动子外，通过筛选植物cDNA文库中选择性表达或最好在靶组织中表达的基因，然后确定启动子区，可以鉴定其它启动子以用于本发明。

最好是，所使用的启动子在除其它组织外的所有分生组织中具有相对高的表达，因为现在已知膦酸除草剂可以被转运到该类型植物组织中并在其中积累。另一方面，可以用嵌合基因的组合来累积性地导致必需的转酰基酶总体表达水平，以产生除草剂耐性表型。提供相对较高水平表达的启动子可以引起所需蛋白在植物中产生至范围为0.1毫克/克鲜重植物组织、至0.5毫克/克鲜重植物组织、至1.0毫克/克鲜重植物组织、至2.0毫克/克鲜重植物组织或更高的水平。在田间遗传上等基因作物中所需蛋白的植物内水平可以跨越一定的范围，但一般水平为在特定样品中所分析的所有植株平均值的70%内、更优选在50%内、甚至更优选在25%内。

如有需要，可以对用于本发明的DNA构建体(即嵌合植物基因)中的启动子进行修饰，以影响其控制特征。例如，可以将CaMV35S启动子连接于拟南芥(*Arabidopsis thaliana*)核酮糖-1.5-二磷酸羧化酶小亚基基因(ssRUBISCO)，所述基因代表在缺乏光时表达ssRUBISCO的部分、以产生在叶片中有活性、但在根中无活性的启动子。所产生的嵌合启动子可以如本文所述使用。对于该描述的目的，短语“CaMV35S”启动子因此包括CaMV35S启动子的变异，例如通过与操纵基因区连接、随机或控制诱变等衍生的启动子。此外，可以改变所述启动子，使其含有多个“增强子序列”以有助于提高基因表达。



Kay 等(1987)已经报道了这样的增强子序列的实例。

由本发明 DNA 构建体产生的一种 RNA 也含有 5'非翻译前导序列。这种序列可以得自选择以表达所述基因的启动子, 并且可以对其进行特异性修饰, 以便增加所述 mRNA 的翻译。所述非翻译或 5'非翻译前导序列(NTR 或 UTR)可以得自非相关的启动子或编码序列。例如, 所述 5'非翻译区也可以得自病毒 RNA、得自合适的真核生物基因或得自合成基因序列。本发明不限于在以下实施例之一中提出的构建体, 其中所述非翻译区得自伴随所述启动子序列的 5'非翻译序列。可用于本发明中的植物基因前导序列的实例是小麦叶绿素 a/b 结合蛋白(cab)前导序列和矮牵牛热激蛋白 70 (hsp70)前导序列(Winter, 1988)。

对于在单子叶植物中的最佳表达, 在所述 DNA 表达构建体中也应该包括内含子。该内含子通常置于非翻译序列中的 mRNA 5'端附近。该内含子可以得自但不限于一组内含子, 包括玉米 hsp70 内含子(Brown 等, 美国专利第 5,424,412 号; 1995)或水稻 *Act1* 内含子(McElroy 等, 1990)。

当在质粒或其它多核苷酸构建体中包括一个以上的表达盒时, 包含一种 DNA 分子的第一表达盒通常含有一个组成型启动子、一个编码草甘膦氧化还原酶(GOX)的结构 DNA 序列和一个 3'非翻译区。包含一种 DNA 分子的第二表达盒通常含有一个组成型启动子、一个编码 N-乙酰基转移酶(能够与 AMPA 反应产生 N-乙酰基-AMPA)的结构 DNA 序列和一个 3'非翻译区。也设想了包含一种 DNA 分子的另外的表达盒。例如, 在本发明设想的植物中也可以表达下述基因以及提供增强的除草剂耐性的基因: 编码杀虫或杀真菌活性、耐干旱或耐热、抗生素类化合物、药用化合物或试剂(例如肿瘤抑制蛋白或抗体组分)、生物聚合物、其它商业上有用的化合物等的基因。已经描述了在植物细胞中有活性的许多组成型启动子。用于组成型表达或者 GOX 或者 N-乙酰基转移酶合适启动子包括但不限于花椰菜花叶病毒(CaMV) 35S 启动

子(Odell 等, 1985)、玄参花叶病毒(FMV) 35S (Sanger 等, 1990)、甘蔗杆状 DNA 病毒启动子(Bouhida 等, 1993)、鸭趾草黄斑驳病毒启动子(Medberry 和 Olszewski 1993)、来自核酮糖 1,5-二磷酸羧化酶小亚基(ssRUBISCO)的光诱导型启动子(Coruzzi 等, 1984)、水稻胞质丙糖磷酸异构酶(TPI)启动子(Xu 等, 1994)、拟南芥属的腺嘌呤磷酸核糖转移酶(APRT)启动子(Moffatt 等, 1994)、水稻肌动蛋白 1 基因启动子(Zhong 等, 1996)以及甘露碱合酶和章鱼碱合酶启动子(Ni 等, 1995)。所有这些启动子已经用来产生各种类型的植物可表达重组 DNA 构建体。已经用这些启动子和其它启动子中的许多通过来自大肠杆菌的报道基因(例如 *uidA* (β -葡糖醛酸糖苷酶)基因)的表达, 进行了组成型启动子的比较分析(Li 等, 1997; Wen 等, 1993)。

可以选择用于含一种 DNA 分子的所述第二盒的启动子, 以控制或限制需要细胞死亡的特异性表达。在一个优选实施方案中, 所述启动子将能够指导仅在或主要在对于植物存活或植物生存力重要的组织中表达, 同时限制含一种 DNA 的所述第二盒在其它非必需组织中的表达。例如, 分化为花粉发育组织或终端组织的组织, 例如花粉本身、花药的绒毡层细胞层或花药组织。另一方面, 能够调节特定细胞和组织类型中基因表达的植物启动子是众所周知的。在本发明实施方案中最优选的植物启动子是在雄性生殖组织发育期间或花粉中特异性表达的启动子, 并且所述启动子的表达水平足以产生与包含一种 DNA 分子的所述第一表达盒组成型启动子转录的有义 RNA 互补的抑制性 RNA 分子。这些类型的启动子的实例包括 TA29 烟草绒毡层特异性启动子(Mariani 等, 1990)、来自矮牵牛的 PA1 和 PA2 查耳酮二氢黄酮异构酶启动子(van Tunen 等, 1990)、来自甘蓝(*Brassica oleracea*)的 SLG 基因启动子(Heizmann 等, 1991)和来自番茄的 LAT 基因启动子(Twell 等, 1991)。

已经分离出来自水稻的花药和花粉特异性启动子。实例包括 Osg6B 启动子, 已经表明它驱动 β -葡糖醛酸糖苷酶基因在转基因水稻

的未成熟花药中表达。在小穗、叶片或根的其他组织中没有检测到活性(Yokoi 等, 1997)。已经表明来自水稻的 PS1 花粉特异性启动子在水稻花粉中特异性地表达 β -葡糖醛酸糖苷酶基因(Zou 等, 1994)。已经鉴定出在水稻花药绒毡层中特异性地表达的另外的水稻基因(Tsuchiya 等, 1994; Tsuchiya 等, 1997)。例如, 通过构建 cDNA 文库以鉴定花药特异性克隆, 可以进行主要在水稻花药发育期间表达的其它基因的分离(Qu 等)。

本领域技术人员会认识到用于分离在植物起作用的启动子的方法, 其中所述启动子来自多种基因或基因家族的成员, 这些基因或基因家族成员在例如根、茎、分生组织、叶片、花、果实、花粉或涉及产生花粉的植物细胞类型的特定植物组织中高表达(Stinson 等, 1987; Brown 和 Crouch, 1990; McCormick 等, 1989)。组织特异性启动子的其它实例包括玉米聚半乳糖醛酸外切酶基因的启动子(Dubald 等, 1993)和 Zmc13 mRNA 启动子(Hanson 等, 1989)。已经表明优先在番茄花粉中表达的启动子是 LAT52 和 LAT59 启动子(Twell 等, 1991)。在美国专利第 5,470,359 号中公开了玉米 pZtap 启动子序列的一部分(psgB6-1)。

本发明的重组 DNA 分子通常包含一种有效或操作性地连接于以下组分的启动子: 一个编码 5'非翻译区的 DNA 序列; 一个植物内含子的 DNA 序列; 一个编码叶绿体转运肽(CTP)的结构序列; 一个编码改进的除草剂耐性的基因的 DNA 编码序列和一个 3'非翻译区。

所述 5'非翻译前导序列可以得自选择以表达所述异源 DNA 序列的启动子, 如有需要, 可以对其进行特异性地修饰, 以便增加 mRNA 的翻译。5'非翻译区也可以得自病毒 RNA、得自合适的真核生物基因或得自合成基因序列。本发明不限于其中所述非翻译区得自伴随所述启动子序列的 5'非翻译序列的构建体。所述前导序列也可以得自非相关启动子或编码序列。

植物有效的重组 DNA 分子的 3'非翻译区含有一个在植物中起作用以使得腺苷酸核苷酸添加到所述 RNA 的 3'末端的聚腺苷酸化信

号。所述 3'非翻译区可以得自在植物细胞中表达的各种基因。胭脂碱合酶 3'非翻译区(Fraley 等, 1983)、豌豆 ssRUBISCO 的 3'非翻译区(Coruzzi 等, 1994)、大豆 7S 种子贮存蛋白基因的 3'非翻译区(Schuler 等, 1982)和豌豆 ssRUBISCO 基因的豌豆小亚基通常因具有这种能力而被应用。含有土壤杆菌肿瘤诱导(Ti)质粒基因的聚腺苷酸化信号的 3'转录的非翻译区也是合适的。

适用于在单子叶植物中表达的植物内含子的实例包括例如玉米 hsp70 内含子、水稻肌动蛋白 1 内含子、玉米 ADH 1 内含子、拟南芥属 SSU 内含子、拟南芥属 EPSPS 内含子、矮牵牛 EPSPS 内含子和本领域技术人员已知的其它内含子。

可能特别有利的是指导赋予除草剂耐性的蛋白定位至亚细胞区室, 例如定位至线粒体、内质网、液泡、叶绿体或其它质体区室。通过在其氨基末端包含一个叶绿体转运肽(CTP), 可以将蛋白导向叶绿体。先前已经鉴定出天然存在的、作为含有氨基末端叶绿体引导肽(将所述前体引导至叶绿体输入机器)的较大的前体蛋白合成的叶绿体定向蛋白, 这是本领域众所周知的。叶绿体引导肽通常由位于叶绿体细胞器内的特异性内切蛋白酶切割, 由此从所述前体将所述定向的成熟酶、最好是活性酶释放出到叶绿体环境中。适合于指导所述除草剂耐性基因或转酰基酶基因产物引导至植物细胞的叶绿体或质体的肽的编码序列的实例包括矮牵牛 EPSPS CTP、拟南芥属 EPSPS CTP2 和内含子以及本领域技术人员已知的其它序列。这样的引导序列可供用于将所需的已表达蛋白转移至其最有效起作用的细胞结构中, 或通过将所需的已表达蛋白转移至所需表型功能所必需的细胞过程集中的细胞区域。已经发现叶绿体引导肽在草甘膦抗性植物的选择中特别有用(Barry 等, 美国专利第 5,463,175 号; Barry 等, 美国专利第 5,633,435 号)。草甘膦通过抑制在叶绿体内发生的芳族氨基酸生物合成而起作用, 以杀伤所述细胞。因此, 将所述抗性基因产物集中于叶绿体中, 提供了增强的对所述除草剂的抗性。本文的实施例可供用于也被引导至或定位

至叶绿体中并在其中集中的转酰基酶。叶绿体引导肽的具体实例是本领域众所周知的，包括拟南芥核酮糖二磷酸羧化酶小亚基 *asr1A* 转运肽、拟南芥 EPSPS 转运肽和玉蜀黍(*Zea mays*)核酮糖二磷酸羧化酶小亚基转运肽。在此中已经发挥作用以将异源蛋白定位至叶绿体的一种 CTP 得自拟南芥核酮糖二磷酸羧化酶小亚基 *ats1A* 转运肽。编码本文使用该转运肽变体的多核苷酸序列提供了天然转运肽氨基酸序列加上重复的转运肽切割位点，在本文中已经表明可用于将活性重组转酰基酶配置到叶绿体(SEQ ID NO:9)。

将植物有效除草剂耐性或除草剂抗性基因定位至叶绿体或质体的一种替代方法包括叶绿体或质体转化。可以产生其中仅线粒体 DNA 或叶绿体 DNA 已经改变以掺入本申请中设想的分子的重组植物。在叶绿体中起作用的启动子是本领域已知的(Hanley-Bowden 等, Trends in Biochemical Sciences 12:67-70, 1987)。例如 Daniell 等(美国专利第 5,693,507 号; 1997)和 Maliga 等(美国专利第 5,451,513 号; 1995)已经描述获得含已经插入异源 DNA 的叶绿体的细胞的方法和组合物。

AMPA 在植物中积累可以引起植物毒性症状，它在表型上表现为叶片褪绿、生长停滞、不育和死亡，尽管不是所有这些症状都在每种植物种中表现出来。本文已经发现，AMPA 分子通过转酰基进行酶修饰，产生 N-酰基-AMPA，提供了克服 AMPA 植物毒性效应的方法。分析 AMPA 转化为 N-酰基-AMPA 的一种方法涉及在一定体积水性反应中提供^[14C]标记的 AMPA 作为所述转酰基酶的一种底物，提供酰基-CoA 作为该酶的另一底物，并且如本文实例中描述的通过在阴离子交换柱上进行 HPLC，从 N-酰基-^[14C]-AMPA 产物分离^[14C]标记的 AMPA 底物。令人惊讶的是，已经表明所述转酰基酶能够利用其它酰化-CoA 化合物作为底物，以将所述 AMPA 底物转酰基化。特别是，已经表明丙酰-CoA 是体外转酰基反应的反应性特别强的底物，产生 N-丙酰-^[14C]-AMPA。当采用 AMPA 作为酰基受体底物时，较大的酰化-CoA 化合物，即丁酰-CoA 或甲基丙二酸单酰-CoA 和其碳链长度大于

C₃ 的共价连接于 CoA 的其它有机分子证明在转酰基反应中效率较低。虽然有这种信息，本领域技术人员仍然会认识到，与 PhnO 包含的底物特异性相比，因氨基酸序列与本文鉴定的 PhnO 或 PhnO 样酶同源而大致相关的其它转酰基酶，将在所述 AMPA 转酰基酶反应中具有相似 5 的底物特异性。这些其它酶在概念上也 在本文所述发明的范围和 精神内。例如，由各种各样的在对植物毒性化合物(例如 AMPA)转酰基中 可以用作底物的酰基-CoA 和酰基载体蛋白化合物，介导脂肪酸生物合成。能够用脂肪酸中间体进行 AMPA 转酰基的转酰基酶可以想象地通过消除 AMPA 植物毒性而提供植物保护。例如 PhnO 的、能够转 10 酰基的酶在种类繁多的含 CP 键并另外含有一个 CN 键的毒性化合物的 解毒中可能是有用的。

用包含 *phnO* 基因序列或 *phnO* 样基因序列的一种或更多种表达载体转化细菌、酵母或真菌细胞、植物细胞或全株的方法和组合物是本说明书的其它方面。转基因细菌、酵母或真菌细胞、植物细胞或得自 15 这种转化方法的植物或得自这种转基因植物的后代和种子也是本发明的 其它实施方案。

转化细菌和酵母或真菌细胞的方法是本领域众所周知的。通常，转化方法与用来转化其它细菌(例如大肠杆菌)或酵母(例如酿酒酵母 (*Saccharomyces cerevisiae*))的熟知的方法类似。植物细胞的 DNA 转化 20 方法包括但不限于土壤杆菌介导的植物转化、原生质体转化、将基因 转移到花粉中、注射到生殖器官中、注射到未成熟胚中、质体或叶绿体转化和微粒轰击。这些方法中的每一种方法具有不同的优点和缺点。因此，将基因引入特定植物种的一种特定方法对于另一种植物种可能不是最有效的，但本领域技术人员熟知哪种方法可用于特定植物 25 种。

有许多将转化 DNA 区段引入细胞中的方法，但不是所有的方法都适合于将 DNA 传递至植物细胞。认为合适的方法事实上包括可以将 DNA 引入细胞中的任何方法，例如通过土壤杆菌感染、双元细菌人工

染色体(BIBAC)载体(Hamilton 等, 1996)、例如通过以下方法直接传递 DNA: PEG 介导的原生质体转化(Omirulleh 等, 1993)、干燥/抑制介导的 DNA 摄入、电穿孔、与碳化硅纤维一起搅拌、DNA 包被粒子的加速等。在某些实施方案中, 加速方法是优选的, 包括例如微粒轰击等。

5 将 DNA 引入细胞中的技术是本领域技术人员熟知的。已经描述了用于将基因传递至细胞中的 4 种常用方法: (1)化学方法(Graham 和 van der Eb, 1973; Zatloukal 等, 1992); (2)物理方法, 例如微注射(Capecchi, 1980)、电穿孔(Wong 和 Neuman, 1982; Fromm 等, 1985; 美国专利第 5,384,253 号)和基因枪(Johnston 和 Tang, 1994; Fynan 等, 1993; 10 Luthra 等, 1997); (3)病毒载体(Clapp, 1993; Lu 等, 1993; Eglitis 和 Anderson, 1988a; 1988b); 和(4)受体介导的机制(Curiel 等, 1991; 1992; Wagner 等, 1992)。

已经发表了主要利用根癌土壤杆菌的、用于棉花(美国专利第 5,004,863 号; 美国专利第 5,159,135 号; 美国专利第 5,518,908 号)、大豆(美国专利第 5,569,834 号; 美国专利第 5,416,011 号; McCabe 等 15 (1988); Christou 等(1988))、芸苔属(*Brassica*) (美国专利第 5,463,174 号)、花生(Cheng 等(1996); De Kathen 和 Jabobsen (1990))的转化双子叶植物并获得转基因植物的方法。

也已经报道了利用电穿孔、微粒轰击和土壤杆菌的单子叶植物转 20 化。在 天门冬(Bytebier 等(1987))、大麦(Wan 和 Lemaux (1994))、玉米(Rhodes 等(1988); Ishida 等(1996); Gordon-Kamm 等(1990); Fromm 等(1990); Koziel 等(1993); Armstrong 等(1995))、燕麦(Somers 等(1992))、鸭茅(Horn 等(1988))、水稻(Toriyama 等(1988); Park 等(1996); Abedinia 等(1997); Zhang 和 Wu (1988); Zhang 等(1988); Battraw 和 Hall (1990); 25 Christou 等(1991); Park 等(1996))、黑麦(De La Pena 等(1987))、甘蔗(Bower 和 Birch (1992))、苇状羊茅(Wang 等(1992))和小麦(Vasil 等(1992); Weeks 等(1993))中已经达到转化和植株再生。在 Davey 等(1986)中也讨论了用于单子叶植物转化和植株再生的技术。

也可以产生其中仅线粒体或 DNA 或叶绿体 DNA 已经改变以掺入本申请中设想的分子的重组植物。在叶绿体中起作用的启动子是本领域已知的(Handley-Bowden 等, Trends in Biochemical Sciences 12:67-70, 1987)。Daniell 等, 美国专利第 5,693,507 号(1997)和 Maliga 等(美国专利第 5,451,513 号; 1995)已经描述获得含已经插入异源 DNA 的叶绿体的细胞的方法和组合物。已经用异源 DNA 转化、改变核基因组和叶绿体或质体基因组的重组植物也在本发明范围内。

本发明公开了包含编码 AMPA 转酰基酶的多核苷酸序列的 DNA 构建体。本文公开了用于鉴定和分离编码在 AMPA 的 N-酰化中起作用的肽的异源基因的方法。用于构建并在植物中表达合成基因的方法是本领域技术人员熟知的, 并且详细描述于美国专利第 5,500,365 号, 用于单子叶植物的方法特别描述于美国专利第 5,689,052 号。本发明设想了在单子叶植物和双子叶植物的转化中, 单独应用 AMPA 转酰基酶基因或结合编码 GOX 介导的草甘膦降解酶的基因应用 AMPA 转酰基酶基因。为了增强这些基因的表达, 本发明提供 DNA 构建体, 所述 DNA 构建体包含编码这些类型蛋白(所述蛋白被定位至植物细胞质)的多核苷酸序列以及编码质体引导肽、位于编码所述 AMPA 转酰基酶和/或 GOX 蛋白的多核苷酸序列上游的序列。

一方面, 本发明提供的核苷酸信息可供用于制备有能力与选定的本文公开的多核苷酸的基因序列特异性杂交的相对短的 DNA 序列。在这些方面, 根据所选定的编码 AMPA 转酰基酶多肽的多核苷酸序列例如 SEQ ID NO:1、SEQ ID NO:2 和 SEQ ID NO:3 中所示序列的考虑, 制备合适长度的核酸探针。也可以根据所选定的编码质体引导肽的多核苷酸序列例如 SEQ ID NO:9、SEQ ID NO:11、SEQ ID NO:13 和 SEQ ID NO:14 中所示序列的考虑, 制备这样的核酸探针。这样的核酸探针与编码 AMPA 转酰基酶多肽或质体引导肽序列的基因序列特异性杂交的能力, 使得它们在许多实施方案中特别有用。最为重要的是, 所述探针可以用于种类繁多的检测特定样品中互补序列的存在的测

定。

在某些实施方案中，最好使用寡核苷酸引物。用本发明的多核苷酸设计这样的引物的序列，以用于利用 PCR™ 技术对来自任何合适生物体的 AMPA 转酰基酶基因的特定序列进行检测、扩增或突变。所述方法也可以用来对编码质体引导肽的多核苷酸的特定序列进行检测、扩
5 增或突变。与编码本发明的 AMPA 转酰基酶多肽和质体引导肽的多核苷酸相关的基因区段，也可以采用这样的引物，通过 PCR™ 进行扩增。

为了提供按照本发明的某些优点，用于杂交研究或测定的优选核酸序列包括：与顺式邻接编码 AMPA 转酰基酶的多核苷酸序列(例如
10 SEQ ID NO:5 或 SEQ ID NO:6 所示的序列)、或编码 AMPA 转酰基酶的多核苷酸序列的至少长度为 14-30 等的连续核苷酸大致互补的序列；或与编码质体引导肽的序列的至少长度为 14-30 等的连续核苷酸大致互补的序列。所谓“大致互补”是指在一种多核苷酸在序列中与靶多核苷酸序列优选约 70% 互补、或更优选约 80% 互补、或甚至更优选约 90% 互补、或最优选约 99-100% 互补。
15

长度至少为 14 个核苷酸的大小有助于确保该片段具有足够的长度，以形成稳定并且选择性的双链体分子。一般优选在长度为 14 个碱基以上区段范围内具有互补序列的分子。为了提高所述杂交体的稳定性和选择性，并因此提高所获得的特定杂交分子的质量和程度，人们
20 一般优选设计具有 14-20 个核苷酸或需要时可更长的基因互补序列的核酸分子。通过例如直接用化学方法，例如亚磷酸胺化学；应用核酸复制技术，例如美国专利第 4,683,195 号和第 4,683,202 号的 PCR™ 技术，所述专利具体通过引用结合到本文中；或通过从含有合成插入片段和合适的限制位点的重组质粒中切下所选择的 DNA 片段，可以容易地制备这样的片段。
25

本发明也设想了包含本发明多核苷酸的表达载体。因此，在一个实施方案中，表达载体是分离并纯化的 DNA 分子，它包含有效连接于编码本发明多肽的编码区的一个启动子，所述编码区有效连接于转录

终止区，由此所述启动子驱动所述编码区的转录。所述编码区可以包括一种编码AMPA转酰基酶的区段或序列以及编码质体引导肽的区段或序列。所述包含表达载体的DNA分子也可以含有一种植物功能性内含子，也可以含有其它植物功能性元件，例如编码非翻译序列(UTL)的序列和作为转录或翻译的增强子起作用的序列。

本文所用的术语“操作性连接的(operatively linked)”或“有效连接的(operably linked)”是指作为启动子起作用的序列与编码区连接，其连接方式使得该编码区的转录受该启动子的控制和调节。将启动子操作性连接于编码区以调节上游和下游的方法是本领域熟知的。

优选的植物转化载体包括衍生自根癌土壤杆菌Ti质粒的载体以及 Herrera-Estrella (1983)、 Bevan (1983)、 Klee (1985)和欧洲专利申请号 EP 0120516 (每个文献具体通过引用结合到本文中)公开的那些载体。另外，定向至叶绿体或质体转化的植物优选转化载体包括在美国专利第 5,693,507 号(1997)、美国专利第 5,451,513 号(1995)、 McBride 等 (1995)、 Staub 等(1995a)、 Staub 等(1995b)和 WO 95/24492 (每个文献具体通过引用结合到本文中)中公开的那些载体。

当本发明的表达载体将用于转化植物时，选择能够在该特定植物种中驱动表达的启动子。在不同植物种中有功能的启动子也是本领域熟知的。可用于在植物中表达所述多肽的启动子是如描述的诱导型启动子、病毒启动子、合成启动子或组成型启动子的那些启动子(Odell 等, 1985)和/或时序调节、空间调节和时空调节的那些启动子。优选的启动子包括增强的 CaMV35S 启动子和 FMV35S 启动子。

以双链 DNA 形式存在的基因定位至植物核基因组的表达涉及由 RNA 聚合酶从所述 DNA 的编码链转录信使 RNA (mRNA)，并且随后在细胞核内加工所述 mRNA 初级转录物。来自叶绿体或质体内表达的基因也产生在翻译之前不进一步加工的 mRNA 转录物。在任一事件中，DNA 转录出 mRNA 均受称为“启动子”的 DNA 区域的调节。包含所述启动子的 DNA 表现为一种碱基序列，该序列发送信号使 RNA

聚合酶与该 DNA 结合并启动采用所述 DNA 链之一作为模板的转录 mRNA，以制备相应的 RNA 链。所选定的特定启动子应该能够引起 AMPA 转酰基酶编码序列的足够表达，以导致产生除草剂耐性或除草剂抗性有效量的、定位至所需胞内位置的转酰基酶蛋白。

5 在植物组织中结构基因可以由各种各样的启动子驱动。启动子可以是近组成型的(near-constitutive)(即它们在所有组织中均驱动转基因的转录)，例如 CaMV35S 启动子；或影响双子叶植物或单子叶植物的组织特异性或发育特异性的启动子。当启动子为近组成型启动子，例如为 CaMV35S 或 FMV35S 时，在许多转化植物组织和大多数植物器
10 官(例如愈伤组织、叶片、种子、茎、分生组织、花和根)中发现多肽表达的增强。增强的或重复形式的 CaMV35S 和 FMV35S 启动子在本发明的实施中特别有用(Kay 等, 1987; Rogers, 美国专利 5,378,619)。

 本领域技术人员会认识到，有许多启动子在植物细胞中有活性，并且已经在文献中进行了描述。这样的启动子可以得自植物或植物病
15 毒，包括但不限于胭脂碱合酶(NOS)启动子和章鱼碱合酶(OCS)启动子(它们载于根癌土壤杆菌的肿瘤诱导质粒上)、花椰菜花叶病毒(CaMV) 19S 和 35S 启动子、来自核酮糖 1,5-二磷酸羧化酶(ssRUBISCO, 一种非常丰富的植物多肽)小亚基的光诱导型启动子、水稻 *Act1* 启动子和玄参花叶病毒(FMV) 35S 启动子。所有这些启动子已经用来创造各种类型
20 的 DNA 构建体，所述构建体已经在植物中表达(参见例如 McElroy 等, 1990, 美国专利 5,463,175)。

 另外，也可能最好采用含有组织特异性启动子的植物整合型载体，促使例如 AMPA 转酰基酶的基因在植物的特定组织中表达，所述 AMPA 转酰基酶提高除草剂耐性或除草剂抗性。具体的靶组织可以包
25 括叶片、茎、根、块茎、种子、果实等，并且所选择的启动子应该具有所需的组织和发育特异性。因此，应该通过选择具有所需组织表达能力和接近的启动子强度的启动子，并选择在所述靶组织中产生所需转酰基酶活性的转化体，使启动子功能最佳化。在异源结构基因在植

物中表达方面，这种来自转化体库的选择方法是常规使用的，因为在含有同一异源基因的转化体之间由于在所述植物基因组中基因插入位点而有变异(通常称为“位置效应”)。除已知在植物细胞中引起 DNA 转录(组成型或组织特异性的)的启动子外，可以通过在植物 cDNA 文库中筛选选择性或优选在靶组织中表达的基因，然后确定启动子区，
5 来鉴定其它启动子以用于本发明中。叶绿体或质体功能性启动子是本领域已知的(Hanley-Bowden 等, Daniell 等, Maliga 等)。

其它典型的组织特异性启动子是玉米蔗糖合成酶 1 (Yang 等, 1990)、玉米乙醇脱氢酶 1 (Vogel 等, 1989)、玉米集光复合体(Simpson, 1986)、玉米热激蛋白(Odell 等, 1985)、豌豆小亚基 RuBP 羧化酶(Poulsen 等, 1986; Cashmore 等, 1983)、Ti 质粒甘露碱合酶(McBride 和 Summerfelt, 1989)、Ti 质粒胭脂碱合酶(Langridge 等, 1989)、矮牵牛查耳酮异构酶(Van Tunen 等, 1988)、菜豆富甘氨酸蛋白 1 (Keller 等, 1989)、CaMV 35s 转录物(Odell 等, 1985)和马铃薯 patatin (Wenzler 等, 1989)启动子。优选的启动子是花椰菜花叶病毒(CaMV 35S)启动子和 S-E9 小亚基 RuBP 羧化酶启动子。
10

如有需要，可以对用于本发明 DNA 构建体中的启动子进行修饰，以影响其控制特性。例如，可以将 CaMV35S 启动子连接于 ssRUBISCO 基因代表在缺乏光时表达 ssRUBISCO 的部分，以产生在叶片有活性而在根中无活性的启动子。为了这种描述，短语“CaMV35S”启动子因此包括 CaMV35S 启动子的变异，例如由于与操纵基因区连接、随机诱变或控制诱变等得到的启动子。此外，可以将所述启动子改变，以含有多个“增强子序列”，以有助于提高基因表达。这类增强子序列的实例已经由 Kay 等(1987) 报道。叶绿体或质体特异性启动子是本领域已知的(Daniell 等, 美国专利第 5,693,507 号; 其通过引用结合到本文中)。例如可得自例如以下的叶绿体基因的启动子：得自菠菜或豌豆的 *psbA* 基因、得自玉米的 *rbcL* 和 *atpB* 启动子区和 rRNA 启动子。任何叶绿体或质体有效启动子均在本发明范围内。
20
25

由用组织特异性启动子转化的植物细胞产生的本发明转基因植物，可以与由用不同的组织特异性启动子转化的植物细胞产生的第二转基因植物杂交，以产生在一个以上的特定组织中表现出转化效应的杂种转基因植物。

5 由本发明 DNA 构建体产生的 RNA 也可以含有 5' 非翻译前导序列 (5'UTL)。该序列可以得自选择用来表达所述基因的启动子，并且可以对其进行特异性地修饰，以增加所述 mRNA 的翻译。5' 非翻译区也可以得自病毒 RNA、合适的真核生物基因或合成的基因序列。本发明不
10 限于其中非翻译区得自伴随所述启动子序列的 5' 非翻译序列的构建体。可用于本发明的一种植物基因前导序列是矮牵牛热激蛋白 70 (hsp70) 前导序列 (Winter 等, 1988)。

 5'UTL 当定位至所述 DNA 序列时能够调节转录起始位点和编码
 序列起点之间的基因表达。已经进行了前导序列的编译(compilation),
 以预测最适或亚适序列，并产生“共有序列”和优选的前导序列(Joshi,
15 1987)。设想优选的前导序列包括包含预测指导所连接的结构基因最佳
 表达的序列的那些前导序列，即包括可以增加或保持 mRNA 稳定性并
 防止翻译不正确起始的优选共有前导序列。按照本说明书，这类序列
 的选择是本领域技术人员已知的。最优选由在植物、特别是在玉米中
 高表达的基因衍生的序列。一个特别有用的前导序列可以是矮牵牛
20 HSP70 前导序列。

 按照本发明，设计用来特异性增强所述多肽在所述转化植物中表
 达的表达载体，可以包括编码质体或叶绿体引导肽的某些区域，在本
 文中以各种形式缩写为 CTP、CTP1、CTP2 等，每种代表一种不同
 或变异的引导肽序列。这些区域当与某些 GOX 或 AMPA 转酰基酶蛋
25 白序列结合时便于涉及所编码蛋白的转录、翻译和表达的细胞过程完
 全被利用。这种引导肽以多种方式起作用，例如通过将所表达的蛋白
 转移至其最为有效操作的细胞结构，或通过将所表达的蛋白转移至表
 达所必需的细胞过程集中的细胞区域。CTP 的应用也可以提高形态正

常植株的回收率以及提高可以回收转基因植物的频率。

已经发现叶绿体引导肽在草甘膦抗性选择标记系统中特别有用。在该系统中，转化以表达赋予草甘膦抗性的蛋白的植物与将所述肽引导至植物细胞的叶绿体的 CTP 一起转化。草甘膦抑制导致包括氨基酸和维生素在内的芳族化合物生物合成的莽草酸途径。具体地说，草甘膦通过抑制 5-烯醇丙酮酰-3-磷酸莽草酸合酶(EPSP 合酶和 EPSPS)，抑制磷酸烯醇丙酮酸和 3-磷酸莽草酸转化为 5-烯醇丙酮酰-3-磷酸莽草酸。引入编码 EPSPS 的转基因允许所述植物细胞抵抗草甘膦的效应，特别是当所述转基因编码草甘膦不敏感性 EPSPS 酶时。因此，由于除草剂草甘膦的作用是通过中断芳族氨基酸的生物合成、特别是中断所述细胞叶绿体中的芳族氨基酸的生物合成而杀伤所述细胞，所以 CTP 通过将细胞表达的草甘膦抗性酶集中在叶绿体中，即细胞的靶细胞器中，而提供对所述除草剂的提高的抗性。典型的除草剂抗性酶包括 EPSPS 和草甘膦氧化还原酶(GOX)基因(参见 Comai, 1985, 美国专利第 4,535,060 号，其具体通过引用全部结合到本文中)。

CTP 可以将蛋白导向叶绿体和其它质体。例如，所述靶细胞器可以是造粉体。本发明优选的 CTP 包括导向叶绿体以及其它质体的那些 CTP。优选 CTP 的具体实例包括玉米 RUBISCO SSU 蛋白 CTP 和功能相关的肽，例如拟南芥 RUBISCO 小亚基 CTP 和拟南芥 EPSPS CTP。这些 CTP 例如为分别示于 SEQ ID NO:9、SEQ ID NO:11、SEQ ID NO:13 和 SEQ ID NO:14 中的多核苷酸序列和氨基酸序列。

也可以产生其中仅线粒体或叶绿体 DNA 已经被改变以掺入本申请中设想的分子的重组植物、细胞、种子和其它植物组织。在叶绿体中起作用的启动子是本领域已知的(Hanley-Bowden 等, Trends in Biochemical Sciences 12:67-70, 1987)。获得含插入异源 DNA 的叶绿体的细胞的方法和组合物已经描述于美国专利第 5,693,507 号(1997)。McBride 等(WO 95/24492)公开了编码 Cry1A δ -内毒素蛋白的基因在烟草植株叶绿体基因组中的定位和表达。

本发明的一个典型实施方案包括赋予植物除草剂耐性或除草剂抗性的酶或蛋白的编码基因的质体或叶绿体定向或质体或叶绿体定位。质体或叶绿体引导序列已经从许多核编码的植物基因中分离出来，并且已经显示出引导胞质合成的蛋白输入质体或叶绿体(Keegstra
5 和 Olsen 的综述, 1989)。本领域熟知的各种各样的质体引导序列可以用于实施本发明，它们包括但不限于 ADPGPP、EPSP 合酶或 ssBUBISCO。另外，用于单子叶作物的质体引导序列(肽和核酸)可以包含内含子序列的基因组编码片段以及在编码的质体引导序列中的重复(duplicated)蛋白酶切位点。

10 用于双子叶作物的优选 CTP 序列在本文中称为(SEQ ID NO:9)，它包含一个含叶绿体引导肽序列的基因组编码片段，所述叶绿体引导肽序列得自拟南芥的 EPSP 合酶基因，其中豌豆 ssRUBISCO CTP 的转运肽切割位点取代了天然的 EPSP 合酶 CTP 切割位点(Klee 等, 1987)。

为了使在单子叶植物中的表达最佳化，在所述 DNA 表达构建体
15 中也可以包含内含子。这样的内含子通常置于非翻译序列中 mRNA 5' 端附近。这种内含子可以得自但不限于以下一组内含子：玉米热激蛋白(HSP) 70 内含子(美国专利 5,424,412; 1995)、水稻 *Act1* 内含子(McElroy 等, 1990)、*Adh* 内含子 1 (Callis 等, 1987)或蔗糖合酶内含子(Vasil 等, 1989)。

20 本发明基因的定位至植物核基因组的 3' 非翻译区也含有一个聚腺苷酸化信号，它在植物中起作用以引起腺苷酸核苷酸添加至所述 mRNA 的 3' 端。RNA 聚合酶转录核基因组编码 DNA 序列一直到聚腺苷酸化发生的位点。通常，位于聚腺苷酸化位点下游数百个碱基对的 DNA 序列用来终止转录。那些 DNA 序列在此称为转录终止区。转录
25 的信使 RNA (mRNA) 的有效聚腺苷酸化需要那些区。优选的 3' 区的实例是(1)土壤杆菌肿瘤诱导(Ti)质粒基因(例如胭脂碱合酶(NOS)基因)的含聚腺苷酸化信号的 3' 转录的非翻译区，和(2)植物基因例如豌豆核酮糖-1,5-二磷酸羧化酶小亚基基因(在此命名为 E9)的 3' 端(Fischhoff 等,

1987)。构建体将通常包括目的基因与用作终止转录的信号并且在构建体中计划用于核基因组表达的 3'端 DNA 序列，便于所产生的 mRNA 聚腺苷酸化。最优选的 3'元件设想为得自根癌土壤杆菌胭脂碱合酶基因的那些 3'元件(*nos* 3'端) (Bevan 等, 1983)、根癌土壤杆菌章鱼碱合酶基因 T7 转录物的终止子和马铃薯或番茄的蛋白酶抑制剂 I 或 II 基因的 3'端。如有需要，可以再包括调节元件例如 TMV Ω 元件(Gallie 等, 1989)。

按照本发明且如上所述，编码赋予植物除草剂耐性或除草剂抗性特征的酶的叶绿体或质体定位基因，不需要提供转录终止和聚腺苷酸化信号的序列，而是可能仅需要所述基因 3'端的转录终止信息。对于引入叶绿体或质体、或引入叶绿体或质体基因组中的编码序列，mRNA 转录的终止与细菌基因表达领域中熟知的方法相似。例如，通过干环结构或类似于依赖于 *rho* 的序列的结构，可以在或者多顺反子或单顺反子序列中终止转录。

转录增强子或增强子重复可以用来增加表达。这些增强子通常发现位于在真核细胞中起作用的启动子中的转录起始处的 5'，但通常可以以正向或反向插入所述编码序列的 5'或 3'。增强子的实例包括得自 CaMV 35S 启动子、章鱼碱合酶基因(Ellis 等, 1987)、水稻肌动蛋白基因和非植物真核生物(例如酵母; Ma 等, 1988)启动子的元件。

在本发明的某些实施方案中，内部核糖体结合位点(IRES)元件的应用用来产生多基因或多顺反子信息。IRES 元件能够绕过依赖于 5'甲基化帽翻译的核糖体扫描模型，并在内部位点开始翻译(Pelletier 和 Sonenberg, 1988)。已经描述了得自小 RNA 病毒科两个成员(脊髓灰质炎病毒和脑心肌炎病毒)的 IRES 元件(Pelletier 和 Sonenberg, 1988)，以及得自哺乳动物信息的 IRES 元件(Macejak 和 Sarnow, 1991)。IRES 元件可以连接于异源可读框。多个可读框可以一起转录，每个可读框由一个 IRES 分隔，产生多顺反子信息。借助所述 IRES 元件，核糖体可以达到每个可读框，进行有效的翻译。采用单启动子/增强子可以有效

表达多基因，以转录一个单一信息。

任何异源可读框均可以连接于 IRES 元件。这包括以下蛋白的基因：分泌蛋白、由独立的基因编码的多亚基蛋白、胞内蛋白或膜结合蛋白和选择标记。这样，可以用一种构建体和一种选择标记，将几种蛋白的表达同时工程化到一个细胞中。

计划利用叶绿体或质体特异性转录和翻译机器在叶绿体或质体内表达的构建体可以含有或者单顺反子或多顺反子序列。

选择哪种表达载体以及选择哪种与多肽编码区最终操作性连接的启动子直接取决于所需的功能特性，例如蛋白表达的位置和时间以及待转化的宿主细胞。在构建重组 DNA 分子的领域中有固有的众所周知的限制。然而，可用于实施本发明的载体能够指导与其操作性连接的多肽编码区的表达。

可用于在高等植物中表达基因的典型的载体是本领域众所周知的，并且包括衍生自描述的根癌土壤杆菌的肿瘤诱导(Ti)质粒(Rogers 等, 1987)的载体。然而，已知几种其它植物整合型载体系统在植物中起作用，包括描述的 pCaMVCN 转移控制载体(Fromm 等, 1985)。pCaMVCN (可得自 Pharmacia, Piscataway, NJ)包括 CaMV35S 启动子。

在优选实施方案中，用来表达所述多肽的载体包括一个在植物细胞中有效的选择标记，最好是抗药性选择标记。一个优选的抗药性标记是其表达导致卡那霉素抗性的基因，即已描述的含有胭脂碱合酶启动子、Tn5 新霉素磷酸转移酶 II (*nptII*)和胭脂碱合酶 3'非翻译区的嵌合基因(Rogers 等, 1988)。

用于制备表达载体的方法是本领域众所周知的。用来转化植物的表达(转化)载体和制备那些载体的方法描述于美国专利 4,971,908、4,940,835、4,769,061 和 4,757,011 (每个文献具体通过引用结合到本文中)。可以修饰那些载体，以使其包括按照本发明的编码序列。

已经开发了多种方法以通过互补粘性末端或平端将 DNA 操作性地连接于载体。例如，可以将互补同聚物序列段加至待插入的 DNA

区段以及加至载体 DNA。然后通过互补同聚物尾之间的氢键连接所述载体和 DNA 区段，形成重组 DNA 分子。

5 编码有能力赋予细胞除草剂抗性酶活性提高的多肽的编码区，最好是单独的或结合编码 GOX 酶或 GOX 功能等同体的基因的编码 AMPA 转酰基酶的多核苷酸或这种多核苷酸的功能等同体。按照这样的实施方案，也优选包含 SEQ ID NO:3、SEQ ID NO:7 或 SEQ ID NO:19 的 DNA 序列的编码区。

10 已经表明结合质体引导肽编码基因成功地转化植物、以足够的除草剂防护水平表达所述 AMPA 转酰基酶的特定的 AMPA 转酰基酶编码基因，是包含在所述质粒载体中的那些基因。含质体引导序列的优选质粒包括：pMON17261、pMON10151、pMON10149、pMON32570、pMON32571、pMON32572、pMON32573、pMON32926、pMON32931、pMON32932、pMON32936、pMON32938、pMON32946、pMON32947、pMON32948 和
15 pMON32950。这些质粒含有编码 SEQ ID NO:9、SEQ ID NO:11、SEQ ID NO:13 和 SEQ ID NO:14 中所示引导序列的多核苷酸序列。公开了包含连接于编码序列的植物有效启动子的表达盒，某些表达盒具有或不具有 5' 非翻译序列和/或内含子序列，其中所述编码序列含有连接于植物有效终止序列的至少一个 AMPA 转酰基酶或转乙酰基酶，特别
20 是在 SEQ ID NO:23、SEQ ID NO:25、SEQ ID NO:27、SEQ ID NO:29 和 SEQ ID NO:31 中提出的表达盒。

25 本文所述的研究已经鉴定出在植物中增强 AMPA 转酰基酶表达的方法，所述转酰基酶当加入感病植物的核基因组、质体或叶绿体基因组中时赋予也表达 GOX 或相似基因的植物抵抗草甘膦和相关除草剂的保护。美国专利 5,500,365 (具体通过引用结合到本文中)描述了用于合成植物基因以优化所合成基因编码的蛋白的表达水平的方法。该方法涉及对外源转基因结构基因序列的修饰，以使其更“象植物”，并因此更有可能被所述植物翻译和表达。用于最好在单子叶植物中增

加转基因表达的相似方法公开于美国专利 5,689,052 (其具体通过引用结合到本文中)。农学、园艺、观赏和其它经济或商业上有用的植物可以按照本文所述的方法制备。

5 这类植物可以共表达所述 AMPA 转酰基酶基因和/或 GOX 基因以及以下其它物质: 抗真菌、抗细菌或抗病毒的致病相关肽、多肽或蛋白; 杀虫蛋白; 赋予除草剂抗性的其它蛋白; 和涉及提高植物制品质量或植物农艺性能的蛋白。在植物中同时共表达异源多种蛋白是有利的, 因为它可以利用一种以上的作用方式来防治植物病害, 或改进所述植物或由所述植物代谢产生的产物的品质。

10 设想了可能希望引入包含一种以上基因的大 DNA 序列。利用细菌或酵母人工染色体(分别为 BAC 或 YAC)或甚至植物人工染色体, 可以有助于这种序列的引入。例如, Hamilton 等(1996)公开了应用 BAC 进行土壤杆菌介导的转化。

15 最后, 用于引入单子叶植物基因组中的最理想的 DNA 序列可以是编码所需性状(例如增产)的同源基因或基因家族, 并且将其引入置于新启动子或增强子等、或许甚至同源或组织特异性(例如根颈/根鞘、轮生体、茎、穗梗(earshank)、籽粒或叶片特异性)启动子或控制元件的控制之下。事实上, 设想了本发明的具体应用可以是产生包含被以组织特异性方式导向的转基因的转化体。例如, 除草剂抗性或除草剂耐性基因可以在所述植物的生殖组织中特异性地表达, 或可以以负方式
20 在所述组织中特异性地受调节, 这可以提供用于增强那些组织的除草剂耐性或敏感性的方法。这样的调节控制方法可以提供用于调节转基因逃逸到环境中或用于防止有产权的或许可使用的知识产权或工业化财产(commercialized property)的违法使用的方法。

25 用于在转基因植物中组织特异性地导向基因表达的载体通常将包括组织特异性启动子, 也可以包括其它组织特异性控制元件, 例如增强子序列。按照本说明书, 在某些植物组织中指导特异性的或增强的表达的启动子是本领域技术人员已知的。

