

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

C12N 15/54

C12N 15/82 C12N 9/10

C12N 5/00 C12P 17/06



[12] 发明专利申请公开说明书

[21] 申请号 00808784.9

[43] 公开日 2003 年 4 月 2 日

[11] 公开号 CN 1408024A

[22] 申请日 2000.4.14 [21] 申请号 00808784.9

[30] 优先权

[32] 1999. 4. 15 [33] US [31] 60/129,899

[32] 1999. 7. 30 [33] US [31] 60/146,461

[86] 国际申请 PCT/US00/10368 2000.4.14

[87] 国际公布 WO00/63391 英 2000.10.26

[85] 进入国家阶段日期 2001.12.11

[71] 申请人 卡尔根尼有限公司

地址 美国加利福尼亚州

[72] 发明人 B·萨维季 M·W·拉斯纳

J·D·韦斯

D·波斯特 - 比特恩米勒

[74] 专利代理机构 北京纪凯知识产权代理有限公司

代理人 程 伟

权利要求书 4 页 说明书 39 页 序列表 38 页
附图 26 页

[54] 发明名称 编码参与生育酚合成的蛋白质的核
酸序列

[57] 摘要

提供了生产改变了生育酚含量和组成的植物和种子的核酸序列和方法。 这些方法对提高植物中生育酚水平及在宿主植物细胞中产生希望的生育酚成分具有特殊的用途。

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

1. 一种编码异戊烯转移酶的分离的核酸序列。
2. 如权利要求 1 中所述的分离的核酸序列，其中所述异戊烯转移酶选自直链异戊烯转移酶和芳香族异戊烯转移酶。
3. 如权利要求 1 所述的所述的分离的 DNA 序列，其中所述核酸序列由真核细胞源分离得到。
4. 如权利要求 3 所述的分离的 DNA 序列，其中所述真核细胞源选自哺乳动物、线虫、真菌及植物细胞。
5. 如权利要求 4 中所述的 DNA 编码序列，其中所述异戊烯转移酶蛋白来自于拟南芥属。
6. 如权利要求 5 所述的 DNA 编码序列，其中所述异戊烯转移酶蛋白由选自图 1 的序列编码。
7. 如权利要求 4 所述的 DNA 编码序列，其中所述异戊烯转移酶蛋白来自于谷类。
8. 如权利要求 7 所述的 DNA 编码序列，其中所述异戊烯转移酶蛋白由包含图 3 所示序列的 EST 的序列编码。
9. 如权利要求 4 所述的 DNA 编码序列，其中所述异戊烯转移酶蛋白来源于大豆。
10. 如权利要求 9 所述的 DNA 编码序列，其中所述异戊烯转移酶蛋白由包含图 2 和图 9 所示序列的 ESTs 的序列编码。

11.如权利要求 1 所述的分离的 DNA 序列， 其中所述核酸序列从原核细胞源分离得到。

12.如权利要求 11 所述的所述的分离的 DNA 序列， 其中所述原核细胞源为集胞蓝细菌属。

13.一种核酸构建体， 包含作为有效结合组分的宿主细胞中的功能性转录起始区、 编码异戊烯转移酶的核酸序列和转录终止区。

14.如权利要求 13 所述的核酸构建体， 其中所述编码异戊烯转移酶的核酸序列由真核生物和原核生物的生物体中获得。

15.如权利要求 14 所述的核酸构建体， 其中所述编码异戊烯转移酶的核酸序列由植物源获得。

16.如权利要求 15 所述的核酸构建体， 其中所述编码异戊烯转移酶的核酸序列由拟南芥属、 大豆及谷物获得。

17.如权利要求 13 所述的核酸构建体， 其中所述编码异戊烯转移酶的核酸序列由集胞蓝细菌属获得。

18.一种植物细胞， 包含权利要求 13 所述的构建体。

19.一种改变宿主细胞中生育酚含量的方法， 包括： 用一构建体转化该宿主细胞， 该构建体包含作为有效结合组分的宿主细胞中的功能性转录起始区、 编码异戊烯转移酶的核酸序列和转录终止区。

20.如权利要求 19 所述的方法， 其中所述宿主细胞选自原核细胞和真核细胞。

21.如权利要求 20 所述的方法， 其中所述原核细胞是集胞蓝细菌属

22.如权利要求 20 所述的方法，其中所述真核细胞是植物细胞。

23.如权利要求 22 所述的方法，其中所述植物细胞由拟南芥属、大豆及谷物获得。

24.一种在宿主细胞中生产目的生育酚化合物的方法，该方法包括获得转化的宿主细胞，该宿主细胞在其基因组中含有并表达：

含有编码异戊烯转移酶的 DNA 序列的构建体，该构建体有效地连接到宿主细胞中的转录起始区，

其中所述异戊烯转移酶参与生育酚的合成。

25.如权利要求 24 所述的方法，其中所述宿主细胞选自原核细胞和真核细胞。

26.如权利要求 25 所述的方法，其中所述原核细胞是集胞蓝细菌属。

27.如权利要求 24 所述的方法，其中所述真核细胞是植物细胞。

28.如权利要求 27 所述的方法，其中所述植物细胞选自拟南芥属、大豆及谷物。

29.一种提高宿主细胞内生育酚产物的生物合成量的方法，该方法包括用一构建体转化该宿主细胞，该构建体包含作为有效结合组分的宿主细胞中的功能性转录起始区、编码异戊烯转移酶的核酸序列和转录终止区。

30.如权利要求 29 所述的方法，其中所述宿主细胞选自原核细胞和真核细胞。

31.如权利要求 30 所述的方法，其中所述原核细胞是集胞蓝细菌属。

32.如权利要求 30 所述的方法，其中所述真核细胞是植物细胞。

33.如权利要求 32 所述的方法，其中所述植物细胞由拟南芥属、大豆及谷物获得。

编码参与生育酚合成的蛋白质的核酸序列

引言

本申请要求美国在先申请 60/129, 899 (1999 年 4 月 15 日提交) 和美国在先申请 60/146, 461 (1999 年 7 月 30 日提交) 的优先权。

技术领域

本发明涉及核酸与氨基酸的序列和构建体, 及其相关方法。

背景

类异戊二烯类化合物广泛地存在于所有生物体中。植物合成多种系列的类异戊二烯类化合物, 超过 22, 000 种之多 (Connolly 和 Hill(1992) Dictionary of Terpenoids, Chapman and Hall, New York, NY)。在植物中, 类异戊二烯类化合物在特定的细胞功能诸如固醇的生成中起重要作用, 促成真核细胞膜的构造以及存在于泛醌和质体醌、生长调节剂 (如脱落酸、赤霉素、油菜素类固醇) 或光合色素叶绿素和类胡萝卜素的开链聚异戊二烯化合物的侧链。尽管其它的植物性类异戊二烯类化合物的生理作用 (比如, 多种系列的次级代谢产物的生理作用) 并不明显, 但是已知有些类异戊二烯类化合物在介导对不同的环境挑战的适应性反应中起关键作用。尽管类异戊二烯类化合物的结构和功能具有显著的多样性, 但是它们都源自单一种代谢前体, 即异戊烯二磷酸 (IPP) (Wright, (1961) Annu.Rev. Biochem. 20:525-548; 以及 Spurgeon 和 Porter, (1981) Biosynthesis of Isoprenoid Compounds., Porter 和 Spurgeon 编辑 (John Wiley, New York) 第 1 卷, 1-46 页)。

在高等植物的叶绿体中, 存在着源自类异戊二烯途径的若干独特和相互关联的生化途径, 它们导致次级代谢产物 (包括生育酚) 的生成。生育酚不仅在植物中起着至关重要的作用, 而且从哺乳动物的营养角度来看也是重要的。在质体中, 生育酚占总醌库的量达到 40%