也设想了通过将组成型表达的基因(在所有组织中)和仅在不需要所述基因产物的那些组织中表达的反义基因共同引入,可以功能性地实现组织特异性表达。例如,可以引入编码大肠杆菌 AMPA 转酰基酶的基因,使得采用花椰菜花叶病毒的 35S 启动子将其在所有组织中表达。或者,水稻肌动蛋白启动子或双子叶植物或单子叶植物种的组蛋白启动子也可以用于组成型基因表达。此外,设想了组合得自一种以上启动子的元件的启动子可能是有用的。例如,美国专利 5,491,288 公开了将花椰菜花叶病毒启动子与组蛋白启动子组合。因此,采用例如玉米醇溶蛋白启动子在玉米籽粒中表达 AMPA 转酰基酶基因的反义转录物,可能防止所述转酰基酶在种子中累积。因此,在既表达 GOX 又表达转酰基酶的植物中,应用草甘膦除草剂可能导致种子组织不能成熟。相反,对所述 GOX 基因的反义抑制可能实现相同的结果。最好是,在特定组织中抑制所述转酰基酶可能是更为有利的,特别是当已经证明特定组织不耐受 AMPA 或相关化合物时。本发明人特别设想了,一种相似的策略可以与本发明一起使用,以指导可筛选标记或选择标记在种子组织中的表达。

另一方面,人们可能希望获得按照本发明使用的新组织特异性启动子序列。为此,人们可以首先从所关注的组织中分离 cDNA 克隆,并例如采用 RNA 印迹法鉴定在该组织中特异性表达的那些克隆。理想的是,人们有可能鉴定出不以高拷贝数存在、但所述基因产物在特定组织中相对丰富的基因。可以采用本领域技术人员已知的分子生物学技术,将相应的基因组克隆的启动子和控制元件定位。

设想了仅需要在特定条件下在转基因植物中表达某些基因。例如,建议赋予对环境胁迫因素例如干旱的抗性的某些基因仅需要在真正的胁迫条件下表达。进一步设想了这类基因在植物的整个发育期间的表达可能具有有害效应。已知存在着大量对环境应答的基因。例如,由光通过植物光敏素介导调节某些基因,例如编码核酮糖二磷酸羧化酶小亚基的 *rbcS* 的表达。其它基因由次级刺激物诱导。例如,脱落酸

(ABA)的合成由某些环境因素诱导, 包括但不限于水胁迫。已经表明许多基因由 ABA 诱导(Skriver 和 Mundy, 1990)。也预期可能仅需要在实际存在除草剂的条件下表达赋予对应用除草剂的抗性的基因。因此, 对于某些所需性状, 将需要在转基因植物中诱导型地表达基因。

5 在本发明的某些实施方案中, 提出仅需要在植物发育的某个时期在转基因植物中表达基因。发育的定时通常与组织特异性基因表达相关。例如, 在传粉后约 15 天在胚乳中启动玉米醇溶蛋白贮存蛋白的表达。

也设想了在细胞内使 DNA 插入特异性地定向可能是有用的。例如, 10 将所引入的 DNA 导向细胞核、特别是导向到一种植物染色体内的精确位置中, 以便达到定点整合可能是有用的。例如, 使用以取代该细胞中现有基因、或互补根本无功能或根本不存在的基因的基因通过转化引入可能是有用的。

也设想了用本发明表达载体转化的植物。也设想了得自这样的转化细胞或转基因细胞的转基因植物。本领域技术人员会认识到, 可以 15 采用本领域熟知的方法, 将含本发明结构编码序列的嵌合植物基因插入植物基因组中。用于植物细胞的 DNA 转化的这样的方法包括土壤杆菌介导的植物转化、应用脂质体、采用病毒或花粉的转化、电穿孔、原生质体转化、将基因转移到花粉中、注射到生殖器官中、注射到未 20 成熟胚和微粒轰击。这些方法中的每一种方法具有不同的优点和缺点。因此, 将基因引入特定植物品系的一种特定方法对于另一种植物品系可能不是最有效的, 但众所周知哪种方法可用于特定植物品系。

有许多将转化 DNA 区段引入细胞中的方法, 但不是所有的方法都适合于将 DNA 传递至植物细胞。认为合适的方法事实上包括可以将 25 DNA 引入细胞中的任何方法, 例如通过根癌土壤杆菌和相关的土壤杆菌菌株感染、例如通过以下方法直接传递 DNA: PEG 介导的原生质体转化(Omirulleh 等, 1993)、干燥/抑制介导的 DNA 摄入、电穿孔、与碳化硅纤维一起搅拌、DNA 包被粒子的加速等。在某些实施方案中,

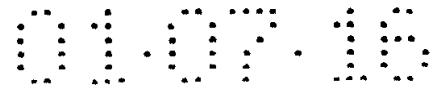
加速方法是优选的，包括例如微粒轰击等。

将 DNA 引入细胞中的技术是本领域技术人员熟知的。已经描述了用于将基因传递至细胞中的 4 种常用方法：(1)化学方法(Graham 和 van der Eb, 1973); (2)物理方法，例如微注射(Capecchi, 1980)、电穿孔
5 (Wong 和 Neumann, 1982; Fromm 等, 1985)和基因枪(Johnston 和 Tang, 1994; Fynan 等, 1993); (3)病毒载体(Clapp, 1993; Lu 等, 1993; Eglitis 和 Anderson, 1988a; 1988b); 和(4)受体介导的机制(Curiel 等, 1991; 1992; Wagner 等, 1992)。

将短高压电脉冲施加到多种动物和植物细胞上，导致在质膜中形成
10 纳米大小的孔。DNA 或者通过这些孔，或者由于伴随孔封闭的膜组分重分配，而被直接吸收到细胞质中。电穿孔可能非常有效，并且既可以用于瞬时表达所克隆的基因，又可以用来建立携带整合拷贝目的基因的细胞系。与磷酸钙介导的转染和原生质体融合相反，电穿孔时常产生携带一个整合拷贝、或最多几个整合拷贝的外源 DNA 的细胞
15 系。

通过电穿孔的方法引入 DNA 是本领域技术人员众所周知的。为了通过电穿孔实现转化，人们可以利用或者易碎的组织例如细胞的悬浮培养物、或胚胎发生愈伤组织，或者人们可以直接转化未成熟胚或其它构成组织。人们可以通过将选定细胞暴露于果胶降解酶(果胶溶
20 酶)，或者以控制的方式机械制造伤口使所述细胞对转化更加敏感，部分降解其细胞壁。则这种细胞可以通过电穿孔进行 DNA 转移的受体，可以在该阶段进行电穿孔，然后根据新掺入的 DNA 的性质，通过合适的选择或筛选方案鉴定转化细胞。

将转化 DNA 区段传递到植物细胞的再一种有利方法是微粒轰
25 击。在该方法中，可以用核酸包被粒子，然后通过推进力将其传递至细胞中。典型的粒子包括由钨、金、铂等构成的粒子。采用这些粒子，所述小金属粒子表面上的 DNA 穿过细胞壁被携带到胞质中(Klein 等, 1987; Klein 等, 1988; Kawata 等, 1988)。金属粒子穿透数层细胞，因此



离、组织距离和氮气压之类的物理参数。人们也可以通过修改影响受体细胞生理状态和可能因此影响转化和整合效率的条件，使创伤减少因素(TRF)最小化。例如，可以调节渗透状态、组织水合作用和受体细胞的传代培养阶段或细胞周期，以使转化最佳。根据本说明书，其它
5 常规调节的实施是本领域技术人员已知的。

粒子介导的转化方法是本领域技术人员众所周知的。美国专利 5,015,580 (其具体通过引用结合到本文中)描述了采用这种技术转化大豆。

土壤杆菌介导的转移是将基因引入植物细胞中的广泛适用的系
10 统，因为可以将所述 DNA 引入全部植物组织中，由此绕过由原生质体再生完整植株的需要。应用土壤杆菌介导的植物整合型载体将 DNA 引入植物细胞是本领域熟知的。参见例如已描述的方法(Fraley 等, 1985; Rogers 等, 1987)。在美国专利 5,004,863 (其具体通过引用结合到本文中)中描述了用土壤杆菌介导的转移对棉花植株进行基因工程；在美国
15 专利 5,349,124 (其具体通过引用结合到本文中)中描述了莴苣植株的相似转化；以及在美国专利 5,416,011 (其具体通过引用结合到本文中)中描述了大豆的土壤杆菌介导的转化。此外，Ti-DNA 的整合是一种相对精确的方法，几乎不导致重排。转移的 DNA 区由边界(border)序列限定，通常将间插 DNA 插入植物基因组中，如已描述的(Spielmann 等,
20 1986; Jorgensen 等, 1987)。

现代土壤杆菌转化载体能够在大肠杆菌以及土壤杆菌中复制，提供方便的操作，如已描述的(Klee 等, 1985)。此外，近来在用于土壤杆菌介导的基因转移的载体方面的技术进展已经改进了所述载体中基因和限制位点的布置，以便于构建能够表达各种多肽编码基因的载体。
25 已描述的载体(Rogers 等, 1987)具有方便的多接头区，多接头区侧翼为用于指导所插入多肽编码基因表达的启动子和聚腺苷酸化位点，这些载体适用于本发明目的。另外，含有致瘤性(armed)和非致瘤性(disarmed) Ti 基因的土壤杆菌可以用于转化。在土壤杆菌介导的转化有效的那些

植物品种中，所述基因转移由于其易得和限定的特性而成为选择的方法。

5 土壤杆菌介导的叶圆盘和其它组织(例如子叶和下胚轴)的转化看来限于土壤杆菌天然感染的植物。土壤杆菌介导的转化在双子叶植物中是最有效的。看来几乎没有单子叶植物是土壤杆菌的天然宿主，尽管如已描述的用土壤杆菌载体已经在天门冬中产生了转基因植物 (Bytebier 等, 1987)。最近也已经用土壤杆菌转化了其它单子叶植物。在该组中包括玉米(Ishida 等)和水稻(Cheng 等)。

10 用土壤杆菌转化法产生的转基因植物通常在一条染色体上仅含有单个基因。这种转基因植物对于所加入的基因可以称为杂合的。然而，因为使用词语“杂合的”通常暗示在一对染色体的第二条染色体的同一基因座上存在互补基因，而在含一个所加入基因的植物中没有这种基因，所以人们认为这种植物更准确的名称是独立分离子，因为所加入的外源基因在有丝分裂和减数分裂期间独立地分离。

15 当所述植物作为杂种商品化时，例如玉米，可能优选独立分离子。在这种情况下，使含有该基因的独立分离子与另一植株杂交，以形成目的基因杂合的杂种植物。

20 另一种优选是所加入的结构基因纯合的转基因植物，即含有两个所加入基因的转基因植物，一对染色体的每条染色体上的同一基因座都有一个基因。通过使含有单个所加入基因的独立分离子转基因植株进行有性杂交(自交)，使所产生的某些种子萌发，并分析所产生的植株的目的基因活性和孟德尔遗传性，表明相对于对照(天然的非转基因植物)或独立分离子转基因植物为纯合性，可以获得纯合的转基因植物。

25 可以使两种不同的转基因植物交配，以产生含有两个独立分离的所加入的外源基因的后代。使合适的子代自交可以产生两个所加入的编码目的多肽的外源基因均纯合的植物。也设想了与亲本植株回交和与非转基因植物异型杂交(out-crossing)。

采用基于磷酸钙沉淀法、聚乙二醇处理、电穿孔和这些处理的组合的方法，可以完成植物原生质体转化(参见例如 Potrykus 等, 1985; Lorz 等, 1985; Fromm 等, 1985; Uchimiya 等, 1986; Callis 等, 1987; Marcotte 等, 1988)。

5 这些系统对不同植物种质的应用取决于由原生质体再生该特定植物品种的能力。描述了用于由原生质体再生谷类作物的说明性方法(参见例如 Fujimura 等, 1985; Toriyama 等, 1986; Yamada 等, 1986; Abdullah 等, 1986)。

10 为了转化不能由原生质体成功再生的植物种质，可以利用将 DNA 引入完整细胞或组织的其它方法。例如，如已描述的，可以实现由未成熟胚或外植体再生谷类作物(Vasil, 1988)。

15 未修饰的细菌基因通常在转基因植物细胞中表达差。植物密码子选择比单细胞生物例如细菌更加非常类似人类和的其它高等生物的密码子选择。几篇报道已经公开了改进重组基因在植物中表达的方法(Murray 等, 1989; Diehn 等, 1996; Iannacone 等, 1997; Rouwendal 等, 1997; Futterer 等, 1997; 和 Futterer 和 Hohn, 1996)。这些报道公开了各种方法，用于根据植物密码子频率表将编码序列工程化，以代表更有效翻译的序列；采用避免可疑的聚腺苷酸化或富 A/T 结构域或内含子剪接共有序列的重组序列，改进密码子第三碱基位置偏倚。

20 美国专利 5,500,365 (其具体通过引用结合到本文中)描述了合成植物基因以优化所合成基因编码的蛋白的表达水平的优选方法。该方法涉及修饰外源转基因的结构基因序列，以使其更“象植物”，并因此更有可能被所述单子叶植物或双子叶植物翻译和表达。然而，美国专利 5,689,052 中公开的方法提供了最好是在单子叶植物中的转基因表达的增强，其通过引用结合到本文中。简而言之，按照 Brown 等的方法，降低编码所需蛋白的多核苷酸序列中罕用和半罕用单子叶植物密码子的频率，并用更优选的单子叶植物密码子取代。在单子叶植物
25 中由修饰的多核苷酸序列编码的所需多肽累积的增加是提高优选密码

子频率的结果, 所述优选密码子频率的提高是通过以连续 6 个核苷酸的片段分析所述编码序列, 并基于就单子叶植物中最稀少的 284、484 和 664 六聚体的出现频率而论的所述六聚体出现的频率, 改变所述序列来实现。此外, Brown 等公开了通过应用降低罕用密码子频率的方法以及以下方法增加重组基因表达: 在所述核苷酸序列中减少聚腺苷酸化信号和内含子剪接位点的出现, 除去所述核苷酸序列中自身互补的序列, 并用非自身互补的核苷酸取代这种序列, 同时维持编码所述多肽的结构基因, 并降低所述核苷酸序列中 5'-CG-3'二核苷酸对出现的频率。按顺序进行这些步骤, 并且有累积效应, 产生对于所需多肽中存在的大多数氨基酸而言含有单子叶植物优先利用的更优选的单子叶植物密码子的核苷酸序列。

因此, 可以采用诸如本文公开的转化方法, 通过转化植物在那些植物中增加编码目的多肽的基因量。特别是, 叶绿体或质体转化可以在含这些亚细胞器结构的组织中产生以至多约 10,000 拷贝/细胞存在的所需编码序列(McBride 等, *Bio/Technology* 13:362-365, 1995)。

也可以如已描述的(Zhou 等, 1983; Hess, 1987)通过将 DNA 直接转移到花粉中, 将 DNA 引入植物中。如已描述的方法(Pena 等, 1987), 通过将所述 DNA 注射到植物生殖器官中, 可以获得多肽编码基因的表达。也可以将 DNA 直接注射到未成熟胚的细胞中, 并如描述的方法(Neuhaus 等, 1987; Benbrook 等, 1986)通过干燥的胚的再水合, 将 DNA 引入细胞中。

在实现将外源 DNA 传递至受体细胞后, 获得转基因植物的下一步一般涉及鉴定转化细胞, 用于进一步培养和植株再生。如本文所述, 为了提高鉴定转化体的能力, 人们可能需要使用选择标记基因或可筛选标记基因作为可表达的目的基因, 或除可表达的目的基因外, 可以使用所述标记或可筛选标记基因。在这种情况下, 于是人们通常应该通过将所述细胞暴露于一种或更多种选择剂来分析潜在转化的细胞群体, 或人们应该筛选具有所需标记基因性状的细胞。

转化细胞鉴定方法的一个典型实施方案涉及将转化的培养物暴露于选择剂，例如代谢抑制剂、抗生素、除草剂等。已经被转化并稳定整合了赋予对所用的选择剂抗性的标记基因的细胞，将在培养物中生长并分裂。敏感细胞经不起进一步培养。优选标记基因的一个实例是赋予草甘膦抗性。当将该基因用作选择标记时，推定的转化细胞培养物用草甘膦处理。处理后，将可得到转基因细胞用于进一步培养，而不能得到敏感的或非转化的细胞。该方法详细地描述于美国专利 5,569,834，其具体通过引用结合到本文中。优选选择标记系统的另一实例是新霉素磷酸转移酶(*nptII*)抗性系统，该系统赋予对抗生素卡那霉素的抗性，如美国专利 5,569,834 (其具体通过引用结合到本文中)中所述。另外，在用该系统转化后，当用卡那霉素处理时，将可得到转化细胞用于进一步培养，而不能得到非转化细胞。再一优选选择标记系统涉及应用赋予对巴龙霉素抗性的基因构建体。该类型选择标记系统的应用描述于美国专利 5,424,412 (其具体通过引用结合到本文中)。

另一优选的选择标记系统涉及应用本发明设想的基因。特别是，可以将 *phnO* 基因或编码 AMPA 转酰基酶的大致相似的基因用作选择标记。通过在 AMPA 存在下培养所述细胞，可以从没有掺入重组分子的细胞群体中选择已经在其基因组中具有所引入的重组 DNA 分子的植物细胞。本领域技术人员会认识到这种选择标记系统具有的优于先前的选择标记系统的具体优点。用于整合到植物基因组中的重组 DNA 中的选择标记减少整合的靶 DNA 量，因为所述选择标记也将用于由转化植物细胞产生的植株中的改善的除草剂耐性或改善的除草剂抗性。这种选择标记也提供以前未知的另一种标记系统，特别是在常常仅有数目有限的可利用的选择标记的大田中提供这种标记系统。

壮观霉素是一种质体或叶绿体蛋白合成的抑制剂，但不是由核基因组编码的胞质核糖体的蛋白合成的抑制剂，通过利用叶绿体或质体对壮观霉素的敏感性，简化 transplastonomic 选择(质体或叶绿体转化事件的选择)。壮观霉素防止光合作用所需的叶绿体蛋白的积累，因此可

以基于在其颜色方面的差异辨别抗壮观霉素转化植物细胞：抗性转化细胞是绿色的，而敏感细胞由于质体蛋白合成受抑制则是白色的。用合适的细菌 *aad* 基因转化叶绿体或质体，或用编码壮观霉素抗性质体或叶绿体功能核糖体 RNA 的基因转化叶绿体或质体，提供了选择和维持 transplastomic 事件的方法(Maliga, Trends in Biotechnology 11:101-106, 1993)。

还设想了可筛选和可选择标记的组合将可用于鉴定转化细胞。在某些细胞或组织类型中，例如草甘膦或卡那霉素的选择剂可能或者未提供足够的杀伤活性以清楚地识别转化细胞，或者可能引起对转化体以及同样对非转化体的显著的非选择性抑制，因此使得所述选择技术不很有效。提出用一种生长抑制化合物例如草甘膦在低于引起 100% 抑制的那些浓度的浓度下进行选择，然后筛选例如卡那霉素的可选标记基因表达的生长中的组织，将使得人们可以从经不起单独选择的细胞或组织类型中回收转化体。我们提出选择和筛选的组合可能使得人们能够在更多细胞类型和组织类型中鉴定转化体。本发明的转酰基酶的可得性可以消除通过提供额外的选择方法进行组合选择和筛选的必要性。

由或者单一植物原生质体或者各种外植体发育或再生植株是本领域熟知的(Weissbach 和 Weissbach, 1988)。这种再生和生长过程通常包括以下步骤：选择转化细胞、将那些个体化细胞培养至通常的胚胎发育阶段至生根的小植株阶段。转基因胚和种子的再生是相似的。此后将所产生的转基因生根苗种植在合适的植物生长介质例如土壤中。

得自叶片外植体、含有由土壤杆菌引入的编码目的多肽的外来、外源基因的植株的发育或再生，可以采用本领域熟知的方法完成，如已描述的方法(Horsch 等, 1985)。在该方法中，在选择剂存在下，并于诱导被转化的植物品系苗再生的培养基中培养转化体，如已描述的方法(Fraley 等, 1983)。特别是，美国专利 5,349,124 (其具体通过引用结合到本文中)详细描述了创造表达杂种晶体蛋白的基因转化莴苣细胞

和由其产生的植株，其中所述杂种晶体蛋白赋予这种植株对鳞翅目幼虫的杀虫活性。

5 该方法通常在 2-4 个月内产生苗，然后将这些苗转移至合适的含所述选择剂和防止细菌生长的抗生素的促进生根培养基。然后将在所述选择剂存在下生根形成小植株的苗移植至土壤或其它介质中，以允许产生根。这些步骤根据所用的具体植物品系而不同，这种变化是本领域众所周知的。

10 最好是，让所述再生植株自花授粉以提供纯合的转基因植物，或使得自所述再生植株的花粉与农学上重要的种子生长的植株(最好是近交系)杂交。相反，用得自这些重要系植株的花粉给再生植株传粉。采用本领域技术人员熟知的方法，培育本发明的含有所需多肽的转基因植物。

15 在一个实施方案中，本发明的转基因植物因此具有增加量的编码 AMPA 转酰基酶多肽的编码区，所述多肽也可以与质体引导肽一起表达。优选的转基因植物是独立分离子，并且可以将该基因及其活性传递给其后代。更优选的转基因植物是该基因纯合的，并且当有性交配时将该基因传递至其所有后代。可以将得自转基因植物的种子种植在大田或温室中，使所得的性成熟转基因植物自花传粉，以产生纯育植株。这些植株的后代成为纯育系，评价其所述转酰基酶转基因表达以及改善的除草剂耐性，特别是当所述转酰基酶转基因与编码 GOX 酶的基因一起共表达时。

20

按照本发明的基因和转酰基酶不仅包括本文公开的全长序列，而且也包括保留本文所具体例举序列的特征性改善的除草剂防护活性的这些序列的片段或融合蛋白。

25 本领域技术人员应该明显看出，可以通过几种方法鉴定并获得 AMPA 转酰基酶基因和肽。特定的基因或其部分可以得自培养物保藏机构，或例如利用基因机合成构建。可以采用制备点突变的标准技术容易地构建这些基因的变异物。此外，采用市售的外切核酸酶或内切

核酸酶，按照标准方法，可以制备这些基因的片段。例如，诸如 *Bal31* 的酶或定点诱变可以用来从这些基因的末端有规则地切下核苷酸。此外，采用多种其它限制酶可以获得编码活性片段的基因。蛋白酶可以用来直接获得这些转酰基酶的活性片段。

5 采用本文提供的内容，也可以从大肠杆菌菌株和/或 DNA 文库中分离出等同的 AMPA 转酰基酶和/或编码这些转酰基酶的基因。例如，本文公开和要求保护的抗转酰基酶的抗体可以用来从蛋白混合物中鉴定并分离出其它转酰基酶。具体地说，可以针对本文公开的转酰基酶产生抗体，并且用这些抗体通过免疫沉淀、柱免疫纯化、酶联免疫测定(ELISA)或蛋白质印迹法，特异性地鉴定出等同的 AMPA 转酰基酶。

10 用于鉴定本发明的肽和基因的另一方法是通过应用寡核苷酸探针。这些探针是具有可检测标记的核苷酸序列。正如本领域众所周知的，如果所述探针分子和靶核酸样品中的序列通过在这两种分子之间形成强结合而杂交，则可以合理地假设所述探针和靶样品含有基本上相同的多核苷酸序列。所述探针的可检测标记提供了以已知方式确定杂交是否已发生的手段。这种探针分析提供了本发明 AMPA 转酰基酶基因的快速鉴定方法。

15 采用标准方法，利用 DNA 合成仪可以合成用作按照本发明探针的核苷酸区段。在所述核苷酸区段作为探针的应用中，用本领域技术人员已知的任何合适的标记来标记所述特定的探针，所述标记包括放射性标记和非放射性标记。典型的放射性标记包括 ^{32}P 、 ^{125}I 、 ^{35}S 等。采用 DNA 酶和 DNA 聚合酶，通过常规切口平移反应，可以从与所述 DNA 样品互补的核苷酸序列构建用放射性同位素标记的探针。然后将探针和样品在杂交缓冲液中混合，并于合适的温度下保持直至发生退火。此后，洗涤下膜上无关的物质，留下样品和结合的探针分子，通常通过放射自显影和/或液体闪烁计数检测和定量。

25 非放射性标记包括：例如配体，如生物素或甲状腺素；以及酶，例如水解酶或过氧化物酶；或各种化学发光剂，例如萤光素，或荧光

化合物如荧光蛋白、罗丹明、德克萨斯红及其衍生物等。所述探针也可以在两端用不同类型的标记物来标记以易于分离，例如在上述末端采用同位素标记来标记，而在另一端采用生物素标记，或可以在两端用具有重叠吸收谱和发射谱的不同荧光发射体来标记。

5 双链体的形成和稳定性取决于杂交体两条链之间的大致互补，如上所述，可以容许一定程度的错配。因此，本发明的探针包括所述序列的突变(单突变和多突变)、缺失、插入以及它们的组合，其中所述突变、插入和缺失允许与所述目的靶多核苷酸形成稳定的杂交体。以许多方式，采用目前本领域一般技术人员已知的方法，也许通过将来
10 可能已知的其它方法，可以在给定的多核苷酸序列中产生突变、插入和缺失。

 所列的所述探针的可能变异是部分由于遗传密码的丰余性引起的。由于遗传密码的丰余性，对于用来制备蛋白的大多数氨基酸而言，可以使用一种以上的编码核苷酸三联体(密码子)。因此，不同的核苷
15 酸序列可能编码一种特定的氨基酸。因此，大肠杆菌 AMPA 转酰基酶和肽和质体引导肽的氨基酸序列以及编码它们的多核苷酸，可以由编码所述蛋白或肽的同一氨基酸序列的等同核苷酸序列制备。因此，本发明包括这种等同的核苷酸序列。此外，反向或互补序列是本发明的一个方面，可以由本领域技术人员容易地使用。另外，已经表明，已
20 鉴定结构和功能的蛋白可以通过改变氨基酸序列来构建，只要这种改变不改变蛋白的二级结构(Kaiser 和 Kezdy, 1984)。因此，本发明包括本文描述的氨基酸序列的突变体，其中所述突变体不改变所述蛋白的二级结构，或如果结构被改变，则基本上保留所述生物活性。此外，如本发明中所讨论的，本发明也包括带有编码 AMPA 转酰基酶的基因
25 和/或编码质粒引导肽的基因的整个基因或其部分的生物的突变体。这样的突变体可以采用本领域技术人员熟知的技术制备。例如，可以用 UV 辐射制备宿主生物的突变体。同样，这样的突变体可以包括也可以采用本领域熟知的方法制备的无孢子宿主细胞。

5 定点诱变是可用于通过对编码这种肽的结构基因进行特异性诱
 变来制备单个的、新的特有的有用肽或生物功能等同蛋白或肽的技
 术。所述技术还提供制备并测试序列变异体的现成可用的能力，其中
 所述变异体例如通过以下方法来改变基因的编码序列而制备：将一个
 或多个核苷酸序列变化引入所述 DNA 中用以产生新的或有用的限制
 10 性内切核酸酶切割识别序列，或以用以改变所述编码序列，使得基因
 的密码子和 G/C 百分比代表特定属或物种更常用的密码子和 G/C 百分
 比。定点诱变允许通过利用特异性诱变寡核苷酸序列产生缺失、插入
 或置换突变，所述特异性诱变寡核苷酸序列包含所需突变的 DNA 序
 15 列。诱变寡核苷酸通常提供具有足够大小的引物序列以及在所需突变
 靶位点的两侧形成稳定双链体的序列复杂性。通常最好是长约 17-25
 个核苷酸的引物，并且在所需突变靶位点两侧具有重叠的约 5-10 个残
 基。

15 一般而言，定点诱变技术是本领域熟知的，如各种出版物例举
 的。正如人们将认识到的，该技术通常使用以单链和双链形式存在的
 噬菌体载体。可用于定点诱变的典型载体包括载体例如 M13 噬菌体。
 这些噬菌体容易从市场上获得，并且其用法一般是本领域技术人员众
 所周知的。双链质粒也常规用于定点诱变，通常含有丝状噬菌体的复
 20 制起点，所述复制起点在辅助噬菌体的存在下，允许由所述质粒载体
 合成单链 DNA。

25 一般而言，通过首先获得单链载体或将在序列中包括突变靶位点
 的双链载体的两条链解链，进行按照本文的定点诱变。一般合成制备
 带有所需突变序列的诱变寡核苷酸引物。然后让该诱变引物在所述突
 变靶位点与单链载体一起退火，并经过 DNA 聚合酶例如大肠杆菌聚合
 酶 I Klenow 片段处理，以便完成含突变链的合成。因此，形成其中一
 30 条链编码原始非突变序列而第二条链带有所需突变的杂双链体。然后
 用这种杂双链体载体转化合适的细胞例如大肠杆菌细胞，并选择包括
 含所述诱变引物序列代表的突变的重组载体的克隆。

作为产生潜在有用的种类的方法，提供采用定点诱变制备编码选定肽的 DNA 区段的序列变体，所述制备不是限制性的，因为有可以获得肽序列变体及编码其的 DNA 序列的其它方法。例如，编码所需肽序列的重组载体可以用诱变剂例如羟胺处理，以获得序列变体。这样的方法可以顺利地改变所述蛋白的生物化学和生物物理特性或其作用方式。这些包括但不限于：1)增加 AMPA 转酰基酶的形成，2)提高蛋白稳定性或减少蛋白酶降解，3)增加对底物的识别和结合，4)改善酶动力学，和 5)由于上述任何原因或所有原因而增加 N-乙酰基-AMPA 的形成。

可以在本发明肽的结构和编码它们的 DNA 区段中进行修饰和变化，并且仍获得编码具有所需特性的蛋白或肽的功能分子。本文设想的生物功能等同的肽、多肽和蛋白应该与本文公开的 AMPA 转酰基酶氨基酸序列的序列或其中相应的部分具有至少约 40%至约 65%序列相似性，优选具有约 66%至约 75%的序列相似性，更优选具有约 76%至约 85%相似性，最优选具有约 86%至约 90%或更高的序列相似性。

以下的讨论基于改变蛋白的氨基酸，以产生等同的或甚至改进的第二代分子。在本发明的具体实施方案中，设想突变的 AMPA 转酰基酶蛋白可用于改善或增强所述蛋白在植物中的表达，并因此在植物细胞中提高或改善所述 AMPA 转酰基酶活性和/或所述重组转基因的表达。根据表 1 中给出的密码子，通过在双子叶植物、更优选在单子叶植物中改变所述 DNA 序列的密码子，可以完成所述氨基酸改变。

表 1

氨基酸			密码子					
丙氨酸	Ala	A	GCA	GCC	GCG	GCU		
半胱氨酸	Cys	C	UGC	UGU				
天冬氨酸	Asp	D	GAC	GAU				
谷氨酸	Glu	E	GAA	GAG				
苯丙氨酸	Phe	F	UUC	UUU				
甘氨酸	Gly	G	GGA	GGC	GGG	GGU		
组氨酸	His	H	CAC	CAU				
异亮氨酸	Ile	I	AUA	AUC	AUU			
赖氨酸	Lys	K	AAA	AAG				
亮氨酸	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU
甲硫氨酸	Met	M	AUG					
天冬酰胺	Asn	N	AAC	AAU				
脯氨酸	Pro	P	CCA	CCC	CCG	CCU		
谷氨酰胺	Gln	Q	CAA	CAG				
精氨酸	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU
丝氨酸	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU
苏氨酸	Thr	T	ACA	ACC	ACG	ACU		
缬氨酸	Val	V	GUA	GUC	GUG	GUU		
色氨酸	Trp	W	UGG					
酪氨酸	Tyr	Y	UAC	UAU				

5 例如，在蛋白结构中的某些氨基酸可以用其它氨基酸取代，而不明显损失与例如抗体的抗原结合区或底物分子上的结合位点的结构的相互作用结合能力。由于正是蛋白的相互作用能力和特性限定了该蛋白的生物学功能活性，因此在蛋白序列中，当然还有其作为基础的DNA编码序列中，可以进行某些氨基酸序列取代，依然获得具有相似特性的蛋白。因此本发明人考虑到，在所公开的组合物的肽序列或编

码所述肽的相应的 DNA 序列中, 可以进行各种改变, 而不明显损失其生物利用性或生物活性。

在制备这种改变时, 可以考虑氨基酸的亲水指数。亲水氨基酸指数在赋予蛋白相互作用性生物功能方面的重要性是本领域普通了解的 (Kyte 和 Doolittle, 1982, 通过引用结合到本文中)。普遍认为, 所述氨基酸的相对亲水特性对所产生的蛋白的二级结构产生影响, 所述二级结构进而又限定了该蛋白与其它分子的相互作用, 所述其它分子例如为酶、底物、受体、DNA、抗体、抗原等。

根据氨基酸的疏水性和电荷特性, 已经确定了每种氨基酸的亲水指数 (Kyte 和 Doolittle, 1982), 它们是: 异亮氨酸(+4.5); 缬氨酸(+4.2); 亮氨酸(+3.8); 苯丙氨酸(+2.8); 半胱氨酸/胱氨酸(+2.5); 甲硫氨酸(+1.9); 丙氨酸(+1.8); 甘氨酸(-0.4); 苏氨酸(-0.7); 丝氨酸(-0.8); 色氨酸(-0.9); 酪氨酸(-1.3); 脯氨酸(-1.6); 组氨酸(-3.2); 谷氨酸(-3.5); 谷氨酰胺(-3.5); 天冬氨酸(-3.5); 天冬酰胺(-3.5); 赖氨酸(-3.9); 和精氨酸(-4.5)。

本领域已知某些氨基酸可以被具有相似亲水指数或计分的其它氨基酸取代, 仍产生具有相似生物活性的蛋白, 即仍获得生物功能等相同的蛋白。在制备这样的改变时, 优选亲水指数在 ± 2 以内的氨基酸取代, 特别优选亲水指数在 ± 1 以内的氨基酸的取代, 甚至更特别优选亲水指数在 ± 0.5 以内的氨基酸的取代。

本领域也知道, 根据亲水性可以有效地进行相似氨基酸的取代。美国专利 4,554,101 (其通过引用结合到本文中) 叙述了取决于其相邻氨基酸亲水性的蛋白的最大局部平均亲水性与所述蛋白的生物特性相关。

如美国专利 4,554,101 中详细描述, 已经确定了氨基酸残基的以下亲水性值: 精氨酸(+3.0); 赖氨酸(+3.0); 天冬氨酸(+3.0 \pm 1); 谷氨酸(+3.0 \pm 1); 丝氨酸(+0.3); 天冬酰胺(+0.2); 谷氨酰胺(+0.2); 甘氨酸(0); 苏氨酸(-0.4); 脯氨酸(-0.5 \pm 1); 丙氨酸(-0.5); 组氨酸(-0.5);

半胱氨酸(-1.0); 甲硫氨酸(-1.3); 缬氨酸(-1.5); 亮氨酸(-1.8); 异亮氨酸(-1.8); 酪氨酸(-2.3); 苯丙氨酸(-2.5); 色氨酸(-3.4)。

5 人们理解, 一种氨基酸可以被具有相似亲水性值的另一氨基酸取代, 并且仍获得生物学等同的、特别是免疫学等同的蛋白。在这样的改变中, 优选亲水性值在 ± 2 以内的氨基酸取代, 特别优选亲水性值在 ± 1 以内的氨基酸的取代, 甚至更特别优选亲水性值在 ± 0.5 以内的氨基酸的取代。

10 如上概述的, 因此氨基酸取代一般基于氨基酸侧链取代基的相对相似性, 例如其疏水性、亲水性、电荷、大小等。考虑各种上述特征的典型取代基是本领域技术人员熟知的, 并且包括: 精氨酸和赖氨酸; 谷氨酸和天冬氨酸; 丝氨酸和苏氨酸; 谷氨酰胺和天冬酰胺; 以及缬氨酸、亮氨酸和异亮氨酸。

15 本领域技术人员知道, 编码异源蛋白的多核苷酸当加入转基因植物的核 DNA 中时表达通常较差(Diehn 等的综述, 1996)。最好是, 基本上如美国专利 5,500,365 和 5,689,052 (其具体通过引用结合到本文中) 中所述, 设计编码目的异源蛋白的核苷酸序列。可用于表达的核苷酸序列的实例包括但不限于 SEQ ID NO:3、SEQ ID NO:7、SEQ ID NO:11 和 SEQ ID NO:19。

20 在所述基础多肽序列中的氨基酸取代基可以选自所述天然存在的氨基酸所属类别的其它成员。可以将氨基酸分为以下 4 组: (1)酸性氨基酸; (2)碱性氨基酸; (3)中性极性氨基酸; 和(4)中性非极性氨基酸。这些不同组中的代表性氨基酸包括但不限于: (1)酸性(带负电)氨基酸, 例如天冬氨酸和谷氨酸; (2)碱性(带正电)氨基酸, 例如精氨酸、组氨酸和赖氨酸; (3)中性极性氨基酸, 例如甘氨酸、丝氨酸、苏氨酸、半胱氨酸、胱氨酸、酪氨酸、天冬酰胺和谷氨酰胺; (4)中性非极性(疏水性)氨基酸, 例如丙氨酸、亮氨酸、异亮氨酸、缬氨酸、脯氨酸、苯丙氨酸、色氨酸和甲硫氨酸。

25 通过用这些组中的一种氨基酸取代同一组中的另一种氨基酸, 可

以进行基础多肽序列中的保守氨基酸改变。所述编码核苷酸序列(基因、质粒 DNA、cDNA 或合成 DNA)因此将具有相应的碱基置换,使得它编码 AMPA 转酰基酶的生物功能等同形式。

以下实施例描述本发明的优选实施方案。此中的权利要求书范围内的其它实施方案对于通过考虑本文公开的说明书或本发明的实施的本领域技术人员是显而易见的。本说明书和实施例仅打算考虑为示例性的,本发明的范围和精神由在实施例之后的权利要求书指明。在实施例中,所有百分比均以重量为基础给出,除非另有说明。

10

实施例

实施例 1

该实施例描述 N-氨基甲基磷酸(AMPA)对植物愈伤组织的生长抑制效应,而 N-乙酰-氨基甲基磷酸在体外培养条件下对植物愈伤组织没有抑制作用。

15

表达细菌 GOX 基因并且也暴露于草甘膦的某些重组植物物种,可以表现出通过诸如褪绿、花蕾脱落和育性降低的症状表现出的植物毒性效应。先前没有确定这些症状的基础。先前的研究已经表明,表达 GOX 的植物将草甘膦代谢为 AMPA 和乙醛酸(美国专利第 5,463,175 号)。乙醛酸容易被植物代谢,然而, AMPA 在植物组织中持续存在,可能是例如褪绿、矮化或其它不希望有的效应的植物毒性效应的原因。先前已经表明,无色杆菌属(*Achromobacter*)菌种 LBAA 能够将 AMPA 酶促修饰为 N-乙酰 AMPA (美国专利第 5,463,175 号)。所述无色杆菌属数据结合植物毒性数据表明, AMPA 在植物中的 N-酰化可以提供对褪绿和其它不希望有的效应的有效解除。因此,将烟草愈伤组织暴露于 AMPA 和暴露于 N-乙酰 AMPA,以便确定这些化合物中任一种是否表现出与在表达 GOX 并暴露于草甘膦的植物中观察到的相似的细胞毒性效应。

20

25

由 MS104 平板(MS 盐 4.3 g/l、蔗糖 30 g/l、B5 维生素 500X 2

ml/l、NAA 0.1 mg/l 和细菌培养用琼脂 1.0 mg/l)上的野生型烟草“Samsun”烟草栽培品种(*Nicotiana tabacum* cv. “Samsun”)的叶片小片,产生烟草愈伤组织。将愈伤组织加样到具有或不具有 AMPA 以及具有或不具有 N-乙酰 AMPA 的平板上。平板含有浓度为 0.1 mM 或 0.4 mM 的 AMPA 或 N-乙酰 AMPA。将平板温育至多 3 周,并定期进行监测。

不含 AMPA 或 N-乙酰 AMPA 的对照平板上的愈伤组织,如预期的以正常速率生长、再生根和苗。在 AMPA 存在下的愈伤组织严重受抑制。没有观察到生长,显示在这些浓度下 AMPA 的植物毒性效应。在含 N-乙酰 AMPA 的平板上的愈伤组织不受抑制,并且形成与对照愈伤组织生长相似的根和苗。该结果表明,作为 GOX 介导的草甘膦代谢副产物的 AMPA 可能导致植物中所观察到植物毒性。该结果也表明从遗传转化愈伤组织中选择植物的改进方法的可能性以及增强草甘膦除草剂抗性的可能方法。

实施例 2

该实施例描述在细菌无色杆菌(*Achromobacter* sp.)菌株 LBAA 中通过 GOX 酶水解进行的草甘膦的降解,导致 AMPA 和 N-乙酰-AMPA 的产生。

先前已经表明,GOX 介导的草甘膦降解产生乙醛酸和 AMPA (Barry 等, US 5,463,175)。也表明无色杆菌菌株 LBAA 由于草甘膦降解而产生 AMPA 和乙醛酸。按照以下方法,在草甘膦上生长的无色杆菌菌株 LBAA 的静息细胞中表征了草甘膦降解途径。在含有葡萄糖、葡萄糖酸和柠檬酸作为碳源、并且含有硫胺和酵母提取物(0.01%)以供应微量需求、而且含有 0.2 mM 草甘膦作为磷源的 DF3S 培养基中生长的 100 ml LBAA 培养物的细胞,以 200 Klett 单位的细胞密度收获,用 20 ml DF3S 培养基洗涤 2 次,并将 20 ml 细胞的等同物再悬浮于 100 l 含有 [¹⁴C]草甘膦(2.5 ml 52 mCi/mmol, Amersham; CFA.745)的相同培养基

中。将细胞混合物于 30 °C 在震荡下温育，并且以不同的间隔取出 20 ml 样品。将样品离心，以从肉汤上清液中分离所述细胞。上清液和细胞沉淀均通过 HPLC 进行分析。

5 通过配有放射性同位素标记检测的强阴离子交换(SAX) HPLC 分析以这种方式制备的样品，以测定其^[14C]-AMPA 和 N-乙酰-^[14C]-AMPA 的水平。用 Waters WISP 自动注射器注射样品。用 MACS2 - Monsanto 自动色谱数据采集系统收集色谱图形和定量数据。Spherisorb S5 SAX, 250 mm X 10mm 柱或 Alltech 5 微米, 250mmX10mm SAX 柱用于该分析。所用的溶剂命名为溶液 A 和溶液 B。溶液 A 含有 4% 甲醇中的用 H₃PO₄ 调至 pH 2.0 的 0.005M KH₂PO₄。溶液 B 含有 10 4% 甲醇中的用 H₃PO₄ 调至 pH 2.0 的 0.10 M KH₂PO₄。每种样品的运转时间由分级梯度程序组成，洗脱液流速为 3 ml/分钟，而闪烁液(商标为 ATOMFLOW, No. NEN-995, 得自 Packard Instruments)的流速为 9 ml/分钟。鉴别所分析的每种样品中的^[14C]-AMPA 与 N-乙酰-^[14C]-AMPA 的 HPLC 溶剂分布型表示为时间 0 至 5 分钟 100% 溶剂 A，然后 5 分钟至 15 分钟时 100% 溶剂 B，然后至 20 分钟 100% 溶剂 A，此时准备 15 柱子以接受另一样品。

将细胞沉淀首先再悬浮于通过加入 0.65 N HCl 制成酸性的 DF3S 培养基中，煮沸 5 分钟，然后短暂离心以提供用于 HPLC 分析的溶液 20 相。上清液在 HPLC 分析之前进行相似的处理。一个酸化的草甘膦对照也进行 HPLC 分析，草甘膦的保留时间(RT)测定为 10.8 分钟。温育 2 小时后保留在上清液中的草甘膦峰中的放射性量已经降低至原始水平的约 33%，表明草甘膦被大量代谢。发现约 3% 的所述草甘膦在细胞内。与甲胺标准共洗脱、RT 为 6 分钟的物质占上清液中原始放射性 25 量的约 5%，占细胞内容物中鉴定的原始放射性量的约 1.5%。

在随后一个实验中进一步阐述 GOX 介导的草甘膦降解途径，其中将^[14C]AMPA 的代谢与如上所述在静息细胞中^[14C]草甘膦的代谢进行比较，所述静息细胞以 165 Klett 单位收获，并且以每 100 ml DF3S

培养基 15 ml 细胞的相当量再悬浮。通过 HPLC 分析样品，样品由如上所述酸化且处理的全培养物组成。发现暴露于 ^{14}C 草甘膦 2 小时的培养物在保留时间为 14.7 分钟的甲胺/N-乙酰-甲胺峰中有 25%的所述标记，12.5%为保留时间为 6 分钟的 AMPA，有 30%在保留时间为 13.2 分钟的峰中，30%为保留时间为 10.8 分钟的草甘膦。对暴露于 ^{14}C -AMPA 2 小时的培养物的分析表明，发现 15%的所述标记为 N-乙酰-甲胺/甲胺，59%为 AMPA，而 18%在 13.2 分钟的峰中。于 13.2 分钟洗脱出的物质通过负离子电喷雾质谱被鉴定为 N-乙酰-AMPA。该结果表明强离子在 m/e 152 和 m/e 154，正如对于分子量为 153 道尔顿的该化合物所预期的。 m/e 154 离子是由同位素 ^{14}C 原子产生的。N-乙酰-甲基- ^{14}C -AMPA 由 N-甲基- ^{14}C -AMPA 产生，后者是 ^{14}C -AMPA 制剂中的一种已知的杂质。

这些数据表明，无色杆菌属菌株 LBAA 中的草甘膦降解途径由草甘膦水解进行至 AMPA，然后 AMPA 假定通过去磷酸化步骤转化为产物甲胺，假定通过某些先前未知的酰基转移步骤转化为 N-乙酰-AMPA。然后少量的 N-乙酰-AMPA 转化为 N-乙酰-甲胺。根据当用氨基磷酸作为唯一磷酸源时在大肠杆菌中鉴定的产物，已经推导出一个相似的酰化步骤(Avila 等, 1987)。

20 实施例 3

该实施例描述了大肠杆菌中 AMPA 酰基转移酶活性的鉴定。

Avila 等(1987)在研究 C-P 键裂开时，鉴定出在大肠杆菌体内用作唯一磷酸源的多种氨基磷酸底物的代谢的去磷酸化的生物降解产物。他们的研究表明，AMPA 在大肠杆菌 K-12 中不是酰化的底物。另外，Avila 等关心在大肠杆菌中磷酸的 C-P 键裂开时 N-联化学取代的效应，并鉴定出衍生自某些氨基磷酸代谢的 N-乙酰化产物。Avila 等也证明，“野生型”大肠杆菌 K12 菌株与野生型大肠杆菌 B 菌株不同，不能用磷酸作为磷酸源。因此，考虑到如实施例 1 中表明的 AMPA 对

愈伤组织的植物毒性效应和如实施例 2 中表明的 GOX 介导的草甘膦降解产生 AMPA， Avila 等中的大肠杆菌数据表明，在能够将氨甲基磷酸(AMPA)转化为 N-乙酰-AMPA 的某些细菌菌种中可能存在一种酶或途径。具有那些特性的酶或途径如果在植物中表达，可能在用草甘

5 膦处理时赋予表达 GOX 的植物相当大的优势。

为了测试这一点，在含低磷酸培养基上培养适于在 AMPA 上生长的大肠杆菌 K-12 菌株，以便获得用于分析能够就 AMPA N-酰化的酶存在的细胞裂解液。 *phn* (*mpu*)操纵子在大肠杆菌 K-12 中是隐蔽的，因为它有 8 个碱基对的插入，引起 *phnE* 基因中的移码突变。这种移码

10 使 PhnE 失活，并且对该操纵子内 *phnE* 下游的其它基因的翻译产生极性效应，导致这种突变体不能利用磷酸作为磷酸源(Makino 等, J. Bacteriol. 173:2665-2672, 1991)。对自发衍生的突变的选择恢复 *phn* 操纵子的功能(*phn*+或 *mpu*+). 因此，适于在 AMPA、甲基磷酸或乙基磷酸上生长的 K-12 菌株含有这样的有效自发衍生的突变。

简而言之，将一等份的大肠杆菌 K-12 菌株 JM101 (*mpu*-)的新鲜 L 液体培养物平板接种 MOPS (Neidhardt 等, 1974)完全琼脂培养基上，该培养基含有 25mg/ml 氨基酸、 10 mg/ml 维生素 B1 [硫胺]、 0.2%葡萄糖和 1.5% DIFCO “纯化的”琼脂与作为唯一磷酸源的氨甲基磷酸(AMPA; 0.2 mM; Sigma Chemical Co., St. Louis, MO)，并于 37 °C 培养 3

15 天。挑出在该培养基上产生的菌落，并且在含有或者 AMPA 或者甲基磷酸(Alfa)作为唯一磷酸源的 MOPS 完全琼脂上划线接种。从在含有两种磷酸的培养基上等同均一生长的那些菌落中选择一个菌落，命名为大肠杆菌 JM101 *mpu*+, 并且进一步命名为大肠杆菌菌株 GB993。

当大肠杆菌在磷酸源缺乏或有限的培养基中生长时，诱导 *phn* 操纵子。因此，当在 MOPS 基本培养基上生长时，将大肠杆菌 GB993 与亲代 JM101 菌株进行比较。在相同的条件下，培养 GB993 及其 *mpu*-亲代菌株 JM101，只是可得到的磷酸量不同或补充有 AMPA 的磷酸量不同。于 37 °C 连续震荡下，在 250 ml 侧臂锥形瓶中的含有 0.1 mM 或

25

5 mM 磷酸或补充约 0.2 mM AMPA 的 0.1 mM 磷酸的 pH 7.0 MOPS 培养基(5 ml 10X MOPS 盐、 0.5 ml 1 mg/ml 硫胺、 0.5 ml 20%葡萄糖, 用 dH₂O 调至 50 ml)中, 以双份培养 50 ml 培养物。一般让培养物生长至约 220 Klett 单位, 通过离心沉淀细胞, 将细胞再悬浮于 1.5 ml 10 mM Tris/1 mM DTT 中, 并且以 1,000 psi 通过 2 次弗氏压碎器进行裂解。将裂解液离心以除去碎片, 使上清液通过用 50 mM Tris pH 7.0 平衡的 G-50 柱。表 2 表明以该方式生长的细胞培养物的结果。

表 2.