生育酚和生育三烯酚（不饱和生育酚的衍生物）是众所周知的抗氧化剂，在保护细胞免受自由基损害以及预防多种疾病（包括心脏病、癌症、白内障、视网膜病、老年性痴呆和神经变性）中起重要的作用，而且已经表明它们对关节炎症状以及在抗衰老中具有有益的效应。维生素 E 被用于鸡饲料中，以提高鸡肉的货架寿命、外观、香味和氧化稳定性，以及将母育酚从饲料转移到鸡蛋中。已经表明，维生素 E 对家畜动物的正常生育是必需的，并改善家畜动物的总体表现和增强其免疫力。在动物饲料中补充维生素 E 还可赋予奶制品氧化稳定性。

天然生育酚作为补充物，在过去的三年里，对其需求一直在以 10-20% 的速度稳定地增长。目前，对天然生育酚的需求超过了供给，已知天然生育酚的生物活性，高于合成得到的生育酚的外消旋混合物。天然生育酚都是 d-立体异构体，而合成 α -生育酚是八种 d, l- α -生育酚异构体的混合物，只有其中的一种（占 12.5%）与天然 d- α -生育酚完全相同。天然 d- α -生育酚，与其它天然生育酚或生育三烯酚相比，具有最高的维生素 E 活性（1.49 IU/mg）。合成 α -生育酚的维生素 E 活性为 1.1 IU/mg。1995 年，世界市场对生育酚精制原料的需求为 10.2 亿美元；其中合成物质占市场的 85-88%，剩下的 12-15% 为天然物质。天然生育酚和生育三烯酚的最佳来源物为植物油和谷类产品。当前，大部分的天然维生素 E 是产自 γ -生育酚，而 γ -生育酚又出自豆油加工，随后通过化学修饰可将其转化为 α -生育酚（ α -生育酚表现出最高的生物活性）。

不通过化学修饰而提高植物中生育酚和生育三烯酚（特别是那些能够直接使用的更希望得到的化合物）水平的方法，将有助于本领域，这是由于这些分子表现出更好的功能性和生物可利用性。。

另外，希望得到这样的方法，其可增加宿主植物细胞中其它类异戊二烯衍生化合物的产量。而且，还需要在宿主植物细胞中产生特定的类异戊二烯化合物的方法。

发明概述

本发明涉及异戊烯转移酶（PT），更具体地，涉及 PT 的多核苷酸和多肽。本发明的多核苷酸和多肽包括源自真核细胞来源物的那些多核苷酸和多肽。

因而，本发明的一个方面涉及编码异戊烯转移酶蛋白质的分离的多核苷酸序列。更具体地，本发明提供了编码来自细菌和植物源的 PT 蛋白质的分离的核酸序列。

本发明的另一方面涉及一些寡聚核苷酸，包括部分或全部 PT 编码序列。

本发明的又一个方面提供了可用于转录或者转录及翻译（表达）异戊烯转移酶的重组 DNA 构建体。更具体地，本发明提供了能够在宿主细胞中转录或者转录及翻译的构建体。

在本发明的另一个方面中，提供了在宿主细胞或其后代中产生异戊烯转移酶的方法。更具体地，用可以用于转录或者转录及翻译异戊烯转移酶的 DNA 构建体对宿主细胞进行了转化或转染。含有异戊烯转移酶的重组细胞也是本发明的一部分。

另一方面，本发明涉及使用多核苷酸和多肽序列改变宿主细胞（尤其是宿主植物细胞）中的生育酚含量的方法。在本文中构思了含有如此改变的生育酚含量的植物细胞。

通过表达异戊烯转移酶而得到的被修饰的植物、种子和油也被认为是本发明的一部分。

附图的简要说明

图 1.提供了排列于 ATPT2、ATPT3、ATPT4、ATPT8 和 ATPT12 之间的氨基酸序列，通过 Clustal W 分析得到。

图 2 提供了表达构建体 pCGN10800 的示意图。

图 3 提供了表达构建体 pCGN10801 的示意图。

图 4 提供了表达构建体 pCGN10803 的示意图。

图 5 提供了表达构建体 pCGN10806 的示意图。

图 6 提供了表达构建体 pCGN10807 的示意图。

图 7 提供了表达构建体 pCGN10808 的示意图。

图 8 提供了表达构建体 pCGN10809 的示意图。

图 9 提供了表达构建体 pCGN108010 的示意图。

图 10 提供了表达构建体 pCGN10811 的示意图。

图 11 提供了表达构建体 pCGN10812 的示意图。

图 12 提供了表达构建体 pCGN10813 的示意图。

图 13 提供了表达构建体 pCGN10814 的示意图。

图 14 提供了表达构建体 pCGN10815 的示意图。

图 15 提供了表达构建体 pCGN10816 的示意图。

图 16 提供了表达构建体 pCGN10817 的示意图。

图 17 提供了表达构建体 pCGN10819 的示意图。

图 18 提供了表达构建体 pCGN10824 的示意图。

图 19 提供了表达构建体 pCGN10825 的示意图。

图 20 提供了表达构建体 pCGN10826 的示意图。

图 21 给出了 Clustal W 分析得到的集胞蓝细菌属 (*Synchocystis*) 序列敲除间的氨基酸序列排列。

图 22 给出了拟南芥属中 ATPT2、ATPT3、ATPT4、ATPT8 和 ATPT12 蛋白序列的氨基酸序列, 以及集胞蓝细菌属中 slr1736、slr0926、sll1899、slr0056 和 slr1518 氨基酸序列。

图 23 给出了野生型集胞蓝细菌属菌株 6803 和集胞蓝细菌属 slr1736 敲除的集胞蓝细菌属制备物的酶分析结果。

图 24 给出了包含 pCGN10822 的转基因拟南芥属种子提取物 HPLC 分析资料的条线图, pCGN10822 提供了从 napin 启动子起始向有义方向的 ATPT2 序列的表达。给出了 α -, γ -, 及 δ -生育酚的图谱, 同时也给出了 22 种转化细胞、一种非转化(野生型)对照组的总生育酚图谱。

图 25 给出了用 pCGN10803 (35S-ATPT2, 向反义方向)、pCGN10802 (1625 行, napin ATPT2 向有义方向)、pCGN10809 (1627 行, 35S-ATPT3 向有义方向) 转化的植物拟南芥属种子提取物的 HPLC 分析条线图, 同时也给出了非转化(野生型)对照组及经空载体转化对照组的图谱。

发明的详细描述

本发明尤其提供了改变(如增加和减少)生育酚水平和/或调整其在宿主细胞内含量的组成和方法。本发明特别提供了多核苷酸、多肽和在宿主植物细胞中调整生育酚含量的方法。

本发明提供了参与直链和芳香基化合物异戊烯化的多聚核苷酸及多肽序列。本文所用的直链异戊烯基转移酶包括编码参与直链化合物异戊烯化的蛋白质序列,该直链化合物包括但不限于牻牛儿苗牻牛儿苗焦磷酸和法尼基焦磷酸。本文所用的芳香基异戊烯基转移酶包括编码参与芳香基化合物异戊烯化的蛋白质序列,该芳香基化合物包括但不限于甲基萘醌、泛醌、叶绿素及尿黑酸。本发明中的异戊烯基转移酶优选尿黑酸的异戊烯化的蛋白质序列。

高等植物中 α -生育酚的生物合成涉及到尿黑酸和叶绿基焦磷酸缩合形成 2-甲基-6 叶绿基苯醌,其可以通过环化和随后的甲基化形成多种生育酚(Fiedler 等, 1982, *Planta*, 155:511-515, Soll 等, 1980, *Arch. Biochem. Biophys.* 204:544-550, Marshall 等, 1985, *Phytochem.*, 24:1705-1711, 这些文献的全部内容在此引入作为参考)。Normiset 等鉴定和描述了,拟南芥属 *dps2* 突变在生育酚和质体醌-9 的积聚物中是不存在的。进一步的遗传学和生物化学分析显示 *PDS2* 编码的蛋白质是造成尿黑酸异戊烯化的原因,这在生育酚合成中可能是限速步骤,该基因现在仍未被分离。因此,本发明的一个方面是提供参与尿黑酸异戊烯化的多聚核苷酸和多肽。