10

磷酸底物对细胞生长的影响

	菌株					
	JM101	JM101	JM101	GB993	GB993	GB993
	0.1 mM 磷酸	5 mM 磷酸	0.2 mM AMPA	0.1 mM 磷酸	5 mM 磷酸	0.2 mM AMPA
生长期(hrs)	48	29	54	48	29	54
收获密度 (Klett 单位)	155	240	-	140	244	185

-表明没有可测量的生长

15

用 HPLC 分析测定培养基和细胞裂解液中是否存在任何 AMPA 酰基转移酶活性。该分析监测 [¹⁴C]AMPA 向 N-乙酰-[¹⁴C]AMPA 的转化。一般而言, 将由 16.5 mg 乙酰-CoA、 250 μl 2M Tris pH 7.5、 4.5 ml dH₂O 和 [¹⁴C]AMPA (30 mM)组成的 100 μl 2X 测定溶液与 25-75 μl 裂解液和 0.5 M MgCl₂ 和 MnCl₂ 各 1 μl 混合, 并用 dH₂O 调至 200 μl。将分析物于 37 °C 温育 30 分钟, 并用 200 μl 90-100 mM NaOAc (乙酸钠) pH 4.4 的乙醇溶液猝灭, 然后立即如上所述通过 HPLC 分析或贮存于 -20 °C。仅由在 AMPA 或 0.1 mM 磷酸补充的培养基存在下生长的培养物得到的 GB993 裂解液样品表现出明显的 AMPA 酰基转移酶活性。该结果表明, 编码能够将 AMPA N-酰化的酰基转移酶的基因存在于 GB993 中, 并且在低磷酸条件下生长时其表达受调节。因此, 该酶活

20

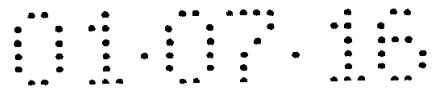
性的编码序列看来是 *pho* 调节子的部分, 并且可能存在于 *phn* 操纵子中。

实施例 4

5 该实施例描述编码能够将 AMPA N-酰化的酶的大肠杆菌 *phn* 操纵子基因的鉴定。

实施例 3 表明, 在大肠杆菌裂解液中观察到的 AMPA 酰基转移酶活性可能由 *phn* 操纵子中的基因编码。先前已经克隆了大肠杆菌 B 和大肠杆菌 K-12 中的完整的 *phn* 操纵子, 并且已经对大肠杆菌 B 中的所述操纵子进行测序(Wanner 等, Chen 等)。已经表明, 大肠杆菌 K-12 *phn* 操纵子的 DNA 序列与公布的大肠杆菌 B 的 *phn* 操纵子的 DNA 序列相同, 只是在 *phnE* 基因有 8 个碱基对的插入(Wanner 等)。含不同量的来自两种细菌中任一种的遗传背景的 *phn* 操纵子基因的克隆是容易得到 (Wanner 等, Chen 等, Dr. J.W. Frost Purdue University)。用含不同量 JM101 *phn* 操纵子 DNA 的质粒转化 JM101(*mpu*-), 以便测试质粒定位的 *phn* 基因, 该基因当表达时, 赋予 JM101 利用 AMPA 作为唯一磷酸源的能力。

从 J. Frost (Dr. J.W. Frost, Department of Chemistry, Purdue University, West Lafayette, Indiana 47907)获得的、在本文中命名为 pF 的质粒含有一个大肠杆菌 K-12 8 kb *EcoRI* 片段, 该片段编码 *phn* 操纵子基因 *phnG* 至 *phnQ*。在 *phnG* 编码区的 5'端存在一个 *NcoI* 位点。用 *EcoRI* 和 *NcoI* 消化质粒 pF, 释放含有基因 *phnG* 至 *phnI* 的 2 kb *NcoI-EcoRI* 片段、和一个长度约 6 kb 的第二 *NcoI-EcoRI* 片段, 后一片段含有基因 *phnJ* 至 *phnQ*。凝胶纯化每种片段, 并将其连接到一个克隆表达载体中, 其连接方式将允许从质粒载(plasmid borne)诱导型启动子表达存在于每种 *NcoI-EcoRI* 片段中的 *phn* 操纵子基因。将所述 2 kb 片段插入载体 pMON7258 内的 *NcoI-EcoRI* 位点中, 所产生的质粒命名为 p58-1, 载体 pMON7258 是除了多接头区外与 pUC118 相同的一种



正选择克隆载体(Viera 等, Methods Enzymol. 153:3, 1987)。所述 2 kb 片段在 p58-1 中的方向允许从该载体内的 *lac* 启动子表达 *phnG- phnI* 基因。将所述 6 kb *EcoRI-NcoI* 片段插入相似的正选择载体 pMON7259 的 *NcoI* 和 *EcoRI* 位点中, 产生命名为 pMON17195 的质粒。pMON7259 除多接头区外与 pUC119 相同, 它含有一个方向与 pMON7258 中的多克隆位点相反的多克隆位点, 并且也允许从 *lac* 启动子表达 *phnJ- phnQ* 基因。将 p58-1 和 pMON7259 转化到大肠杆菌 K12 (*mpu-*) 菌株 JM101 中, 并以氨苄青霉素抗生素抗性选择维持。也将 pMON7259 和 pF 转化到 JM101 中分别作为阴性对照和阳性对照。

每种转化体的培养物在补充有 2%酪蛋白氨基酸、硫胺和 0.2%葡萄糖的 M9 液体肉汤培养基中于 37 °C 震荡生长过夜, 然后以 1:50 稀释到 250 ml 侧臂锥形瓶中的 50 ml 新鲜预温热的同一组成的培养基中。将培养物于 37 °C 震荡培养, 直至在 Klett-Summerson 分光光度计上通过 #2 绿色滤光片测定细胞密度达到约 80-100 Klett 单位。通过加入 100 微升 500 mM IPTG, 使得 IPTG 的终浓度约为 1 mM, 诱导来自质粒 *lac* 启动子的表达。让诱导期生长期进行 2 小时。表 3 表明来自 1:50 稀释液通过 2 小时诱导期的每种培养物的细胞密度分布型。

表 3.

含不同 *phn* 质粒的 JM101 培养物的诱导分布型

培养物/质粒	IPTG	I ₀	I ₁	I ₂
pMON7259	+	13	75	222
p58-1	+	15	70	212
pMON17195	+	15	90	220
pF	+	17	97	290
pF	-	15	-	260

I₀ 表示为 1:50 稀释度时间点的细胞培养物密度; I₁ 表示为加入 IPTG 时的细胞培养物密度; I₂ 表示为收获时的细胞培养物密度。

通过在 Beckman J2 离心机中以 10,000 rpm 于 4 °C 离心 10 分钟,

收获每种培养物中的细胞。细胞沉淀在冰冷的 154 mM NaCl 溶液中洗涤 1 次, 然后再悬浮于 1.5 ml 提取缓冲液(50 mM Tris-HCl pH 7.5, 1 mM DTT, 50 mM Tris-HCl pH 7.5)中。以 1000 psi 通过 2 次弗氏压碎器将细胞悬浮液破碎。产生的裂解液在 EPPENDORF™ 5402 型微量离心机中以 14,000 rpm 于 4 °C 离心 15 分钟, 以除去碎片。将每种澄清的裂解液转移到新鲜的预冷管中, 用 50 mM Tris-HCl pH 7.5 将提取物的体积调至 2.5 ml。用 25 ml 50 mM Tris-HCl pH 7.5 平衡 PD10 柱子, 然后将每种样品加样到脱盐柱上。用 50 mM Tris-HCl pH 7.5 将每种洗脱出的样品调至 3.5 ml。将每种样品分配到分析管中, 与试剂混合, 以便如表 4 中所示分析 AMPA 酰基转移酶活性的存在。

表 4.

表达 *phn* 基因的细菌裂解液的分析条件

样品	IPTG	提取物 体积*	50 mM Tris 体积*	2X 分析 混合物体积*	总体积*
pMON7259	+	25	75	100	200
pMON7259	+	100	0	100	200
p58-1	+	25	75	100	200
p58-1	+	100	0	100	200
pMON17195	+	25	75	100	200
pMON17195	+	100	0	100	200
pF	+	25	75	100	200
pF	+	100	0	100	200
pF	-	25	75	100	200
pF	-	100	0	100	200
-	na	-	100	100	200

*所有体积均以微升计

15 每种样品的混合物的组成, 以质粒内容物命名, 如用于 AMPA 酰基转移酶分析法制备。

每种混合物于 37 °C 温育 30 分钟, 用等体积(200 微升) 90-100 mM NaOAc (乙酸钠) pH 4.4 的乙醇溶液猝灭, 如果不立即通过如上所述的

HPLC 分析, 则于-20 °C 贮存过夜。每种裂解液的未使用部分或者贮存于 4 °C, 或者与甘油混合至 10%(体积)且于-20 °C 贮存。

经过 AMPA 转酰基酶分析的每种裂解液样品, 如上所述通过 HPLC 分析其^[14C]APMA 和酰化^[14C]AMPA 的存在。结果示于表 5 中。

5

表 5.

细菌裂解液 AMPA 转化为乙酰-AMPA 的 HPLC 分析

样品	%乙酰 AMPA	%AMPA
pMON7259	无数据	无数据
pMON7259	8	92
p58-1	5	95
p58-1	13	87
pMON17195	100	0
pMON17195	100	0
pF	61	39
pF	97	3
pF	52	48
pF	90	10
-	-	100

每种样品的 HPLC 分析的结果, 表明作为两个峰中^[14C]混合总量的百分比的^[14C]AMPA 或乙酰-^[14C]AMPA 的相对量。

10 该数据表明, 含分离自 pMON17195 中 pF 的 6 kb *NcoI-EcoRI* 片段的质粒含有一个或多个基因, 所述基因在 IPTG 诱导大肠杆菌的 *mpu-*菌株中的 *lac* 启动子时, 引发能够将所述分析混合物中可利用的所有^[14C]AMPA 转化为乙酰-^[14C]AMPA 的酰基转移酶活性的产生。通过限制缺失分析, 进一步限定 AMPA N-酰化所需的一个或多个基因。

15 构建含有或者来自大肠杆菌 B 或者来自大肠杆菌 K-12 的 *phn* 操纵子的不同区段的质粒, 以进一步描绘涉及在 *mpu-*大肠杆菌 JM101 中表达时赋予 AMPA 酰基转移酶活性的一个或多个 *phn* 操纵子基因的性质。pMON7333 含有 pMON17195 等同的大肠杆菌 DNA 插入, 但

却是在 pUC119 中，并且是一个含有野生型 *phn* 操纵子基因 *phnG* 至 *phnQ* 的单一的大肠杆菌 B 菌株 *HindIII* 片段。通过将来自 pMON7333 的 5,713 个碱基对的 *NcoI* 至 *EcoRI* 大肠杆菌 B DNA 片段克隆到 pMON7259 中，构建 pMON15020，pMON15020 含有基因 *phnJ* 至 *phnQ*。通过将来自 pMON17195 的 1,686 个碱基对的 *EcoRI* 至 *SalI* 片段插入含有大肠杆菌 K-12 基因 *phnO*、*phnP* 和 *phnQ* 的正选择克隆表达载体 pBlueScriptSP (Invitrogen) 中，构建 pMON15022。通过将来自 pMON17195 的 1,820 个碱基对的 *SalI* 片段缺失，留下大肠杆菌 K-12 *phn* 操纵子基因 *phnJ* 和 *phnK*、*phnL* 的 5' 端以及全部的 *phnO*、*phnP* 和 *phnQ*，构建 pMON15023。

将质粒 pMON17195、pMON15020、pMON15022、pMON15023 和 pMON7259 转化到 *mpu*-大肠杆菌 K-12 菌株 JM101 中，并通过氨苄青霉素抗生素选择来维持。这些转化体中的每种转化体的过夜培养物在抗生素选择下培养，并以 1:50 稀释到新鲜的如上所述的 M9 培养基中，并且在震荡下于 37 °C 在 250 ml 侧臂锥形瓶中温育至细胞密度约为 100 Klett 单位。如实施例 3 所述，用 IPTG 诱导每种培养物，将培养物在震荡下再温育 2 小时。通过在 Beckman J2 离心机中以 4,000 RPM 于 4 °C 离心 10 分钟，收获细胞。细胞沉淀用 50 ml 154 mM NaCl 洗涤 1 次，并于 -20 °C 贮存。

将细胞沉淀再悬浮于如实施例 3 所述的 1.5 ml 提取缓冲液中，并通过以 1000 psi 通过弗氏压碎器 2 次破碎细胞。破碎的细胞悬浮液在 Eppendorf 微量离心机 5402 型中以 14,000 rpm 于 4 °C 离心 5 分钟。将澄清的连接于倾至在冰上预冷的新管中，加入提取缓冲液将总体积调至 2.5 ml。这些样品在用 25 ml 50 mM Tris-HCl pH 7.5 预平衡的 PD10 柱上脱盐，并用 3.5 ml 50 mM Tris-HCl pH 7.5 洗脱。然后将样品如上所述进行 AMPA 酰化分析，于 37 °C 温育 30 分钟，然后用 200 微升 90.9 mM NaOAc pH 4.4 猝灭。用于该分析中的每种样品的体积记录在表 6 中。所有体积均代表所用的每种溶液的微升数。

表 6.

表达来自质粒的 *phn* 基因的细菌裂解液的分析条件

质粒	提取物	50 mM Tris	2X 分析 混合物	总体积
-	-	100	100	200
pMON 17195	25	75	100	200
pMON 17195	100	-	100	200
pMON 15020	75	75	100	200
pMON 15020	100	-	100	200
pMON 15022	75	75	100	200
pMON 15022	100	-	100	200
pMON 15023	75	75	100	200
pMON 15023	100	-	100	200
pMON 7259	75	75	100	200
pMON 7259	100	-	100	200

每种样品的混合物的组成，以质粒内容物命名，如用于 AMPA 酰基转移酶分析法制备。

5 猝灭的样品如上所述经过 HPLC 分析。表 7 描述了每种样品的 HPLC 分析结果，表明作为两个峰中^[14C]混合总量的百分比的^[14C]AMPA 或乙酰-^[14C]AMPA 相对量。

表 7.

细菌裂解液 $[^{14}\text{C}]$ -AMPA 转化为乙酰- $[^{14}\text{C}]$ -AMPA 的 HPLC 分析

样品	提取物 体积	% $[^{14}\text{C}]$ - AMPA	%乙酰- $[^{14}\text{C}]$ - AMPA	总% $[^{14}\text{C}]$
-	-	100	-	100
pMON 17195	25	66	34	100
pMON 17195	100	26	74	100
pMON 15020	75	-	100	100
pMON 15020	100	-	100	100
pMON 15022	75	-	100	100
pMON 15022	100	-	100	100
pMON 15023	75	-	100	100
pMON 15023	100	-	100	100
pMON 7259	75	87	13	100
pMON 7259	100	72	28	100

每种样品的 HPLC 分析, 表明作为两个峰中 $[^{14}\text{C}]$ 混合总量的百分比的 $[^{14}\text{C}]$ AMPA 或乙酰- $[^{14}\text{C}]$ AMPA 相对量。

表 7 中的数据表明, AMPA 酰化活性得自由 *phnO*、*phnP* 和 *phnQ* 组成的 *phn* 操纵子可读框, *phnO*、*phnP* 和 *phnQ* 是 pMON15022 中仅存的 *phn* 基因。诱导时提供 AMPA 酰化活性的其它质粒也至少含有 *phnO*、*phnP* 和 *phnQ* 基因, 这提供了强有力的证据, 表明观察到的活性是这些基因产物中的一种或多种的结果。因此, 基于 *phnO*、*phnP* 和 *phnQ* 基因序列构建了另外的质粒, 以便确定哪个或哪些基因是酰化功能所需的。

有时在文献中已知细菌酰基转移酶(acylase)、转酰基酶(transacylase)和酰基转移酶(acyltransferase)基因。大多数为小的 15-25 kDa 蛋白。因此, 在大小比较的基础上, 仅 *phnO* 和 *phnQ* 基因产物属于该类别。然而, 基于与 GENBANK、SWISSPROT 和 EMBL 数据库中的其它蛋白的相似性比较, 预测的 *phnO* 基因产物看来最接近地类似具有酰基转移酶活性的其它蛋白。例如, 大肠杆菌 PhnO 蛋白与 Wohlleben 等(Mol. Gen. Genet. 217:202-208, 1989)中描述的庆大霉素乙

酰转移酶-3-I 很好地匹配。含有单一 6.0 kb *NcoI-EcoRI* 片段上的大肠杆菌 B *phn* 操纵子基因 *phnJ* 至 *phnP* 的 pMON15020 用 *SalI* 和 *EcoRI* 消化, 释放含有 *phnO*、*phnP* 和 *phnQ* 基因的 2.0 kb 片段。将该 2 kb 片段切下, 并从 0.7% TAE 琼脂糖凝胶中纯化, 用 T4 DNA 聚合酶处理, 以切下 3' 突出端, 然后用 Klenow 和脱氧核苷酸三磷酸酶(dXTP) 处理, 以提供平端, 然后连接到 pBlueScriptSP 的 *EcoRV* 位点中, 产生质粒 pMON15024。pMON15024 用 *NdeI* 和 *EcoRI* 消化, 缺失含有 *phnP* 大部分和所有 *phnQ* 编码序列的 1200 个碱基对的片段。仍含有 *phnO* 基因的剩余的 pMON15024 质粒片段, 在二脱氧核苷酸存在下用 Klenow 片段 DNA 聚合酶、按照生产商的说明进行处理, 以便补平暴露于限制酶消化产生的 3' 端, 然后将其连接在一起, 产生质粒 pMON15027。pMON15027 仅含有 3' 侧翼为小部分 *phnP* 的 *phnO* 基因。将得自 pMON15024 的 1200 个碱基对的 *NdeI* 至 *EcoRI* 片段克隆到 pMON2123 中, 产生 pMON15026, 后者含有 3' 侧翼为 *phnQ* 的 *phnP* 基因 3' 的 2/3。将质粒 pMON15024、15026 和 15027 导入 *mpu*-JM101 中, 在生长并诱导后如上所述分析转化体细胞裂解液中 AMPA 酰基转移酶活性的存在。仅 pMON15024 和 pMON15027 表现出酰基转移酶活性, 表明 *phnO* 基因产物负责 AMPA 的酰化。

用热循环法, 产生仅含 *phnO* 基因且带有方便的侧翼限制性内切核酸酶位点以用于进一步克隆操作的 DNA 片段。根据公布的 *phnO* 基因和侧翼序列, 由 Midland Certified Reagents Co. (Midland Texas) 合成合成寡核苷酸引物, 以便扩增 *phnO* 基因(Chen 等, *J. Biol. Chem.* **256**: 4461-4471, 1990)。命名为 AATPCR6 的序列 AAACACCATGGCTGCTTGTG (SEQ ID NO: 5) 代表与 *phnO* 基因模板链同源的合成寡核苷酸。SEQ ID NO: 5 的 5' 腺苷残基对应于公布的 *phn* 操纵子序列的碱基对 13,955, 紧接位置 13,962-13964 的 *phnO* ATG 起始密码子的 5' (Chen 等, *J. Biol. Chem.* **256**: 4461-4471, 1990)。根据公布的 *phnO* 序列, SEQ ID NO: 5 在位置 13,965 掺入了一个单碱基对

错配，表示为 C 至 G 的转换，这在位置 2 产生一个丙氨酸密码子，取代了脯氨酸密码子，也产生了一个跨越 ATG 起始密码子的唯一的 *NcoI* 限制位点。命名为 AATPCR7 的序列

GTGACGAATTCGAGCTCATTACAGCGCCTTGGTGA (SEQ ID NO: 6)

5 代表与 *phnO* 基因的编码链同源的合成寡核苷酸。SEQ ID NO: 6 的 3' 腺苷残基对应于公布的 *phn* 操纵子的碱基对 14,380 (Chen 等, J. Biol. Chem. **256**: 4461-4471, 1990)。SEQ ID NO: 6 中位置编号 19 的胸苷对应于公布的 *phnO* 序列的位置 14,396 的腺苷(Chen 等)。SEQ ID NO: 6 的一部分重叠天然 *phnO* 终止密码子，在紧接该天然终止密码子的 3' 和邻近引入一个符合读框的第二终止密码子，也在这些终止密码子的 3' 引入一个唯一的 *EcoRI* 和 *SacI* 限制位点。

pMON15024 用作在标准热扩增反应中扩增 *phnO* 基因的模板。简而言之，制备 100 微升反应样品，它含有 0.1 ng 模板 DNA、反应缓冲液、200 pM 每种引物、200 mM dNTP、1.25 U Taq DNA 聚合酶，并且用矿物油覆盖。该反应样品经过 35 个循环：94 °C 1 分钟、50 °C 2 分钟和 72 °C 3 分钟，通过在溴化乙锭染色的 0.7% TAE 琼脂糖凝胶上分析 5 微升反应样品测定，这导致扩增 459 个碱基对的 DNA 产物。用标准方法，在用 *NcoI* 和 *EcoRI* 限制性内切核酸酶消化所述 459 个碱基对的扩增产物样品后，从 1% TAE 琼脂糖凝胶上纯化出 444 个碱基对的产物。将所述 444 个碱基对的产物连接到 pMON7259 中匹配位点中，产生 pMON15028。分析从含 pMON15028 的 JM101 的 IPTG 诱导培养物如上所述制备的细胞裂解液中 AMPA 酰基转移酶活性的存在，并将其与含 pMON15027 的培养物进行比较。结果是不可区分的，因此证实，*phnO* 编码能够将 AMPA 酰化的酶。另外，该结果表明，由于用 AATPCR6 寡核苷酸引物(SEQ ID NO: 5)热扩增而引入所述基因编码序列中的该蛋白中的 P2A 突变，当在大肠杆菌中表达时对所产生的 PhnO 蛋白的酰基转移酶活性没有影响。

实施例 5

该实施例描述了针对 PhnO 肽的多克隆抗体的产生。

对 *phnO* 基因产物的进一步研究需要应用针对 PhnO 蛋白的抗体。因此，在大肠杆菌 JM101 中超量产生 PhnO，以用作免疫原，以在注射到山羊中时刺激产生抗体。将含有质粒 pMON15028 中的 P2A 突变的 *phnO* 基因在 *NcoI* 至 *EcoRI* DNA 片段上引入质粒 pMON17061 中，产生 pMON15032。pMON15032 中的 *phnO* 表达在邻近噬菌体 T7 基因 10L 核糖体结合序列的大肠杆菌 *recA* 启动子的控制之下。使细胞生长至对数中期，并通过用 50 mg 茶啉酮酸粉溶于 1 ml 0.1 N NaOH 中的储液将茶啉酮酸加入所述培养物中至约 50 ppm，进行诱导。培养物在诱导条件下于 37 °C 维持 12 小时。如实施例 3 中所述收获细胞，并在磷酸缓冲盐溶液中进行超声处理。测定经诱导的大肠杆菌裂解液里约 23% 的总可溶性蛋白是 PhnO，而通过 SDS-PAGE 和考马斯蓝染色判断，约 60% 的总 PhnO 蛋白释放到可溶相中。通过制备型 SDS-PAGE 进一步纯化该蛋白，提供足够量的 PhnO，以用于产生与 PhnO 或相关的 AMPA 转酰基酶蛋白结合或免疫原性反应的抗体。简而言之，根据大小，在 15% SDS-PAGE 凝胶中将所述 PhnO 蛋白与其它蛋白分离。切下含有所述 PhnO 蛋白的凝胶条，称重，并用 polytron 在体积等于凝胶条质量的磷酸缓冲盐溶液(PBS, pH 7.0)中匀浆。将匀浆液与等体积完全弗氏培养基混合，直至获得胶体混合物。用 8 ml 该混合物的接种物，首次注射到山羊中。注射后 2 周，收集 50 ml 放血样品，通过离心从血液固体分离出血清。4 周时，在 50% 不完全弗氏佐剂的胶体混合物中给予凝胶纯化的 PhnO 蛋白的增强注射，在 6 周时获得第二次放血样品。

用来自第二次放血的血清，根据对 PhnO 蛋白特异性的足够的抗体效价的存在进行筛选。将含 pMON15032 的 JM101 细胞的提取物进行蛋白质印迹分析。该提取物中的蛋白浓度通过 Bradford 分析测定为约 55 mg/ml，对采用这种同一提取物的先前的考马斯染色的凝胶进行

光密度计扫描,表明约 23%的总细胞蛋白是 PhnO。将该提取物在 PD10 柱上脱盐,用 10 mM Tris pH 7.5 洗脱,并用等体积 2X SDS 样品缓冲液稀释。用 1X 样品缓冲液制备连续稀释液,并加样到 15% SDS PAGE 凝胶的各孔中。将除大肠杆菌 PhnO 提取物外的另外的样品与含每泳道另外 10 微克蛋白的烟草样品蛋白提取物混合。用烟草叶片蛋白提取物筛选抗植物蛋白的交叉反应性抗体的存在。按照大小,通过以恒定的 7.5 mA 于 4 °C 电泳 14 小时,分离蛋白,并将该凝胶在 0.5 安培下,在 Tris-甘氨酸转移缓冲液中电印迹 1 小时,电印迹到 MSI 0.45 微米硝化纤维素滤膜上。然后用 TBST (Tris, BSA, NaCl, Tween-20, Short Protocols in Molecular Biology, 第 3 版, Wiley and Sons, Pub.)于室温下封闭膜 2 小时,于室温下与第二次放血血清的 1:500 稀释液一起温育 45 分钟,在 TBST 中洗涤 2 次,再与碱性磷酸酶缀合的兔抗山羊 IgG (Boehringer Mannheim Biochemicals, Inc.)一起温育 45 分钟,用 TBST 洗涤 3 次,然后用碱性磷酸酶缓冲液洗涤 1 次,最后与含有 NBT 和 BCIP 的标准显色溶液温育 2.5 分钟。通过用大量蒸馏水洗涤膜,终止反应。该抗体能够在少至 50 纳克大肠杆菌提取物中检测到 PhnO 蛋白,而与一半的该样品中另外的植物蛋白的存在无关。另外,在两组样品中检测到非常少的交叉反应性条带,表明当用这种蛋白质印迹法测试时,该血清样品含有非常少的与或者大肠杆菌或者烟草植物蛋白交叉反应的 IgG。

也利用用于产生能够与 PhnO 蛋白特异性结合或免疫原性反应的抗体的一种替代来源。将 *phnO* 基因置于含有金属结合氨基酸编码序列(His6)的市售载体(Invitrogen)中,使 His6 位于所述 *phnO* 编码序列上游并与其符合读框。将 His6-*phnO* DNA 序列在 *NcoI* 至 *EcoRI* 片段上插入大肠杆菌表达载体 pMON6235 中,并处于大肠杆菌阿拉伯糖操纵子 *araBAD* 启动子的控制之下,产生质粒 pMON32909。当用阿拉伯糖诱导含有 pMON32909 的大肠杆菌 W3110 时,产生 His6-PhnO 蛋白,按照生产商的说明,将该蛋白在金属亲和柱上纯化。

用与上述用于山羊类似的免疫方法, 将 His 标记的纯化 His6-PhnO 蛋白标准注射到 6 只新西兰白兔中。也表明在这些兔中产生的抗血清特异性地结合 PhnO 蛋白, 并且与其它大肠杆菌细菌蛋白或烟草植物蛋白没有交叉反应性。

5

实施例 6

该实施例用氨甲基膦酸和乙酰-CoA 作为酶分析中的底物、并且通过终点动力学分析测量, 描述 AMPA 转酰基酶的特性。

10 测定 PhnO 酶对于底物氨甲基膦酸和乙酰-CoA 的表观 K_m (K_m) 和 V_{max} (V_{max})。所述 PhnO K_m 和 V_{max} 的测定通过终点动力学分析来进行, 测定在不同底物浓度下消耗每种底物的酶速度, 并且将酶速度的倒数对底物浓度的倒数作图, 产生酶促动力学的 Lineweaver-Burk 图。如实施例 2 中所述, 用如实施例 4 中产生的、从 pMON15032 表达 *phnO* 的大肠杆菌的脱盐粗制裂解液中的酶, 监测 [^{14}C]-AMPA 向 N-乙酰- ^{14}C -AMPA 的转化。用 Bradford 的方法, 测定每 ml 提取物的总蛋白, 表明总蛋白约为 22.5 mg/ml。对从这些裂解液中分离 PhnO 蛋白的考马斯染色的 SDS-聚丙烯酰胺凝胶的光密度扫描, 表明 PhnO 占总蛋白的约 23%, 因此, 确定该细胞提取物每 ml 含有约 5.2 mg PhnO 蛋白。在测定 PhnO 对于 AMPA 的表观 K_m 和 V_{max} 的第一个分析中, 15 [^{14}C]-AMPA 浓度范围为 2-38 mM。将酶反应物于 37 °C 温育 5 分钟, 用 1 体积的 100 mM 乙酸钠(NaOAc) pH 4.4 的乙醇溶液猝灭。通过 HPLC 分析样品, 测定转化为 N-乙酰- ^{14}C -AMPA 的 [^{14}C]-AMPA 的量。每组反应物的分析条件和结果示于表 8 中。

20

表 8.

对于 AMPA 底物的 PhnO 酶促动力学

样品	S ¹	%转化 ²	速度 ³	1/S	1/V	V/S
1	200	39.5	79	1.0	0.0127	79.00
2	400	35.1	140	0.5	0.0071	70.00
3	800	32.9	263	0.25	0.0038	65.75
4	1200	26.8	322	0.166	0.0031	53.67
5	1600	26.2	426	0.125	0.0023	53.25
6	2000	22.1	442	0.100	0.0023	44.20
7	2400	19.2	461	0.083	0.0022	38.42
8	2800	17.6	493	0.071	0.0020	35.21
9	3200	17.3	554	0.063	0.0018	34.63
10	3600	14.5	522	0.056	0.0019	29.00
11	4000	13.6	544	0.050	0.0018	27.20
12	6000	12.7	762	0.033	0.0013	25.15
13	7600	10	760	0.026	0.0013	19.76

1 - 反应中以 nm (纳摩尔) 计的 AMPA 底物浓度

2 - 根据相对于样品中剩余的 [¹⁴C]-AMPA 量的生成的 N-乙酰-[¹⁴C]-AMPA 的百分比测量的转化百分比

3 - 以单位计的每 mg 蛋白每分钟转化为 N-乙酰-AMPA 的 AMPA (nm) 的酶速度

根据表 8 的 1/V 对 1/S 数据的 Lineweaver-Burk 图表明, PhnO 对于作为底物的 AMPA 的表观 K_m 约为 9 mM, 表观 V_{max} 约为 824 U/mg 蛋白。

- 5 在相似的实验中测定 PhnO 对于底物乙酰-CoA 的表观 K_m。在获得终点动力学的几次尝试后, 确定在约 30 mM 的 AMPA 浓度和约 1-10 ng 的酶量下, 转化数太低以致于不可靠。用氘标记的乙酰-CoA 尝试一种替代的方法。所述标记的比活比用 [¹⁴C] 约高 40 倍, 达到灵敏度增加, 允许测定 PhnO 对于乙酰-CoA 的表观 K_m。 [³H]-乙酰-CoA (Amersham, Inc.) 的比活是 360 mCi/mg 或 250 μCi/ml。通过弱离子交换 HPLC 色谱, 监测由 PhnO 介导的从 [³H]-乙酰-CoA 至 [³H]-乙酰-AMPA 的酰基转移作用, 调节乙酰-CoA 和乙酰-AMPA 的保留时间, 使得这些化合物间隔约 3 分钟。通过将 KH₂PO₄ 缓冲液 (pH 5.5) 的浓度调至 40 mM, 并且在 AX100 弱离子交换柱上的流速为 1 ml/分钟, 而达到这一

点。使每种样品与 PhnO 和 30 mM AMPA 于 37 °C 反应 5 分钟，用 100 mM NaOAc pH 4.4 的乙醇溶液猝灭，然后通过 HPLC 分析。在每个反应中 [³H]-乙酰-CoA 底物范围为 25 μM 至 1.3 mM，每个反应中还有约 5 ng PhnO、50 mM Tris pH 7.5、1 mM MnCl₂、1 mM MgCl₂ 和 30 mM AMPA。通过 HPLC 分析样品，以测定 N-[³H]-乙酰-AMPA 的产生量和 [³H]-乙酰-CoA 的剩余量。这些反应的分析条件和结果示于表 9 中。

表 9.

乙酰-CoA 供体底物的 PhnO 酶促动力学

样品号	[乙酰-CoA] ¹	速度 ²	1/[S] ³	1/V ⁴	V/S ⁵
1	25	34	0.0400	0.0294	1.3600
2	50	66	0.0200	0.0152	1.3200
3	75	94	0.0133	0.0106	1.2533
4	100	125	0.0100	0.0080	1.2500
5	125	150	0.0080	0.0067	1.2000
6	150	173	0.0066	0.0058	1.1533
7	175	193	0.0057	0.0052	1.1029
8	200	219	0.0050	0.0046	1.0950
9	225	240	0.0044	0.0042	1.0667
10	250	259	0.0040	0.0039	1.0360
11	375	339	0.0027	0.0030	0.9040
12	390	287	0.0026	0.0035	0.7359
13	520	331	0.0019	0.0030	0.6365
14	650	352	0.0015	0.0028	0.5415
15	780	372	0.0013	0.0027	0.4769
16	910	397	0.0011	0.0025	0.4363
17	1040	411	0.0009	0.0024	0.3952
18	1170	425	0.0008	0.0024	0.3632
19	1300	434	0.0007	0.0023	0.3338

1 - 以 μM 单位计的底物浓度

2 - 根据每单位时间 [³H]-乙酰-AMPA 中的 [³H] 掺入量测定的酶速度

3 - 底物浓度的倒数

4 - 速度倒数

5 - 速度与底物浓度之比

10

根据表 9 的 1/V 对 1/S 数据的 Linweaver-Burk 图表明，PhnO 对于作为底物的乙酰-CoA 的表观 K_m 在 375-390 μM 之间，表观 V_{max} 约为 824 U/mg 蛋白。

用从 pMON15032 表达 *phnO* 的大肠杆菌的粗制裂解液中的酶，测定所述 PhnO 酶活性的大致 pH 范围。测定该酶在一定 pH 值范围内由含乙酰-CoA 和 AMPA 的混合物产生 N-乙酰 AMPA 的能力。反应在 MES/MOPS/Tricine 缓冲液中进行，该缓冲液平衡至 4.5-9.0 的 pH 值，实际 pH 值范围为 5.2-9.0。简而言之，将 95 微升合适的缓冲液与 100 微升如实施例 4 中所述的 2X 分析混合物和 5 微升含约 400 ng/μl 的 PhnO 蛋白的脱盐大肠杆菌裂解液混合。反应物于 37 °C 温育 5 分钟，用 100 mM NaOAc pH 4.4 的乙醇溶液猝灭，然后如实施例 4 中所述通过 HPLC 进行分析。结果示于表 10 中。

10

表 10.

PhnO 酶 pH 分布型

缓冲液 pH	¹ 模拟反应 pH	² %转化	N-乙酰 CoA (nmole)	³ 速度 (nmole/min/μg)
5.0	5.23	3.7	222	22.2
5.5	5.62	3.9	234	23.4
6.0	5.92	4.2	252	25.2
6.5	6.47	13.3	798	79.8
7.0	7.0	27.0	1620	162.0
7.5	7.48	32.0	1920	192.0
8.0	8.05	34.3	2058	205.8
8.5	8.46	33.5	2010	201.0
9.0	9.0	33.9	2034	203.4

1 - 表明对于给定的每种初始缓冲液 pH 值混合所有试剂后真正的 pH 值

2 - 如表 9 中 Km 和 Vmax 一样进行测定

3 - 如表 9 中 Vmax 一样进行测定

结果表明，采用 AMPA 和乙酰-CoA 作为底物的最佳 PhnO 转酰基酶活性的 pH 约为 pH 8.0。然而，在约 6.5 至至少 9.0 的 pH 范围内 PhnO 利用乙酰-CoA 作为乙酰供体，有效地将 AMPA 转化为 N-乙酰-AMPA。

15

用纯化的 PhnO 蛋白进行另外的实验，以进一步表征该酶对于酰基-CoA 酰基供体化合物的底物优选范围。本文已经确定，至少一种底物酰基供体或离去基团可以是二碳酸化合物，例如化合物乙酰-CoA 中的乙酰部分。不知道当与 PhnO 转酰基酶和作为酰基受体分子的 AMPA 反应时，由不同碳链长度组成的什么样范围的酰基分子能够或可能作为来自酰基-CoA 酰基供体的离去基团起作用。因此，开发了与实施例 2 中所述的分析类似的 HPLC 分析，以测定该酶将酰基基团从酰基-CoA 酰基供体转移至 [¹⁴C]-AMPA 的能力范围。

在于 37 °C 进行萘啶酮酸诱导 3 小时后，从 1 升从 pMON15032 表达重组 *phnO* 基因的大肠杆菌 JM101 的 Luria Bertani 肉汤培养物中纯化 PhnO。通过离心收获细胞，并将其再悬浮于 40 ml 冷 Tris 缓冲液(0.1 M Tris-HCl pH 8)中并置于冰上。使细胞悬浮液成为 1 mM DTT 和 0.5 mM PMSF。以 1,100 psi 通过预冷的弗氏压碎器 2 次，将悬浮液裂解，于 4 °C 以 12,000 g (在 Sorvall SA600 转子中 10,000 rpm)离心 40 分钟，然后置于冰上。将澄清的上清液倾至新的 15 ml 聚丙烯管中。将样品再次分为 2 个等份，并于 -80 °C 维持直至进一步用于 PhnO 蛋白的纯化。用以上实施例 2 中所述的 HPLC 法，分析 20 微升所述可溶性部分的酶活性，只是在加入酸终止分析后，将样品贮存于 -80 °C。按照生产商的说明制备 Sephacryl S200 柱，并用含 20 mM Tris pH 8.0 和 0.5 mM MgCl₂ 的溶液平衡。在冰上解冻后，将全部总可溶性提取物铺在柱床顶部。从柱流出液收集 40 个 9 ml 的流分，在 15% SDS-PAGE 凝胶上分离后，采用抗 PhnO 抗血清通过蛋白质印迹法分析 30 微升的每个流分。此外，用实施例 2 中所述的方法，分析 30 微升的每个流分的 AMPA 酰基转移酶活性。合并表现出酰基转移酶活性并且对应于阳性蛋白质印迹数据的样品。这些样品由该实施例中的流分 7 至流分 19 代表，将其混合到 100 ml 体积中，分配到各含有 10 ml 体积的 10 个管(ten 10 tubes)中，并贮存于 -80 °C 以备进一步的使用。

用阴离子交换色谱测定离开在 Sephacryl S200 分离期间共洗脱的

其它污染蛋白的 PhnO 的洗脱分布图。将来自混合 PhnO 活性流分的一个管在冰上解冻，并注射到用缓冲液 A (1 升的 20 mM Tris-HCl pH 8.0 Mili-Q 蒸馏去离子水)和 B (1 升 20 mM Tris-HCl pH 8.0, 1 M NaCl)预平衡的 5/5 Mono-Q 柱中。将含有 PhnO 活性蛋白的样品注射到该柱中，并收集 1 毫升的流分。将含 PhnO 样品加样后，该柱以 1.8 ml/分钟缓冲液 A 的流速洗涤 5 分钟。5 分钟时，将缓冲液 B 以 0.5 ml/min 加入流动体积中达 4 分钟。缓冲液 B 在 10 分钟时直线上升至流动体积的 22%，在 12 分钟时上升至 30%，于 13 分钟时上升至 36%，于 14 分钟时上升至 41%，于 15 分钟时上升至 46%，于 16 分钟时上升至 74%，于 16-22 分钟时上升至 100%，此时终止缓冲液 B 流动，缓冲液 A 以 100%再起始，以平衡该柱。对来自从 Mono-Q 柱收集的各个流分的 10 微升体积，如实施例 2 中所述通过蛋白质印迹法分析转酰基酶活性。合并表现出阳性 AMPA 酰基转移酶活性并与蛋白质印迹数据相关的流分，并将其作为纯化的蛋白样品维持。用该纯化的 PhnO 蛋白的样品，测定酶的酰基供体底物的专一性。

如下制备酶反应物。100 微升反应物由 50 mM Tris-HCl pH 8.0、1 mM MgCl₂、3 微升 1.3 mM [¹⁴C]-AMPA (115,392 dpm/μl)、0.1 mM 或 1 mM 酰基-CoA 酰基供体和 2.5 微升纯化的酶样品组成。制备分析预混合物，在每个 100 微升的反应物中使用来自所述预混合物的 45 微升。该 45 微升预混合物样品由 40 微升蒸馏去离子水、2 微升 50 mM MgCl₂ 和 3 微升 1.3 mM [¹⁴C]-AMPA (115,392 dpm/μl)组成。通过在微量离心管中将 40 微升 125 mM Tris-HCl pH 8.0、2.5 微升蛋白样品和 10 微升酰基-CoA 酰基供体化合物于室温下混合，起始反应。每种酰基-CoA 酰基供体化合物作为 1 mM、5 mM 或 10 mM 原料储液制备。然后将每个管与 45 微升含有 [¹⁴C]-AMPA 受体底物的分析预混合物混合，轻轻混合，然后转移至 30 °C 水浴中达 5 分钟。加入 4 微升 1M HCl 终止每个反应，进行涡旋混合，并置于冰上或贮藏于 -20 °C 直至通过 HPLC 分析 [¹⁴C]-AMPA 或相关化合物的存在。

5 进行HPLC分析采用: 配有481 波长最大值UV检测器和闪烁泵、Phenomenex PHENOSPHERE 5 微升 80 埃 SAX-二氧化硅 HPLC 柱(250X4.6 mm, 3500 PSI 最大压力)的 Waters 510 双泵 HPLC 系统; 缓冲液 A, 由 5 mM KH₂PO₄、 2%甲醇组成, 并用 H₃PO₄调至 pH 2.0; 缓冲液 B, 由 200 mM KH₂PO₄、 2%甲醇组成, 并用 H₃PO₄调至 pH 2.0; 和 HAZARD Atomflow (Packard), 含有 64% 1,2,4-三甲基苯、 7.5%磺基琥珀酸 dicotyl 钠、 3.5%磺基琥珀酸二戊酯钠和 6%聚氧乙烯(4)十二烷基醚。用于每种样品分析的 HPLC 梯度条件与实施例 2 中描述的条件相似, 有微小的变化。表 11 中提供了流速。

10

表 11.

HPLC 梯度条件				
时间(min)	流动 (ml/min)	%A	%B	流速 ¹
0.0	1	100	0	3
2.0	1	100	0	3
5.0	1	50	50	3
15.0	1	0	100	3
17.0	1	0	100	3
17.3	1	100	0	3
21.0	1	100	0	3
21.3	0.1	100	0	3

1 - 闪烁液流速, 以 ml/min 计

15 如上所述制备酰基-CoA 酰基供体化合物的储液, 将它们列于如下: 乙酰-CoA 钠、正丙酰-CoA 锂、戊二酰-CoA 锂、甲基丙二酸单酰 CoA 锂、巴豆酰-CoA 锂、异丁酰-CoA 锂、琥珀酰-CoA 钠、甲基巴豆酰-CoA 锂、正戊酰-CoA 锂和去磺基-CoA 锂。所有化合物均得自 Sigma Chemical Company, St. Louis, MO。通过测定转化为某些其它^[14C]-化合物(例如 N-乙酰-^[14C]-AMPA)的^[14C]-AMPA HPLC 色谱峰面积的百分比、与反应期间产生的 N-乙酰-^[14C]-AMPA 的量, 测定所述纯化的酶

20

将 CoA 连接的酰基部分转移至 $[^{14}\text{C}]$ -AMPA 的活性百分比, 其中 $[^{14}\text{C}]$ -AMPA 与 1 mM 乙酰-CoA 是确定为 100%参比的 PhnO 的底物。结果示于表 12 中。

5

表 12.

AMPA 转酰基酶对于酰基-CoA 酰基供体底物的酶促效率

酰基-CoA 酰基供体	$[^{14}\text{C}]$ -AMPA %转化 ¹	%活性
乙酰-CoA 0.1 mM	79.2	79.2
乙酰-CoA 0.5 mM	98.7	98.7
乙酰-CoA 1 mM	100.00	100.00
丙酰-CoA 0.1 mM	78.2	78.2
丙酰-CoA 0.5 mM	97.8	97.8
丙酰-CoA 1 mM	100.00	100.00
戊二酰-CoA 0.1 mM	0.81	0.81
戊二酰-CoA 0.5 mM	0.00	0.00
戊二酰-CoA 1 mM	0.57	0.57
甲基丙二酸单酰-CoA 0.1 mM	1.11	1.11
甲基丙二酸单酰-CoA 0.5 mM	2.08	2.08
甲基丙二酸单酰-CoA 1 mM	2.21	2.21
巴豆酰-CoA 0.1 mM	0.80	0.80
巴豆酰-CoA 0.5 mM	0.00	0.00
巴豆酰-CoA 1 mM	0.00	0.00
异丁酰-CoA 0.1 mM	2.10	2.10
异丁酰-CoA 0.5 mM	0.20	0.20
异丁酰-CoA 1 mM	0.00	0.00
琥珀酰-CoA 0.1 mM	5.06	5.06
琥珀酰-CoA 0.5 mM	3.38	3.38
琥珀酰-CoA 1 mM	1.56	1.56
甲基巴豆酰-CoA 0.1 mM	0.00	0.00
甲基巴豆酰-CoA 0.5 mM	0.00	0.00
甲基巴豆酰-CoA 1 mM	0.99	0.99
戊酰-CoA 0.1 mM	0.24	0.24
戊酰-CoA 0.5 mM	0.00	0.00
戊酰-CoA 1 mM	0.33	0.33
去磺基-CoA 0.1 mM	0.95	0.95
去磺基-CoA 0.5 mM	1.25	1.25
去磺基-CoA 1 mM	0.52	0.52

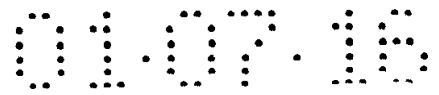
1 - 转化为某些其它 $[^{14}\text{C}]$ -化合物例如 N-乙酰- $[^{14}\text{C}]$ -AMPA 的 $[^{14}\text{C}]$ -AMPA HPLC 色谱峰面积的百分比, 与反应期间产生的 N-乙酰- $[^{14}\text{C}]$ -AMPA 的量, 其中 $[^{14}\text{C}]$ -AMPA 与 1 mM 乙酰-CoA 是确定为 100%参比的 PhnO 的底物

5 这些结果表明, PhnO 酶能够有效地利用酰基-CoA 连接的化合物, 所述化合物具有一个酰基, 并且碳链长度不超过 3 个以进行对 AMPA 进行酰基转移。碳链长度比丙酰长并且不宽或庞大的其它化合物, 例如甲基丙二酸单酰-CoA、异丁酰-CoA 和琥珀酰-CoA 化合物也是有效的酰基-CoA 酰基供体, 但酶促效率较低。

实施例 7

该实施例描述了 AMPA 酰基转移酶蛋白的体外表达以及将其定向到分离的叶绿体中。

10 许多叶绿体定位蛋白由核基因作为前体表达, 并由叶绿体转运肽(CTP)导向叶绿体。CTP 在涉及定向蛋白输入到叶绿体中的步骤期间被除去。这样的叶绿体蛋白的实例包括核酮糖-1,5-二磷酸羧化酶(RUBISCO)的小亚基(SSU)、5-烯醇-丙酮酰莽草酸-3-磷酸(EPSPS)、铁氧还蛋白、铁氧还蛋白氧化还原酶、集光复合体蛋白 I 和蛋白 II 以及
15 硫氧还蛋白 F。已经在体内和体外证明, 利用具有 CTP 的融合体可以将非叶绿体蛋白导向叶绿体, 并且 CTP 序列足以将蛋白导向叶绿体(Della-Cioppa 等, 1987)。5-烯醇丙酮酰莽草酸-3-磷酸合成酶(EPSPS)定位在叶绿体中, 并且在植物中是草甘膦的靶。已经发现, 将草甘膦氧化还原酶导向叶绿体, 为植物提供对草甘膦的耐性, 尽管定位至胞
20 质的 GOX 也能够提供这种耐性。一般而言, 重组 GOX 酶定位至叶绿体。GOX 介导的草甘膦代谢产生 AMPA, 已经表明 AMPA 具有植物毒性。本文中已经表明, PhnO 能够将 AMPA N-酰化, 并且 N-乙酰-AMPA 没有植物毒性。因此, 可能有必要将植物中的 AMPA 失活。这假定 AMPA 酰基转移酶可以在植物中作为活性酶表达, 并且这样的酰
25 基转移酶能够被输入到叶绿体中并保留酶活性。鉴于实施例 1 中描述的 AMPA 的植物毒性, 将 AMPA 酰基转移酶基因导入植物表达载体中, 以测试在植物中的表达。另外, 也对酰基转移酶输入叶绿体进行了测试。



将编码叶绿体引导肽的 DNA 序列连接至编码 AMPA 酰基转移酶的 DNA 序列的 5', 并且与其符合读框地连接。用 *Bgl*III 和 *Nco*I 限制性内切核酸酶, 从 pMON17058 切下编码拟南芥核酮糖-1-二磷酸羧化酶小亚基叶绿体转运肽(CTP, SEQ ID NO:9)的 DNA 序列, 并将其插入 pMON15028 中的互补限制位点, 产生 pMON15029, 使得所述 CTP 编码序列连接 pMON15028 中的 *phnO* 编码序列的 5' 并与其符合读框地连接。所产生的 pMON15029 中的嵌合 *phnO* 基因能够产生叶绿体定向 PhnO 蛋白。将含有得自 pMON15029 的所述 CTP-PhnO 编码序列的 *Eco*RI-*Bgl*III DNA 盒 SEQ ID NO:11 插入 pBlueScript KS(-)中的 *Eco*RI 和 *Bam*HI 位点中, 产生 pMON15036。pMON15036 中的 CTP-PhnO 编码序列可以在体外转录/翻译系统中从噬菌体 T3 启动子表达。构建一个相似的植物瞬时表达质粒 pMON15035, 但没有叶绿体引导序列。从 pMON15028 切下仅含有 *phnO* 编码序列的 *Eco*RI-*Bgl*III DNA 片段, 并将其插入 pBlueScript KS(+)中的 *Eco*RI 和 *Bam*HI 位点中, 使得 PhnO 在体外转录/翻译系统中可以从噬菌体 T7 启动子产生。从 pMON15028 切下编码 PhnO 的 *Nco*I-*Eco*RI DNA 序列, 并将其插入 pMON17061, 产生 pMON15032。pMON15032 可供用于从大肠杆菌 *recA* 启动子表达 *phnO*。从 pMON15028 切下编码 PhnO 的 *Bgl*III-*Eco*RI DNA 片段, 并将其插入 pBlueScript KS(-), 产生 pMON15033。pMON15033 可供用于从大肠杆菌 *lac* 启动子表达 *phnO*。从 pMON15029 切下编码 CTP-PhnO 的 *Bgl*III-*Eco*RI DNA 片段, 并将其插入 pBlueScript KS(-)中的匹配位点, 可供用于从 pMON15034 的大肠杆菌 *lac* 启动子表达叶绿体定向 PhnO 蛋白。

将 pMON15032、pMON15033 和 pMON15034 导入大肠杆菌 JM101 中。如上所述将培养物培养和诱导, 只是通过加入 0.1 M NaOH 中的 50 ppm 萘啶酮酸诱导来自含 pMON15032 的细胞的表达。由每种培养物制备澄清的裂解液, 并如上所述进行 AMPA 酰基转移酶分析, 以便测定 AMPA 酰基转移酶活性的存在。所有裂解液均含有相当大量

的高于对照水平的酰基转移酶活性。更为重要的是，从 pMON15034 表达的 CTP-PhnO 肽(SEQ ID NO:12)看来保留了全部酰基转移酶活性。

pMON15035 (PhnO)和 pMON15036 (CTP-PhnO)在体外用来产生
5 [35S]-甲硫氨酸标记的 PhnO 蛋白，以用于叶绿体输入分析。简而言之，
在 *Short Protocols In Molecular Biology*, 第3版, (Ausubel 等编辑, Wiley
& Sons Pub., (1995), 其通过引用结合到本文中)中描述了用于体外转
录和翻译的方法。在 100 微升反应中，用 *Hind*III 限制性内切核酸酶将
10 约 20 微克质粒 DNA 完全消化。在体外转录反应中用 20 微升所述质粒
消化物或约 4 微克线性化质粒 DNA 产生 mRNA，用于在后来的翻译
反应中产生 PhnO 或 CTP-PhnO 蛋白产物。转录反应物由以下物质组
成：20 微升线性化质粒 DNA、20 微升 5X 转录缓冲液(200 mM
15 TrisHCl pH 8.0, 40 mM MgCl₂, 10 mM 亚精胺和 250 mM NaCl)、20 微
升 5X 核糖核苷三磷酸混合物(ATP、CTP、UTP 各 5 mM, 5 mM 二鸟
苷三磷酸(G-5'ppp5'-G)TP, 5 mM GTP)、10 微升 0.1 M 二硫苏糖醇
(DTT)、10 微升 RNasin™ (一种胰腺核糖核酸酶抑制剂混合物，得自
Promega)、4 微升 RNA 聚合酶(T7 或 T3, New England Biolabs, Inc.)和
蒸馏去离子水，至 100 微升。每种反应物于 37 °C 温育 1 小时。在 1.4%
琼脂糖甲醛凝胶上分析 4.5 微升每种反应物，以确保每种反应物产生
20 合适的 RNA 模板用于以后的翻译步骤。

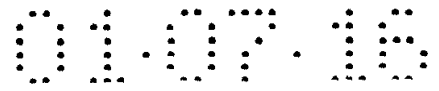
用 20 微升所述转录反应物，产生 [35S]-甲硫氨酸标记的 PhnO 蛋
白，以用于叶绿体输入分析。简而言之，将 RNA 与 6 微升无甲硫氨酸
的氨基酸混合物水溶液、15 微升 [35S]-甲硫氨酸(1400 Ci/mmol,
Amersham)和 200 微升兔网织红细胞裂解液混合。将这些反应物于 37
25 °C 温育 2 小时，并置于干冰上进行贮存。在 15% SDS-PAGE 凝胶上分
析 10 微升每种反应的样品。将凝胶真空干燥，并直接置于 KODAK™
X-O-MAT™ 胶片的乳胶一侧上，进行放射自显影。结果表明，每种质
粒产生足够量的 PhnO (pMON15035)和 CTP-PhnO (pMON15036)预测

分子量的相应的肽，其产生量足以在输入分析中测试叶绿体中的摄入。

按照 Edelman 等, *Methods in Chloroplast Molecular Biology*, Elsevier Biomedical Press, 第 86 章(1982)所述, 从一个去脉纹的 (deveined)直立莴苣的叶球(head)分离完整的叶绿体。制备 1 升研磨缓冲液(GR 缓冲液)储液(2 mM NaEDTA, 1 mM MgCl₂, 1 mM MnCl₂, 50 mM HEPES-KOH pH 7.5 和 0.33 mM 山梨醇)。恰好在使用之前, 将 890 mg 抗坏血酸加入 900 ml GR 缓冲液储液中。将一个撕开的去脉纹直立莴苣叶球与 900 ml GR 缓冲液混合, 并且通过在韦林氏搅切器中以高速混合 3 次, 每次 3 秒, 将其捣碎(emacerate)。浆液通过 4 层 Miracloth 过滤, 滤液在 SORVALL™ GS-3 转子中于 4 °C 以 5,000 RPM 离心 10 分钟。倾去上清液, 用玻璃棒将沉淀再悬浮于 4 毫升 GR 缓冲液中。通过 Percoll 梯度离心, 分离叶绿体。通过将 16 ml PBF-Percoll 与 4 ml 5X 缓冲液(10 mM EDTA, 5 mM MgCl₂, 5 mM MnCl₂, 250 mM HEPES-KOH, 30 克山梨醇, 490 mg 抗坏血酸钠, 85.5 mg 谷胱甘肽, 用 ddH₂O 调至 100 ml)混合, 制备 80% Percoll。通过将 8 ml PBF-Percoll 与 4 ml 5X 缓冲液和 8 ml ddH₂O 混合, 制备 40% Percoll 溶液。通过将 10 ml 40% Percoll 铺在 10 ml 80% Percoll 上, 在 30 ml Corex 管中制备 Percoll 梯度。通过将再悬浮的叶绿体铺在 percoll 梯度上, 并打开制动器, 在 SS-34 SORVALL™悬挂式转子中以 9,500 RPM 于 4 °C 离心 10 分钟, 分离叶绿体。破裂的叶绿体保持在上层中, 将其吸去。完整的叶绿体位于 40/80% Percoll 梯度的界面, 将其取出到新的 30 ml COREX™管中。分离的叶绿体用 GR 缓冲液洗涤 2 次, 并且在每次洗涤后, 关闭制动器, 在 SS-34 转子中于 4 °C 以 6,000 RPM 离心 10 分钟, 收集叶绿体。通过用玻璃棒轻轻搅拌, 将经分离、洗涤的叶绿体再悬浮于 1 ml 无菌 50 mM HEPES-KOH pH 7.7、330 mM 山梨醇中, 然后测定浆液的叶绿素浓度。将 5 ml 80%丙酮溶液加入 20 微升所述叶绿体浆液中, 轻轻地进行涡旋混合。所产生的混合物通过 Whatman™ #1 滤纸过滤到培养管中。

于 645nm 和 663nm 对 80% 的丙酮空白测定滤液的吸光度。按照公式 #1: $[\text{叶绿素 } \mu\text{g/ml}] = [A_{645} + [A_{663} * (8.02)]]$, 通过测量以 $\mu\text{g/ml}$ 计的叶绿素的量并乘以再悬浮叶绿体的体积(公式#2), 该体积在该实施例中为 5 ml。计算以微克计的叶绿素浓度。因此在所测定的样品中以 $\mu\text{g/ml}$ 计的叶绿素浓度等于在公式#2 中测定的数值除以样品测量体积, 该体积在该实施例中为 20 μl 。在该实施例中, A_{645} 测定为 0.496, 而 A_{663} 测定为 1.0814。因此, 在所测量的样品中叶绿素的浓度是 4.67 $\mu\text{g}/\mu\text{l}$ 。用 HEPES-KOH pH 7.7、330 mM 山梨醇溶液将叶绿体浆液中叶绿素的浓度调至 4.0 $\mu\text{g}/\mu\text{l}$, 并且将所产生的叶绿体悬浮液贮存在黑暗冰上。

典型的 300 微升摄入实验含有 5 mM ATP、8.3 mM 未标记的甲硫氨酸、322 mM 山梨醇、58.3 mM HEPES-KOH (pH 8.0)、50 微升网织红细胞裂解液翻译产物和完整的叶绿体(约 200 微克叶绿素)。将摄入混合物于室温下, 在直接置于采用 150 瓦灯泡、设定在最大光度的光纤发光体前面的 10X75 mm 玻璃试管中轻轻震荡。在 0、5、10 和 15 分钟时取出 2 个独立的 70 微升每种摄入混合物样品。用离心法, 一个样品在 150 微升聚乙烯管中的 100 微升硅酮油梯度上, 以 11,000 X g 离心 30 秒, 然后立即在干冰中冷冻。在这些条件下, 完整的叶绿体在硅酮油层下形成沉淀, 而含有网织红细胞裂解液的温育介质仍飘浮在界面表面上。另一样品用蛋白酶(1/10 体积或 0.25 mg/ml 胰蛋白酶和胰凝乳蛋白酶混合物各 7 微升)在冰上处理 30 分钟, 然后进行硅酮油分离并在干冰上冷冻。然后将叶绿体沉淀再悬浮于 50-100 微升裂解缓冲液(10 mM HEPES-KOH pH 7.5, 1 mM PMSF, 1 mM 苯甲脒, 5 mM ϵ -氨基-正己酸和 30 微克/ml 抑酶肽), 并以 15,000 X g 离心 20 分钟, 以沉淀类囊体膜。来自该离心的澄清的上清液(基质蛋白)和 1 等份来自每种摄入实验的网织红细胞裂解液温育介质与等体积的 2X SDS-PAGE 样品缓冲液混合, 并在 15% SDS-PAGE 凝胶上分析, 干燥并如上所述曝光于胶片。暴露于 $[^{35}\text{S}]$ -甲硫氨酸标记的 CTP-PhnO 的叶绿体含有 $[^{35}\text{S}]$ -标记的蛋白, 其大小与预测的 CTP-加工型式的 PhnO 一致, 而暴露于甲



硫氨酸标记的 PhnO 的叶绿体缺乏标记蛋白。输入到叶绿体中的标记蛋白也抗蛋白酶。这些结果表明，PhnO 当与质体引导肽序列融合时，可以被导向叶绿体。

5 实施例 8

该实施例描述用 AMPA 酰基转移酶转化的植物的鉴定和表征。

采用本领域众所周知的多种植物转化方法中的任一种，已经成功地转化了种类繁多的植物种。特别是，根癌土壤杆菌介导的植物转化是目前使用的优选方法，然而，基因枪法(ballistic methods)通过微粒轰击增加裸露 DNA 直接至植物细胞的传递，基因枪法在产生重组转化植物方面也是非常有效的。另外，涉及应用脂质体、电穿孔、增加游离 DNA 摄入的化学药品以及应用病毒或花粉转化的方法是可以用来将本发明的 DNA 构建体插入植物细胞中的替代方法。可以通过实施本发明转化的植物包括但不限于：玉米、小麦、棉花、水稻、大豆、甜菜、
10 canola、亚麻、大麦、油料种子油菜、向日葵、马铃薯、烟草、番茄、苜蓿、莴苣、苹果、杨树、松树、桉树、金合欢、杨树、美国枫香、radiata pine、火炬松、云杉、柚木、苜蓿、三叶草和其它饲料作物、草坪草、油棕、甘蔗、香蕉、咖啡、茶树、可可、苹果、胡桃、杏仁、葡萄、花生、pulses、矮牵牛、万寿菊、长春花、秋海棠、天竺葵、
15 三色堇、凤仙花、燕麦、高粱和小米。用于本发明的 DNA 分子可以是天然或天然存在的基因，或者是由有用的多核苷酸序列构建的嵌合基因，所述有用的多核苷酸序列包括启动子、增强子、翻译或非翻译前导序列、编码信号肽的序列、编码转运肽的序列、结构基因、结构基因的融合体、终止子、内含子、反向重复序列或同向重复序列、接头和聚腺苷酸化序列。在本发明中设想的 DNA 序列包括单链和双链多核苷酸序列、线性序列和共价闭环多核苷酸序列、质粒、杆粒、粘粒、细菌人工染色体(BAC)、酵母人工染色体(YAC)以及病毒 DNA 和 RNA
20 序列。土壤杆菌介导的植物转化的考虑中，合适的植物转化载体包括

衍生自根癌土壤杆菌的 Ti 质粒的那些载体以及例如 Herrera-Estrella (1983), Bevan (1984), Klee (1985)和 EPO 公布号 125,516 (Schilperoort 等)中公开的那些载体。除衍生自土壤杆菌的 Ti 质粒或毛根诱导(Ri)质粒的植物转化载体外,如上所述的替代方法可以用来将本发明的 DNA
5 构建体插入植物细胞中。

用于植物转化的质粒一般由在别处、特别是在美国专利第 5,463,175 号(Barry 等, 1995) (其通过引用结合到本文中)中已经描述的载体进行构建。构建质粒, 并采用 *Tn7* 氨基糖苷腺苷酰基转移酶抗性 (*aad* 基因, 通常称为链霉素/壮观霉素或 Spc/Str 抗性)在大肠杆菌中维持, *Tn7* 氨基糖苷腺苷酰基转移酶抗性也是在土壤杆菌中选择并维持
10 的决定因子。本领域熟知用于大肠杆菌中的其它质粒的维持和选择标记也可以使用, 它们基本上由新霉素磷酸转移酶、庆大霉素乙酰转移酶和 β 内酰胺酶基因组成, 它们可以单独使用或组合存在于单一复制子或载体上。质粒一般含有衍生自广宿主范围质粒 RK2 的复制起点 *oriV* 以及衍生自质粒 pBR322 的 *ori322* 和 *bom* (分别是用于在大肠杆菌中维持的复制起点以及接合转移的流动性(mobility)的基础)序列。
15

将编码 AMPA 酰基转移酶的 *phnO* 基因插入植物转化载体的表达盒中。这些盒以 5' 至 3' 的顺序一般顺序含有以下元件: 包含一个植物有效(operable)启动子的序列、一个编码叶绿体或质体转运肽的序列、
20 包含在多接头内的一个或多个克隆位点和一个植物功能性 3' 非翻译区。表达盒通常被构建为含有邻接所述盒结构域的特有的限制位点, 使得可以从一个质粒切下所述完整的盒, 并将其置于其它相似构建的质粒载体中。优选由 8 个碱基对识别序列组成的限制位点, 本发明中的大多数盒至少在一端邻接一个 *NotI* 限制性内切核酸酶识别位点。优选的启动子是玄参花叶病毒启动子 P-FMV (Gowda 等, 1989)、花椰菜花叶病毒 35S 启动子 CaMV 35S (Odell 等, 1985)或增强的 CaMV 35S 启动子(美国专利第 5,196,525; Kay 等, 1987)。在文献中已经描述了在
25 植物细胞中有活性的许多其它启动子。这样的启动子可以得自植物或

植物病毒，包括但不限于：胭脂碱合酶(NOS)和章鱼碱合酶(OCS)启动子(它们载于一般在根癌土壤杆菌毒性或无毒性菌株中发现的肿瘤诱导质粒上)、花椰菜花叶病毒(CaMV) 19S 启动子、鸭趾草黄斑驳病毒(comalina yellow mottle virus)启动子、甘蔗杆状 DNA 病毒启动子、花生褪绿条死病病毒启动子、水稻肌动蛋白启动子和光诱导型核酮糖 1,5-二磷酸羧化酶小亚基启动子(ssRUBISCO)。这些启动子可以用来产生各种类型的 DNA 构建体，所述 DNA 构建体可用于在植物中的基因表达(参见例如 Barry 等，美国专利第 5,463,175 号)。由于其组成型性质而设想的特别理想的启动子是花椰菜花叶病毒 35S (CaMV35S)启动子和玄参花叶病毒 35S (FMV35S)启动子，先前已经表明，它们在大多数植物器官中产生高水平的表达。也设想了可以指导组织特异性表达或定向表达的其它启动子，例如在组织如叶片、分生组织、花、果实和有生殖特性的器官中表达。另外，也设计了嵌合启动子。也使用了胭脂碱合酶基因(NOS 3')和豌豆核酮糖二磷酸羧化酶合酶 E9 基因(E9 3') 3'非翻译终止和聚腺苷酸化序列。

由 AMPA 酰基转移酶结构基因插入启动子下游并位于编码叶绿体引导肽的序列和 3'非翻译序列之间构成的表达盒，一般存在于植物转化载体上。表达盒一般在所述盒的两端一端邻接胭脂碱型 T-DNA 右边界区，另一端邻接左边界区，这两种边界区衍生自 pTiT37 (Fraley 等，1985)。某些植物转化载体仅含有右边界区，这是起始从土壤杆菌至宿主细胞的 T-DNA 转移所需要的。大多数植物转化载体也含有一个如上所述以及在美国专利第 5,463,175 号中描述的 GOX (草甘膦氧化还原酶)基因。当插入植物基因组中时，由这些载体表达的 GOX 酶一般被定向至叶绿体。

将植物转化载体带动转移到携带非致瘤性(disarmed) Ti 质粒 pTiC58 (pMP90RK) (Koncz 和 Schell, 1986)的 ABI 土壤杆菌属菌株 A208 中。该 Ti 质粒不携带诱导冠瘿形成的 T-DNA 植物激素基因。采用辅助质粒 pRK2013 (Ditta 等，1980)，通过三亲本接合系统，将该植

物载体交配到 ABI 中。另一方面, 通过如 Mattanovich 等描述的电穿孔(*Efficient transformation of Agrobacterium spp. by electroporation*, *Nucleic Acids Res.* (1989), 17(16), 6747, 其通过引用结合到本文中), 可以将所述植物转化质粒导入 ABI 菌株中。当植物组织与 ABI:植物载体接合体(conjugate)一起温育时, 该重组载体通过非致瘤性 pTiC58 质粒编码的 *vir* 功能转移至植物细胞。理想的是, 该重组载体在 T-DNA 右边界区打开, 右边界序列和左边界序列之间的 DNA 被定向转移并且插入宿主植物基因组中, 尽管可以转移并插入完整的重组植物转化载体序列。pTiC58 Ti 质粒不转移至植物细胞, 但保留在土壤杆菌供体中。

可以从已经用功能性 AMPA 酰基转移酶结构基因转化的植物细胞或植物组织再生重组植株。再生步骤方法的选择不是关键, 可得到用于豆科(Leguminosae)(苜蓿、大豆、三叶草等)、伞形科(Umbelliferae)(胡萝卜、芹菜、欧洲防风)、十字花科(Cruciferae)(甘蓝、萝卜、油菜籽(rapeseed)等)、葫芦科(Cucurbitaceae)(甜瓜和黄瓜)、禾本科(Gramineae)(小麦、水稻、玉米等)、茄科(Solanaceae)(马铃薯、烟草、番茄、胡椒)和各种花卉作物(floral crops)宿主的合适的方案。参见例如 Ammirato, 1984; Shimamoto, 1989; Fromm, 1990; 和 Vasil, 1990。也可以在含 AMPA 的培养基上选择已经用 AMPA 酰基转移酶转化的重组植物。通过如实施例 1 所述根据 AMPA 毒性进行筛选, 本领域一般技术人员可以容易地确定用于任何特定宿主的 AMPA 的合适抑制浓度。另一方面, 当将 AMPA 酰基转移酶转化到先前用 GOX 转化并且根据在草甘膦上的生长经过选择的植物中时, 可以用或者 AMPA 或者草甘膦作为选择组分, 以选择表达足够水平的 AMPA 酰基转移酶的转化事件。草甘膦必须以在其它情况下对表达 GOX 并且根据在草甘膦上的生长经过选择的重组植物有抑制作用的水平来应用, 因为 AMPA 水平的增加可能是由于 GOX 介导的草甘膦降解而产生的。在表达重组 GOX 酶的植物中, 已经表明暴露于增加水平的草甘膦诱导叶片的黄化

或褪绿、生长停滞特征和不育。与 GOX 表达协同表达或结合表达的 AMPA 酰基转移酶可以克服这些有害效应。用 AMPA 作为植物转化选择标记来作为草甘膦选择的替代方法也是可能的。

5 烟草

烟草植株用 *phnO* 基因转化。烟草叶圆盘转化法利用来自约 1 月龄的叶片的健康组织。用 10% CLOROXTM 加上表面活性剂将表面消毒 15-20 分钟后，将叶片在无菌水中冲洗 3 次。用无菌纸张钻孔器在叶圆盘上钻孔，将其面向上置于 MS104 培养基(4.3 g/l MS 盐, 30 g/l 蔗糖, 2 ml/l 500X B5 维生素, 0.1 mg/l NAA 和 1.0 mg/l BA)上, 预培养 1 天。然后
10 后用 1:5 稀释的含有主题载体的非致瘤性土壤杆菌属 ABI 的过夜培养物(最终培养物密度于 550 nm 下测定约为 0.6 OD)接种叶圆盘。通过将叶圆盘与培养基一起置于无菌离心管中, 进行接种。30-60 秒后, 将液体排干, 并且在无菌滤纸之间将叶圆盘吸干。然后将叶圆盘面向上
15 置于 MS104 饲养平板上的圆滤纸上, 培养 2-3 天。在该共培养期后, 将叶圆盘转移, 仍面向上转移至含有 MS104 培养基的选择平板。2-3 周后, 愈伤组织形成, 从叶圆盘分离出各个愈伤组织块。当苗大到足以与茎区别时, 将苗干净地从愈伤组织上切下。将苗置于具有选择的无激素生根培养基(MSO: 4.3 g/l MS 盐, 30 g/l 蔗糖和 2 ml/l 500X B5 维
20 生素)。根在 1-2 周内形成。最好在在生根的苗进行任何叶片愈伤组织分析, 同时仍保持无菌。将生根的苗置于土壤中, 维持在高湿度环境(即: 塑料容器或塑料袋)中。通过逐渐将苗暴露于周围湿度条件下来锻炼(harden off)苗。

选择三个烟草转化事件用于进一步分析, 这些事件命名为系
25 33476、36779 和 37235。用 pMON17226 (Barry 等, 美国专利第 5,463,175 号, 1995)产生含有 FMV-CTP-GOX 基因构建体的植物系 33476。系 36779 和 37235 用 pMON17261 产生, 后者是衍生自 pMON17226 的质粒, 它除含有 FMV-CTP-GOX 外, 还含有一个含

FMV-CTP-PhnO 基因序列(SEQ ID NO:11)的 *NotI* 盒。所述 *NotI* 盒如下构建。将由 SEQ ID NO:9 代表的编码 CTP 的序列作为 *BglII-NcoI* 片段从 pMON17058 切下, 将其插入 pMON15028 中, 形成由 SEQ ID NO:11 代表的序列, 在该序列中 CTP 编码序列位于 SEQ ID NO:7 内代表的
5 PhnO 编码序列的上游并且符合其读框。所产生的构建体命名为 pMON15029。将 CTP-PhnO 编码序列在 *BglII-SacI* 片段上从 pMON15029 切下, 并与 pMON17063 片段结合, 产生 pMON15038。用限制性消化拆开 pMON17063, 产生构建 pMON15038 必需的部分。用 *SacI* 和 *HindIII* 消化 pMON17063, 产生其中插入一个启动子片段和
10 CTP-PhnO 序列的载体骨架。在单独的反应中也用 *HindIII* 和 *BglII* 消化 pMON17063, 产生含有 FMV 启动子序列的一个片段。将该启动子片段和 CTP-PhnO 片段在一个反应中与所述载体骨架片段连接在一起, 产生 pMON15038, 后者含有一个带有由 FMV 启动子表达并且下游侧翼为一个 NOS E9 3' 转录终止和聚腺苷酸化序列的叶绿体定向 PhnO 肽
15 的编码序列的 *NotI* 盒。将该 *NotI* 序列从 pMON15038 切下, 并插入 pMON17241 的唯一 *NotI* 位点中, 产生 pMON17261, pMON17261 含有: 从 FMV 启动子表达并且下游侧翼为 E9 3' 序列的叶绿体定向 GOX 编码序列、与 CTP-PhnO 编码序列和表达盒。预期得自该载体的转化事件不仅抗草甘膦, 而且也提供对 AMPA 植物毒性的抗性。对衍生自
20 pMON17261 的系 36779 和 37235 通过 PCR 分析其草甘膦氧化还原酶和 AMPA 酰基转移酶的编码基因的存在, 通过蛋白质印迹分析其 GOX 酶和 PhnO 酶的存在, 以及通过 HPLC 分析其由于 GOX 介导的 [¹⁴C]-草甘膦降解产生的代谢物的存在。

系 33476 作为衍生自 pMON17226 的转化事件获得, 选择其作为
25 “仅 GOX” 对照。系 36779 和 37235 在暴露于草甘膦后表现出不同的表型, 选择其作为在用 pMON17261 转化后产生的草甘膦抗性事件。系 37235 在暴露于草甘膦后变为褪色或黄化, 表型相似于仅 GOX 的系 33476。然而, 系 36779 没有表现出这种褪色效应。从每种这些事件

的叶片组织以及 wt Samsun 烟草叶片提取 DNA，进行 PCR 以测定转化 *phnO* 基因是否存在。

从转化的烟草系分离的基因组 DNA 用作 PCR 反应中的模板 DNA，将反应产物与野生型 Samsun 烟草进行比较。PCR 反应由 50 微升的总体积组成，含有 10X 扩增缓冲液、1.5 mM MgCl₂、各为 1 mM 的脱氧核苷酸混合物、50-100 ng 基因组 DNA、终浓度各为 16.8 pM 的引物和 1.5 单位 AmpliTaq DNA 聚合酶(Cetus/Perkin Elmer)。引物(按定单由 GENOSYS 合成)由 SEQ ID NO:21 和 SEQ ID NO:22 中叙述的序列组成。SEQ ID NO:21 是 20 个碱基对的序列，它能够引发 P2A *phnO* 基因序列(SEQ ID NO:7)的合成，并与该基因中编码序列的前 20 个核苷酸杂交。SEQ ID NO:22 也是 20 个碱基对的序列，但它能够引发从末端编码序列到所述结构编码区内的 *phnO* 基因的合成，并且能够与 *PhnO* 编码序列的末端 20 个核苷酸杂交。扩增条件的组成是：97 °C 1 分钟、60 °C 2 分钟和 72 °C 2 分钟的 3 个循环，然后是 94 °C 1 分钟、60 °C 2 分钟和 72 °C 2 分钟的 37 个循环，然后一般是 4 °C 的浸泡。一般通过 1% TAE 琼脂糖凝胶电泳分析 10 微升样品，以将相关条带与残留的引物分离。当用溴化乙锭染色所述产物凝胶时，通过相对于 *Hind*III 消化的 λ 分子量标记的迁移位置判断约为 432 个碱基对的 *phnO* 基因扩增产物，仅在系 33779 提取物中出现，表明在该系中存在 *phnO* 基因。

在生长室条件下自交后，获得来自 R₀ 转化事件的种子。将 R₀ 种子晒干并种植以产生 R₁ 后代。通过在每个叶脉上点上 2 微升样品(50 微升 [¹⁴C]-草甘膦 Na 盐, 517,000 dpm/μg, 0.42 μg/μl 与 10 微升甘油混合)，将 5 叶期 R₁ 后代的源叶片(source leaves)暴露于 [¹⁴C]-草甘膦。根据叶片上叶脉数，每个叶片接受几个点。3 天后，将另外 15 个 2 微升的点加样到每个叶片上。2 周后，将 5 个 2 微升的点加样到每株植株上两个叶片中的每个叶片上。这些叶片是新叶片，而不是最初应用草甘膦的老叶片。在这最后一次加样后 5 天后，从每株植株的两个浸透

(sink)叶片取 300 毫克组织样品。将每株植株的样品在单独的 1 ml 体积的去离子水中匀浆，在微量离心机中以 9,000 RPM 离心，收集水性体积，并贮存在冰上。如实施例 2 中所述，通过 HPLC 分析提取物中^[14C]标记的代谢物的存在。得自系 33476 (GOX)的提取物仅含^[14C]-AMPA。得自系 37235 的提取物含有非代谢的^[14C]-草甘膦以及微量但可测量量的^[14C]-AMPA。在得自系 36779 的提取物中仅观察到 N-乙酰-^[14C]-AMPA。这些结果与 PCR 的数据一致，这表明系 36779 含有至少一个拷贝的 *phnO* 基因。另外，在暴露于草甘膦后在系 36779 中没有褪色效应，这与存在功能性 GOX 和 PhnO 酶并且缺乏可检测的^[14C]-AMPA 一致。

棉花

将重组 *phnO* 基因转化到 Coker 312 品种的棉花(*Gossypium hirsutum* L.)中。用双边界二元质粒载体，通过土壤杆菌介导的植物转化产生草甘膦耐性棉花系，所述双边界二元质粒载体含有以下之一：
(1) *gox*，无色杆菌菌株 LBAA 基因，编码草甘膦代谢酶草甘膦氧化还原酶(GOX)，(2) *gox* 基因和编码 PhnO 的大肠杆菌 *phnO* 基因，或(3) *gox/phnO* 双基因构建体与编码 5-烯醇丙酮酰莽草酸-3-磷酸合酶(EPSPS)的土壤杆菌属菌株 CP4 基因。所有载体均能够在根癌土壤杆菌和大肠杆菌两种宿主中复制，并且含有一个氨基糖苷腺苷酰转移酶基因(*aad*)，该基因赋予对氨基糖苷例如壮观霉素或链霉素的抗性，并且为质粒维持提供一种方法。

pMON17241 含有一个重组基因，该基因由连接于拟南芥(*Arabidopsis thaliana*)核酮糖-1,5-二磷酸羧化酶小亚基(SSU1A)基因序列 5'的 35S FMV 启动子组成，所述 SSU1A 基因序列编码质体或叶绿体引导肽(Timko 等, 1988)，它与 *gox* 基因编码序列在翻译上融合，并且连接至来自豌豆核酮糖-1,5-二磷酸羧化酶基因的命名为 E9 的 3'非翻译区的 3'。

pMON17213 是一种双基因植物转化载体, 含有包含以下元件的表达盒: (1)一个 35S FMV 启动子, 连接于拟南芥 EPSPS 叶绿体引导肽的编码序列, 后者与菌株 CP4 EPSPS 编码序列符合读框地连接, CP4 EPSPS 编码序列连接于 E9 3'非翻译区的 3'; 和(2)一个 35S FMV 启动子, 连接于一个编码质体引导肽的 SSU1A 基因序列, 所述 SSU1A 基因序列与 GOX 编码序列符合读框地连接, 而 GOX 编码序列连接于一个 NOS 3'终止序列的 3'.

上述的 pMON17261 是一种双基因植物转化载体, 含有包含以下元件的表达盒: (1)一个 FMV 35S 启动子, 连接于一个 SSU1A 叶绿体引导肽编码序列, 所述叶绿体引导肽编码序列与 GOX 编码序列符合读框地连接, 而 GOX 编码序列的下游侧翼为 E9 3'非翻译区; 和(2)一个增强的 35S 启动子, 连接于一个 SSU1A 叶绿体引导肽编码序列(SEQ ID NO:9), 后者与 PhnO 编码序列(SEQ ID NO:7)符合读框地连接, PhnO 编码序列连接于一个 NOS 3'序列的 3'.

pMON10151 是一种双基因植物转化载体, 含有包含以下元件的表达盒: (1)一个 FMV 35S 启动子, 连接于一个 SSU1A 叶绿体引导肽编码序列(SEQ ID NO:9), 所述叶绿体引导肽编码序列与 PhnO 编码序列(SEQ ID NO:7)符合读框地连接, 而 PhnO 编码序列的下游侧翼为一个 NOS 3'序列; 和(2)一个 FMV 35S 启动子, 连接于一个 SSU1A 叶绿体引导肽编码序列, 后者与 GOX 编码序列符合读框地连接, GOX 编码序列的下游侧翼为一个 NOS 3'序列.

pMON10149 是一种三基因植物转化载体, 含有包含以下元件的表达盒: (1)一个 FMV 35S 启动子和一个矮牵牛 HSP70 5'非翻译前导序列, 连接于一个 SSU1A 叶绿体引导肽编码序列, 所述叶绿体引导肽编码序列与 EPSPS 编码序列符合读框地连接, 而 EPSPS 编码序列的下游侧翼为 E9 3'终止和聚腺苷酸化序列; (2)一个 FMV 35S 启动子, 连接于一个 SSU1A 叶绿体引导肽编码序列(SEQ ID NO:9), 所述叶绿体引导肽编码序列与 PhnO 编码序列(SEQ ID NO:7)符合读框地连接,

PhnO 编码序列的下游侧翼为 NOS 3' 序列; 和(3)一个增强的 35S CaMV 启动子, 连接于一个 SSU1A 叶绿体引导肽编码序列, 后者与 GOX 编码序列符合读框地连接, 而 GOX 编码序列的下游侧翼为一个胭脂碱合酶 3' 聚腺苷酸化序列(NOS 3')。

5 在大肠杆菌 K12 菌株中装配质粒载体, 并将其交配到非致瘤性 ABI 土壤杆菌菌株中。按照 Umbeck 等(1987)和 Umbeck (美国专利第 5,159,135 号(1992), 其通过引用结合到本文中)描述的改进方法, 用氨基糖苷抗性的土壤杆菌菌株转化 Coker 312 衍生的下胚轴区段 (section), 只是按照 Trolinder 和 Goodin (1987)描述的改进方法进行植
10 株再生。对草甘膦抗性的选择产生几个棉花愈伤组织系, 随后通过对基因组 DNA 进行 PCR, 测定它们含有从土壤杆菌转移的编码 EPSPS、GOX 或 PhnO 的相应基因。另外, 通过蛋白质印迹分析, 确定这些相同的愈伤组织系表达所述所需基因。在植株再生后, 回收含有指定编码序列的全株棉花植株。

15 当每英亩应用 48 盎司草甘膦至 6-7 叶期时, 先前鉴定的、用包含 EPSPS 和 GOX 编码基因的双基因草甘膦抗性盒转化的植株, 被确定为抗草甘膦, 然而观察到叶片严重褪色。这种植物毒性效应推定是因为 GOX 介导的草甘膦降解而形成 AMPA 引起的。为了测试这一点, 将 AMPA 以三种不同的比率喷洒到野生型 Coker 312 植株上。以每英
20 亩 640 盎司应用草甘膦后 4 天时以及每英亩应用 64 盎司后 8 天时, 在植株中观察到叶片褪色和生长停滞。这些结果表明, 在 EPSPS/GOX 转化棉花植物系中观察到的植物毒性效应是植物中 GOX 介导的 AMPA 产生的结果, 并且这种植物毒性效应可以通过将 AMPA 酰基转移酶与 GOX 共表达而得到消除。为了测试这一点, 将仅表达 GOX 或
25 GOX + EPSPS 或者结合 PhnO 表达的棉花植株用 [¹⁴C]-草甘膦处理, 在施药后 7 天在叶片组织中监测同位素标记的草甘膦的代谢。

在用 pMON10149 转化后, 选择 Coker 312 草甘膦抗性重组棉花系 4416 作为草甘膦抗性棉花系, pMON10149 是一种三基因根癌土壤

杆菌介导的双边界植物转化载体，含有叶绿体定向的 EPSPS、GOX 和 PhnO，每种独立地从独立的 35S 启动子表达。由 R2 种子产生几株 4416 R3 植株。3 叶期或 4 叶期的每株小植株的一个叶片用已经用 [14C]-草甘膦(代号 C-2251)强化的 ROUNDUP ULTRA™市售除草剂混合物(Lot No. GLP-9701-7428-F)的混合物处理。表明 ROUNDUP ULTRA™是 30.25%(重量)草甘膦酸(glyphosate acid)，而[14C]-草甘膦的放射化学纯度为 97.3%，比活为 36.36 mCi/mmol。处理溶液由约 38 μl 组成，含有 1.60×10^6 dpm 与比活为 1.713×10^3 dpm/μg 草甘膦酸的 [14C]-草甘膦。局部应用后 3 天或 7 天，用水冲洗经处理的叶片，将其在液氮中冷冻，由用刮勺将其破碎，然后用 TEKMAR™ tissuemizer，在 10 ml 水中研磨。用 1N HCl 将叶片提取物调至 pH 3.5-4.0，如实施例 2 中所述，通过带有液体闪烁管形瓶收集和检测的 HPLC (HPLC/LSC)，分析约 4-8000 dpm 中 [14C]-代谢物的存在。也提取在局部应用后出现的包括分生组织和新叶片在内的新生长物，并分析 [14C]-代谢物。结果示于表 13 中。

表 13. 草甘膦抗性棉花中的 [14C]-草甘膦代谢

系 4416 植株#	% 草甘膦处理的叶片 提取物中 [14C] 代谢物...*			% 新生长物提取物中 [14C] 代谢物...*		
	草甘膦	AMPA	N-乙酰- AMPA	草甘膦	AMPA	N-乙酰- AMPA
MD03	55.2	2.5	37.4	nd**	nd	93.4
MD04	94.6	2.1	1.7	97.9	nd	nd
A01	48.6	2.1	44.7	0.9	0.2	95.8
A02	67.3	2.0	29.1	0.7	0.2	96.5
A03	48.8	2.0	43.4	1.2	nd	94.0
A04	19.4	1.6	73.9	1.5	nd	94.0
A05	59.9	2.2	31.1	2.2	0.2	95.2
A06	38.2	nd	60.9	1.5	0.2	93.5
A07	64.1	nd	26.8	1.4	0.5	93.9
A08	90.9	2.0	1.9	91.2	2.5	1.9

* 按每种样品中通过 HPLC/LSC 观察到的总 [14C] 同位素的百分比计的

[¹⁴C]-草甘膦、 [¹⁴C]-AMPA 和 N-乙酰-[¹⁴C]-AMPA。

** nd 表示通过 HPLC/LSC 没有检测到所述代谢物。

5 对水冲洗的经草甘膦处理的叶片的分析表明, 在所测试的 10 株植株中的 8 株中存在显著水平的 N-乙酰-[¹⁴C]-AMPA。这些水平代表从经处理的叶片提取的同位素的 27-74%。其余的活性几乎全部是 [¹⁴C]-草甘膦。非常少的 [¹⁴C] 同位素以 [¹⁴C]-AMPA 存在。其余 2 株植株代谢草甘膦的能力非常有限, 表现为所述叶片上或叶片中剩余高水平的 [¹⁴C]-草甘膦。这些植株中的一株在处理 7 天后也表现出矮化迹象, 10 显示出草甘膦的植物毒性。

对所测试的 10 株植株中的新生长物的分析表明, 所存在的主要形式的 [¹⁴C] 标记代谢物是 N-乙酰-[¹⁴C]-AMPA, N-乙酰-[¹⁴C]-AMPA 高于所述样品中总放射性同位素的 90%。相反, 在其余 2 株植株中的高于 90% 的所述同位素是 [¹⁴C]-草甘膦形式, 这与对来自这 2 株植株处理 15 叶片的提取物的分析一致。

也研究了重组棉花系 4268 (GOX/PhnO) 和 3753 (EPSPS/GOX) 中 [¹⁴C]-草甘膦的代谢。该项研究中的植株如以上关于棉花系 4416 所述, 通过将数小滴以 [¹⁴C]-草甘膦强化的 ROUNDUP ULTRA 应用至 3-4 叶期每株植株的一片叶片上, 进行处理。收获处理的叶片, 用水冲洗, 20 然后研磨和提取, 如上所述通过 HPLC 分析提取物中 [¹⁴C]-草甘膦、 [¹⁴C]-AMPA 和 N-乙酰-[¹⁴C]-AMPA 的存在。也提取并分析了在施药后出现的新生长物, 包括分生组织和新叶片。结果示于表 14 中。

表 14. 草甘膦抗性棉花中的¹⁴C-草甘膦代谢

植株	% 草甘膦处理的叶片 提取物中 ¹⁴ C代谢物...*			% 新生长物提取物中 ¹⁴ C代谢物...*		
	草甘膦	AMPA	N-乙酰- AMPA	草甘膦	AMPA	N-乙酰- AMPA
GOX/PhnO 植株						
B01	76.7	3.0	14.0	3.4	1.0	89.9
B02	63.9	4.8	25.0	1.1	1.5	91.5
B03	54.4	3.2	36.4	0.8	nd	94.7
B04	58.3	5.7	28.9	1.1	1.2	91.0
EPSPS/GOX 植株						
C01	59.8	26.6	nd	3.72	85.7	nd
C02	92.7	2.1	0.8	92.8	0.8	nd
C03	81.2	10.7	nd	13.5	72.0	1.9
C04	86.2	6.4	1.0	13.9	76.2	nd