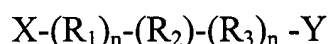
分离的多聚核苷酸和多肽

本发明的第一个方面涉及分离的异戊烯基转移酶多核苷酸,该多核苷酸包括编码本发明多肽的多核苷酸,这些多肽具有选自序列表列出的推定氨基酸序列。本发明还涉及其它与这些多核苷酸序列及其变异体紧密相关的多核苷酸序列。

本发明提供了一种多核苷酸序列,其整个序列与在序列表中列出的每个编码序列完全相同。本发明还提供了成熟多肽或其片段的编码序列,以及与其它编码序列诸如那些编码导肽或分泌序列、前-、原-或前原-蛋白质序列处于同一读框中的成熟多肽或其片段的编码序列。

该多核苷酸还可包含非编码序列，例如包括但不限于非编码的 5' 和 3' 序列，诸如转录但不翻译的序列、终止信号、核糖体结合位点、稳定 mRNA 的序列、内含子、多腺苷酸化信号和编码所添加氨基酸的添加编码序列。例如，可以将一个标记序列包括在内，以促进融合多肽的纯化。本发明的多核苷酸还包括含有结构基因和控制基因表达的天然结合序列的多核苷酸。

本发明还包括下式的多核苷酸：



其中，位于 5' 末端的 X 是氢，位于 3' 末端的 Y 是氢或金属， R_1 和 R_3 为任何核酸残基， n 是一个介于 1-3000 之间的整数（优选介于 1-1000 之间）， R_2 是本发明的核酸序列，尤其是选自序列表列出的核酸序列并优选 SEQ ID NOs: 1、3、5、7、8、10、11、13-16、18、23、29、36 和 38 中的那些核酸序列。在上式中， R_2 是有方向性的，其 5' 末端残基位于左边，结合于 R_1 ，而其 3' 末端残基位于右边，结合于 R_3 。通过每组 R (R 大于 1) 表示的任何核酸残基片段，可以是杂聚物，也可以是均聚物，优选为杂聚物。

本发明还涉及本文所述的多核苷酸的变异体，它们编码本发明的多肽的变异体。那些本发明的多核苷酸片段的变异体，可被用于合成本发明的全长多核苷酸。优选的实施方案为编码多肽变异体的多核苷酸，其中对本发明的多肽序列的 5-10，1-5，1-3，2，1 或无氨基酸残基被以任何组合替代、添加或缺失。特别优选沉默的替代、添加和缺失，因为它们不会改变该多核苷酸或多肽的性能或活性。

本发明的进一步优选的实施方案为：在多核苷酸全长内与编码本发明多肽的多核苷酸有至少 50%、60% 或 70% 相同性的多核苷酸，以及与这种多核苷酸互补的多核苷酸。更优选含有其全长与编码本发明多肽的多核苷酸及其互补多核苷酸有至少 80% 相同性的区域的多核苷酸。在这一点上，尤其优选其全长与编码本发明多肽的多核苷酸及其互补多核苷酸有至少 90% 相同性，特别优选那些有至少 95% 相同性的多核苷酸序列。进一步地，非常优选那些有至少 97% 相同性的多核苷酸序列，而且尤其非常优选那些有至少 98% 和 99% 相同性的多核苷酸序列，其中最优选那些有至少 99% 相同性的多核苷酸序列。

优选的实施方案为这样的多核苷酸，它们所编码的多肽基本上保留了与成熟多肽（由序列表中列出的多核苷酸编码）相同的生物学功能或活性。

本发明进一步涉及可与上述序列杂交的多核苷酸。更具体地，本发明涉及在严格条件下可与上述多核苷酸杂交的多核苷酸。本文所用的术语“严格条件”和“严格杂交条件”是指在序列间具有至少 95%（优选至少 97%）相同性的情况下，一般会发生杂交反应。严格杂交条件的一个例子是，在含有 50% 甲酰胺、5×SSC（150mM NaCl, 15mM 柠檬酸三钠）、50mM 磷酸钠（pH7.6）、5×Denhardt's 溶液、10% 硫酸葡聚糖和 20 毫克/毫升变性剪切鲑精 DNA 的溶液中，在 42℃ 温育过夜，随后在 0.1×SSC 中于大约 65℃ 冲洗杂交载体。其它的杂交和冲洗条件是公知的，并且在 Sambrook 等人的分子克隆：实验室手册（Molecular Cloning: A Laboratory Manual），Cold Spring Harbor, NY（1989）第二版，尤其是第 11 章中有例子。

本发明还提供一种基本上由一种多核苷酸序列组成的多核苷酸，这样的多核苷酸可如此获得：在严格杂交条件下，用在多核苷酸序列表中列出的多核苷酸序列或其片段序列作为探针，筛选含有该种多核苷酸序列的完整基因的合适的文库；然后分离所述的核苷酸序列。用于获得这样一种多核苷酸的片段包括例如本文中所述的探针和引物。

正如本文所讨论的关于本发明的多核苷酸测定法，例如，本发明的多核苷酸可被用作 RNA、cDNA 或基因组 DNA 的杂交探针，以分离全长 cDNA 或编码多肽的基因组克隆，以及分离与序列表中列出的多核苷酸具有高度序列相似性的其它基因的 cDNA 或基因组克隆。这样的探针一般含有至少 15 个碱基。优选地，这样的探针具有至少 30 个碱基并可以具有至少 50 个碱基。尤其优选的探针将具有 30-50 个碱基，其中包括 30 和 50 个在内。

含有或含于序列表中列出的多核苷酸序列的每一种基因，其编码区可以利用序列表中提供的 DNA 序列筛选，以合成寡核苷酸探针。然后，使用与本发明的基因的序列具有序列互补性的标记寡核苷酸，对 cDNA、基因组 DNA 或 mRNA 的文库进行筛选，以鉴定可与探针杂交的文库中的成员。例如，制备与异戊烯基转移酶 EST 序列相应的

合成的寡核苷酸。以寡核苷酸为引物，以聚合酶链式反应（PCR）技术获得异戊烯基转移酶基因的 5'端和 3'端序列。或者，在可以根据特定的异戊烯基转移酶肽制备出低简并性寡核苷酸的情况下，也可以将这样的探针直接用于筛选基因文库，以获得异戊烯基转移酶基因序列。尤其是，筛选噬菌体载体中的 cDNA 文库有助于这种方法，因为杂交的背景较低。

一般地，使用核酸探针所获得的异戊烯基转移酶序列显示出，靶异戊烯基转移酶序列与用作探针的编码序列之间 60-70% 的序列相同。但是，也可以获得低至 50-60% 序列相同性的长序列。核酸探针可以是长的核酸序列片段，也可以是较短的寡核苷酸探针。当使用较长的核酸片段作为探针时（大于大约 100 bp），为了从靶样品中获得与探针序列具有 20-50% 偏差（即，50-80% 的序列相同性）的序列，在进行筛选时可以将杂交的严格性降低。寡核苷酸探针可以比编码异戊烯基转移酶的完整核酸序列短许多，但至少也应该为大约 10 个，优选至少为大约 15 个，更优选至少为大约 20 个核苷酸。当使用较短的区域作探针时，与较长的区域相比，希望能达到更高的序列相同性程度。因此，希望鉴定出高度保守的氨基酸序列区，以设计寡核苷酸探针，用于探测和回收其它相关的异戊烯基转移酶基因。较短的探针对于聚合酶链式反应（PCR）常常特别有用，并且在可以鉴定出高度保守序列的情况下尤其如此（参见 Gould 等。，PNAS USA（1989）86: 1934-1938）。