* 按每种样品中通过 HPLC/LSC 分析后观察到的总¹⁴C同位素标记的代谢物的百分比计的¹⁴C-草甘膦、¹⁴C-AMPA 和 N-乙酰-¹⁴C-AMPA。

** nd 表示通过 HPLC/LSC 没有检测到所述代谢物。

5 在所有 4 个系 4268 植株(GOX/PhnO; B01-B04)的处理叶片中存在显著水平的 N-乙酰-¹⁴C-AMPA。相反,在得自系 3753 植株(EPSPS/GOX; C01-04)的提取物中没有可检测的 N-乙酰-¹⁴C-AMPA。这些植株中的 3 株在处理的叶片提取物中含有显著水平的¹⁴C-AMPA,其范围为 6-27%。一个系 3753 的植株缺乏¹⁴C-草甘膦至 N-乙酰-¹⁴C-AMPA 的转化,该植株也表现出矮化。

10 来自系 4268 植株的新生长物提取物中的 90-95%的所述¹⁴C同位素测定为 N-乙酰-¹⁴C-AMPA 形式。然而,来自系 3753 植株中的 3 株的新生长物提取物中的 72-86%的所述¹⁴C同位素测定是¹⁴C-AMPA,¹⁴C-草甘膦构成这些组织中剩余的所述同位素。得自系 3753 植株编号 C02 的 93%的所述同位素测定是¹⁴C-草甘膦,这与在施药叶片中缺乏草甘膦代谢以及所观察到的矮化一致。另外,在处理后的所有

系 3753 的植株的生长区均变色并且黄化, 但随时间的推移得到改善。到收获时, 新生长叶片变为花斑。

这些结果与所示植物中存在活性 *gox* 和 *phnO* 基因产物一致。GOX 和 PhnO 蛋白以预测的方式将草甘膦代谢为 AMPA 和 N-乙酰-AMPA, 通过 HPLC 和表型观察判断, 系 4268 植物提取物提供一种与系 4416 植物提取物观察到的相似的代谢模式。在这两个系中, 应用 [¹⁴C]-草甘膦后在新生长组织提取物中的主要 [¹⁴C] 产物是 N-乙酰-¹⁴C-AMPA。通过应用草甘膦后在系 3753 中的植物叶片的变色而观察到的植物毒性, 与缺乏 AMPA N-酰基转移酶活性相关。相反, 在 4416 和 4268 植物系中存在 AMPA N-酰基转移酶活性导致在系 3753 植物中没有观察到的植物毒性效应。

canola

用载体 pMON17138 和 pMON17261 转化 canola 植株, 获得了表现出草甘膦耐性的许多转化 canola 植物系。按照 Barry 等描述的方法(美国专利第 5,633,435 号)转化植物。简而言之, 如已描述的方法, 在控制的生长室条件下培育欧洲油菜 Westar 栽培品种(*Brassica napus* cv Westar)植株。取下恰好在抽薹之前植株的 4 个末端节间或抽薹过程中但在开花之前的植株, 在 70% v/v 乙醇中表面消毒 1 分钟, 然后在 2% w/v 次氯酸钠表面消毒 20 分钟, 然后用无菌蒸馏去离子水冲洗 3 次。连有叶片的茎可以在潮湿的塑料袋中冷藏直至消毒之前 3 天。用 Redco 切菜机 200, 维持基部末端的方向, 将 6-7 个茎节切成 5 mm 圆盘。用 1 毫升 ABI 根癌土壤杆菌菌株 A208 接种茎圆盘(外植体), 该菌株含有如上所述制备的重组植物转化质粒。将外植体基部一侧向下置于含有 0.1 X 标准 MS 盐、B5 维生素、3%蔗糖、0.8%琼脂 pH 5.7、1 mg/l BA (6-苄基腺嘌呤)的培养皿中。在平板上铺上 1.5 ml 含有 MS 盐、B5 维生素、3%蔗糖 pH 5.7、4 mg/l 对氯苯氧乙酸、0.005 mg/l 细胞分裂素的培养基, 并用无菌滤纸覆盖。

共培养 2.3 天后, 将外植体转移至深培养皿平板(7 个外植体/平板), 所述平板含有 MS 盐、 B5 维生素、 3%蔗糖、 0.8%琼脂 pH 5.7、 1 mg/l BA、 500 mg/l 羧苄青霉素、 50 mg/l 头孢噻肟、 200 mg/l 卡那霉素或 175 mg/l 庆大霉素用于选择, 3 周后将其转移至新鲜培养基上, 每个平板 5 个外植体。在 25 °C、 具有连续光照(冷白光)的生长室中培养外植体。再过 3 周后, 从外植体上切下苗, 开始叶片愈伤组织再形成测定(recallusing assay), 以证实 R₀ 苗的改进。将 3 个微小片的叶片组织置于愈伤组织再形成培养基上, 该培养基含有 MS 盐、 B5 维生素、 3%蔗糖、 0.8%琼脂 pH 5.7、 5 mg/l BA、 0.5 mg/l 萘乙酸(NAA)、 500 mg/l 羧苄青霉素、 50 mg/l 头孢噻肟、 200 mg/l 卡那霉素或庆大霉素或 0.5 mM 草甘膦。叶片分析物在生长室中与外植体培养

5 相同条件下培养。再过 3 周后, 记录叶片愈伤组织再形成测定的除草剂耐性(愈伤组织或绿色叶片组织)或感病性(sensitivity)(褪色)。

10

将每个苗茎在切下时浸入 ROOTONE , 置于 2 英寸含有 Metro-MIX 350 的花盆中, 并在约 3 周的锻炼期期间维持在 24 °C、 16/8 小时光周期、 400 uE/平方米/秒(HID 灯)下生长室封闭的潮湿环境中。

15

质粒 pMON17138 是土壤杆菌介导的单边界植物转化载体, 它通过在链霉素或壮观霉素上选择而维持在该细菌中。 pMON17138 含有一个 Ti 右边界, 该右边界邻接需要转移到植物基因组中的遗传元件的 3'端。该载体含有两个植物有效的表达盒。一个表达盒由以下元件组成: 一个驱动新霉素磷酸转移酶基因(*nptII*)的花椰菜花叶病毒 35S 启动子和该启动子下游侧翼的一个胭脂碱合酶 3' 转录终止和聚腺苷酸化序列(NOS 3')。另一表达盒由位于豌豆核酮糖二磷酸羧化酶小亚基转录终止和聚腺苷酸化序列上游的玄参花叶病毒启动子(描述于 Rogers, 美国专利第 5,678,319 号)组成。一个叶绿体定向的草甘膦氧化还原酶(GOX)编码序列插入该启动子和豌豆 3' 序列之间。

20

25

质粒 pMON17261 是一种与 pMON17138 相似的土壤杆菌介导的双边界植物转化载体。与 pMON17138 中相同的一个叶绿体定向的

GOX 编码盒存在于 Ti 右边界下游和另一植物有效表达盒的上游, 所述另一植物表达盒由连接于 NOS 3' 序列的玄参花叶病毒启动子(P-FMV)组成。一个叶绿体定向的 PhnO 编码序列插入所述第二 P-FMV 和 NOS3' 序列之间。

5 用 Barry 等(美国专利第 5,633,435 号)中描述的草甘膦喷洒试验, 评估得自采用 pMON17261 和 pMON17138 的转化事件的 R₁ 植株。

玉米

也已经将 AMPA 酰基转移酶基因插入黑墨西哥甜玉米细胞中, 在
10 愈伤组织中检测到该基因的表达和草甘膦抗性。按照 Barry 等(美国专利 5,463,175 号)中描述的方法转化愈伤组织。用各种质粒将编码 GOX 和 EPSPS 的草甘膦抗性基因结合 AMPA 酰基转移酶基因导入玉米细胞中。这些质粒在所用的启动子、所用的叶绿体或质体引导肽序列、所用的非翻译前导序列、内含子的存在与否以及所用的 3' 终止子类型
15 方面相互不同, 然而, 所有质粒都含有合成衍生的编码含 P2A 突变的 PhnO 的 AMPA 酰基转移酶基因。该合成基因由 Monsanto 合成的 3 个较小的多核苷酸序列构建, 并且由 Stratagene, Inc. La Jolla, CA 鉴定存在所需的 DNA 编码序列和氨基酸序列翻译。该非天然存在的基因由 3 个较小的序列装配, 这三个序列由 SEQ ID NO:16、SEQ ID NO:17 和
20 SEQ ID NO:18 组成, 其中所述全装配的基因由 SEQ ID NO:19 表示, 存在于用于玉米愈伤组织转化的每种所述质粒中。所述非天然存在的基因编码序列基于美国专利第 5,500,365 号中 Fishhoff 等描述的方法建立的, 其中用单子叶植物优选密码子取代大肠杆菌优选的密码子。所述全装配的基因编码全长 PhnO 蛋白, 该蛋白除用 SEQ ID NO:5 和 SEQ
25 ID NO:6 通过 PCR 引入的 P2A 突变外与所述天然蛋白序列相同, 所述突变将合适的限制性内切核酸酶识别位点工程到该编码序列的侧翼末端。用于产生玉米愈伤组织数据的质粒以及在邻接所述 AMPA 酰基转移酶编码序列的遗传元件方面的差异示于表 15 中。

表 15. 玉米愈伤组织转化质粒和相关的遗传元件

质粒	相关的遗传元件*
pMON32926	[Pe35S / I-Zm.Hsp70 / CTP / phnO / T-At.Nos] → GOX → EPSPS
pMON32931	[Pe35S / I-Zm.Hsp70 / phnO / T-At.Nos] → GOX → EPSPS
pMON32932	[Pe35S / I-Zm.Hsp70 / CTP / phnO / T-At.Nos] → GOX → EPSPS
pMON32936	[P-Os.Act1 / I-Os.Act1 / CTP / phnO / T-At.Nos] → GOX → EPSPS
pMON32938	[P-Os.Act1 / I-Os.Act1 / CTP / phnO / T-At.Nos] → GOX → EPSPS
pMON32946	[Pe35S / L-Ta.Cab / CTP / phnO / T-Ta.Hsp70] → GOX → EPSPS
pMON32947	[Pe35S / L-Ta.Hsp70 / CTP / phnO / T-Ta.Hsp70] → GOX → EPSPS
pMON32948	EPSPS → [Pe35S / I-Zm.Hsp70 / CTP / phnO / T-At.Nos] → GOX
pMON32950	EPSPS → [Pe35S / I-Zm.Hsp70 / CTP / phnO / T-At.Nos] → GOX
pMON32570	EPSPS → [Pe35S / L-Ta.Cab / I-Os.Act1 / CTP / phnO / T-Ta.Hsp70] → GOX
pMON32571	EPSPS → [Pe35S / L-Ta.Cab / I-Os.Act1 / CTP / phnO / T-Ta.Hsp70] → GOX
pMON32572	EPSPS → [Pe35S / L-Ta.Hsp70 / I-Os.Act1 / CTP / phnO / T-Ta.Hsp70] → GOX
pMON32573	EPSPS → [Pe35S / L-Ta.Cab / I-Os.Act1 / CTP / phnO / T-Ta.Hsp70] → GOX

* 包含在 PhnO 表达盒内的遗传元件如每个质粒中所示。元件以它们出现在所述质粒中的顺序显示，如果有的话，还显示邻接所述 PhnO 表达盒的编码除草剂抗性的其它基因的存在。

→显示邻接所述 PhnO 表达盒的每个基因或多个基因的转录方向。各个元件均在正文中描述。

使用的启动子包括 CaMV e35 S 启动子和水稻肌动蛋白启动子(P-Os.Act1)。使用的内含子包括得自植物基因例如玉米 Hsp70 (I-Zm.Hsp70)和水稻肌动蛋白(I-Os.Act1)的那些内含子。使用的非翻译前导序列包括小麦叶绿素 a/b 结合蛋白(L-Ta.Cab)和玉米 Hsp70 (L-Zm.Hsp70)。使用的终止和聚腺苷酸化序列包括根癌土壤杆菌 NOS 3' (T-At.Nos)和小麦 Hsp70 (T-Ta.Hsp70)。在所有 PhnO 表达盒中使用同一叶绿体引导序列，用 SEQ ID NO:9 表示。

用¹⁴C]-草甘膦代谢测定，来确定转化的玉米愈伤组织是否含有这些酶的功能形式。该测定开发用来筛选大量的玉米愈伤组织样品。愈伤组织得自 Monsanto Company 和 Dekalb Seed Company 玉米转化小组。在表 16 中单独命名为愈伤组织系“19nn-nn-nn”的 Monsanto 愈伤组织样品由 HI II X B73 玉米胚产生。用完全共价闭环的重组植物转化载体质粒 DNA，或用在胚分离后 25-50 天从这样的质粒分离的线性 DNA 片段，轰击愈伤组织样品。轰击后 8-14 周鉴定转化系。将这些系每 2 周在新鲜培养基上传代培养，当用于所述代谢测定中为 5-7 月

龄。Dekalb 愈伤组织系 OO、OR、OW、OX 和 OY 得自 HI II x AW 胚。所有系的命名均对应于用来如表 16 的说明注释的方法对愈伤组织进行基因枪转化的重组质粒或线性片段。

从 Monsanto Radiosynthesis 小组获得 1.5 mM 水溶液中的 4.5 mCi N-磷酸基- ^{14}C -甲基甘氨酸(^{14}C -草甘膦), 其比放射性为 39.4 mCi/mM (5.2 × 10⁵ dpm/μg)。该样品以代号 C-2182.2 鉴别。通过将 2.5 ml ^{14}C -草甘膦(3.3 × 10⁸ dpm)与 2.5 ml 玉米愈伤组织生长培养基(N6 培养基)和 5.0 mg Mon 0818 表面活性剂混合, 制备通过 0.2 微米 Acrodisk (Gelman no. 4192)过滤除菌的储液。在所述产生的施药溶液(dose solution)中 ^{14}C -草甘膦是 0.75 mM。N6 培养基由 Chu 等(1975)描述, 并且采用得自 Sigma Chemical Company, St. Louis, MO 的盐和和维生素制备。Mon 0818 表面活性剂是一种乙氧基化牛油脂肪胺, 这是用于 Roundup 除草剂中的表面活性剂。将该施药溶液如实施例 2 中所述经过 HPLC 分析。结果示于图 1 描述的色谱图中。分辨出 3 个放射性峰, 最大的峰对应于草甘膦(11.3 min, 98.8%)。在该施药溶液中也存在对应于 ^{14}C -AMPA (5.8 min, 0.16%)和未鉴定的物质(10.2 min, 1.0%)的杂质峰。在该施药溶液中不存在对应于 N-乙酰- ^{14}C -AMPA 的峰。用这些试剂制备另外两种施药溶液, 每种施药溶液根据上述制备方法, 按比率放大 3 倍至 15 ml 体积。

合成 N-乙酰- ^{14}C -AMPA, 用作保留时间 HPLC 标准。将 1 ml 吡啶和 2 ml 乙酸酐加入 20 ml 螺口培养管中, 并且在冰上冷却。将 0.1 ml ^{14}C -AMPA (6.2 × 10⁶ dpm, 代码 C-2127.2)的水溶液加入该冷却的溶液中。然后从冰浴中取出培养管, 将其温至约 50-60 °C。在约 30 分钟后取出 10 μl 样品, 将其与 0.5 ml 水混合, 并且按照上述 HPLC 方法进行分析。没有检测到 ^{14}C -AMPA, 然而鉴定出两个新放射性峰; 一个峰在 13.9 分钟(68%), 另一个峰在 15.4 分钟(32%)。分离在 13.9 分钟洗脱出的物质样品, 并且通过负离子电喷雾质谱分析。结果显示于 m/e 152 和 154 的强离子, 正如对分子量为 153 道尔顿的该化合物

所预期的； m/z 154 离子是由同位素 ^{14}C 原子产生的。没有分离出在 15.4 分钟洗脱出动放射性峰。然而，在一个独立的 HPLC 实验中，表明它与合成的 N-乙酰-N-甲基-AMPA 共洗脱。先前已经表明， N-甲基- ^{14}C -AMPA 是起始 ^{14}C -AMPA 物质中的一种杂质。

5 在无菌条件下，将玉米愈伤组织样品转移至无菌 48 孔 COSTAR™ 细胞培养组件(clusters) (cat. No. 3548)的各孔中。没有称量各个愈伤组织样品。然而，在几种情况下，测定了 48 孔板中愈伤组织样品的总重量。各个愈伤组织样品的平均重量通常约为 200-250 mg。在每个测定中，包括一个非转化的愈伤组织样品 HI II X B73 作为对照。向每个愈伤组织样品加入 50 μl 含有 3.3×10^6 dpm ^{14}C -草甘膦的施药溶液。48
10 孔板用封口膜密封，并置于含有一张提供潮湿氛围的湿纸巾的塑料袋中。封闭塑料袋，将其置于 25 $^{\circ}\text{C}$ 的黑暗抽屉中达 10 天。随后将每个愈伤组织样品转移至贴标签的微量离心管(VWR, 1.7-ml, cat. No. 20170-620)。每管加入 1.0 ml 去离子水，将所述试管封闭并置于圆形
15 20 管浮动式微量离心机机架(Nalge cat. no. 5974-1015)中。让这些微量离心管浮在沸水中 30 分钟，用涡旋混合器震摇，然后用 Fisher 牌微量离心机离心 5 分钟。取出 120 μl 上清液样品，用于如下所述通过 HPLC 分析。用 Waters WISP 自动注射器注射样品。获得所分析的每种样品的色谱图形，通过将每种样品中放射性洗脱峰下的面积外推至总
20 ^{14}C ，获得定量信息。图 2 显示在观察的放射性代谢物 ^{14}C AMPA、 ^{14}C 草甘膦和 N-乙酰- ^{14}C -AMPA 和鉴定为 N-乙酰-N-甲基- ^{14}C -AMPA 杂质的标准的混合物的 HPLC 图形。

对于大多数分析，通常用 SPHERISORB™ S5 SAX 250 mm x 10mm 柱完成 HPLC 分析。某些样品在提供相似性能的 ALLTECH™ 5
25 微米 250 x 10 mm SAX 柱上分析。制备两种溶剂，溶剂 A 由用 H_3PO_4 调至 pH 2.0 的 0.005 M KH_2PO_4 组成，并含有 4% 甲醇。溶剂 B 由用 H_3PO_4 调至 pH 2.0 的 0.10 M KH_2PO_4 组成，也含有 4% 甲醇。洗脱剂流速设定在 3 ml/min，闪烁液的流速设定在 9 ml/min，使用

ATOMFLOW™闪烁液(No. NEN-995, 得自 Packard Instruments)。所有柱子溶剂步骤(solvent step)都是线性的, 注射和柱子溶剂流速如实施例 2 中所示。在 20 分钟时制备柱子用于另一次注射。

5 将得自 359 个转化玉米系的愈伤组织样品与 50 μ l 等份的 ^{14}C -草甘膦施药溶液混合, 并且在黑暗中温育 10 天。每种温育后的愈伤组织样品与其附着的施药物质以及 1 ml 水一起转移至 1.7 ml 微量离心管中, 将每个离心管置于沸水中。该步骤使得细胞裂解, 释放出可溶性胞内化合物, 包括任何同位素标记化合物, 例如草甘膦、AMPA 和 N-乙酰-AMPA。在方法开发期间确定了, 如果温育后愈伤组织用水彻底
10 冲洗, 则冲洗掉 85-95%的所述放射性, HPLC 分析表明, 在冲洗液中实际上所有的所述放射性都是由 ^{14}C -草甘膦产生的, 没有由 ^{14}C -代谢物所致的。在这些实验中, 经冲洗的愈伤组织得到含除 ^{14}C -草甘膦外的 ^{14}C -代谢物的提取物。这表明, 冲洗液中的放射性如果不是仅由未吸收的表面 ^{14}C -草甘膦产生的话, 也主要是由其产生的。重要的是,
15 当考虑到转化为代谢物的剂量的百分比相当低时, 将其考虑在内, 因为百分比的计算包括大量的未吸收表面放射性。方法开发研究也表明, 仅仅在水中将温育的愈伤组织煮沸释放出的放射性活性与通过常规研磨/提取方法所释放的放射性活性一样多。进行多个实验, 表明煮沸(oiling)没有改变所述代谢物分布型。合理化方法使得有可能一次分
20 析大量的样品(例如 96 个)。表 16 显示了在 HPLC 分析后获得的产生最高水平的 N-乙酰- ^{14}C -AMPA 或 ^{14}C -AMPA 的愈伤组织样品的代表性数据。GOX 加上 AMPA 酰基转移酶转化的、草甘膦处理的愈伤组织提取物样品的代表性色谱图示于图 3 中。

表 16.

产生各种量的 AMPA 或 N-乙酰-AMPA 的转化玉米愈伤组织系

产生 N-乙酰-[¹⁴ C]-AMPA 的愈伤组织			产生 [¹⁴ C]-AMPA 的愈伤组织		
愈伤组织*	用...转化	N-乙酰-[¹⁴ C]-AMPA 的百分比**	愈伤组织*	用...转化	[¹⁴ C]-AMPA 的百分比**
1978-05-02	pMON32570	0.27	1980-28-03	pMON32571	2.89
1978-08-01	pMON32570	0.94	OR523	pMON32926	2.00
1978-20-02	pMON32570	0.57	OR534	pMON32926	5.00
1978-21-02	pMON32570	0.23	OR537	pMON32926	2.00
1978-22-01	pMON32570	0.90	OR539	pMON32926	5.08
1978-24-02	pMON32570	1.80	1971-08-01	pMON32932	2.64
1978-35-01	pMON32570	0.22	1971-27-03	pMON32932	3.63
1980-01-01	pMON32570	0.27	OO505	pMON32932	2.73
1980-03-01	pMON32571	0.22	OO509	pMON32932	2.86
1981-28-01	pMON32571	0.25	OO510	pMON32932	2.34
1981-02-01	pMON32572	0.65	OO512	pMON32932	2.31
1981-03-01	pMON32572	0.74	OO514	pMON32932	1.98
1981-18-01	pMON32572	0.22	OO535	pMON32932	2.88
1981-23-01	pMON32572	0.48	OO538	pMON32932	2.70
1981-24-02	pMON32572	0.29	OO539	pMON32932	1.97
1981-32-02	pMON32572	1.08	OO553	pMON32932	3.56
1977-05-03	pMON32573	0.39	OO576	pMON32932	3.49
OR516	pMON32926	1.91	OO579	pMON32932	2.85
1972-14-01	pMON32931	0.40	1986-17-01	pMON32936	2.29
1972-32-01	pMON32931	0.75	1986-18-03	pMON32936	3.05
1972-33-01	pMON32931	0.55	1986-18-04	pMON32936	2.15
OO544	pMON32932	0.28	1986-28-02	pMON32936	2.06
1986-06-01	pMON32936	0.30	1983-12-02	pMON32938	2.41
1986-08-01	pMON32936	1.13	1983-31-01	pMON32938	2.90
1986-08-03	pMON32936	0.70	1985-03-02	pMON32946	2.51

5

10

表 16. (续表)

产生 N-乙酰-[¹⁴ C]-AMPA 的愈伤组织			产生 [¹⁴ C]-AMPA 的愈伤组织		
愈伤组织*	用...转化	N-乙酰- [¹⁴ C]-AMPA 的百分比**	愈伤组织*	用...转化	[¹⁴ C]-AMPA 的百分比**
1986-12-01	pMON32936	0.33	1985-38-01	pMON32947	1.99
1986-18-02	pMON32936	0.40	OX512	pMON32948	2.43
1986-18-03	pMON32936	0.51	OX533	pMON32948	3.91
1986-18-04	pMON32936	1.09	OX556	pMON32948	12.11
1986-22-04	pMON32936	0.64	OX504	pMON32950	2.25
1983-11-01	pMON32938	0.21	OX511	pMON32950	2.53
OW534	pMON32946	0.77	OX528	pMON32950	2.58
OW542	pMON32946	0.85	OX532	pMON32950	2.24
1985-26-01	pMON32947	0.60	OX534	pMON32950	4.02
1985-26-03	pMON32947	0.71	OX535	pMON32950	2.34
1985-11-04	pMON32952	0.37	OX540	pMON32950	5.57

*所有系均用基因枪法转化。以 19xx-yy-zz 命名的系用分离的质粒线性片段转化。分离线性片段以便将其从质粒骨架结构中分离。

**测定的作为样品中放射性总量的百分比测定的 N-乙酰-[¹⁴C]-AMPA 或 [¹⁴C]-AMPA 峰的放射性百分比，包括如正文中描述的残留的 [¹⁴C]-草甘膦。

5 所测试的 359 个愈伤组织样品中的 19 个样品产生含有水平明显高于其它愈伤组织样品的 N-乙酰-[¹⁴C]-AMPA 的提取物。愈伤组织 OR516 在这一方面最强，在 2 个月期间对其分析 5 次，提供的数值范围为 0.50-4.54% (平均 1.91%)。尚不了解在不同时间形成的 N-乙酰-[¹⁴C]-AMPA 的百分比分布相对产的基础。在 OR516 的 5 次分析中的 4 次中，存在的 N-乙酰-[¹⁴C]-AMPA 百分比高于 [¹⁴C]-AMPA 的百分比，表明有效地将 [¹⁴C]-AMPA 转化为 N-乙酰-[¹⁴C]-AMPA。产生 N-乙酰-[¹⁴C]-AMPA 第二最有效的愈伤组织是 1978-24-02，它是除 OR516 外唯一的在其提取物中所含有的 N-乙酰-[¹⁴C]-AMPA 高于 [¹⁴C]-AMPA 的其它愈伤组织。所测试的 359 个愈伤组织样品中的 100 个产生所含有的 [¹⁴C]-AMPA 水平明显高于其它愈伤组织样品的提取物。OX556 是 [¹⁴C]-AMPA 的最高程度的产生者，产生的 [¹⁴C]-AMPA 是该项研究中任何其它愈伤组织的所述代谢物的两倍以上。对照愈伤组织 HI II X

B73 不含插入基因，没有产生可检测水平的 N-乙酰-[¹⁴C]-AMPA，并且仅产生背景水平的 [¹⁴C]-AMPA。该结果表明，AMPA 酰基转移酶在玉米中的表达有效地将由于 GOX 介导的草甘膦降解而产生的 AMPA 转化为 N-乙酰-AMPA。

5

小麦

已经表明，GOX 介导的草甘膦降解产生 AMPA，而且先前已经表明 AMPA 是植物毒性效应的来源。因此，如实施例 2 中所述，测定小麦植株暴露于化合物 AMPA 或 N-乙酰-AMPA 的效应，以便观察小麦对这些化合物中任一种的任何敏感性或不敏感性。任何植物毒性效应的观察可能表明，GOX 介导的草甘膦代谢可能对小麦属(*Triticum*) 物种有害。

10

在小麦胚萌发测定中，将小麦未成熟胚暴露于不同浓度的 AMPA 或 N-乙酰-AMPA。制备 MMSO 基础培养基(base media)，该培养基含有 40 克/升麦芽糖、2 克/升 GELRITE™、MS 盐和和维生素。将盐、维生素和麦芽糖溶于 3500 ml 水中，将 pH 调至 5.8。将 500 ml 和 1 克 GELRITE™一起分装到一个独立的瓶中，并且高压灭菌 17 分钟。在培养基冷却至约 45 °C 后，将 AMPA 或 N-乙酰-AMPA 加至限定的浓度。将该混合物在无菌条件下注入 6 个方形 Sundae 杯中，并且让其固化。

15

从 20 天大的幼苗(在花药形成后)分离未成熟小麦胚，将其接种到各个 MMSO 培养基中。每个 Sundae 杯含有 9 个未成熟胚。AMPA (0、0.1、0.15、0.2、0.25、0.3 和 1.0 mM)或 N-乙酰-AMPA (0、0.1、0.3、1.0 和 3.0 mM)的每个浓度使用 3 个独立的平板。将 Sundae 杯温育 10 天，测定根和苗的长度并将其进行比较。结果示于表 17 中。

20

25

表 17.

AMPA 和 N-乙酰 AMPA 对萌发中的苗和根长度影响的比较

磷酸化合物	苗(cm)	根(cm)
AMPA (mM)		
0.00	12.6 ± 2.6	7.0 ± 1.9
0.10	11.7 ± 2.5	8.0 ± 2.0
0.15	11.3 ± 2.1	6.3 ± 1.7
0/20	9.2 ± 1.8	4.6 ± 2.1
0.25	8.5 ± 1.8	3.1 ± 1.6
0.30	6.6 ± 1.8	2.6 ± 1.6
1.00	0.9 ± 0.1	0.4 ± 0.1
N-乙酰-AMPA		
0.00	12.6 ± 2.6	7.0 ± 1.9
0.10	12.0 ± 2.4	5.9 ± 1.4
0.30	11.7 ± 3.5	5.2 ± 1.2
1.00	12.2 ± 3.2	5.4 ± 1.5
3.00	11.2 ± 2.6	5.9 ± 1.6

AMPA 在 0.2 mM 的浓度下对于未成熟胚的生长和伸长基本上没有抑制。然而, 高于 0.2 mM 的浓度对于苗和根的伸长有严重的抑制作用, 表明 AMPA 也可能对小麦有植物毒性, 将单子叶作物物种的性质作为一个整体考虑, 可能对其它单子叶作物以及草坪草有植物毒性。当 AMPA 水平高于 0.20 mM 时, 未成熟胚的萌发显著受影响。1.00 mM AMPA 消除了小麦未成熟胚的萌发。相反, N-乙酰-AMPA 不抑制苗, 在该实验中测试的任何浓度下仅对根的伸长有轻度抑制作用。测试的最高 N-乙酰-AMPA 浓度高于所测定的 AMPA 的最低非抑制性浓度的 10 倍。当 N-乙酰-AMPA 浓度低于 3.0 mM 时, 对于未成熟胚的萌发没有显著影响。该结果表明, 在小麦中 AMPA 的 N-酰化可能防止由于 GOX 介导的草甘膦除草剂代谢而产生的 AMPA 的植物毒性。

按照 Zhou 等的方法(Plant Cell Reports 15:159-163, 1995), 产生重组草甘膦耐性小麦植株。简而言之, 将春小麦 - 普通小麦 Bobwhite 栽

培品种(*Triticum aestivum* cv Bobwhite)用作靶转化系。将原种植株种植在由高强度放电灯(Sylvania, GTE Products Corp., Manchester, NH)提供的 800 微焦耳/平方米/秒、16 小时光周期的环境控制的生长室中。昼/夜温度为 18/16 °C。从开花后 14 天的植株收集未成熟颖果。无菌解剖出未成熟胚, 将其培养在 MMS2 培养基上, 这是一种 Murashige 和 Skoog (Physio. Plant 15:473-497, 1962)基本培养基, 补充有 40 克/升麦芽糖和 2 毫克/升 2,4-D。在某些实验中, 使用 CM4 培养基。CM4 培养基含有是 MMS2 培养基, 但仅含 0.5 毫克/升 2,4-D, 并且包括 2.2 毫克/升氨基吡啶酸。未成熟胚于 26 °C 黑暗中培养。

培养开始后 5 天, 在按照 Russell 等的方法(In Vitro Cell Devel. Biol., 28P:97-105, 1992), 在轰击前 4 小时, 将未成熟胚转移至含有 0.35 M 甘露醇的渗透处理 CM4 培养基。将 30-40 个胚置于每个平板的中心, 在 DuPont PDS1000 装置中进行轰击。按照 Sanford 等的方法(Particle Sci. Technol., 5:27-37, 1987)将质粒 DNA 吸附到 1 μm 钨粒子上。对距离阻挡板(stopping plate) 13 mm 处的胚轰击 2 次。将 100 μm 不锈钢筛网紧靠阻挡板下面放置。

对渗透培养基轰击处理后 16 小时后, 将胚转移至 MMS2 或 CM4 培养基。在 1 周延迟期后, 将所述胚转移至含有 2 mM 草甘膦的 MMS2 或 CM4 培养基。在选择培养基上 9-12 周的愈伤组织增殖后, 将愈伤组织转移至含有 0.2 mg/l 2,4-D 和 0.1 mM 草甘膦的 MMS0.2 再生培养基上。将得自再生培养基的苗转移至无 2,4-D、但含有 0.02 mM 草甘膦的 MMS0 上。

将草甘膦耐性 R₀ 植株以及 R₁ 后代转移至 15 厘米直径的花盆中, 并且在如上所述的环境控制的室中培育。2 周后, 在喷洒室中用 3 ml/升 ROUNDUP (41%有效成分, Mansanto Company)喷散植株, 喷洒室设计用于模拟 0.6 公斤草甘膦/公顷的大田剂量施药。在喷洒后不同阶段, 观察并记录损伤症状。

按照 Shure 等的方法(Cell 35:225-233, 1983), 从 R₀ 和 R₁ 后代的叶

片组织分离基因组 DNA。15 微克基因组 DNA 用 *Bg*III 限制性内切核酸酶消化，并且在 0.8% 琼脂糖凝胶上分离。按照 Sambrook 等描述的标准方法(Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, 1989)，将所述 DNA 转移至 Hybond N 膜(Amersham) 5 上。独立地探测所述膜中编码 EPSPS 和 GOX 的基因的存在。通过 *Bg*III 限制性内切核酸酶消化，从 pMON19574 中释放出含有 EPSPS 基因的 3.4 kb DNA 片段和含有 GOX 基因的 4.8 kb DNA 片段，通过 0.7% 琼脂糖凝胶电泳将其分离，并用 Stratagene PRIME-IT II 随机引物标记试剂盒，用 [³²P] dCTP 标记。将探针分别标记至比活为 3×10^9 计数/分钟/ 10 微克和 1.3×10^9 计数/分钟/微克。在含 50% 甲酰胺、5X SSC、5X Denhardt's、0.5% SDS 和 100 微克/毫升 tRNA 的溶液中将膜于 42 °C 杂交 14 小时。最终的洗涤条件是 0.1% SSC 和 0.1% SDS，于 60 °C 15 分钟。

用来自 R₀ 植株叶片组织的粗制蛋白提取物进行 EPSPS 和 GOX 蛋 15 白分析，按照 Bradford 的方法(Anal. Biochem. 72:248-256, 1976)定量测定总蛋白。用 ELISA 法定量测定所述提取物中存在的 EPSPS 和 GOX 蛋白的百分比，并以可提取的总蛋白的百分比来计算。

在开花后 20 天，分离 R₀ 转基因植株和 Bobwhite 对照植物的未成 20 熟胚，将其在含有 0.02 mM 草甘膦的 MMS0 培养基上培养，用于萌发试验。10 天后记录萌发的和未萌发的胚，并且通过 3:1 分离的 χ^2 检定分析数据。将来自萌发试验的耐性植株转移至土壤中，如上所述用 3 毫升/升 ROUNDUP 喷洒。

用带有草甘膦抗性基因的 5 种质粒如上所述转化小麦未成熟胚。 pMON19338 含有一个核苷酸盒，该核苷酸盒编码与土壤杆菌菌株 CP4 25 EPSPS 酶序列符合读框的矮牵牛 EPSPS 叶绿体转运肽。将该核苷酸盒插入连接于玉米 HPS70 内含子序列 3' 的花椰菜花叶病毒增强的 35S 启动子的下游和胭脂碱合酶 3' 转录终止和聚腺苷酸化序列的上游。便利的限制位点位于所述内含子序列和 3' 终止序列之间，用于遗传元件的

插入。 pMON19643 与 pMON19338 相同, 只是用 GOX 酶编码序列取代土壤杆菌 EPSPS 酶编码序列。 pMON19574 与 pMON19338 相同, 但另外含有一个与 pMON19643 中相同的叶绿体定向的草甘膦氧化还原酶表达盒, 该表达盒位于 EPSPS 表达盒下游并且紧接所述 EPSPS 表达盒。 pMON32570 与 pMON19574 的相似之处在于: 存在编码叶绿体定向的 EPSPS 和叶绿体定向的 GOX 的表达盒, 然而, pMON32570 在所述 EPSPS 和 GOX 表达盒之间也存在一个编码叶绿体定向的 AMPA 酰基转移酶的表达盒。在 pMON19574 中存在而在其它质粒中不存在的其它元件也值得提及。例如, 小麦主要叶绿素 a/b 结合蛋白基因 5'非翻译前导序列存在于 EPSPS 表达盒和 AMPA 酰基转移酶表达盒两者中的所述增强的 35S 启动子和内含子之间(McElroy 等, Plant Cell 2:163-171, 1990)。此外, 存在小麦 hsp17 基因 3'转录终止和聚腺苷酸化序列, 以取代 EPSPS 和 AMPA 酰基转移酶两个表达盒的胭脂碱合酶 3'序列。用以上描述的基因枪转化法, 所有质粒均产生重组草甘膦耐性小麦植株。然而, 仅能够表达 GOX 或能够表达 GOX 和 AMPA 酰基转移酶的质粒都不产生重组草甘膦耐性小麦植株, 或产生经历生长停滞、表型分离异常和不育的问题的植物, 对其不作进一步的分析。用所述质粒进行基因枪转化后获得的数据示于表 18 中。

表 18. 小麦基因枪转化数据

草甘膦耐性基因	外植体编号	转基因事件 编号	转化效率 ¹
GOX	120	0	0
GOX + PhnO	434	6	1.4
EPSPS	120	6	5.0
EPSPS + GOX	120	1	0.8
EPSPS + PhnO + GOX	10,068	314	3.1

1 - 转化效率基于从实验组合产生的外植体总群体鉴定出的转基因事件的百分比, 其中特定的载体构建体已经被轰击到未成熟胚中。

使由这些转化产生的转化的草甘膦抗性植株自交，让其产生 R1 种子，用 R1 种子产生 R1 植株。通过在 3 叶期喷洒草甘膦后 R1 植株的敏感性判断，在 R1 植株中草甘膦抗性一般以 3:1 的预测比率分离。让草甘膦抗性 R1 植株自交，并让其产生 R2 种子。从多种不同的草甘膦抗性系萌发 R2 种子，以产生 R2 草甘膦抗性植株，对所述植株如上所述施用¹⁴C-草甘膦。施用草甘膦后 48 小时，收获植物叶片和茎组织，如上所述提取水溶性化合物，并且如实施例 2 中所述通过 HPLC 分析¹⁴C-草甘膦代谢物的存在。从该柱洗脱出的¹⁴C同位素标记峰下的总面积相加，以提供所分析的每种样品的 100% ¹⁴C-化合物鉴定的基线。结果示于表 19 中。

表 19. 小麦植株提取物中的草甘膦代谢¹

样品和草甘膦耐性基因	植株系 编号	[¹⁴ C]- 草甘膦	[¹⁴ C]- AMPA	乙酰-[¹⁴ C]- AMPA	[¹⁴ C]-其它 ⁴
标准 ²	na	30	26	31	13
	na	29	24	29	18
	na	35	29	36	0
生长培养基 ³	na	60	32	0.2	8
	na	48	25	2	25
	na	87	7	0	6
EPSPS	24756	43	25	1	31
	24756	53	46	0	1
	25397	61	38	0	1
	25397	37	19	1.2	43
	25397	64	20	0	16
EPSPS + PhnO + GOX	27249	6	7	85	2
	27249	14	12	61	13
	27249	5	24	33	38
	25462	48	21	0	31
	25462	44	5	0	51
	25462	54	35	0	11
	26281	48	14	17	21
	26281	64	11	13	12
	26281	38	7	7	48
	28598	20	7	5	68
28598	25	7	5	63	
Bobwhite	na	74	26	0	0
	na	17	15	0	32
	na	34	24	0	42

1 - 在如实施例 1 中所述, 在施用 [¹⁴C]-草甘膦后, 通过 HPLC 分析植株的组织提取物, 并且将曲线图每个峰下的面积相加, 提供每种样品 100% [¹⁴C]-化合物鉴定的基础。

2 - 含有大约等 [¹⁴C] 摩尔比的每种已知草甘膦代谢相关化合物的标准溶液。

3 - 包括 [¹⁴C]-草甘膦的生长培养基; 先前已经表明草甘膦通过光解过程降解为 AMPA, AMPA 可以在某些酰基化合物 (MSL-0598) 存在下自动酰化。

4 - 采用所述公开的色谱法分辨出的未表征的 [¹⁴C]-标记化合物。草甘膦的保留时间约为 9.6 分钟, AMPA 的保留时间约为 5.4 分钟, N-乙酰-AMPA 的保留时间约 12.5 分钟, [¹⁴C]-草甘膦样品中的主要 [¹⁴C]-标记杂质的保留时间约为 4.7 分钟。

标准溶液含有大约等摩尔比的化合物草甘膦、AMPA 和 N-乙酰-AMPA 中的每种以及由于这些同位素标记化合物的化学合成而存在的多种杂质。加入 [¹⁴C]-草甘膦的生长培养基经过与小麦植株经受的相同

5

条件的处理，即将所述培养基暴露于植物接受的入射光强。正如所预期的，观察到草甘膦光降解为 AMPA，看来小百分比的 AMPA 转化为酰基-AMPA，这可能是由于在生长培养基中暴露于其它酰化化合物的结果。先前已经观察到通过暴露于可见光草甘膦光降解为作为主要降解产物的 AMPA(Lund-Hoeie 等, *Photodegradation of the herbicide glyphosate in water*. Bull. Environ. Contam. Toxicol. **36**:723-729, 1986)。

5 用仅 EPSPS (EPSPS-only)质粒转化的重组小麦植株没有由^[14C]草甘膦产生^[14C]-AMPA 或乙酰-^[14C]-AMPA。观察到的^[14C]-AMPA 和微量的乙酰-^[14C]-AMPA 是在生长培养基对照中由于光降解而观察到的范围内。

10 用^[14C]-草甘膦处理的非重组 Bobwhite 对照植物也没有产生 AMPA 或乙酰-AMPA。用含有能够表达 EPSPS、PhnO 和 GOX 的基因的所述三基因构建体质粒转化的植物产生可变的的结果。约 1/3 的这些植株看来有效地将草甘膦转化为乙酰-AMPA，表明所述 GOX 和 PhnO 酶存在并且有功能。DNA 印迹分析证明，所述转基因整合到小麦基因组中，并且传递给以后的后代。

15 用抗 EPSPS、抗 GOX 或抗 PhnO 抗血清以检测所述三基因转化的植物提取物中的这些蛋白的蛋白质印迹分析，提供了关于所述可变的^[14C]-草甘膦代谢观察的基础的进一步的深入了解。蛋白质印迹分析表明，所有所述系均产生 EPSPS，然而仅系 27249 产生 GOX 蛋白和 PhnO 蛋白。这一结果与表 19 中的数据一致，

20 表明系 27249 有效地将^[14C]-草甘膦代谢为乙酰-^[14C]-AMPA。该植株系也没有表现出与其它系相关的矮化、部分可育或改变的分离表型。这些结果表明，GOX 和 AMPA 酰基转移酶在表达重组 EPSPS 的小麦植株中共表达提供了提高的除草剂耐性。

25 实施例 9

该实施例描述了用 *phnO* 基因转化烟草叶绿体。

可以产生其中仅线粒体或叶绿体 DNA 改变以掺入本申请设想的分子的重组植物。在叶绿体中有功能的启动子是本领域已知的

(Hanley-Bowden 等, Trends in Biochemical Sciences 12:67-70, 1987)。获得含有已经插入异源 DNA 的叶绿体的细胞的方法和组合物例如已经由 Daniell 等(美国专利第 5,693,507 号; 1997)和 Maliga 等(美国专利第 5,451,513 号; 1995)描述。可以构建含有一个可以从中产生酰基转移酶蛋白的表达盒的载体。表达盒可以含有一个叶绿体有效启动子序列, 该启动子序列驱动例如 *phnO* 基因的表达, 所述 *phnO* 基因的构建方式与本文的其它多核苷酸大致相同, 采用 PCR 方法、限制性内切核酸酶消化和连接等构建。叶绿体可表达基因将提供: 一个启动子和一个 5' 非翻译区, 来自异源基因或叶绿体基因例如 *psbA*, 这将保证编码酰基转移酶蛋白的 DNA 序列在叶绿体中转录和翻译; 一个编码酰基转移酶蛋白的 DNA 序列; 和一个转录和翻译终止区, 例如叶绿体基因的 3' 反向重复区, 能够稳定所表达的编码酰基转移酶蛋白的 mRNA。在叶绿体中表达可能增强基因产物的积累。含有叶绿体或质体的宿主细胞可以用所述表达盒转化, 然后可以培养所产生的含有所述转化叶绿体的细胞, 以表达所述酰基转移酶蛋白。表达盒除含有所述酰基转移酶基因外, 也可以包括抗生素、除草剂耐性或其它选择标记基因。所述表达盒可以邻接可以促进所述表达盒稳定整合、特别是通过同源重组整合到叶绿体基因组的得自叶绿体 DNA 的 DNA 序列。另一方面, 所述表达盒可以不整合, 但由于包括一个得自叶绿体 DNA 的复制起点, 将能够保证例如异源 *phnO* 或其它酰基转移酶基因在叶绿体中复制。

可以由含有转化叶绿体的细胞产生植株, 然后可以让植株生长以产生种子, 由所述种子可以产生另外的植株。这样的转化法优于核基因组转化, 特别是当通过整合到叶绿体基因组而实现叶绿体转化时, 因为叶绿体基因一般是母性遗传的。这提供了环境“更安全的”转基因植物, 实际上消除了逃逸到环境中的可能性。此外, 可以对叶绿体转化多次, 以产生表达多种所需重组蛋白的功能性叶绿体基因组, 而已经表明, 当需要多种基因时, 核基因组转化相当有限。采用叶绿体或质体转化因此避免了分离事件。与植物核基因组表达不同, 叶绿体

或质体中的表达可以仅由一个启动子起始，并且继续通过多顺反子区，由单一 mRNA 产生多种肽。

5 可以以构建其它植物转化载体的大致相同的方式产生所述表达盒。可以将植物叶绿体有效 DNA 序列插入细菌质粒中，并连接于表达所需基因产物(例如 PhnO 蛋白或其它相似的酰基转移酶)的 DNA 序列，使得所述酰基转移酶蛋白在叶绿体中产生，消除了核调节基因的核基因调节、加帽、剪接或聚腺苷酸化、或者叶绿体或质体引导肽序列的需要。对于包含或者合成构建或者为直接衍生自大肠杆菌基因组的天然基因的 *phnO* 或相似的酰基转移酶基因的表达盒，可以将其插入构建用于叶绿体或质体转化的载体的限制位点中。所述盒在上游可以邻接叶绿体或质体功能启动子，在下游可以邻接叶绿体或质体功能转录和翻译终止序列。可以用众所周知的同源重组法，将所产生的盒掺入叶绿体或质体基因组中。

10

15 另一方面，可以通过采用自主复制型质粒或能够在叶绿体或质体中繁殖的其它载体，获得叶绿体或质体转化。实现该方法的一种方法是利用叶绿体或质体复制起始所需的叶绿体或质体基因组的部分，作为在转化叶绿体或质体中维持所述质粒或载体的手段。通过将叶绿体或质体基因组随机克隆到也含有一个叶绿体或质体选择标记基因的标准细菌载体中，然后转化叶绿体或质体，并且在合适的选择培养基上选择转化细胞，可以容易地鉴定能够稳定复制叶绿体或质体外遗传元件的序列。将如本文所述的表达盒导入叶绿体或质体可复制外遗传元件中，将提供一种有效的方法，用于将酰基转移酶基因和蛋白定位至叶绿体或质体。

20

25 鉴于以上所述，人们会明白达到了本发明的几个优点，并且获得了其它有利的结果。因为可以在不偏离本发明的精神和范围的情况下在上述方法和组合物中进行各种改变，所以上述描述中包含的、和附图和序列中显示的所有要点将解释为是说明性的，而不是限制性的。

引用的参考文献:

- Avila 等, J. Am. Chem. Soc. 109:6758-6764, 1987 .
- Berlyn, Microbiol. Molec. Biol. Rev. 62:814-984, 1998 .
- Chen 等, J. Biol. Chem. 265:4461-4471, 1990 .
- 5 Dumora 等, Biochim. Biophys. Acta 997:193-198, 1989 .
- Franz, 草甘膦的发现、研制和化学, *The Herbicide Glyphosate*. E. Grossbard 和 D. Atkinson 编著, Butterworths. 第 3-17 页, 1985 .
- Hanley-Bowden 等, Trends in Biochemical Sciences 12:67-70, 1987 .
- Hilderbrand 等, *The role of phosphonates in living systems*; Hiderbrand, R.L. 编著, 第 5-29 页, CRC Press, Inc. Boca Raton, FL, 1983 .
- 10 Jacob 等, Appl. Environ. Microbiol. 54:2953-2958, 1988
- Jiang 等, J. Bacteriol. 177:6411-6421, 1995 .
- Kishore 等, J. Biol. Chem. 262:12,164-12,168, 1987 .
- Lacoste 等, J. Gen. Microbiol. 138:1283-1287, 1992 .
- 15 Lee 等, J. Bacteriol. 174:2501-2510, 1992 .
- Maier, Phosphorous Sulfur 14:295, 1983 .
- Makino 等, J. Bacteriol. 173:2665-2672, 1991 .
- McGrath 等, Eur. J. Biochem. 234:225-230, 1995 .
- 20 Metcalf 等, J. Bacteriol. 173:587-600, 1991 .
- Metcalf 等, Gene 129:27-32, 1993 .
- Ohtaki 等, Actinomyceteol. 8:66-68, 1994 .
- Pipke 等, Appl. Environ. Microbiol. 54:1293-1296, 1987 .
- Shinabarger 等, J. Bacteriol. 168:702-707, 1986 .
- Tanaka 等, J. Fac. Agr. Kyushu, Univ. 30:209-223, 1986 .
- 25 Wackett 等, J. Bacteriol. 169:710-717, 1987a
- Wackett 等, J. Bacteriol. 169:1753-1756, 1987b
- Wanner 等, FRMS Microbiol. Lett. 100:133-140, 1992 .
- Wanner 等, Biodegradation 5:175-184, 1994 .