本发明的另一方面涉及异戊烯基转移酶多肽。这种多肽包括：在序列表中列出的分离的多肽及其多肽和片段，尤其是那些显示异戊烯基转移酶活性并且与从序列表所列序列中选出的多肽序列具有至少 50%、60% 或 70%，优选至少 80%，更优选至少 90%，最优选至少 95% 相同性的多肽，还包括这些多肽的组成部分，其中多肽的组成部分优选包含至少 30 个氨基酸，更优选包含至少 50 个氨基酸。

“相同性”，正如本领域所熟知的，是两种或更多种多肽序列间或者两种或更多种多核苷酸序列间的相互关系，其可通过比较这些序列而确定。在本领域中，“相同性”也指多肽或多核苷酸序列间的序列相关性程度，其可通过对这些序列一串串地进行比对而确定。“相

同性”可以容易地通过已知的方法计算出，这些方法包括但不限于：在 Computational Molecular Biology, Lesk, A. M.编辑，牛津大学出版社，纽约，1988；Biocomputing: Informatics and Genome Projects, Smith, D. W. 编辑，Academic Press, 纽约, 1993；Computer Analysis of Sequence Data, Part I, Griffin, A. M.和 Griffin, H. G.编辑，Humana Press, New Jersey, 1994；Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987；Sequence Analysis Primers, Gribskov, M.和 Devereux, J.编辑，Stockton Press, 纽约, (1991)；以及 Carillo, H.和 Lipman, D., SIAM J Applied Math., 48: 1073 (1988) 中描述的那些方法。在设计确定相同性的方法时，目的就是给出待测试序列间的最大匹配。而且，确定相同性的方法被编写在可公开获得的程序中。可用于确定两种序列间相同性的电脑程序包括但不限于：GCG (Devereux, J.等, Nucleic Acids Research, 12 (1): 387 (1984))；一组五个 BLAST 的程序，其中的三个用于核酸序列查询 (BLASTN、BLASTX 和 TBLASTX)，两个用于蛋白质序列查询 (BLASTP 和 TBLASTN) (Coulson, Trends in Biotechnology, 12: 76-80 (1994)；Birren 等, Genome Analysis, 1: 543-559 (1997))。BLAST X 程序可以从 NCBI 和其它途径公开获得 (BLAST Manual, Altschul, S.等, NCBI NLM NIH, Bethesda, MD 20894; Altschul, S.等, J. Mol. Biol, 215: 403-410 (1990))。公知知的 Smith Waterman 算法也可用于确定相同性。

多肽序列比较参数一般包括：

算法：Needleman 和 Wunsch, J.Mol.Biol.48:443-453(1970)

比较矩阵：BLOSSUM62，来自 Hentikoff 和 Hentikoff, Proc. Natl. Acad. Sci USA 89:10915-10919(1992)

空位处罚分：12

空位长度处罚分：4

能够采用这些参数的一个程序，可从 Genetics Computer Group, Madison, Wisconsin 公开获得，这个程序被称为“空位”程序。上述参数，连同对末端空位没有处罚分，就是进行肽比较的缺省设置参数。

多核苷酸序列比较的参数包括：

算法：Needleman 和 Wunsch, J. Mol. Biol. 48:443-453(1970)

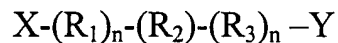
比较矩阵：匹配=+10，错配=0

空位处罚分：50

空位长度处罚分：3

能够采用这些参数的一个程序，可从 Genetics Computer Group, Madison, Wisconsin 公开获得，这个程序被称为“空位”程序。上述参数就是进行核酸比较的缺省设置参数。

本发明还包括下式的多肽：



其中，位于氨基末端的 X 是氢，位于羧基末端的 Y 是氢或金属， R_1 和 R_3 为任何氨基酸残基，n 是一个介于 1-1000 之间的整数， R_2 是本发明的氨基酸序列，尤其是选自序列表中列出的氨基酸序列，优选在 SEQ ID NO: 2、4、6、9、12、17、19-22、24-28、30、32-37 和 39 中提供的序列编码的那些氨基酸序列。在上式中， R_2 是有方向性的，其氨基末端残基位于左边，结合于 R_1 ，而其羧基末端残基位于右边，结合于 R_3 。用一组 R (R 大于 1) 表示的任何氨基酸残基片段，可以是杂聚物，也可以是均聚物，优选为杂聚物。