Wohlleben 等, Mol. Gen. Genet. 217:202-208, 1989 .

引用的专利文献:

- Barry 等, 美国专利第 5,463,175 号, 1995 .
- 5 Barry 等, 美国专利第 5,633,435 号, 1997 .
- Comai , 美国专利第 4,535,060 号, 1985 .
- Daniell 等, 美国专利第 5,693,507 号; 1997 .
- Maliga 等, 美国专利第 5,451,513 号; 1995 .
- McBride 等, WO 95/24492 .

10

说明书附图

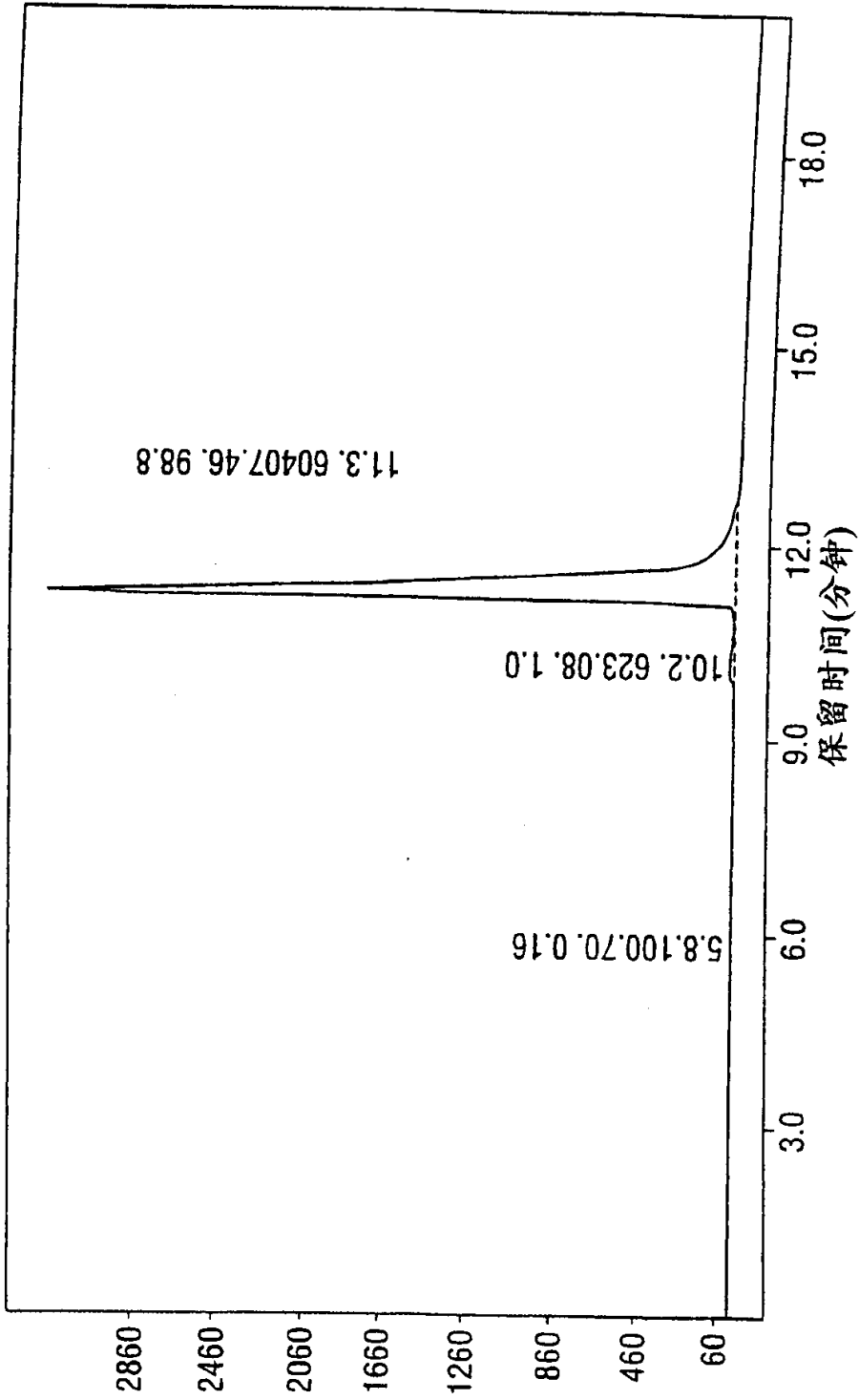


图 1

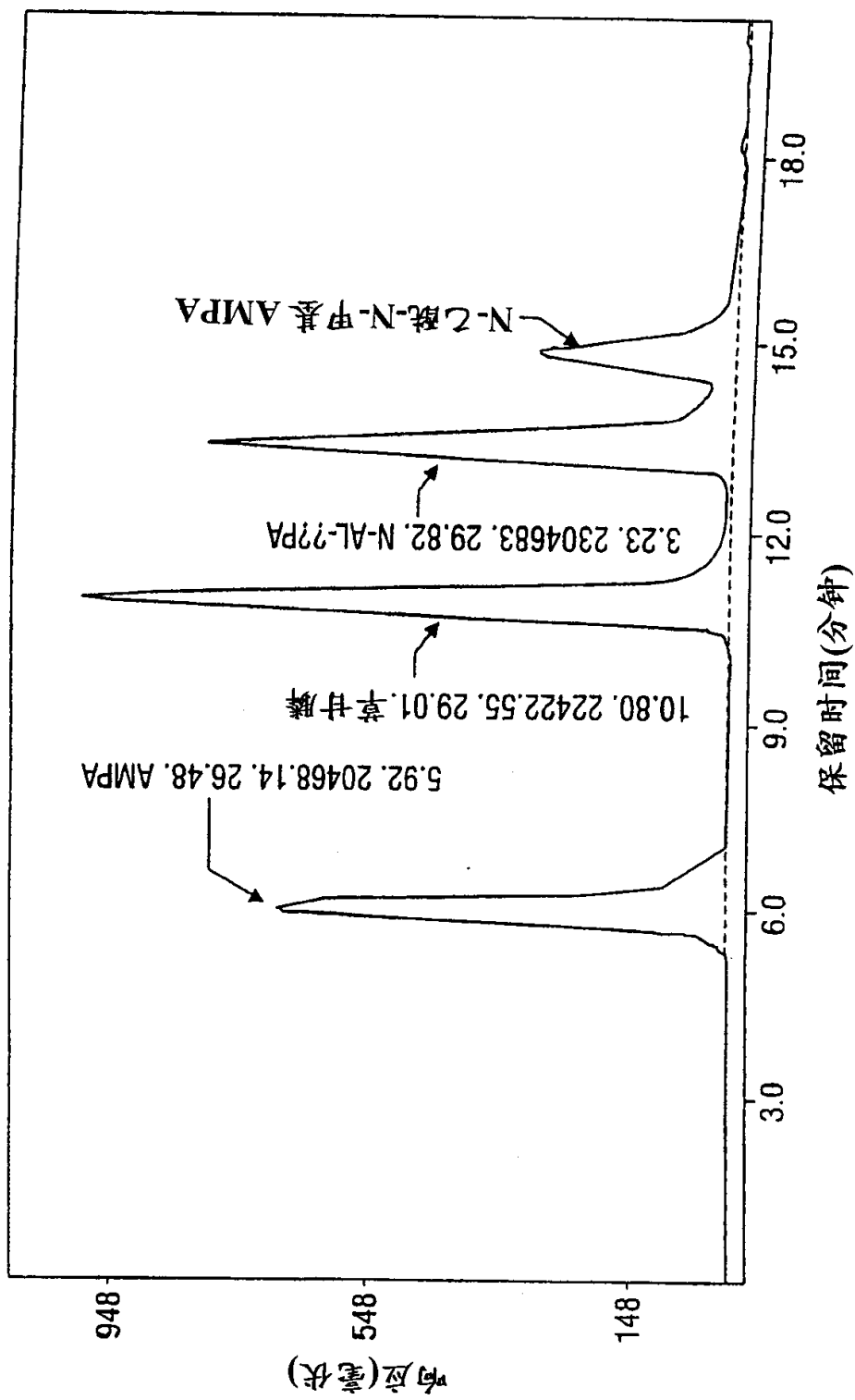


图 2

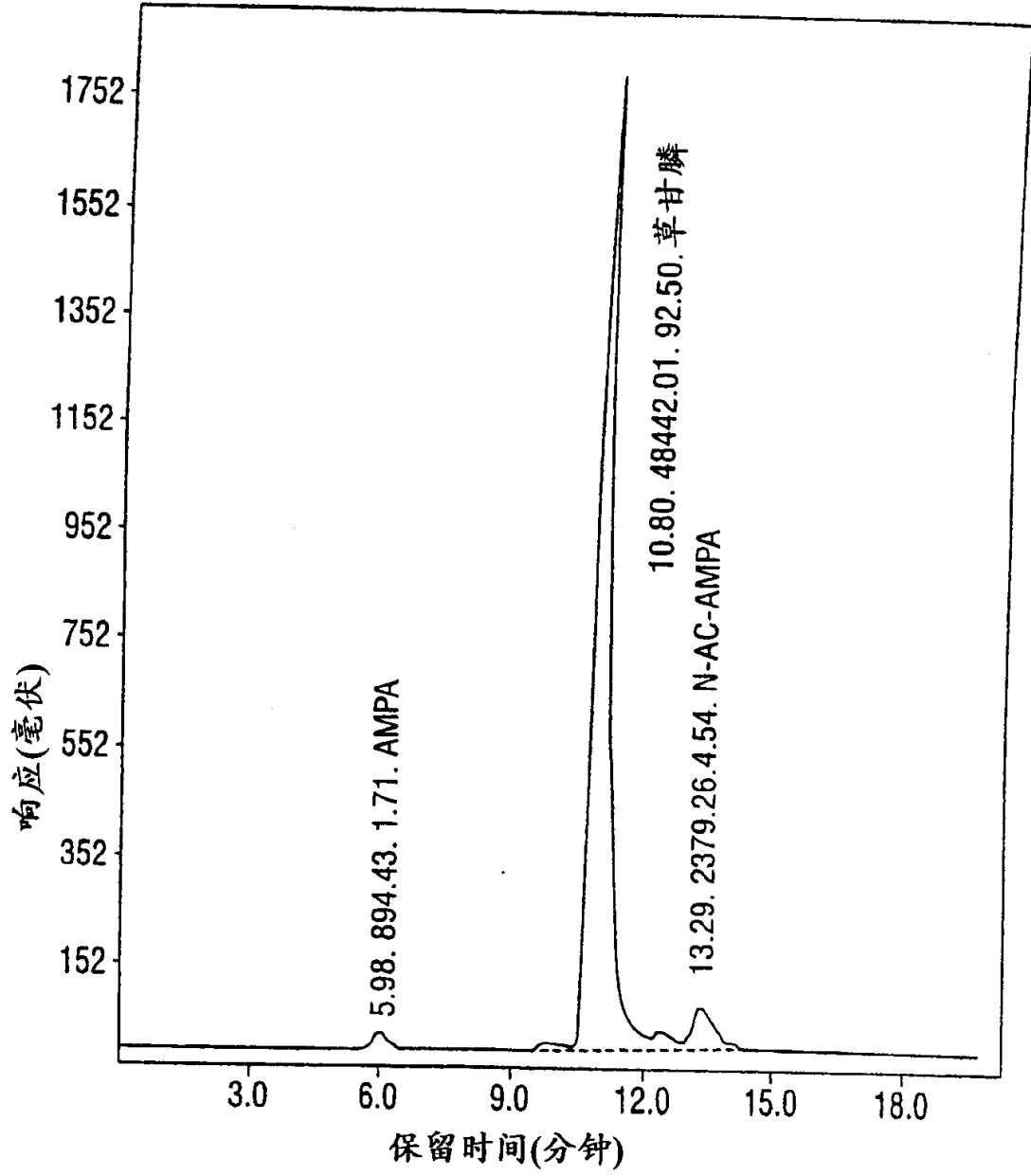


图 3

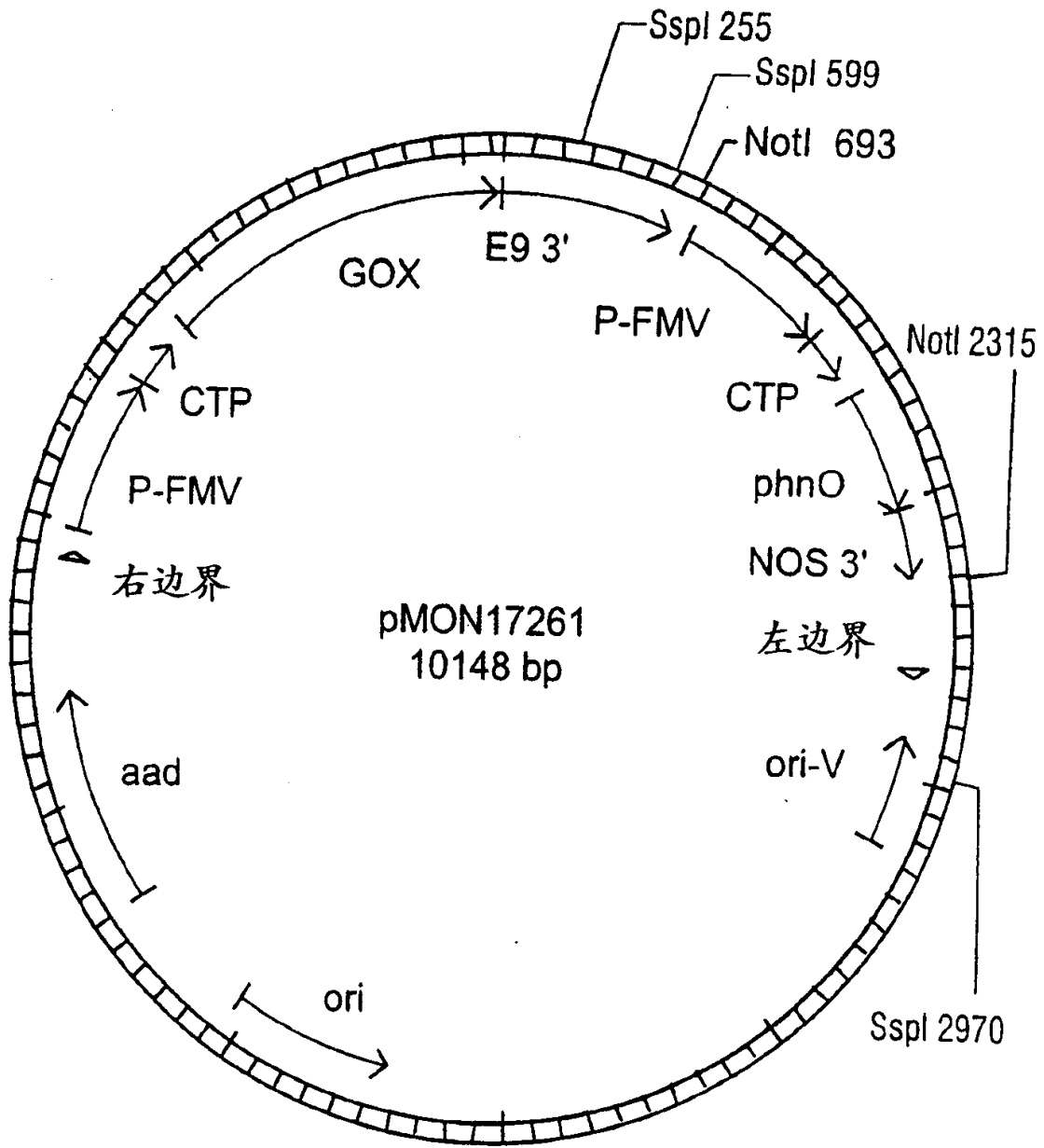


图 4

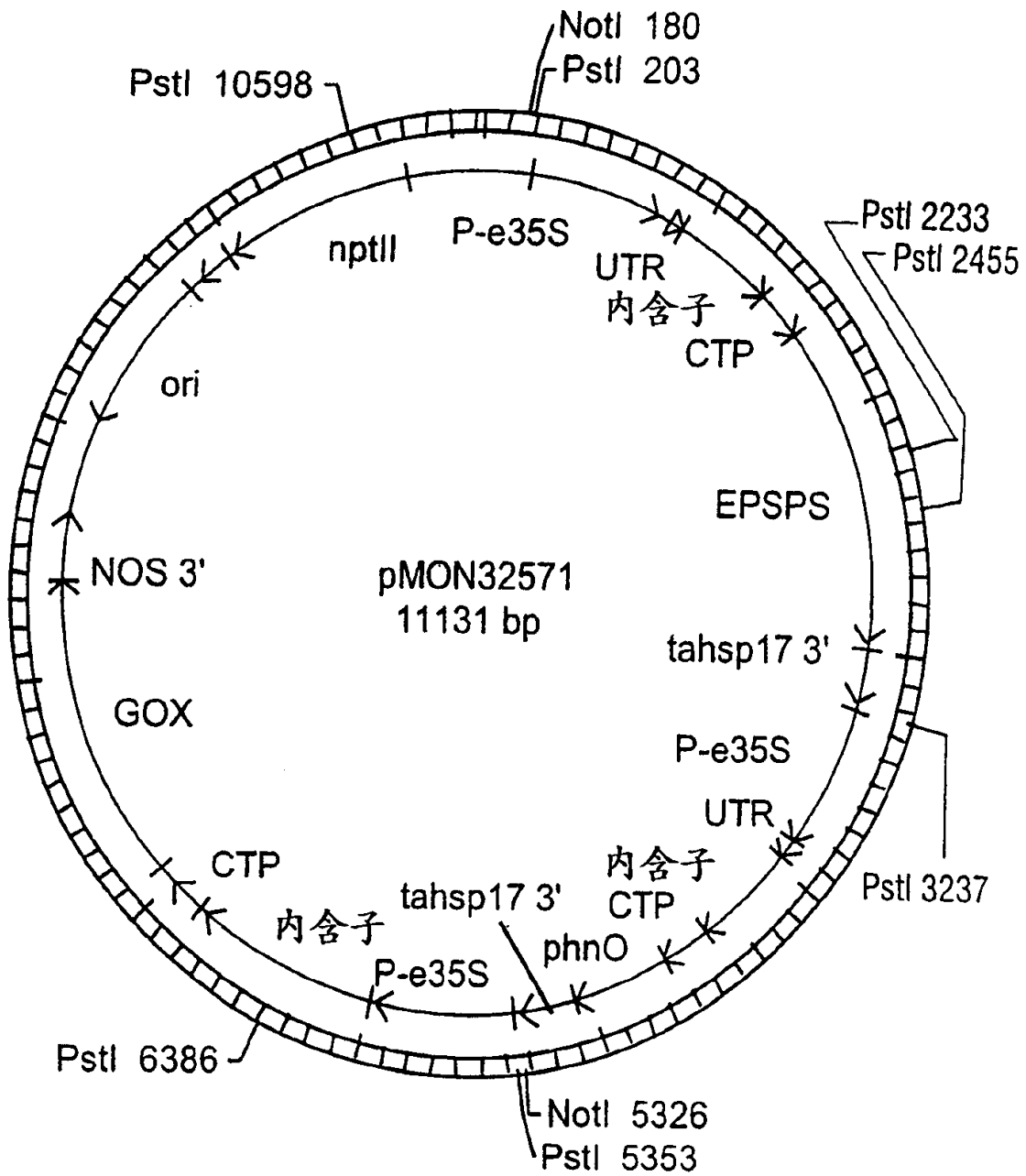


图 5

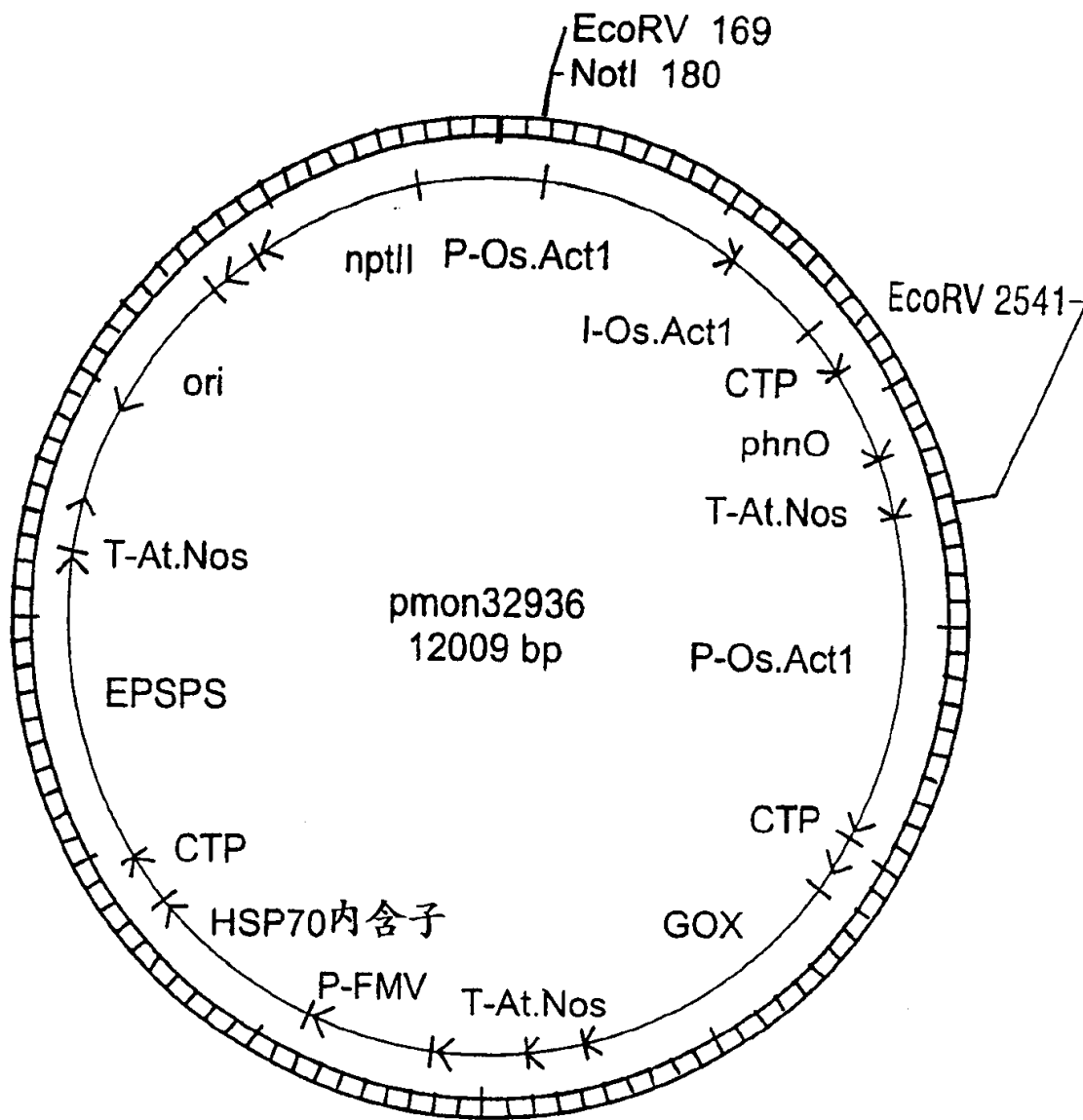


图 6

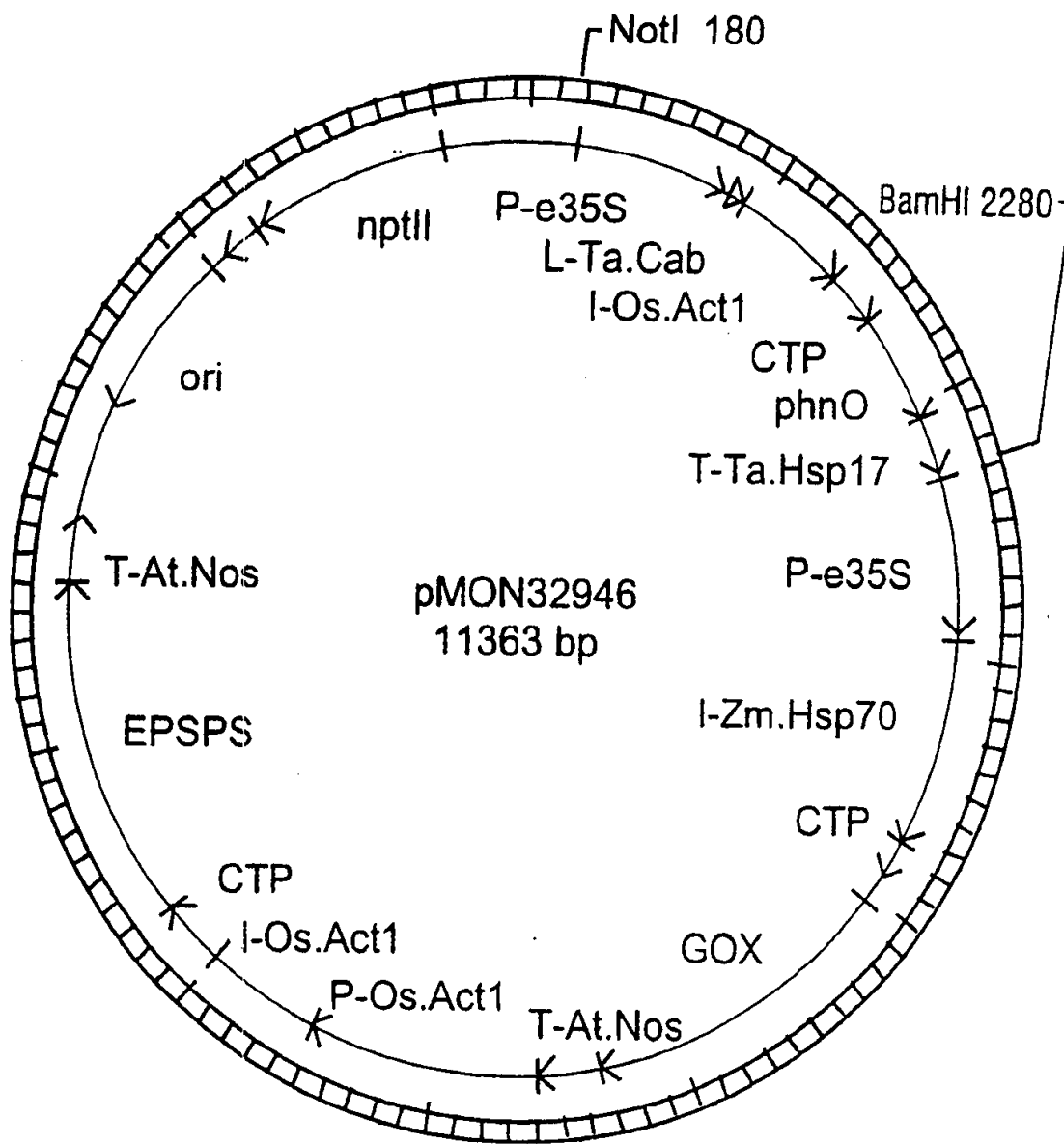


图 7

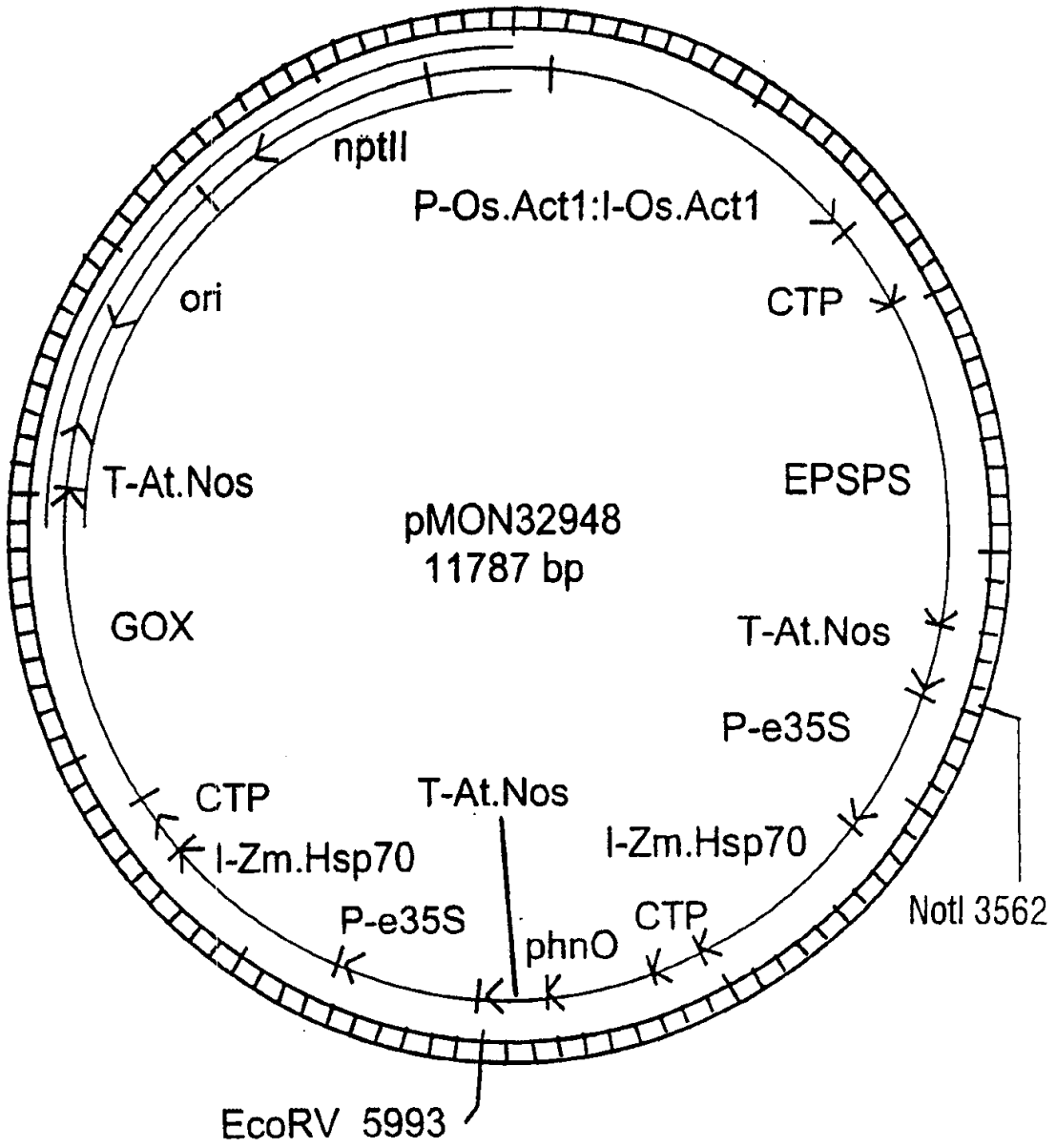


图 8

序 列 表

- <110> Barry, Gerard F.
- 5 <120> 磷酸代谢植物
- <130> 38-21(15303) Phn0
- <140> 未知
- 10 <141> 1999-11-16
- <150> 60/108, 763
- <151> 1998-11-17
- 15 <160> 32
- <170> PatentIn Ver. 2.0
- <210> 1
- 20 <211> 15611
- <212> DNA
- <213> 大肠杆菌(Escherichia coli)
- <400> 1
- 25 ggatccagca tgcagccag ttttccacc attgtcagtc gcaggctaag cggcgcattt 60
- aacatgccgc cgttcgtcca tgtctgaagc tgcacacgcg aaagaagttc ctgcatcagt 120
- 30 cgttcacgaa actgctgctg atgggcttgt ggaaggcggg catcatcgcc ctgcgccaga 180
- tccactaaaa agcggggata aaccgactcc agcatcgcac cggggccgtc cagtaacgtc 240
- ttggtcaata tcgttctgcc gtgaaaagtg tttgaatata atcgcgtaac agctgggcgt 300
- 35 cgggtgaaat ccagccgtga gtcacacag tctgctgcaa ttgctgctgc atcagcctga 360
- ccaccgatc attttgttga cgcagagcca ggctttcgcg taaacgcgtc tgtaattccg 420
- 40 tcaaacatga agcgaactca gcgaaaaaag tattcatgcc tgccgtaaca gattcatcga 480
- cctgctctgc cagaacttta gccatttgtt ggcaataaag atcgacttct gcgcttaatg 540
- ctcgttgcaa cacaactgtaa tcaaccgttt ctgtcgggga tttctcattt ccccgcccc 600
- 45 agtcgggctg attcaaccag cgcgaaaaag tctcacgcac aacgcctaaa cgcgtgctct 660
- gctcgtccgt tgcacactgg cgcgaaatga ctgcactgaa cagctggcga gtgttgaagt 720
- 50 ggggaactac gccgtgaaaa acaggaaaat gaaaccagg acgaaaccct gactcgtca 780

attccatfff gacttggtgc tcaatggggc gaataacatc ggtaaacact cggcaaagg 840
 tggattccag ctoggcaaaa cgcagcgtaa agtcgcgact gatggtgttc tgcgccgtct 900
 5 gtaacagtgt ctcacagcgg gtacgcatct cgcttaacgg ctccgaatca tcctgaaaca 960
 agggggctaa ctgcgcattc agcgcacatct gttggtgacg cagaaagtgg ttggcggagg 1020
 10 tcagggccag ctcgatttca tgtttaatct cgccgctcac ctgcgcctga ttgagttgca 1080
 atagctgcaa actttctctg acctgatgga tattttgccg caattgttca caagcgact 1140
 ttaaccctg cgcacgaaaa tccaggtatt cccgcgctg ctgcgcgtaa ttcaacagtt 1200
 15 tatgcccagc agatcgcaaa gcatacaacg aggcgcttagc gtaagcggca tgaagcaacg 1260
 cctgaattgg ctggcgcaac agcgaatctt cccacaactg atcggcagca tgacgaatat 1320
 gttcgaggtc cgccagatcg gcatgacgcc agcgcctgcc gagcgcggca tgggcaaaat 1380
 20 cttccacca gcgttggtgc tctggcgctg gtaacttacc gttgttgct aactcatggc 1440
 gcgcccgatt cgccaggtag cccacatcg acgacaccgg aaatatctgc tgtggcgtaa 1500
 25 tacagccttt catcagcgtc ccgaaatea gtgcccgcac ctggtcggcg tcgtcaactgt 1560
 tacgatcctg ttgatcgaac ttattgacca gcacatacag cggcaccgat tgccccaccg 1620
 ccaaaatgc ctcacggacc tcttcacgg agatcgattt cagttgcgta taatccagca 1680
 30 ccgccagtac cgccgaggcg cgtgccagct gctggttaag cattttttgc agatgcggtt 1740
 gcccgcttc atttgcccg ggggtatcca gtaacgtcaa ctgaccggga taactctcca 1800
 35 gccccgccag atggacaaac tccacttcaa tcacgggaat atgctcaatg gcggcgtaag 1860
 cagaaaaagg aaaatcgacg tccagcgcct tcgccagtcg cactaaatca ttcaaacttt 1920
 40 tcagacaatg aaaaatagcg tgggcacca gataatattt ttcgaaagcg acgccatfff 1980
 cgatccgctg cataagcga cgcatactt tatctattc cagcacatcg gtcagatgct 2040
 taatcgcga atcacgcagg cgctgttgta attgttgaat taaacaatcg attggcgcga 2100
 45 catgtgaaaa atgcagtacc gtttcctfff gccgggct atggcgaata agcgtcggca 2160
 gcgcagtcat tggcgatta cgattagca gaacctcctg accaacaatg gcattaatgg 2220
 tggttgattt cctgctttc atggtaccga caattgcaag caccatttc agtcgggaaa 2280
 50 ttttacgcaa ctcatctc agcatcgcgt gacgttcggc gatattagc tgactccagg 2340

gtaaagccag ttgtggcgcg tcgtctccgg gtacagagag aggcattttt tccagtaact 2400
 gcaactgttg gcgagaaagc tgtaacaggc gttcagcctc ctgacttaac tcatacaggg 2460
 5 tctgtgtgta catagaaaat tcttccttaa agcaaatttt gttattttat ttagccagat 2520
 tgtttttgag ttctgttttc ggcttttata attactgcaa gaaataattt tatatttagt 2580
 gtgtttgttt ttatcagaat aaataacgtc ttctgatagc tttaaaacgt cagaaagata 2640
 10 aaaatatcat gtgaattaaa aaaagaacaa gtagagcatt aacattatct taaataataa 2700
 atagaggcaa aaagattatt ttctttttgc gtttcctttc aatgaaaac gatcgtcgtc 2760
 15 taaaatcagc agtaccctccg acaaaactcag ggattttgtg tataattgcg gcctttttcg 2820
 gcaatctgcc gttttttggc gcttttgccc tgctgacttt tgaggaaatc cacatgtcat 2880
 taccacactg cccaaaatgc aactccgaat acacttacga agataacggc atgtacatct 2940
 20 gcccggaatg tgcctacgaa tggaacgacg cagaacctgc acaggaaagc gacgagctga 3000
 tcgttaaaga tgctaacggc aatctgctgg ctgacggcga cagcgttacc atcattaaag 3060
 25 atctgaaggt gaaaggtagc tcttcgatgc tgaaaattgg caccaaagtg aaaaacatcc 3120
 gcctggttga aggcgacat aacatcgatt gcaaaatcga cggttttggt ccgatgaaac 3180
 tgaaatctga gtttgtgaaa aagaactgat tgtattgtga tcggtaagcc ggataaggcg 3240
 30 ctgcgcccgc atccggcaac ggtgccagat gcctgatgcg acgcttgcgc gtcttatcag 3300
 gcctacaaat tcccgcaccc tccgtaggcc ggataaggcg tttacgccgc atccggcaac 3360
 35 ggtgccgact gcctgatgcg acgcttgcgc gtcttatcag gcctacaaat tcccgcaccc 3420
 tccgtaggcc ggataaggcg tttacgccgc atccggcaac agtgccaact gcctgatgcg 3480
 acgcttgcgc gtcttatcag gcctacaaat tcccgcaccc tccgtaggcc ggataaggcg 3540
 40 tttacgccgc atccggcaat ggtgccgact gcctgatgcg acgcttgcgc gtcttatcag 3600
 gcctacaaat tcccgcaccc tccgtaggcc ggataaggcg tttacgccgc atccggcaac 3660
 45 agtgccgact gcctgatgcg acgctcgcgc gtcttatcag gccgcctctc atctgtataa 3720
 atttgaact acacttaact ggcttctctt aactgaggtc accatcatgc cgtaagtcc 3780
 ctacctctct tttgccggta actgttccga cgcgattgcc tattatcaac gtacgttggg 3840
 50 cgcggaactg ctctataaaa tcagcttcgg cgaaatgcca aatcagcgc aggacagcgc 3900

cgagaactgc ctttccggaa tgcaatttcc cgataccgcc atcgctcatg ccaacgtgcg 3960
 cattgccgga agcgacatca tgatgagcga tgccatgccg tcaggaaaag ccagctactc 4020
 5 cggctttacg ctgggtgctcg attcgcaaca ggtcgaagaa ggaaaacgct ggtttgacaa 4080
 tcttgccgct aacggaaaaa tcgaaatggc ctggcaggaa actttctggg cgcatggctt 4140
 10 tggcaaagtc accgataaat ttggcgtacc gtggatgatt aatgctgta aacaacaacc 4200
 aacgcaataa cccgccggga ggccccct cccgactgt catcgaattc ccgttaactc 4260
 ttcatctgtt agtcactttt aattaaccaa atcgtcacaa taatccgcca cgatggagcc 4320
 15 acttttttag ggaggctgca tcatgcaaac gattatccgt gtcgagaagc tcgccaaaac 4380
 cttcaatcag catcaggcgc tgcattgctg tgatctgaac attcatcacg gtgaaatggt 4440
 ggctctgctt gggccgctcg gttccggcaa atccaccctt ttacgtcact taagcggtt 4500
 20 gattaccggc gataaatccg ccggcagcca tatcgagctg ctgggccgca cagtccagcg 4560
 cgaaggccgt ctggcgcgag atatccgcaa aagccgcgcc aacaccggt acatcttcca 4620
 25 acaattcaac ctggatgaac gcctgagcgt actggagaac gtgctgattg gcgctcgg 4680
 cagcacgccg ttctggcgca cctgttttag ctggtttacc cgcgagcaga aacaacgcg 4740
 gttacaggcg ctgaccgagc ttggcatggt gcattttgcc catcaacgagc tttccaccct 4800
 30 ctccggcggga cagcagcagc gtgtggcgat tgcccgcgag ctgatgcagc aggccaaggt 4860
 gattctggcc gatgaacca tcgcctcgtt ggaccggaa tccgccgca tcgtgatgga 4920
 35 caccctgagc gacatcaatc agaacgagc catcacctg gtcgtcacgc tgcacaggt 4980
 ggattacgcc ctgcctact gcgaacgat cgtgcctg cgccagggc acgttttcta 5040
 cgacggcagc agccaacagt ttgataacga acgttttgac catctctacc gcagcattaa 5100
 40 tcgcatcgaa gagaacgca aagctgcctg acatccccat cattgaggaa aacgaatgaa 5160
 cgctaagata attgcctcgc tggccttcac cagcatgtc agcctcagca ccctgttaag 5220
 45 cccggcacac gccgaagagc aggaaaaggc gctgaattc ggcatattt caacggaatc 5280
 acagcaaac ctgaaaccg aatggacgcc attcttacag gatattgaga agaagctggg 5340
 cgtgaaggtg aacgccttct ttgcccaga ctacgcagc attatccagg gaatgcgctt 5400
 50 caataaagtg gatatgcctt ggtacggcaa cctgtcggca atggaagcgg tggatcgcgc 5460

caacggccag gtcttcgccc agacggtcgc ggcggatgga tcgccaggtt actggagcgt 5520
 gttgatcgtc aacaaagata gtccgatcaa caacctgaac gatctgctgg cgaagcggaa 5580
 5 agatctcacc ttcggcaatg gcgatcctaa ctccacctct ggcttcctcg tccccggtta 5640
 ctacgtcttc gccaaaaaca atatctccgc cagcgacttc aagcgcaccg tcaacgccgg 5700
 gcatgaaacc aacgcgctgg ccgtcgccaa caagcaggtg gatgtggcga ccaacaacac 5760
 10 cgaaaacctc gacaagctga aaacctccgc gccggagaag ctgaaagaac tgaaagtgat 5820
 ctggaatcg ccgctgatcc caggcgatcc gatcgtctgg cgtaaaaatc tttccgaaac 5880
 15 caccaaagac aagatctacg acttctttat gaattacggc aaaacgccgg aagagaaagc 5940
 ggtgctggaa cgctgggct gggcgcgctt ccgcgcctcc agcgacctgc aactggtgcc 6000
 gattcgccag ctcgcaactgt ttaaagagat gcagggcgtg aaaagcaata aaggactgaa 6060
 20 tgagcaggac aagctggcaa aaaccaccgc gattcaggcg caactggatg acctggaccg 6120
 cctgaacaac gcgctaagcg cgatgagttc ggtgagtaaa gcggtgcagt aaatcgtagg 6180
 25 tcggataaga cgccccggcg tcgcatccga caatgtgcag gcgttgatgc cggatgcggt 6240
 gcaagcacct tatccggcct acagaccgga gccaaacatg caaacatca ccatcgcccc 6300
 acccaagcgc agctggttct cgcttctgag ctgggcccgtt gttctcgccg tgctggtcgt 6360
 30 ctctggcag ggcgcggaaa tggccccgct cacgctgatt aaagacggcg gcaacatgac 6420
 aaccttcgct gccgacttct tcccgccga tttcagccag tggcaggatt acctcaccga 6480
 35 aatggccgtc acgctgcaaa tcgccgtctg gggcaccgcg ctggcggttg ttctctccat 6540
 cccctttggc ctgatgagcg ccgaaaacct ggtgccgtgg tgggtttacc agcccgttcg 6600
 ccgcctgatg gacgcctgcc gcgccattaa cgaaatggtc ttcgcatgc tgttcgtggt 6660
 40 cgccgtcggc ctggaccgt tcgctggcgt gctggcgcta tttatccaca ccaccggcgt 6720
 gctctccaag ctgctttccg aagcggtaga agcaattgaa cctggccccg tggaaggcat 6780
 45 tcgcgccacc ggtccaaca agctcgaaga gatcctctac ggcgtgctgc cgcaggtgat 6840
 gccgtgctg atctctact ccctctatcg cttegaatcc aacgtccgct cggcgaccgt 6900
 cgtcggcatg gtcggcgcgg gcgggatcgg cgtcacctg tgggaagcga ttcgcggttt 6960
 50 ccagttccaa caaacctgac ccctgatggt gcttatcctc gtcacggtca gcctgctgga 7020

tttctctct caacggttgc gtaagcactt tatctgataa gcgaggcatt gatatctatg 7080
 cacttgctta cacatccgac cagctacca acacgctatc aagagatagc cgcaaaactt 7140
 5 gagcaggagc ttcgtcaaca ctaccgctgc ggcgactatc ttcccgccga gcagcaactg 7200
 gcagcgcgct ttgaggtgaa tcgccacacc ctgcgccgcg ccatcgacca actggtggaa 7260
 10 aaaggetggg tacagcgcgc tcagggcgtc ggcggtctgg tgctgatgcg cccgttcgat 7320
 taccgctca acgcccaggc gcgttttagc cagaatctgc tggatcaggg cagccatccc 7380
 accagcgaaa aactgctttc ggtattgctc cccgcgtccg gccacgtcgc tgacgcactg 7440
 15 gggattaccg agggggagaa cgtcatccac ctgcgcacc tgcgtcgggt caacggcgtc 7500
 gcgctctgtt taatcgacca ctacttcgcg gacctaccc tctggccgac gctgcaacgc 7560
 ttcgacagcg gctcgtgca cgattttctg cgcgagcaaa ccggaattgc gctgcgccgc 7620
 20 agccagacgc ggatcagcgc ccgcccgcgc caggccaaag agtgccagcg tcttgaatc 7680
 ccgaatatgt cgcgctgct gtgcgtgctc acccttaacc accgtgacgg tgaagcagc 7740
 25 ccggcggagt actccgtcag cctgacgcgc gccgacatga ttgaattcac tatggagcac 7800
 tgaatgcacg cagataccgc gaccgccag cactggatgt ccgtgctgac gcacagccaa 7860
 ccggctgaac tggcagcag cctgaacgcg ctaaacaatca ccgccgacta tgaggtgatc 7920
 30 cgcgccgctg aaactggcct ggtacagatt caggcgcgga tgggcggcac cggcgaactg 7980
 tttttgccg gcgacgccac gctgaccgc gccgccgtgc gcctgactga cggcacgctc 8040
 35 ggctacagct ggggtgctgg gcgtgataaa cagcacgccg aacgtgctgc gctgattgac 8100
 gcgctgatgc agcaatctcg ccacttcaa aacttatcag aaacccttat tgccccgctg 8160
 gacgctgacc gtatggcacg cattgcccga cgcaggccg aagtgaacgc cagccgggtc 8220
 40 gacttcttta cgatggttcg cggagacaac gcatgacct gaaaccgct tttatgcttc 8280
 ccgtgcagga tgcccagcac agttttctgc gcctgttaaa ggccatgagc gagccgggcg 8340
 45 tgattgctgc cctgcatcag ctcaaacgcg gctggcaacc gctgaatc gccaccacca 8400
 gcgtgctgct gacgctggcc gataacgaca cgcgggtgtg gctttctacc ccattaaata 8460
 acgatatcgt caaccagagc ctgcgttttc ataccaacgc gccgctggtc agccagcccg 8520
 50 aacaggcgac cttcgcggtg acggatgagg cgatttcag cgaacagctc aacgccctt 8580

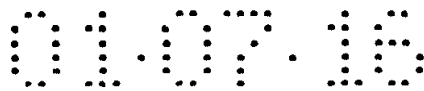
ccaccggcac gcgcgttgcg ccggaagcgg gcgcgacgct gattttacag gtcgccagcc 8640
tgagcggcgg gcgcatgttg cgtctcaccg gcgcgggtat tgccaagaa cgaatgatcg 8700
5 ctccgcagct gccggagtgc attctgcacg aactcaccga gcgcccgcac ccgttcccgc 8760
tcggcatcga cctgatcctg acctgcggcg aacgcctgct ggctattccg cgaaccacgc 8820
atgtggaggt gtgctgatgt acgttgccgt aaaagggggc gaaaaggcga tcgacgccgc 8880
10 ccacgccctg caagagagcc gacgccgggg cgataccgat ttgcctgaac tgagcgtcgc 8940
ccagattgaa cagcagctta acctcgcggt agatcgcgtg atgaccgaag gcggcattgc 9000
15 cgaccgcgaa ctggcggcgc tggcgtgaa acaggccagc ggcgataacg ttgaagcgat 9060
tttctgctg gcgcctacc gcaccacgtt ggcaagctg gcgtaagcg agccgctcga 9120
caccaccggg atgcgtctcg aacgccgat ctccgccgtt tataaagaca ttcccggcgg 9180
20 ccagctgctt ggcccaacct acgactacac ccatgcctg ctcgatttta ccctgctggc 9240
aaacggcgaa gcgccgacgc tgaccaccgc cgacagcgaa caacagccgt cgccgcacgt 9300
25 tttcagcctg ctggcgcgtc aggggctggc gaagtttgaa gaggatagcg gcgcacagcc 9360
ggatgacatc acccgcacgc cgccggttta cccctgctca cgttcttccc gtttgagca 9420
gttgatgcgc ggcgacgaag gctatttgct ggcgctggcc tactccacc agcgtggtta 9480
30 cggacgcaat caccgttcg cgggcgagat ccgcagtgtt tacatcgacg tgtcgattgt 9540
gccggaagag ctgggatttg cggtaaactg cggcgaacta ctgatgaccg agtgtgaaat 9600
35 ggtcaacggt tttatcgacc cgccgatga gccgccgcac ttcacgcgcg gctacgggct 9660
ggtattcggc atgagcgagc gcaaagcgat ggcaatggcg ctggtcgatc gtgcgttgca 9720
ggctccggaa tacggcgagc acgcgacagg cccggcgagc gatgaagagt ttgtgctggc 9780
40 acatgccgac aacgtcgaag ccgcaggctt tgtctcgac ctcaaactcc cccactacgt 9840
cgatttccag gccgaactgg agctactcaa acgtctgcaa caggagaaga accatggcta 9900
45 atctgagcgg ctacaacttt gcctacctg acgagcagac caaacgatg atccgccgcg 9960
ccatcttaaa agcggtgggc atccccggtt atcaggtgcc gtttggcggg cgcgagatgc 10020
cgatgccata cggctgggga accggcggca tacagctcac cgccagcgtg attggcgaaa 10080
50 gcgacgtgct aaagtgatt gaccaggtg cggatgacac caccaacgcc gtgtcgattc 10140

gcaacttctt taagcgcgtg accggggtaa acaccactga acgtacggac gatgcgacgc 10200
 ttatccagac gcgtcaccgc atccccgaaa cgccgctgac cgaagatcag atcattatct 10260
 5 tccaggtgcc aatccccgaa ccgctgcgct ttatcgagcc gcgcgaaacg gaaaccgca 10320
 ccatgcacgc gctggaagag tacggcgtga tgcaggtgaa actgtatgaa gatatcgccc 10380
 gcttcggta tatcgccact acctacgcct atccgggtgaa ggtgaacggg cgctacgtaa 10440
 10 tggaccgctc gccgatcccg aaattcgata acccaaaaat ggacatgatg cccgccctgc 10500
 aactgttcgg cgcggggcgc gagaagcga tctatgcggt gccgccgttt acccgcgtgg 10560
 15 aaagtctcga tttcgacgat cacccgttca ccgttcagca gtgggatgag ccatgcgcca 10620
 tctgcggatc gaccacagc tatcttgatg aagtgggtct ggatgacgcc ggaaaccgca 10680
 tgtttgtctg ctccgatacc gattatgcc gccaacagag cgaggcaaaa aaccaatgaa 10740
 20 tcaaccgtta ctttcggtca ataacctgac ccaccttac gcgccgggca aaggcttag 10800
 cgatgtctct tttgatttat ggccggggga agtgctgggc attgtcgggg aatccgctc 10860
 25 cgggaagacc acgtgctga agtcgatctc cgcgcgcctg acgccgcagc aggggggaaat 10920
 tcaactagag aaccgttcgc tgtatgcaat gagcagggcc gaccgccgtc gcctgctgcg 10980
 taccgaatgg ggcgtggtgc atcagcatcc actcgacggc ctgcgccgcc aggtgtcggc 11040
 30 aggcggcaat atcggcgagc ggctgatggc gaccggggca cgtcattacg gcgatattcg 11100
 tgccaccgcg cagaagtggc tggaagaggt ggagattccc gccaacggga tcgacgacct 11160
 35 gccgaccacc tttccggcg gtatgcagca gcgtttgcag attgcccgca acctggtgac 11220
 gcatccgaag ctggtgttta tggatgaacc gaccggcggg ctggatgtgt cgggtcaggc 11280
 ccgcctgctc gacctgctgc gcggcctggt ggtggagctg aacctcgcgg tggtgattgt 11340
 40 cacccatgat ttaggcgtcg cccgcctgct ggccggaccgt ttgctggtga tgaagcaggg 11400
 gcaagtggtg gagagtgggt taaccgaccg cgtgctcgac gaccgcacatc atccgtatac 11460
 45 acagctgctg gtgtcatcgg ttttcagaa ttgagccggt gccggatgcg gcgtaaacgc 11520
 cttatccggc ctacaaatgc gctccccgta ggtcggataa gacgcgtcag cgtecatcc 11580
 gacaccgaa ccacgaggcg aaaaatgatt aacgtacaaa acgtcagtaa aaccttcac 11640
 50 ctgcaccagc aaaacggcgt gcgcctgccc gtctcaatc gcgcctcgtc caccgtcaac 11700

gcgggcgaat gcgtggtgct ccacggccat tccggcagcg gcaaatcaac tctgctacgc 11760
 tcgctgtacg ccaactatct acccgacgaa ggtcaaatcc agatcaaaca cggtgacgag 11820
 5 tgggtagacc tggtcaccgc gccagcgcgc aaagtgggtg aaatccgcaa aaccaccgtc 11880
 ggctgggtga gccagtttct gcgcgtcadc ccgcgtatct cagcactgga agtgggtgatg 11940
 cagccgctgc tcgataccgg cgttccgcgt gaagcctgcg ccgctaaagc cgcgctctt 12000
 10 ctcaccgcc tgaacgtgcc ggaacgcctg tggcacctgg caccatcgac attttccggt 12060
 ggcgaacagc agcgcgtcaa catcgcccgc ggctttatcg tcgactacc cattctgctg 12120
 15 cttgacgaac ctaccgcctc gctggacgcc aaaaacagcg ccgcggtggt ggaactgatt 12180
 cgcgaagcca aaaccgctgg cgcagccatc gtaggcatct tccatgacga agctgtacgt 12240
 aatgacgtcg ccgaccgct gcaccaatg ggagcctctt catgattatc aataacgta 12300
 20 agctggtgct ggaaaacgag gtgtaagcg gttcgctgga ggtgcagaac ggcgaaatcc 12360
 gcgcctttgc cgaaagccag agccgcctgc cggaggcgat ggacggcgaa ggcggtgctg 12420
 25 tgctgccggg gctgattgag ctgcataccg ataactgga taaattcttc accccgcc 12480
 cgaaagtga ctggcctgcc cactcggcga tgagcagcca cgacgcgctg atggtggcga 12540
 gcggcatcac caccgtactg gatgccgtgg caattggcga cgtgcgcgac ggcgcgatc 12600
 30 ggctggagaa tctggagaag atgatcaacg ccatcgaaga gacgcagaaa cgcgcgctca 12660
 accgcgccga gcaccgtctg catctgcgct gcgaactgcc gcatcacacc acgctgccgc 12720
 35 tgtttgaaaa actggtgcag cgcgagccgg tgacgctggt gtcgctgatg gaccactcgc 12780
 cgggccagcg ccagttgcc aaccgcgaga agtatcgcga atattatcag ggcaaatact 12840
 ccctcactga tgcgcagatg cagcagtacg aagaagagca actggcgctc gccgcacgt 12900
 40 ggtcgcagcc gaatcgcga tccatgcgcg ccctgtgccg cgcgcgaaaa attgcgctt 12960
 ccagccacga tgacgccacc cacgccacg ttgctgaatc tcaccagctt ggcagcgtga 13020
 45 tcgccaatt tcccaccacg ttcgaagcgg cggaagcctc gcgcaagcat ggcataaac 13080
 tgctgatggg cgcgccaat attgtgcgcg gcggctcga ctccggcaac gtggcgcca 13140
 gtgaactggc gcagcttggc ctgctggata tctctcttc cgactactac cccgccagc 13200
 50 tgctcgatgc ggcatttgc gtcgccgatg acgagagcaa ccgctttacg ctgccgagg 13260

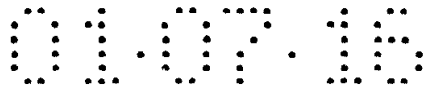
cgggtgaagct ggtgactaaa aatccagcgc aggcgcttaa tctccaggat cgcggggtga 13320
 ttggcgaggg caaacgcgcc gacctggtgc tggcgcatcg caaggacaat catattcata 13380
 5 tgcaccacgt ctggcgctcag ggtaaaaggg tgttctgatg atgggaaaac tgatttggtt 13440
 aatggggccg tccggctccg ggaaagacag cctgctggcg gaactccgcc tgcgggaaca 13500
 aactcagtta ctgggtggcg atcgtacat cacgcgcgat gccagcgcg gaagtgaaaa 13560
 10 ccatatcgcc ctgagcgagc aggagtttt taccgcgcg gggcaaaatc tgttggcctt 13620
 aagctggcac gctaacggtc tgtattatgg cgtcggcgtc gagattgatc tctggctgca 13680
 15 cgccggattc gacgtgctgg tcaacggctc acgcgccat ctgccgcagg cgcgggcg 13740
 ctatcaatcg gcgctgctgc ccgtctgtt acaggtttcg ccggagatcc tccgccagcg 13800
 cctggaaaac cgtggccgtg aaaacgccag tgaattaac gccgcctgg cgcgcgccg 13860
 20 ccgtataact ccacaggatt gccatacgt caacaatgac ggcagcctgc gccagtcggt 13920
 cgacacgctg ctgacgctga tccatcagaa ggagaaacac catgcctgct tgtgagcttc 13980
 25 gcccggccac gcagtacgac accgacgcgg tttacgcgt gatttgtgag ctaaacag 14040
 cggagttga ccaccacg 14100
 30 tgcgctacca tctggcgctg cttgatggcg aagttgtcgg catgatcggc ctgcatttgc 14160
 agtttcatct gcatcatgct aactggatcg gcgaaattca ggagttggtg gtaatgccg 14220
 aggcgcgcgg tctgaacgct ggcagtaagt tactggcgtg ggcagaagaa gaagcccgc 14280
 35 aggccggggc cgaatgacc gaactttcga ccaacgtgaa gcgccacgac gcgcaccgtt 14340
 tctatctgcg cgaaggctac gagcagagcc acttccgctt caccaaggcg ctgtaacatg 14400
 agcctgacct tcacgctcac cggcaccggc ggcgcacagg gcgttccggc atggggctgc 14460
 40 gagtgtgcgg cctgcgccag agcgcggcgc tcgccgagc atgccgcca accgtgcagc 14520
 ggcgtagtga agtttaacga cgcaatcacc ctgatcgacg ccgggctgca cgatctgcc 14580
 45 gatcgtggt cccccgatc gttccagcag tttttgctga cgcattatca tatggatcac 14640
 gtccaggggc tgtttccgct gcgctggggc gttggcgatc cgatcccgtt ttacggccc 14700
 ccgatgaac agggctgcga cgatctgtt aaacatccg gcctgcttga tttcagccac 14760
 50 acggtggaac cgtttgtggt gtttgattg caggggttac aggtcacgcc cctgccgctc 14820

aaccactcaa aactgacctt cggttatctg ctggaaacgg cacacagccg ggtggcgtgg 14880
ctgtctgaca ccgcaggctt gccggaaaaa acgctgaaat ttttacgcaa taatcagccg 14940
5 caggtaatgg tgatggattg cagtcaccgg ccgcgcgegg atgcaccgcg taatcaactgt 15000
gatttaaata ccgtgcttgc gctgaatcag gttatccgct cgccacgggt gattctgacc 15060
catatcagcc accagtttga tgcgtggctg atggaaaacg cactaccgtc agggtttgag 15120
10 gtggggtttg atgggatgga gattgggggtg gcgtgatgag agggaatgtg cgcgctggcc 15180
ccctcacctt aacctctcc ccagaggggc gaggggaccg attgtgctcg atattgaata 15240
15 ttgcgctcgt tttctccctc tccccattgg ggtgaggggc gatgcctgct ccatacccaa 15300
cctcatcgcc catactcacc ttccattctc cgtcttccat cctccagttg ccgacgetcc 15360
tgatcaagct ggcgctggcg atcgtccagc tgctgcggc gatcttcaaa ctggcggcgg 15420
20 cggtcgcat attgtctcgc ccgatcgtcg ctcacttcac gctgccagcc gtggtcgcgc 15480
gaatcttcat agttgaagcg gcgcacgaaa aacgcgaaag cgtttcacga taaatgcgaa 15540
25 aactttagct ttcgcgcttc aaatgaaaca gatgtattaa ttactgcttt ttattcatta 15600
catggggatc c 15611
30 <210> 2
<211> 11672
<212> DNA
<213> 大肠杆菌
35 <400> 2
gaattcccgt taactcttca tctgtagtc acttttaatt aaccaaactg tcacaataat 60
ccgccacgat ggagccactt ttttagggag gctgcatcat gcaaacgatt atccgtgtcg 120
40 agaagctcgc caaaccttc aatcagcacc aggcgctgca tgcggttgat ctgaacattc 180
atcacggtga aatgggtgct ctgcttgggc cgtcgggttc cggaaaatcc acccttttac 240
gtcacttaag cggtttgatt accggcgata aatctgtcgg tagccatata gagctgctgg 300
45 gccgcacagt ccagcgcgaa ggccgcctgg cccgcgatat ccgcaaaagc cgcgcccata 360
ccggctacat attccaacaa ttcaacctgg tgaaccgctt gagcgtactg gagaacgtgc 420
50 tgattggcgc gctcggcagc acgcccctt ggcgcacctg ttttagctgg ttcaccggcg 480
agcagaaaca gcgcgcgcta caggcgtga cccgcgttgg catggtgcat tttgccacc 540



agcgcgtttc caccctctcc ggcggccagc agcaacgtgt ggcgattgcc cgtgcgctga 600
tgcagcaggc gaaagtgatt ctggccgatg aaccctatgc ctcgctggac ccagaatcag 660
5 cgcgcatcgt gatggacacc ctgcgcgaca tcaaccagaa cgacggcatc accgtggtcg 720
tcacgctgca tcaggtggat tacgccctgc gctactgcga acgcatcgtc gccctgcgcc 780
10 aggggcacgt cttctacgac ggcagcagcc aacagtttga taacgaacgt tttgaccatc 840
tctaccgcag cattaaccgc gtcgaagaga acgcgaaagc tgctgacat ccccatcatt 900
gaggaaaacg aatgaacgct aagataattg cctcgtggc cttaccagc atgttcagcc 960
15 tcagcacctt gtaagcccg gcgcacgccg aagagcagga aaaggcgttg aatttcggca 1020
ttatttcaac ggaatcacag caaaacctga aaccgcaatg gacgccgttc ttgcaggata 1080
20 tggagaagaa gctgggcgtg aagtcaacg ctttctttgc cccggactac gcgggcatta 1140
tccaggggat gcgcttcaat aaagtggata tcgcctggta cggcaatctg tcggcgatgg 1200
aagcgggtgga tcgcgccaat gccaggtct tcgccagac gtcgcggcg gatggatgc 1260
25 cgggttactg gagcgtgttg atcgtcaaca aagacagtcc gatcaacaac ctgaacgatc 1320
tgctggcgaa gcgaaagat ctacaccttg gcaatggcga tctaactcc acctctggct 1380
30 tcctcgtccc cggctactac gtcttcgcca aaaacaatat ctccgccagc gacttcaagc 1440
gcaccgtcaa gcgccggcat gaaaccaacg cgctggccgt gcccaacaag caggtggatg 1500
ttgccaccaa caacaccgaa aacctcgaca agctgaaaac ctccgcgcca gagaagctga 1560
35 aagaactgaa ggtgatctgg aagtcgccg tgatcccagg cgatccgatc gtctggcgca 1620
agaatctttc cgaaaccacc aaagacaaga tctacgactt ctttatgaac tacggcaaaa 1680
40 cgccggaaga aaaagcggtg ctggaacgcc tgggctgggc gccattccgc gcttccagcg 1740
acctgcaact ggtgccgatt gccagctcg cgctgtttaa agagatgcag ggcgtgaaaa 1800
gcaataaagg actgaatgag caggacaagc tggcaaaaac caccgagatt caggcgcagc 1860
45 tggatgacct ggaccgcctg aacaacgcgc taagcgcgat gagttcggtg agtaaagcgg 1920
tgcagtaaat cgtaggtcgg ataagacgcc cggcgtcgc atccgacaat gtgcaggcgt 1980
50 tgatccgga tgcggtgcaa gcaccttacc cggcctacag accggagcca aacatgcaaa 2040
ccatcaccat cgccccacc aagcgcagct ggttctcgt tctgagctgg gccgttgtag 2100

tcgccgtggt ggctgctcgc tggcagggcg cggaaatggc cccgcttacg ctgatcaaag 2160
 acggcggcaa catggcgacg ttcgccgcgc acttcttccc gcccgatttc agccagtggc 2220
 5 aggattacct caccgaaatg gccgtcacgc tgcaaatcgc cgtctggggc accgcgctgg 2280
 cgggtggttct ctccatcccc tttggcctga tgagcgcga aacctgggtg ccgtgggtggg 2340
 10 tttaccagcc cgttcgccgc ctgatggacg cctgccgcgc cattaacgaa atggtcttcg 2400
 ccatgctggt cgtggtcgcc gtcggcctcg gcccgttcgc tggcgtgctg gcgtgctggc 2460
 gctgtttatc cacaccaccg gcgtgctctc caagctgctt tccgaagcgg tggaagcgat 2520
 15 tgagcccggc ccggtggaag gcattcgcgc caccggtgcc aacaagctcg aagagatcct 2580
 ctacggcgtg ctgccacagg tgatgccact gctgatctcc tactcctct atcgcttcca 2640
 20 atccaacgtc cgctcggcga ccgtcgtcgg catggtcggc gcaggcggga tcggcgtcac 2700
 cctgtgggaa gcgattcgcg gttccagtt ccaacaacc tgcgccctga tgggtcttat 2760
 catcgtcacg gtcagcctgc tggatttct ctctcaacgg ttgcgtaage actttatctg 2820
 25 ataagcgagg cattgatatc tatgcacttg tctacacatc cgaccagcta cccaacacgc 2880
 tatcaagaga tagccgcaa acttgagcag gagcttcgtc aacactaccg ctgcggcgac 2940
 30 tatcttcccg ccgagcagca actggcagcg cgctttgagg tgaatcgcca cacctgcgc 3000
 cgcgccatcg accaactggt ggaaaaaggc tgggtacagc gccgtcaggg cgtcggcgtg 3060
 ctggtgctga tgcgccggtt cgattaccg ctcaacgcc aggcgcggtt tagccagaat 3120
 35 ctgctggatc agggcagcca tcccaccage gaaaaactgc tttcgggtatt gcgccccgcg 3180
 tccggccacg tcgctgacgc actggggatt accgaggggg agaacgtcat ccacctgcgc 3240
 40 acctgcgtc gtgtcaacgg cgtcgcgctc tgtttaatcg accactactt cgcggacctc 3300
 acctctggc cgacgtgca acgcttcgac ageggctcgc tgcacgattt tctgcgcgag 3360
 caaacggaa ttgcgctcgc ccgagccag acgcggatca gcgcccgcg cggccaggcc 3420
 45 aaagagtgc agcgtcttga aatcccgaat atgtcgcgc tctgtgtgct gcgcacctt 3480
 aaccaccgtg acggtgaaag cagcccggcg gagtactccg tcagcctgac gcgcgccgac 3540
 50 atgattgaat tcaactatga gcaactgatg cacgcagata ccgcgacctg ccagcactgg 3600
 atgtccgtgc tggcgcacag ccaaccggt gaactggcag cacgcctgaa cgcgctaaac 3660



atcaccgccg actatgaggt gatccgcgcc gctgaaactg gcctggtaca gattcaggcg 3720
cggatgggcg gcaccggcga acgttttttt gccggcgacg ccacgctgac ccgcgccgc 3780
5 gtgcgctga ctgacggcac gctcggctac agctgggtgc agggcgctga taaacagcac 3840
gccgaacgct gcgcgctgat tgacgcgctg atgcagcaat ctgccactt tcaaaactta 3900
10 tcagaaacc ttattgcccc gctggacgct gaccgatgg cacgattgc cgcacgccag 3960
gccgaagtga acgccagccg ggtcgacttc tttacgatgg ttcgcggaga caacgcatga 4020
ccctggaac cgcttttatg cttcccgtgc aggatgccc a gcacagttt cgtcgctgt 4080
15 taaaggccat gagcgagccg ggcgtgattg tcgcccgtca tcagctcaaa cgcggctggc 4140
aaccgctgaa tatcgccacc accagcgtgc tgctgacgct ggccgataac gacacgccg 4200
20 tgtgctttc taccataa aataacgata tcgtcaacca gagcctgctg tttcatacca 4260
acgcgccgct ggtcagccag ccggaacagg cgacctcgc ggtgacggat gaggcgattt 4320
ccagcgaaca gctcaacgcc cttccaccg gcaccgccgt tgcgccgaa gcgggtgca 4380
25 cgctgatttt acaggtgcc agcctgagcg gcggacgcat gttgcgctt actggtgcg 4440
gtattgccga agaacgaatg atcgtccgc agctgccgga gtgcattctg cacgaactca 4500
30 ccgagcgcgc gcattccgtc ccgctcggca tcgacctgat cctgacctgt ggcgagcgc 4560
tgctgctat tccggaacc actcatgtgg aggtgtgctg atgtacgtt ccgtgaaagg 4620
ggcgagaag gcgatcgacg ccgcccacgc cctgcaagag agccgacgcc gaggcgatac 4680
35 cgattgccc gaactgagcg tcgcccagat tgaacagcag ctaacctcgc cggtagatcg 4740
cgtgatgacc gaaggcggca ttgccgaccg cgaactggcg gcgctggcgc tgaacaggc 4800
40 cagcggcgat aacgttgaag cgattttcct gctgcgcgcc taccgcacca cgttggcgaa 4860
gctggcggta agcgagccgc tcgacaccac cgggatgctg ctgcaacgcc gtatctccgc 4920
cgtttataaa gacattcccg gcggccagct gcttgccca acctacgact acacccatcg 4980
45 cctgctcgat ttaccctgc tggcaaacgg cgaagcgcgc acgctgacca ccgccgacag 5040
cgaacagcag ccgtcgcgc acgttttcag cctgctggcg cgtcaggggc tggcgaagtt 5100
50 tgaagaggat agcggcgcac agccggatga catcaccgc acgccgccg tttaccctg 5160
ctcacgctcc tccgtttgc agcagttgat gcgcggcgac gaaggctatt tgctggcgt 5220

ggctactcc acccaacgcg gttacgggcg caatcacccg ttcgcaggcg agatccgcag 5280
 5 cggctatatac gacgtgtcga ttgtgccgga agagctggga ttgtcgggtga acgtcggcga 5340
 actgtgatg actgagtgtg aatgggtaa cggttttatc gaccgcgagg gtgagccgcc 5400
 gcacttcacg cgcggctacg ggctgggtgtt cggcatgagc gagcgcaaag cgatggcgat 5460
 10 ggcgctggtc gaccgcgctc tgcaagcccc ggagtacggc gagcacgca caggccccgc 5520
 gcaggatgaa gagttcgtgc tggcacatgc cgacaacgtc gaagccgcag gctttgtctc 5580
 acacctcaaa ctccccact acgtcgattt ccaggccgaa ctggagctac tcaaactct 5640
 15 gcaacaggag cagaacctg gctaattcga gcggctacaa cttgcctac ctgcacgagc 5700
 agaccaaacg catgatccgc cgcgccatct taaaagcggg ggcgatcccc ggttatcagg 5760
 20 tgccgtttgg cgggcgcgag atgccgatgc cgtacggctg gggaaccggc ggcattcagc 5820
 ttaccgccag cgtgattggc gaaagcgagc tgctgaaggt gattgaccag ggcgcggatg 5880
 acaccaccaa cgcggtgtcg attcgcaact tcttcaagcg cgtgaccggg gtaaaccaca 5940
 25 cggaacgtac ggacgatgcg acggttatcc agacgcgtca ccgcatcccc gaaacgccgc 6000
 tgaccgaaga tcagataatt atcttcagg tgccaatccc cgagccgctg cgctttatcg 6060
 30 agccgcgcga aacggaaacc cgcacatgc acgcgctgga agagtacggc gtgatgcagg 6120
 tgaactgta tgaagatac gcccgcttcg gtcaatcgc caccacctac gcctatccgg 6180
 tgaaggtaaa tgggcgctac gtgatggacc cgtgcgcat cccgaaatc gataacccaa 6240
 35 aatggacat gatccccgc ctgcaactgt tcggcgcggg gcgcgagaag cgcattctatg 6300
 cgggtccgcc gtttaccgc gtgaaagtc tcgatttcga cgatcacccg ttcaccgttc 6360
 40 agcagtggga tgagccatgc gccatctcgc gatcgacca cagctatctt gatgaagtgg 6420
 tgctggatga cgccggaac cgcattgttg tctgctccga taccgattat tgccgccaac 6480
 agagcgaggc aaaaaaccaa tgaatcaacc gttactttcg gtcaataacc tgaccacct 6540
 45 ttacgcgccc ggcaaaggt ttagcgatgt ctctttgat ttatggccgg gggaagtgt 6600
 gggcattgtc ggggaatccg gctccgggaa gaccacgctg ctgaagtcga tctccgcgcg 6660
 50 cctgacgccc cagcaggggg aaattcacta cgagaaccgt tcgctgtatg caatgagcga 6720
 ggccgaccgc cgtcgcctgc tgcgtaccga atggggcgtg gtgcatcagc atccactcga 6780

cggcctgcgc cgccagggtg cggcaggcgg caatatcggc gagcggctga tggcgaccgg 6840
 5 ggcacgtcat tacggcgata ttctgtccac cgcgcagaag tggctggaag aggtggagat 6900
 tcccgccaac cggatcgacg acctgccgac caccttttcc ggcggtatgc agcagcgttt 6960
 gcagattgcc cgcaacctgg tgacgcatcc gaagctgggtg tttatggatg aaccgaccgg 7020
 10 cgggctggat gtgtcgggtc aggcccgcct gctcgacctg ctgcgcggcc tgggtggtga 7080
 gctgaacctc gcggtggtga ttgtcaccca tgatttaggc gtcgcccgcc tgctggcgga 7140
 ccgtttgctg gtgatgaagc aggggcaagt ggtggagagt gggttaaccg accgcgtgct 7200
 15 cgacgaccgg catcatccgt atacacagct gctgggtgca tcggttttgc agaattgagc 7260
 cggtgccgga tgcggcgtaa acgccttacc cggcctacaa atgcgctccc cgtaggtcgg 7320
 20 ataagacgcg tcagcgtcgc atccgacacc cgaaccacga ggcgaaaaat gattaacgta 7380
 caaaacgtca gtaaacctt catcctgcac cagcaaaacg gcgtgcgcct gcccgctcctc 7440
 aatcgcgcct cgctcacctg caacgcgggc gaatgcgtgg tgctccacgg ccattccggc 7500
 25 agcggcaaat caactctgct acgctcgtg tacgccaact atctgcccga cgaaggtcaa 7560
 atccagatca aacacggtga cgagtgggta gacctggtca ccgcgccagc gcgcaaagtg 7620
 30 gtggaatcc gaaaaccac cgtcggctgg gtgagccagt ttctgcgctg catcccgcgt 7680
 atctcagcac tggaagtggg gatgcagccg ctgctcgata ccggcgttcc gcgtgaagcc 7740
 tgcgcccgta aagccgcgcg tcttctcacc cgcctgaacg tgccggaacg cctgtggcac 7800
 35 ctggcaccat cgacattttc cgggtggcga cagcagcgcg tcaacatcgc ccgcggcttt 7860
 atcgtcgact accccattct gctgcttgac gaacctaccg cctcgtgga cgccaaaaac 7920
 40 agcgcgcggg tgggtggaact gattcgcgaa gccaaaaccg gtggcgcagc catcgtagcc 7980
 atcttccatg acgaagctgt acgtaatgac gtcgccgacc gcctgcaccc aatgggagcc 8040
 tcttcatgat tatcaataac gttaagctgg tgctggaaaa cgaggtggtg agcggttcgc 8100
 45 tggaggtgca gaacggcgaa atccgcgcct ttgccgaaag ccagagccgc ctgccggagg 8160
 cgatggacgg cgaaggcggc tggtgctgc cggggtgat tgagctgcat accgataatc 8220
 50 tggataaatt cttcaccccg cgcccgaag ttgactggcc tgcccactcg gcgatgagca 8280
 gccacgacgc gctgatgggt gcgagcggca tcaccaccgt actggatgcc gtggcaattg 8340

gcgacgtgcg cgacggcggc gatcggctgg agaactctgga gaagatgac aacgccatcg 8400
 aagagacgca gaaacgcggc gtcaaccgcg ccgagcaccg tctgcatctg cgctgcgaac 8460
 5 tgccgcatca caccacgctg ccgctgtttg aaaaactggt gcagcgcgag ccggtgacgc 8520
 tgggtgctgct gatggaccac tcgcccggcc agcgcagtt cgccaaccgc gagaagtatc 8580
 10 gcgaatatta tcagggcaaa tactccctca ctgatgcgca gatgcagcag tacgaagaag 8640
 agcaactggc gctcgccgca cgctggtcgc agccgaatcg cgaatccatc gccgccctgt 8700
 gccgcgcgcg aaaaattgcg cttgccagcc acgatgacgc caccacgccc cacgttgctg 8760
 15 aatctcacca gcttggcagc gtgatgcgag aatttcccac cacgttcgaa gcggcggaag 8820
 cctcgcgcaa gcatggcatg aacgtgctga tgggcgcgcc gaatattgtg cgcggcggct 8880
 20 cgcaactcgg caacgtggcg gccagtgaac tggcgcagct tggcctgctg gatatcctct 8940
 cttecgacta ctaccccgcc agcctgctcg atgcggcatt tcgctgcgcc gatgaccaga 9000
 gcaaccgctt tacgctgccg caggcgggta agctgggtgac taaaaatcca gcgcagggcg 9060
 25 ttaatctcca ggatcgcggg gtgattggcg agggcaaacg cgccgacctg gtgctggcgc 9120
 atcgcaagga caatcatatt catatcgacc acgtctggcg tcagggtaaa aggggtttct 9180
 30 gatgatggga aaactgattt ggtaaatggg gccgtccggc tccgggaaag acagcctgct 9240
 ggcggaactc cgctgcggg acaaaactca gttactgggt gcgcatcgct acatcacgcg 9300
 cgatgccagc gccggaagtg aaaaccatat cgccctgagc gagcaggagt tttttaccgg 9360
 35 cgcggggcaa aatctgttgg ccttaagctg gcacgctaac ggtctgtatt atggcgtcgg 9420
 cgtcgagatt gatctctggc tgcaagccgg attcgacgtg ctggtcaacg gctcaagcgc 9480
 40 ccatctgccg caggegcggg cgcgctatca atcggcctg ctgcccgtct gtttacaggt 9540
 ttcgcccggag atcctccgcc agcgcctgga aaaccgtggc cgtgaaaacg ccagtgaat 9600
 taacgcccgc ctggcgcgcg ccgcccgcta tactccacag gattgccata cgctcaacaa 9660
 45 tgacggcagc ctgcgccagt cggctgacac gctgctgacg ctgatccatc agaaggagaa 9720
 acaccatgcc tgcttgtgag cttcggccgg ccacgcagta cgacaccgac gcggtttacg 9780
 50 cgctgatttg tgagctaaaa caggcggagt ttgaccacca cgcgtttcgc gtgggtttta 9840
 acgccaatct gcgcgacca aacatgcgct accatctggc gctgcttgat ggcaagttg 9900

tcggcatgat cggcctgcat ttgcagtttc atctgcatca tgtcaactgg atcggcgaaa 9960
 5 ttcaggagtt ggtggtaatg ccgcaggcgc gcggtctgaa cgtcggcagt aagtactgg 10020
 cgtgggcaga agaagaagcc cgccaggccg gggccgaaat gaccgaactt tcgaccaacg 10080
 tgaagcgcca cgacgcgcac cgtttctatc tgcgcgaagg ctacgagcag agccacttcc 10140
 10 gcttcaccaa ggcgctgtaa catgagcctg acctcacgc tcaccggcac cggcggcgca 10200
 cagggcgttc cggcatgggg ctgcgagtgt gggcctgcg ccagagcgcg gcgctcgccg 10260
 cagtatgcc gccaacctg cagcggcgta gtgaagtta acgacgcaat caccctgatc 10320
 15 gagccgggc tgcacgatct cgccgatcgc tggtcgccc gatcgttcca gcagttttg 10380
 ctgacgcatt atcatatgga tcacgtccag gggctgtttc cgctgcgctg gggcgttggc 10440
 20 gatccgatcc cggtttacgg cccgccgat gaacagggtc gcgacgatct gtttaacat 10500
 ccgggcctgc ttgatttcag ccacacggtg gaaccgtttg tgggtttga tttgcagggg 10560
 ttacaggtca cggccctgcc gctcaaccac tcaaaactga ccttcggtta tctgctggaa 10620
 25 acggcacaca gccgggtggc gtggctgtct gacaccgag gtttccgga aaaaacgctg 10680
 aaatttttac gcaataatca gccgcaggta atggtgatgg attgcagtca cccgccgcgc 10740
 30 gcggatgcac cgcgtaatca ctgtgattta aatacctgctc ttgcgctgaa tcaggttate 10800
 cgctgccac gggtgattct gaccatate agccaccagt ttgatgcgtg gctgatggaa 10860
 aacgcactac cgtcagggtt tgagggtggg tttgatggga tggagattgg gttggcgtga 10920
 35 tgagagggaa tgtgcgcgct gccccctca ccctaaccct ctccccagag gggcgagggg 10980
 accgattgtg ctgatattg aatattgcgc tegtittctc cctctccca ttggggtgag 11040
 40 gggcgatgcc tgetccatac ccaacctcat cggccatact catcttccat tctccgctct 11100
 tcatectca gttgccgacg ctctgatca agctggcgct ggcgatcgtc cagctgcctg 11160
 cggegatctt caaactggcg gggcggtcg tcatattgct tgcgccgac gtcgctcact 11220
 45 tcacgctgcc agccgtcgtc gcgcgaatct tcatagtctc gccacggtc agggttataa 11280
 gcgtcattaa tcgcctgctg aatattgcca atggtgtcgt cgataatate ggcctgggcc 11340
 50 ggaacgtgga cagcgtgagc aggggtgaata aaagaatatag cggaaagcgt ttcattagcc 11400
 aacctcaaaa agaaactcta tccacattaa tcattactca tccatgcaag tagtggatga 11460

atctcaattt ctccgctgct ctattgccgt aatcgectcc acgcgttggt gatgacgacc 11520
 gccttcgtac tgtgcgccc gccacgcac cacaatcatt ttgcccagtt cgaggccaac 11580
 5 cactcgtgaa ccaaaagcca gcacgttggt gtcgttatgc tgccgcgaaa gttgcgcgga 11640
 ataaggttcg ctacagacga ccgcgcgaat tc 11672

10 <211> 435
 <212> DNA
 <213> 大肠杆菌

15 <400> 3
 atgcctgctt gtgagcttcg cccggccacg cagtacgaca ccgacgcggt ttacgcgctg 60
 atttgtgagc taaaacaggc ggagtttgac caccacgcgt ttcgcgtggg ttttaacgcc 120
 20 aatctgcgcg acccaaacat gcgctaccat ctggcgtgc ttgatggcga agttgtcggc 180
 atgatcggcc tgcatttga gtttcatctg catcatgtca actggatcgg cgaaattcag 240
 gagttggtgg taatgccgca ggcgcgcggt ctgaacgtcg gcagtaagtt actggcgtgg 300
 25 gcagaagaag aagcccgcca ggccggggcc gaaatgaccg aactttcgac caacgtgaag 360
 cgccacgacg cgcaccgttt ctatctgcgc gaaggctacg agcagagcca cttccgcttc 420
 30 accaaggcgc tgtaa 435

<210> 4
 <211> 144
 35 <212> PRT
 <213> 大肠杆菌

<400> 4
 40 Met Pro Ala Cys Glu Leu Arg Pro Ala Thr Gln Tyr Asp Thr Asp Ala
 1 5 10 15
 Val Tyr Ala Leu Ile Cys Glu Leu Lys Gln Ala Glu Phe Asp His His
 20 25 30
 45 Ala Phe Arg Val Gly Phe Asn Ala Asn Leu Arg Asp Pro Asn Met Arg
 35 40 45
 Tyr His Leu Ala Leu Leu Asp Gly Glu Val Val Gly Met Ile Gly Leu
 50 50 55 60
 His Leu Gln Phe His Leu His His Val Asn Trp Ile Gly Glu Ile Gln

	65		70		75		80									
	Glu	Leu	Val	Val	Met	Pro	Gln	Ala	Arg	Gly	Leu	Asn	Val	Gly	Ser	Lys
					85					90					95	
5	Leu	Leu	Ala	Trp	Ala	Glu	Glu	Glu	Ala	Arg	Gln	Ala	Gly	Ala	Glu	Met
			100						105				110			
10	Thr	Glu	Leu	Ser	Thr	Asn	Val	Lys	Arg	His	Asp	Ala	His	Arg	Phe	Tyr
			115					120					125			
15	Leu	Arg	Glu	Gly	Tyr	Glu	Gln	Ser	His	Phe	Arg	Phe	Thr	Lys	Ala	Leu
		130					135					140				

<210> 5
 <211> 20
 <212> DNA
 <213> 人工序列
 <220>
 <223> 人工序列的描述: 合成寡核苷酸

25 <400> 5
 aaacaccatg gctgcttgtg

30 <210> 6
 <211> 35
 <212> DNA
 <213> 人工序列
 <220>
 <223> 人工序列的描述: 合成寡核苷酸

35 <400> 6
 gtgacgaatt cgagctcatt acagcgcctt ggtga 35

40 <210> 7
 <211> 435
 <212> DNA
 <213> 人工序列

45 <220>
 <223> 人工序列的描述: 非天然存在的核苷酸序列, 编码修饰的PhnO
 蛋白P2A; 于核苷酸位置4 g-c

50 <220>
 <221> CDS
 <222> (1)..(432)

<400> 7
atg gct gct tgt gag ctt cgc ccg gcc acg cag tac gac acc gac gcg 48
Met Ala Ala Cys Glu Leu Arg Pro Ala Thr Gln Tyr Asp Thr Asp Ala
5 1 5 10 15
gtt tac gcg ctg att tgt gag cta aaa cag gcg gag ttt gac cac cac 96
Val Tyr Ala Leu Ile Cys Glu Leu Lys Gln Ala Glu Phe Asp His His
20 25 30
10
gcg ttt cgc gtg ggt ttt aac gcc aat ctg cgc gac cca aac atg cgc 144
Ala Phe Arg Val Gly Phe Asn Ala Asn Leu Arg Asp Pro Asn Met Arg
35 40 45
15
tac cat ctg gcg ctg ctt gat ggc gaa gtt gtc ggc atg atc ggc ctg 192
Tyr His Leu Ala Leu Leu Asp Gly Glu Val Val Gly Met Ile Gly Leu
50 55 60
20
cat ttg cag ttt cat ctg cat cat gtc aac tgg atc ggc gaa att cag 240
His Leu Gln Phe His Leu His His Val Asn Trp Ile Gly Glu Ile Gln
65 70 75 80
25
gag ttg gtg gta atg ccg cag gcg cgc ggt ctg aac gtc ggc agt aag 288
Glu Leu Val Val Met Pro Gln Ala Arg Gly Leu Asn Val Gly Ser Lys
85 90 95
30
tta ctg gcg tgg gca gaa gaa gaa gcc cgc cag gcc ggg gcc gaa atg 336
Leu Leu Ala Trp Ala Glu Glu Glu Ala Arg Gln Ala Gly Ala Glu Met
100 105 110
35
acc gaa ctt tcg acc aac gtg aag cgc cac gac gcg cac cgt ttc tat 384
Thr Glu Leu Ser Thr Asn Val Lys Arg His Asp Ala His Arg Phe Tyr
115 120 125
40
ctg cgc gaa ggc tac gag cag agc cac ttc cgc ttc acc aag gcg ctg 432
Leu Arg Glu Gly Tyr Glu Gln Ser His Phe Arg Phe Thr Lys Ala Leu
130 135 140
45
taa 435

<210> 8
<211> 144
<212> PRT
45 <213> 人工序列

<400> 8
Met Ala Ala Cys Glu Leu Arg Pro Ala Thr Gln Tyr Asp Thr Asp Ala
1 5 10 15
50
Val Tyr Ala Leu Ile Cys Glu Leu Lys Gln Ala Glu Phe Asp His His
20 25 30

Ala Phe Arg Val Gly Phe Asn Ala Asn Leu Arg Asp Pro Asn Met Arg
 35 40 45

5 Tyr His Leu Ala Leu Leu Asp Gly Glu Val Val Gly Met Ile Gly Leu
 50 55 60

His Leu Gln Phe His Leu His His Val Asn Trp Ile Gly Glu Ile
 65 70 75 80

10 Glu Leu Val Val Met Pro Gln Ala Arg Gly Leu Asn Val Gly Ser Lys
 85 90 95

15 Leu Leu Ala Trp Ala Glu Glu Glu Ala Arg Gln Ala Gly Ala Glu Met
 100 105 110

Thr Glu Leu Ser Thr Asn Val Lys Arg His Asp Ala His Arg Phe Tyr
 115 120 125

20 Leu Arg Glu Gly Tyr Glu Gln Ser His Phe Arg Phe Thr Lys Ala Leu
 130 135 140

<210> 9
 <211> 264
 <212> DNA
 <213> 人工序列

<220>
 <223> 人工序列的描述: 转运肽编码序列

<220>
 <221> CDS
 <222> (1)..(264)

35 <400> 9
 atg gct tcc tct atg ctc tct tcc gct act atg gtt gcc tct ccg gct 48
 Met Ala Ser Ser Met Leu Ser Ser Ala Thr Met Val Ala Ser Pro Ala
 1 5 10 15

40 cag gcc act atg gtc gct cct ttc aac gga ctt aag tcc tcc gct gcc 96
 Gln Ala Thr Met Val Ala Pro Phe Asn Gly Leu Lys Ser Ser Ala Ala
 20 25 30

45 ttc cca gcc acc cgc aag gct aac aac gac att act tcc atc aca agc 144
 Phe Pro Ala Thr Arg Lys Ala Asn Asn Asp Ile Thr Ser Ile Thr Ser
 35 40 45

50 aac ggc gga aga gtt aac tgc atg cag gtg tgg cct ccg att gga aag 192
 Asn Gly Gly Arg Val Asn Cys Met Gln Val Trp Pro Pro Ile Gly Lys
 50 55 60

```

aag aag ttt gag act ctc tct tac ctt cct gac ctt acc gat tcc ggt 240
Lys Lys Phe Glu Thr Leu Ser Tyr Leu Pro Asp Leu Thr Asp Ser Gly
65          70          75          80

5   ggt cgc gtc aac tgc atg cag gcc 264
    Gly Arg Val Asn Cys Met Gln Ala
          85

10  <210> 10
    <211> 88
    <212> PRT
    <213> 人工序列

15  <400> 10
    Met Ala Ser Ser Met Leu Ser Ser Ala Thr Met Val Ala Ser Pro Ala
      1          5          10          15

20  Gln Ala Thr Met Val Ala Pro Phe Asn Gly Leu Lys Ser Ser Ala Ala
      20          25          30

    Phe Pro Ala Thr Arg Lys Ala Asn Asn Asp Ile Thr Ser Ile Thr Ser
      35          40          45

25  Asn Gly Gly Arg Val Asn Cys Met Gln Val Trp Pro Pro Ile Gly Lys
      50          55          60

    Lys Lys Phe Glu Thr Leu Ser Tyr Leu Pro Asp Leu Thr Asp Ser Gly
      65          70          75          80

30  Gly Arg Val Asn Cys Met Gln Ala
      85

35  <210> 11
    <211> 696
    <212> DNA
    <213> 人工序列

40  <220>
    <223> 人工序列的描述: CTP-AMPA乙酰基转移酶编码序列和氨基酸序列翻译

    <220>
    <221> CDS
45  <222> (1)..(696)

    <400> 11
    atg gct tcc tct atg ctc tct tcc gct act atg gtt gcc tct ccg gct 48
    Met Ala Ser Ser Met Leu Ser Ser Ala Thr Met Val Ala Ser Pro Ala
    1          5          10          15

    cag gcc act atg gtc gct cct ttc aac gga ctt aag tcc tcc gct gcc 96

```




His Phe Arg Phe Thr Lys Ala Leu
225 230

5 <210> 12
<211> 232
<212> PRT
<213> 人工序列

10 <400> 12
Met Ala Ser Ser Met Leu Ser Ser Ala Thr Met Val Ala Ser Pro Ala
1 5 10 15

15 Gln Ala Thr Met Val Ala Pro Phe Asn Gly Leu Lys Ser Ser Ala Ala
20 25 30

Phe Pro Ala Thr Arg Lys Ala Asn Asn Asp Ile Thr Ser Ile Thr Ser
35 40 45

20 Asn Gly Gly Arg Val Asn Cys Met Gln Val Trp Pro Pro Ile Gly Lys
50 55 60

Lys Lys Phe Glu Thr Leu Ser Tyr Leu Pro Asp Leu Thr Asp Ser Gly
65 70 75 80

25 Gly Arg Val Asn Cys Met Gln Ala Met Ala Ala Cys Glu Leu Arg Pro
85 90 95

30 Ala Thr Gln Tyr Asp Thr Asp Ala Val Tyr Ala Leu Ile Cys Glu Leu
100 105 110

Lys Gln Ala Glu Phe Asp His His Ala Phe Arg Val Gly Phe Asn Ala
115 120 125

35 Asn Leu Arg Asp Pro Asn Met Arg Tyr His Leu Ala Leu Leu Asp Gly
130 135 140

Glu Val Val Gly Met Ile Gly Leu His Leu Gln Phe His Leu His His
145 150 155 160

40 Val Asn Trp Ile Gly Glu Ile Gln Glu Leu Val Val Met Pro Gln Ala
165 170 175

45 Arg Gly Leu Asn Val Gly Ser Lys Leu Leu Ala Trp Ala Glu Glu Glu
180 185 190

Ala Arg Gln Ala Gly Ala Glu Met Thr Glu Leu Ser Thr Asn Val Lys
195 200 205

50 Arg His Asp Ala His Arg Phe Tyr Leu Arg Glu Gly Tyr Glu Gln Ser
210 215 220

His Phe Arg Phe Thr Lys Ala Leu
225 230

5 <210> 13
<211> 415
<212> DNA
<213> 玉蜀黍 (Zea mays)

10 <220>
<221> N_区
<222> (15)..(163)

15 <220>
<221> 内含子
<222> (164)..(322)

20 <220>
<221> C_区
<222> (323)..(411)

<400> 13
tctagaggat cagcatggcg cccaccgtga tgatggcctc gtcggccacc gccgtcgctc 60

25 cgttctctggg gctcaagtcc accgccagcc tccccgtcgc ccgccgtcc tccagaagcc 120
tcggcaacgt cagcaacggc ggaaggatcc ggtgcatgca ggtaacaaat gcatcctagc 180
tagtagttct ttgcattgca gcagctgcag ctagcgagtt agtaatagga agggaactga 240

30 tgatccatgc atggactgat gtgtgttgcc catcccatcc catcccattt cccaaacgaa 300
ccgaaaacac cgtactacgt gcaggtgtgg ccctacggca acaagaagtt cgagacgtg 360

35 tcgtacctgc cgccgctgtc gaccggcggg cgcacccgct gcatgcaggc catgg 415

40 <210> 14
<211> 174
<212> DNA
<213> 人工序列

45 <220>
<223> 人工序列的描述: 叶绿体或质体转运肽编码序列和氨基酸序列翻译

<220>
<221> CDS
<222> (1)..(174)

50 <400> 14
atg gct tcc tct atg ctc tct tcc gct act atg gtt gcc tct ccg gct 48
Met Ala Ser Ser Met Leu Ser Ser Ala Thr Met Val Ala Ser Pro Ala

	1	5	10	15	
	cag gcc act atg gtc gct cct ttc aac gga ctt aag tcc tcc gct gcc	96			
5	Gln Ala Thr Met Val Ala Pro Phe Asn Gly Leu Lys Ser Ser Ala Ala				
	20 25 30				
	ttc cca gcc acc cgc aag gct aac aac gac att act tcc atc aca agc	144			
	Phe Pro Ala Thr Arg Lys Ala Asn Asn Asp Ile Thr Ser Ile Thr Ser				
10	35 40 45				
	aac ggc gga aga gtt aac tgc atg cag gcc	174			
	Asn Gly Gly Arg Val Asn Cys Met Gln Ala				
	50 55				
15	<210> 15				
	<211> 58				
	<212> PRT				
	<213> 人工序列				
20	<400> 15				
	Met Ala Ser Ser Met Leu Ser Ser Ala Thr Met Val Ala Ser Pro Ala				
	1 5 10 15				
25	Gln Ala Thr Met Val Ala Pro Phe Asn Gly Leu Lys Ser Ser Ala Ala				
	20 25 30				
	Phe Pro Ala Thr Arg Lys Ala Asn Asn Asp Ile Thr Ser Ile Thr Ser				
30	35 40 45				
	Asn Gly Gly Arg Val Asn Cys Met Gln Ala				
	50 55				
35	<210> 16				
	<211> 157				
	<212> DNA				
	<213> 人工序列				
40	<220>				
	<223> 人工序列的描述: 合成的寡核苷酸, 代表432个碱基对的AMPA酰基转移酶基因的碱基对1-157				
45	<400> 16				
	atggcgcgtt gcgagcttcg cccagccacg cagtagcaca cgcagccgt gtacgcgctg	60			
	atctgcgagc tcaagcaggc ggagttcgac caccacgcct tccgcgtggg cttcaacgcc	120			
50	aacctgcgcg accccaacat gcctacat ctggcgc	157			
	<210> 17				

<211> 187
 <212> DNA
 <213> 人工序列

5 <220>
 <223> 人工序列的描述: 合成的寡核苷酸序列, 代表432个碱基对的AMPA酰基转移酶基因的碱基对158-344

<400> 17
 10 tgcttgatgg cgaagtggtc ggcatgatcg gcctgcacct ccagttccac ctgcatcatg 60
 tcaactggat cggcgagatc caggagctgg tcgtgatgcc acaggcgagg ggtctgaacg 120
 tcggcagcaa gctcctggcg tgggccgagg aggaagccag gcaggccgga gccgagatga 180
 15 ccgagct 187

<210> 18
 20 <211> 88
 <212> DNA
 <213> 人工序列

<220>
 25 <223> 人工序列的描述: 合成的寡核苷酸序列, 代表 432 个碱基对的 AMPA 酰基转移酶基因的碱基对 345-432

<400> 18
 30 cagcaccaac gtgaagcgcc acgacgcgca ccgcttctac ctgcgcgaag getacgagca 60
 gagccacttc cgcttcacca aggcgctg 88

35 <210> 19
 <211> 432
 <212> DNA
 <213> 人工序列

40 <220>
 <223> 人工序列的描述: 合成的寡核苷酸, 提供 AMPA 乙酰基转移酶的
 单子叶植物优化编码序列

45 <220>
 <221> CDS
 <222> (1)..(432)

<400> 19
 50 atg gcc gct tgc gag ctt cgc cca gcc acg cag tac gac acc gac gcc 48
 Met Ala Ala Cys Glu Leu Arg Pro Ala Thr Gln Tyr Asp Thr Asp Ala
 1 5 10 15

gtg tac gcg ctg atc tgc gag ctc aag cag gcg gag ttc gac cac cac 96
Val Tyr Ala Leu Ile Cys Glu Leu Lys Gln Ala Glu Phe Asp His His
20 25 30

5 gcc ttc cgc gtg ggc ttc aac gcc aac ctg cgc gac ccc aac atg cgc 144
Ala Phe Arg Val Gly Phe Asn Ala Asn Leu Arg Asp Pro Asn Met Arg
35 40 45

10 tac cat ctg gcg ctg ctt gat ggc gaa gtg gtc ggc atg atc ggc ctg 192
Tyr His Leu Ala Leu Leu Asp Gly Glu Val Val Gly Met Ile Gly Leu
50 55 60

15 cac ctc cag ttc cac ctg cat cat gtc aac tgg atc ggc gag atc cag 240
His Leu Gln Phe His Leu His His Val Asn Trp Ile Gly Glu Ile Gln
65 70 75 80

20 gag ctg gtc gtg atg cca cag gcg agg ggt ctg aac gtc ggc agc aag 288
Glu Leu Val Val Met Pro Gln Ala Arg Gly Leu Asn Val Gly Ser Lys
85 90 95

25 ctc ctg gcg tgg gcc gag gag gaa gcc agg cag gcc gga gcc gag atg 336
Leu Leu Ala Trp Ala Glu Glu Glu Ala Arg Gln Ala Gly Ala Glu Met
100 105 110

30 acc gag ctc agc acc aac gtg aag cgc cac gac gcg cac cgc ttc tac 384
Thr Glu Leu Ser Thr Asn Val Lys Arg His Asp Ala His Arg Phe Tyr
115 120 125

35 ctg cgc gaa ggc tac gag cag agc cac ttc cgc ttc acc aag gcg ctg 432
Leu Arg Glu Gly Tyr Glu Gln Ser His Phe Arg Phe Thr Lys Ala Leu
130 135 140

40 <210> 20
<211> 144
<212> PRT
<213> 人工序列

45 <400> 20
Met Ala Ala Cys Glu Leu Arg Pro Ala Thr Gln Tyr Asp Thr Asp Ala
1 5 10 15
Val Tyr Ala Leu Ile Cys Glu Leu Lys Gln Ala Glu Phe Asp His His
20 25 30
Ala Phe Arg Val Gly Phe Asn Ala Asn Leu Arg Asp Pro Asn Met Arg
35 40 45
50 Tyr His Leu Ala Leu Leu Asp Gly Glu Val Val Gly Met Ile Gly Leu
50 55 60

His Leu Gln Phe His Leu His His Val Asn Trp Ile Gly Glu Ile Gln
 65 70 75 80

5 Glu Leu Val Val Met Pro Gln Ala Arg Gly Leu Asn Val Gly Ser Lys
 85 90 95

Leu Leu Ala Trp Ala Glu Glu Glu Ala Arg Gln Ala Gly Ala Glu Met
 100 105 110

10 Thr Glu Leu Ser Thr Asn Val Lys Arg His Asp Ala His Arg Phe Tyr
 115 120 125

Leu Arg Glu Gly Tyr Glu Gln Ser His Phe Arg Phe Thr Lys Ala Leu
 130 135 140

15

<210> 21
 <211> 20
 <212> DNA
 20 <213> 人工序列

<220>
 <223> 人工序列的描述: 合成的寡核苷酸 PHN1, 用作扩增引物

25 <400> 21
 atggctgctt gtgagcttcg

30 <210> 22
 <211> 20
 <212> DNA
 <213> 人工序列

35 <220>
 <223> 人工序列的描述: 合成的寡核苷酸 PHN2, 用作扩增引物

<400> 22
 cagcgccttg gtgaagcgga

40

<210> 23
 <211> 1630
 <212> DNA
 <213> 人工序列

45 <220>
 <223> 人工序列的描述: 表达盒, 其包含植物有效启动子, 所述启动子连接于一个
 编码 AMPA 乙酰基转移酶的编码序列, 所述编码序列连接于一个转录终止序列

50 <220>
 <221> 启动子
 <222> (33)..(605)

<220>
 <221> 转运肽
 <222> (627)..(892)

5

<220>
 <221> CDS
 <222> (893)..(1324)

10

<220>
 <221> 终止子
 <222> (1350)..(1605)

<400> 23

15

gcggccgcgt tcaagcttga gctcaggatt tagcagcatt ccagattggg ttcaatcaac 60
 aaggtacgag ccatactact ttattcaaat tggatcgcg aaaaccaaga aggaactccc 120
 atcctcaaag gtttgaagg aagaattctc agtccaaagc ctcaacaagg tcagggtaca 180

20

gagtctcaa accattagcc aaaagctaca ggagatcaat gaagaatctt caatcaaagt 240
 aaactactgt tccagcacat gcatcatggt cagtaagttt cagaaaaaga catccaccga 300

25

agacttaaag ttagtgggca tctttgaaag taatcttctc aacatcgagc agctggcttg 360
 tggggaccag acaaaaaagg aatggtgcag aattgttagg cgcacctacc aaaagcatct 420
 ttgcctttat tgcaaagata aagcagattc ctctagtaca agtggggaac aaaataacgt 480

30

ggaaaagagc tgtcctgaca gcccactcac taatgcgtat gacgaacgca gtgacgacca 540
 caaaagaatt ccctctatat aagaaggcat tcattcccat ttgaaggatc atcagatact 600

35

gaaccaatcc ttctagaaga tctccacaat ggcttctctt atgctctctt ccgctactat 660
 ggttgctctt cgggctcagg ccaactatgt cgctccttc aacggactta agtctctcgc 720
 tgccttecca gccaccgca aggctaacaa cgacattact tccatcacia gcaacggcgg 780

40

aagagttaac tgcattcagg tgtggcctcc gattggaaag aagaagtttg agactctctc 840
 ttaccttctt gacctaccg attccggtgg tcgctgcaac tgcattcagg cc atg gct 898
 Met Ala

45

gct tgt gag ctt cgc cgg gcc acg cag tac gac acc gac gcg gtt tac 946
 Ala Cys Glu Leu Arg Pro Ala Thr Gln Tyr Asp Thr Asp Ala Val Tyr
 5 10 15

50

gcg ctg att tgt gag cta aaa cag gcg gag ttt gac cac cac gcg ttt 994
 Ala Leu Ile Cys Glu Leu Lys Gln Ala Glu Phe Asp His His Ala Phe

	20	25	30	
	cgc gtg ggt ttt aac gcc aat ctg cgc gac cca aac atg cgc tac cat			1042
5	Arg Val Gly Phe Asn Ala Asn Leu Arg Asp Pro Asn Met Arg Tyr His			
	35	40	45	50
	ctg gcg ctg ctt gat ggc gaa gtt gtc ggc atg atc ggc ctg cat ttg			1090
10	Leu Ala Leu Leu Asp Gly Glu Val Val Gly Met Ile Gly Leu His Leu			
		55	60	65
	cag ttt cat ctg cat cat gtc aac tgg atc ggc gaa att cag gag ttg			1138
	Gln Phe His Leu His His Val Asn Trp Ile Gly Glu Ile Gln Glu Leu			
		70	75	80
15	gtg gta atg ccg cag gcg cgc ggt ctg aac gtc ggc agt aag tta ctg			1186
	Val Val Met Pro Gln Ala Arg Gly Leu Asn Val Gly Ser Lys Leu Leu			
		85	90	95
20	gcg tgg gca gaa gaa gaa gcc cgc cag gcc ggg gcc gaa atg acc gaa			1234
	Ala Trp Ala Glu Glu Glu Ala Arg Gln Ala Gly Ala Glu Met Thr Glu			
		100	105	110
	ctt tcg acc aac gtg aag cgc cac gac gcg cac cgt ttc tat ctg cgc			1282
25	Leu Ser Thr Asn Val Lys Arg His Asp Ala His Arg Phe Tyr Leu Arg			
		115	120	125
	gaa ggc tac gag cag agc cac ttc cgc ttc acc aag gcg ctg			1324
	Glu Gly Tyr Glu Gln Ser His Phe Arg Phe Thr Lys Ala Leu			
		135	140	
30	taatgagctc ggtaccggat ccaattcccg atcgttcaaa catttgcaa taaagtttct			1384
	taagattgaa tcctgttgcc ggtcttgca tgattatcat ataatttctg ttgaattacg			1444
35	ttaagcatgt aataattaac atgtaatgca tgacgttatt tatgagatgg gttttatga			1504
	ttagagtccc gcaattatac atttaatac gcatagaaaa caaaatatag cgcgcaaac			1564
40	aggataaatt atcgcgcgcg gtgtcatcta tgttactaga tcgggatcg atccccgggc			1624
	ggccgc			1630
45	<210> 24			
	<211> 144			
	<212> PRT			
	<213> 人工序列			
50	<400> 24			
	Met Ala Ala Cys Glu Leu Arg Pro Ala Thr Gln Tyr Asp Thr Asp Ala			
	1	5	10	15

Val Tyr Ala Leu Ile Cys Glu Leu Lys Gln Ala Glu Phe Asp His His
20 25 30

Ala Phe Arg Val Gly Phe Asn Ala Asn Leu Arg Asp Pro Asn Met Arg
5 35 40 45

Tyr His Leu Ala Leu Leu Asp Gly Glu Val Val Gly Met Ile Gly Leu
50 55 60

His Leu Gln Phe His Leu His His Val Asn Trp Ile Gly Glu Ile Gln
10 65 70 75 80

Glu Leu Val Val Met Pro Gln Ala Arg Gly Leu Asn Val Gly Ser Lys
15 85 90 95

Leu Leu Ala Trp Ala Glu Glu Glu Ala Arg Gln Ala Gly Ala Glu Met
100 105 110

Thr Glu Leu Ser Thr Asn Val Lys Arg His Asp Ala His Arg Phe Tyr
20 115 120 125

Leu Arg Glu Gly Tyr Glu Gln Ser His Phe Arg Phe Thr Lys Ala Leu
130 135 140

25

<210> 25
<211> 2122
<212> DNA
<213> 人工序列

30

<220>
<223> 人工序列的描述: 表达盒, 其包含植物启动子, 所述启动子连接于一个
编码 AMPA 乙酰基转移酶的序列, 所述编码序列连接于一个终止序列

35

<220>
<221> 启动子
<222> (6)..(620)

40

<220>
<221> 5' UTR
<222> (645)..(715)

45

<220>
<221> 内含子
<222> (729)..(1178)

<220>
<221> 转运肽
<222> (1179)..(1406)

50

<220>
<221> CDS

<222> (1407).. (1838)

<220>

<221> 终止子

5 <222> (1849).. (2082)

<400> 25

ctgcaggctc gatgtgagac ttttcaaca aggtaatat ccggaacct cctcggattc 60

10 cattgccag ctatctgtca ctttattgtg aagatagtgg aaaaggaagg tggctcctac 120

aatgccatc attgcgataa aggaaaggcc atcgttgaag atgcctctgc cgacagtgg 180

15 cccaaagatg gacccccacc cacgaggagc atcgtggaaa aagaagacgt tccaaccag 240

tcttcaaagc aagtggattg atgtgatggt ccgatgtgag acttttcaac aaagggtaat 300

atccggaaac ctctcggat tccattgccc agctatctgt cactttattg tgaagatagt 360

20 ggaaaaggaa ggtggctcct acaaatgcca tcattgcgat aaaggaaagg ccatcgttga 420

agatgcctct gccgacagtg gtcccaaaga tggaccccca cccacgagga gcatcgtgga 480

25 aaaagaagac gttccaacca cgtcttcaaa gcaagtggat tgatgtgata tctccactga 540

cgtaagggat gacgcacaat cccactatcc ttcgcaagac ctttctctata tataaggaag 600

ttcatttcat ttggagagga cacgctgaca agctgactct agcagatcct ctagaacct 660

30 ctccacaca ctcaagccac actattggag aacacacagg gacaacacac cataagatcc 720

aaggagggcc tccgccgccg ccgtaacca cccgccct ctctctttc tttctcgtt 780

35 ttttttccg tctcggctc gatctttggc cttggtagtt tgggtggcg agaggcggct 840

tcgtgcgcgc ccagatcggg gcgcgggagg ggcgggatct cgcggggaat gggctctcg 900

gatgtagatc tgcgatccgc cgttgttggg ggagatgat gggcgttta aatttcgccg 960

40 tgctaaaca gatcaggaag aggggaaaag ggcactatgg tttatattt tatatattc 1020

tgctgcttcg tcaggcttag atgtgctaga tcttctttc ttctttttgt gggtagaatt 1080

taatccctca gcattgttca tcggtagttt ttcttttcat gatttcgtga caaatgcagc 1140

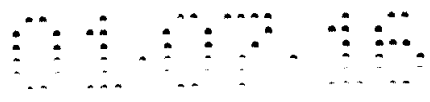
45

ctcgtgcgga gcttttttgt aggtagaagt gatcaacct ggcgcaagtt agcagaatct 1200

gcaatgggtg gcagaacca tctcttatct ccaatctctc gaaatccagt caacgcaaat 1260

50 ctcccttacc ggtttctctg aagacgcagc agcatccacg agcttatccg atttcgtcgt 1320

cgtaggggatt gaagaagagt gggatgacgt taattggctc tgagcttcgt cctcttaagg 1380



tcattgtcttc tgtttccacg gcgtgc atg gcc gct tgc gag ctt cgc cca gcc 1433
Met Ala Ala Cys Glu Leu Arg Pro Ala

5
acg cag tac gac acc gac gcc gtg tac gcg ctg atc tgc gag ctc aag 1481
Thr Gln Tyr Asp Thr Asp Ala Val Tyr Ala Leu Ile Cys Glu Leu Lys
10 15 20 25

10
cag gcg gag ttc gac cac cac gcc ttc cgc gtg ggc ttc aac gcc aac 1529
Gln Ala Glu Phe Asp His His Ala Phe Arg Val Gly Phe Asn Ala Asn
30 35 40

15
ctg cgc gac ccc aac atg cgc tac cat ctg gcg ctg ctt gat ggc gaa 1577
Leu Arg Asp Pro Asn Met Arg Tyr His Leu Ala Leu Leu Asp Gly Glu
45 50 55

20
gtg gtc ggc atg atc ggc ctg cac ctc cag ttc cac ctg cat cat gtc 1625
Val Val Gly Met Ile Gly Leu His Leu Gln Phe His Leu His His Val
60 65 70

25
aac tgg atc ggc gag atc cag gag ctg gtc gtg atg cca cag gcg agg 1673
Asn Trp Ile Gly Glu Ile Gln Glu Leu Val Val Met Pro Gln Ala Arg
75 80 85

30
ggg ctg aac gtc ggc agc aag ctc ctg gcg tgg gcc gag gag gaa gcc 1721
Gly Leu Asn Val Gly Ser Lys Leu Leu Ala Trp Ala Glu Glu Glu Ala
90 95 100 105

35
agg cag gcc gga gcc gag atg acc gag ctc agc acc aac gtg aag cgc 1769
Arg Gln Ala Gly Ala Glu Met Thr Glu Leu Ser Thr Asn Val Lys Arg
110 115 120

40
cac gac gcg cac cgc ttc tac ctg cgc gaa ggc tac gag cag agc cac 1817
His Asp Ala His Arg Phe Tyr Leu Arg Glu Gly Tyr Glu Gln Ser His
125 130 135

45
ttc cgc ttc acc aag gcg ctg taaagatctg aattctgcat gcgtttggac 1868
Phe Arg Phe Thr Lys Ala Leu
140

50
gtatgctcat tcaggttga gccaatgg ttgatgtgtg tgcgagttct tgcgagctg 1928
atgagacatc tctgtattgt gtttctttcc ccagtgtttt ctgtacttgt gtaatcgget 1988
aatcgccaac agattcggcg atgaataaat gagaataaaa ttgttctgat tttgagtgca 2048
aaaaaaaaagg aattagatct gtgtgtgttt tttggatccc cggggcggcc gccccgggtg 2108
gtgagcttct gcag

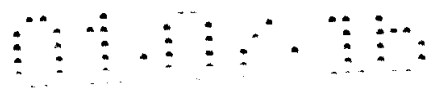
- <210> 26
 <211> 144
 <212> PRT
 <213> 人工序列
- 5
- <400> 26
 Met Ala Ala Cys Glu Leu Arg Pro Ala Thr Gln Tyr Asp Thr Asp Ala
 1 5 10 15
 Val Tyr Ala Leu Ile Cys Glu Leu Lys Gln Ala Glu Phe Asp His His
 20 25 30
 Ala Phe Arg Val Gly Phe Asn Ala Asn Leu Arg Asp Pro Asn Met Arg
 35 40 45
 Tyr His Leu Ala Leu Leu Asp Gly Glu Val Val Gly Met Ile Gly Leu
 50 55 60
 His Leu Gln Phe His Leu His His Val Asn Trp Ile Gly Glu Ile Gln
 65 70 75 80
 Glu Leu Val Val Met Pro Gln Ala Arg Gly Leu Asn Val Gly Ser Lys
 85 90 95
 Leu Leu Ala Trp Ala Glu Glu Glu Ala Arg Gln Ala Gly Ala Glu Met
 100 105 110
 Thr Glu Leu Ser Thr Asn Val Lys Arg His Asp Ala His Arg Phe Tyr
 115 120 125
 Leu Arg Glu Gly Tyr Glu Gln Ser His Phe Arg Phe Thr Lys Ala Leu
 130 135 140
- 10
- 15
- 20
- 25
- 30
- 35
- <210> 27
 <211> 2378
 <212> DNA
 <213> 人工序列
- 40
- <220>
 <223> 人工序列的描述: 表达盒, 其包含一个植物启动子, 所述启动子连接于
 一个内含子、一个编码 AMPA 乙酰基转移酶的序列和一个终止序列
- 45
- <220>
 <221> 启动子
 <222> (28)..(965)
- 50
- <220>
 <221> 内含子
 <222> (966)..(1423)
- <220>

<221> 转运肽
 <222> (1440)..(1667)

5 <220>
 <221> CDS
 <222> (1668)..(2099)

10 <220>
 <221> 终止子
 <222> (2114)..(2369)

<400> 27
 gatatcccta gggcgccgc gttaacaagc ttactcgagg tcattcatat gcttgagaag 60
 15 agagtcggga tagtccaaaa taaaacaaag gtaagattac ctggtcaaaa gtgaaaacat 120
 cagttaaaag gtggtataaa gtaaaatatic gtaataaaa ggtggcccaa agtgaaattt 180
 20 actcttttct actattataa aaattgagga tgtttttgtc ggtactttga tacgtcattt 240
 ttgtatgaat tggtttttaa gtttattcgc ttttggaat gcatactctgt atttgagtcg 300
 ggttttaagt tcgtttgctt ttgtaaatac agagggattt gtataagaaa tatctttaga 360
 25 aaaaccata tgctaatttg acataatttt tgagaaaaat atatattcag gcaattctc 420
 acaatgaaca ataataagat taaaatagct ttccccggt gcagcgcacg ggtatttttt 480
 ctagtaaaaa taaaagataa acttagactc aaaacattta caaaaacaac ccctaaagtt 540
 30 cctaaagccc aaagtctat ccacgatcca tagcaagccc agcccaacce aaccacaacc 600
 agcccacccc agtccagcca actggacaat agtctccaca cccccccact atcaccgtga 660
 35 gttgtccgca cgcaccgcac gtctcgcagc caaaaaaaaa aagaaagaaa aaaaagaaaa 720
 agaaaaaaca gcagggtgggt ccgggtcgtg ggggccggaa acgcgaggag gatcgcgagc 780
 cagcgacgag gccggccctc cctccgcttc caaagaaacg cccccatcg ccaactatata 840
 40 cataaccccc cctctcctcc cataccccca accctaccac caccaccacc accacctcca 900
 cctcctcccc cctcgtgccc ggacgacgag ctctcctccc ctccccctcc gccgccgccc 960
 45 cgccggtaac cacccgccc ctctctctt tctttctcg ttttttttc cgtctcggtc 1020
 tcgatctttg gccttgtag tttgggtggg cgagaggcgg ctctcgtccc cccagatcgg 1080
 tgcgcgggag gggcgggac tcgcggctgg ctctcctccc cgtggatccg gcccgatct 1140
 50 cgccgggaat ggggctctcg gatgtagac tgcgatccgc cgttggtggg gccgatgatg 1200



gggcccttaa aatttccgcc gtgctaaaca agatcaggaa gaggggaaaa gggcactatg 1260
 gtttatattt ttatatattt ctgctgcttc gtcaggctta gatgtgctag atctttcttt 1320
 5 cttctttttg tgggtagaat ttaatccctc agcattgttc atcggtagtt tttcttttca 1380
 tgattcgtga caaatgcagc ctggtgogga cgtttttttg taggtagaag tgatcaacca 1440
 tggcgcaagt tagcagaatc tgcaatgggtg tgcagaacce atctcttate tccaatctct 1500
 10 cgaaatccag tcaacgcaaa tctcccttat cggtttctct gaagacgcag cagcatccac 1560
 gaggcttatcc gatttcgtcg tcgtggggat tgaagaagag tgggatgacg ttaattggct 1620
 15 ctgagcttgc tcctcttaag gtcattgtctt ctgtttccac ggcgtgc atg gcc gct 1676
 Met Ala Ala
 tgc gag ctt cgc cca gcc acg cag tac gac acc gac gcc gtg tac gcg 1724
 20 Cys Glu Leu Arg Pro Ala Thr Gln Tyr Asp Thr Asp Ala Val Tyr Ala
 5 10 15
 ctg atc tgc gag ctc aag cag gcg gag ttc gac cac cac gcc ttc cgc 1772
 25 Leu Ile Cys Glu Leu Lys Gln Ala Glu Phe Asp His His Ala Phe Arg
 20 25 30 35
 gtg ggc ttc aac gcc aac ctg cgc gac ccc aac atg cgc tac cat ctg 1820
 30 Val Gly Phe Asn Ala Asn Leu Arg Asp Pro Asn Met Arg Tyr His Leu
 40 45 50
 gcg ctg ctt gat ggc gaa gtg gtc ggc atg atc gcc ctg cac ctc cag 1868
 Ala Leu Leu Asp Gly Glu Val Val Gly Met Ile Gly Leu His Leu Gln
 55 60 65
 35 ttc cac ctg cat cat gtc aac tgg atc ggc gag atc cag gag ctg gtc 1916
 Phe His Leu His His Val Asn Trp Ile Gly Glu Ile Gln Glu Leu Val
 70 75 80
 gtg atg cca cag gcg agg ggt ctg aac gtc ggc agc aag ctc ctg gcg 1964
 40 Val Met Pro Gln Ala Arg Gly Leu Asn Val Gly Ser Lys Leu Leu Ala
 85 90 95
 tgg gcc gag gag gaa gcc agg cag gcc gga gcc gag atg acc gag ctc 2012
 45 Trp Ala Glu Glu Glu Ala Arg Gln Ala Gly Ala Glu Met Thr Glu Leu
 100 105 110 115
 agc acc aac gtg aag cgc cac gac gcg cac cgc ttc tac ctg cgc gaa 2060
 Ser Thr Asn Val Lys Arg His Asp Ala His Arg Phe Tyr Leu Arg Glu
 120 125 130
 50 ggc tac gag cag agc cac ttc cgc ttc acc aag gcg ctg taaagatctg 2109
 Gly Tyr Glu Gln Ser His Phe Arg Phe Thr Lys Ala Leu



135

140

aattcccgat cgttcaaaca ttggcaata aagtttctta agattgaatc ctgttgccgg 2169

5 tcttgcgatg attatcatat aatttctggt gaattacggt aagcatgtaa taattaacat 2229

gtaatgcatg acgttatitta tgagatgggt ttttatgatt agagtcccgc aattatacat 2289

10 ttaatacgcg atagaaaaca aaatatagcg cgcaaactag gataaattat cgcgcgcggt 2349

gtcatctatg ttactagatc ggggatatc 2378

<210> 28

15 <211> 144

<212> PRT

<213> 人工序列

<400> 28

20 Met Ala Ala Cys Glu Leu Arg Pro Ala Thr Gln Tyr Asp Thr Asp Ala
1 5 10 15

Val Tyr Ala Leu Ile Cys Glu Leu Lys Gln Ala Glu Phe Asp His His
20 25 30

25 Ala Phe Arg Val Gly Phe Asn Ala Asn Leu Arg Asp Pro Asn Met Arg
35 40 45

30 Tyr His Leu Ala Leu Leu Asp Gly Glu Val Val Gly Met Ile Gly Leu
50 55 60

His Leu Gln Phe His Leu His His Val Asn Trp Ile Gly Glu Ile Gln
65 70 75 80

35 Glu Leu Val Val Met Pro Gln Ala Arg Gly Leu Asn Val Gly Ser Lys
85 90 95

Leu Leu Ala Trp Ala Glu Glu Glu Ala Arg Gln Ala Gly Ala Glu Met
100 105 110

40 Thr Glu Leu Ser Thr Asn Val Lys Arg His Asp Ala His Arg Phe Tyr
115 120 125

45 Leu Arg Glu Gly Tyr Glu Gln Ser His Phe Arg Phe Thr Lys Ala Leu
130 135 140

<210> 29

50 <211> 2107

<212> DNA

<213> 人工的

<220>
<223> 人工序列的描述: 表达盒, 其包含植物有效启动子, 所述启动子连接于一个前导序列、内含子、一个编码 AMPA 乙酰基转移酶的序列和一个终止序列

5 <220>
<221> 启动子
<222> (26).. (590)

10 <220>
<221> 5' UTR
<222> (615).. (685)

15 <220>
<221> 内含子
<222> (699).. (1148)

20 <220>
<221> 转运肽
<222> (1149).. (1426)

<220>
<221> CDS
<222> (1427).. (1858)

25 <220>
<221> 终止子
<222> (1869).. (2102)

<400> 29
30 gcggccgcgt taacaagctt ctgcaggtcc gatgtgagac ttttcaacaaggtaatat 60
ccggaacct cctcggattc cattgcccag ctatctgtca ctttattgtg aagatagtgg 120
aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc atcgttgaag 180
35 atgcctctgc cgacagtggc cccaagatg gacccccacc cagaggagc atcgtggaaa 240
aagaagacgt tccaaccacg tcttcaaagc aagtgattg atgtgatggt ccgatgtgag 300
40 acttttcaac aaaggtaat atccggaac ctctcggat tccattgccc agctatctgt 360
cactttattg tgaagatagt ggaaaaggaa ggtggctcct acaaatgcca tcattgcgat 420
aaaggaaagg ccatcgttga agatgcctct gccgacagtg gtcccaaaga tggaccccc 480
45 cccacgagga gcatcgtgga aaaagaagac gttccaacca cgtcttcaa gcaagtggat 540
tgatgtgata tctccactga cgtaaggat gacgcacaat cccactatcc ttcgcaagac 600
50 ccttctctta tataaggaag ttcatttcat ttggagagga cacgctgaca agctgactct 660
agcagatcct ctagaacctt cttccacaca ctcaagccac actattggag aacacacagg 720

gacaacacac cataagatcc aaggaggcc tccgccgcg ccggtacca cccgccct 780
 ctctctttc tttctcgtt ttttttccg tctcgtctc gatctttggc cttggtagt 840
 5 tgggtgggcg agaggcggt tctgtcgcgc ccagatcggg gcgcgggagg ggcgggatct 900
 cgccgggaat ggggctctcg gatgtagatc tgcgatccgc cgttgttggg ggagatgatg 960
 10 gggcgtttaa aatttcgcc tgctaaacaa gatcaggaag aggggaaaag ggcactatgg 1020
 tttatatttt tatatatttc tgctgcttcg tcaggcttag atgtgctaga tctttctttc 1080
 ttctttttgt gggtagaatt taatccctca gcattgttca tcggtagitt ttcttttcat 1140
 15 gatttcgtga caaatgcagc ctctgtcggg gctttttgt aggtagaagt gatcaacat 1200
 ggcccaagtt agcagaatct gcaatgggtg gcagaacca tctcttatct ccaatctctc 1260
 20 gaaatccagt caacgcaaat ctcccttate ggtttctctg aagacgcagc agcatccacg 1320
 agcttatccg atttcgtcgt cgtggggatt gaagaagagt gggatgacgt taattggctc 1380
 tgagcttcgt cctcttaagg tcatgtcttc tgtttccacg gcgtgc atg gcc gct 1435
 25 Met Ala Ala
 tgc gag ctt cgc cca gcc acg cag tac gac acc gac gcc gtg tac gcg 1483
 30 Cys Glu Leu Arg Pro Ala Thr Gln Tyr Asp Thr Asp Ala Val Tyr Ala
 5 10 15
 ctg atc tgc gag ctc aag cag gcg gag ttc gac cac cac gcc ttc cgc 1531
 Leu Ile Cys Glu Leu Lys Gln Ala Glu Phe Asp His His Ala Phe Arg
 20 25 30 35
 35 gtg ggc ttc aac gcc aac ctg cgc gac ccc aac atg cgc tac cat ctg 1579
 Val Gly Phe Asn Ala Asn Leu Arg Asp Pro Asn Met Arg Tyr His Leu
 40 45 50
 40 gcg ctg ctt gat ggc gaa gtg gtc ggc atg atc ggc ctg cac ctc cag 1627
 Ala Leu Leu Asp Gly Glu Val Val Gly Met Ile Gly Leu His Leu Gln
 55 60 65
 ttc cac ctg cat cat gtc aac tgg atc ggc gag atc cag gag ctg gtc 1675
 45 Phe His Leu His His Val Asn Trp Ile Gly Glu Ile Gln Glu Leu Val
 70 75 80
 gtg atg cca cag gcg agg ggt ctg aac gtc ggc agc aag ctc ctg gcg 1723
 50 Val Met Pro Gln Ala Arg Gly Leu Asn Val Gly Ser Lys Leu Leu Ala
 85 90 95
 tgg gcc gag gag gaa gcc agg cag gcc gga gcc gag atg acc gag ctc 1771



<210> 31
<211> 2436
5 <212> DNA
<213> 人工序列

<220>
10 <223> 人工序列的描述: 单子叶植物表达盒, 其包含植物有效启动子, 所述启动子
连接于一个内含子、一个编码 AMPA 乙酰基转移酶的序列和一个终止序列

<220>
15 <221> 启动子
<222> (26).. (640)

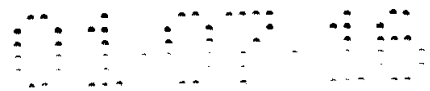
<220>
<221> 内含子
<222> (670).. (1473)

20 <220>
<221> 转运肽
<222> (1498).. (1725)

25 <220>
<221> CDS
<222> (1726).. (2157)

30 <220>
<221> 终止子
<222> (2172).. (2427)

<400> 31
gcggccgcgt taacaagctt ctgcaggtcc gatgtgagac tttcaacaa agggtaatat 60
35 ccggaaacct cctcggattc cattgccag ctatctgtca ctttattgtg aagatagtgg 120
aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc atcgttgaag 180
atgcctctgc cgacagtggc cccaaagatg gacccccacc cacgaggagc atcgttgaaa 240
40 aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatggt ccgatgtgag 300
acttttcaac aaagggtaat atccggaac ctctcggat tccattgccc agctatctgt 360
45 cactttattg tgaagatagt ggaaaaggaa ggtggctcct acaaatgcca tcattgcat 420
aaaggaaagg ccatcgttga agatgcctct gcgacagtg gtcccaaaga tggaccccca 480
cccacgagga gcatcgtgga aaaagaagac gttccaacca cgtcttcaaa gcaagtggat 540
50 tgatgtgata tctccactga cgtaagggat gacgcacaat cccactatcc ttcgcaagac 600



ccttcctcta tataaggaag ttcatttcat ttggagagga cacgctgaca agctgactct 660
 agcagatcta cegtcttcgg tacgcgctca ctccgccctc tgcctttggt actgccacgt 720
 5 tctctgaaat gctctcttgt gtgggtgattg ctgagagtgg tttagctgga tctagaatta 780
 cactctgaaa tcgtgttctg cctgtgctga ttacttgccg tcctttgtag cagcaaaaata 840
 tagggacatg gtagtacgaa acgaagatag aacctacaca gcaatacgag aaatgtgtaa 900
 10 tttgggtgctt agcggatattt atttaagcac atgttgggtgt tatagggcac ttggattcag 960
 aagtttgctg ttaatttagg cacaggcttc atactacatg ggtcaatagt atagggattc 1020
 15 atattatagg cgatactata ataatttgtt cgtctgcaga gcttattatt tgccaaaatt 1080
 agatattcct attctgtttt tgtttgtgtg ctgttaaatt gttaacgcct gaaggaataa 1140
 atataaatga cgaaattttg atgtttatct ctgctccttt attgtgacca taagtcaaga 1200
 20 tcagatgcac ttgttttaaa tattgttgc tgaagaaata agtactgaca gtattttgat 1260
 gcattgatct gcttgtttgt tgtaacaaaa tttaaaaata aagagtttcc tttttgttc 1320
 25 tctccttacc tcctgatggg atctagtatc taccaactga cactatattg cttctcttta 1380
 catacgtatc ttgctcgatg ccttctcct agtggtgacc agtgttactc acatagtctt 1440
 tgctcatttc attgtaatgc agataccaag cggcctctag aggatccagg agcaaccatg 1500
 30 gcgcaagtta gcagaatctg caatgggtgtg cagaacccat ctcttatctc caatctctcg 1560
 aaatccagtc aacgcaatc tccttatcg gtttctctga agacgcagca gcatccacga 1620
 35 gcttatccga tttcgtcgtc gtggggattg aagaagagtg ggatgacgtt aattggetct 1680
 gagcttcgtc ctcttaaggt catgtcttct gtttccacgg cgtgc atg gcc get tgc 1737
 Met Ala Ala Cys

40 gag ctt cgc cca gcc acg cag tac gac acc gac gcc gtg tac gcg ctg 1785
 Glu Leu Arg Pro Ala Thr Gln Tyr Asp Thr Asp Ala Val Tyr Ala Leu
 5 10 15 20

45 atc tgc gag ctg aag cag gcg gag ttc gac cac cac gcc ttc cgc gtg 1833
 Ile Cys Glu Leu Lys Gln Ala Glu Phe Asp His His Ala Phe Arg Val
 25 30 35

50 ggc ttc aac gcc aac ctg cgc gac ccc aac atg cgc tac cat ctg gcg 1881
 Gly Phe Asn Ala Asn Leu Arg Asp Pro Asn Met Arg Tyr His Leu Ala
 40 45 50

ctg ctt gat ggc gaa gtg gtc ggc atg atc ggc ctg cac ctc cag ttc 1929
 Leu Leu Asp Gly Glu Val Val Gly Met Ile Gly Leu His Leu Gln Phe
 55 60 65

5 cac ctg cat cat gtc aac tgg atc ggc gag atc cag gag ctg gtc gtg 1977
 His Leu His His Val Asn Trp Ile Gly Glu Ile Gln Glu Leu Val Val
 70 75 80

10 atg cca cag gcg agg ggt ctg aac gtc ggc agc aag ctc ctg gcg tgg 2025
 Met Pro Gln Ala Arg Gly Leu Asn Val Gly Ser Lys Leu Leu Ala Trp
 85 90 95 100

15 gcc gag gag gaa gcc agg cag gcc gga gcc gag atg acc gag ctc agc 2073
 Ala Glu Glu Glu Ala Arg Gln Ala Gly Ala Glu Met Thr Glu Leu Ser
 105 110 115

20 acc aac gtg aag cgc cac gac gcg cac cgc ttc tac ctg cgc gaa ggc 2121
 Thr Asn Val Lys Arg His Asp Ala His Arg Phe Tyr Leu Arg Glu Gly
 120 125 130

20 tac gag cag agc cac ttc cgc ttc acc aag gcg ctg taaagatctg 2167
 Tyr Glu Gln Ser His Phe Arg Phe Thr Lys Ala Leu
 135 140

25 aattcccgat cgttcaaca tttggcaata aagtttctta agattgaatc ctgttgccgg 2227
 tcttgcatg attatcatat aatttctggt gaattacggt aagcatgtaa taattaacat 2287
 gtaatgcatg acgttattta tgagatgggt ttttatgatt agagtcccgc aattatacat 2347

30 ttaatacgcg atagaaaaca aaatatagcg cgcaaactag gataaattat cgcgcgcggt 2407
 gtcatctatg ttactagatc ggggatatc 2436

35 <210> 32
 <211> 144
 <212> PRT
 <213> 人工序列

40 <400> 32
 Met Ala Ala Cys Glu Leu Arg Pro Ala Thr Gln Tyr Asp Thr Asp Ala
 1 5 10 15

45 Val Tyr Ala Leu Ile Cys Glu Leu Lys Gln Ala Glu Phe Asp His His
 20 25 30

 Ala Phe Arg Val Gly Phe Asn Ala Asn Leu Arg Asp Pro Asn Met Arg
 35 40 45

50 Tyr His Leu Ala Leu Leu Asp Gly Glu Val Val Gly Met Ile Gly Leu
 50 55 60

His Leu Gln Phe His Leu His His Val Asn Trp Ile Gly Glu Ile Gln
 65 70 75 80
 5 Glu Leu Val Val Met Pro Gln Ala Arg Gly Leu Asn Val Gly Ser Lys
 85 90 95
 Leu Leu Ala Trp Ala Glu Glu Glu Ala Arg Gln Ala Gly Ala Glu Met
 100 105 110
 10 Thr Glu Leu Ser Thr Asn Val Lys Arg His Asp Ala His Arg Phe Tyr
 115 120 125
 Leu Arg Glu Gly Tyr Glu Gln Ser His Phe Arg Phe Thr Lys Ala Leu
 15 130 135 140
 20