本发明的多肽包括由多核苷酸编码的分离的多肽，其中的多核苷酸含有选自本文序列表中列出的序列。

本发明的多肽可以是成熟的蛋白质或者可以是融合蛋白质的一部分。

这些多肽的片段和变异体也被认为是本发明的一部分。一个片段就是一个变异体多肽，它所具有的氨基酸序列与前面描述的多肽的氨基酸序列的一部分而不是全部完全相同。这些片段可以独立存在，也可以含于一个更大的多肽中，其中该片段形成一部分或一个区域，最优选是形成单一连续区域。优选的片段为具有生物学活性的片段，它们可以介导本发明多肽的活性，其中包括那些具有相似活性或提高活性或降低活性的片段。本发明还包括那些在动物尤其是人中具有抗原性或免疫原性的片段。

这些多肽的变异体还包括与序列表中列出的序列不同的多肽，这些不同是由保守氨基酸的取代产生的，即一个氨基酸残基被另一个具有相似性质的氨基酸残基取代。一般地，这中取代发生在 Ala、Val、Leu 和 Ile 之间，Ser 和 Thr 之间，Asp 和 Glu 之间，Asn 和 Gln 之间，Lys 和 Arg 之间，或 Phe 和 Tyr 之间。尤其优选其中 5-10、1-5、1-3 个或者 1 个氨基酸被以任何组合取代、缺失或添加的变异体。

那些是本发明多肽片段的变异体，通过肽合成可用于产生相应的全长多肽。因而，可将这些变异体用作生成本发明的全长多肽的中间体。

本发明的多核苷酸和多肽可用于，例如，转化宿主细胞诸如植物宿主细胞，正如本文所进一步讨论的。

本发明还提供了编码多肽的多核苷酸，其中的多肽为成熟蛋白质再加上添加的氨基或羧基末端氨基酸，或者成熟的多肽中的氨基酸（例如，当蛋白质的成熟形式具有多于一个多肽链时）。这种序列可以，例如，在蛋白质从前体到成熟形式的转变加工中起作用，帮助蛋白质转运，缩短或延长蛋白质半寿期，或者在蛋白质的测定或生产中简化操作。预期可使用胞内酶从成熟的蛋白质中除去任何添加的氨基酸。

具有融合到一种或多种前序列（prosequence）的成熟形式的多肽的前体蛋白质，可能是该多肽的无活性形式。一般地，当前序列被除去时，无活性前体被激活。在激活之前可以将一部分或者全部前序列除去。这种前体蛋白质一般称为前蛋白质。

植物构建及应用方法

我们尤其感兴趣的是将核苷酸序列用于重组 DNA 构建体，以指导本发明的异戊烯基转移酶序列在宿主细胞中的转录或转录及翻译（表达）。表达构建体一般含有宿主细胞中的功能性启动子，该启动子有效地连接到本发明的编码异戊烯基转移酶的核酸序列，还含有宿主细胞中的功能性转录终止区。

当序列的排列方式使得第一种核酸序列影响到第二种核酸序列的功能时，这两种核酸序列就是“有效地连接”或“有效地结合”。优

选地，这两种序列是单一连续核酸分子的一部分，更优选地，这两种序列相邻。例如，如果一个启动子调节或介导一个基因在细胞中的转录，那么，所述启动子就与所述基因进行了有效地连接。

本领域的技术人员会认识到，在植物细胞中有许多功能性启动子，并且已在文献中对它们进行了描述。还展望了叶绿体和质体特异性启动子、叶绿体或质体中功能性启动子，以及在叶绿体或质体中有效的启动子。

一组植物功能性启动子为组成型启动子如 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。已将 CaMV35S 和 FMV35S 启动子的增强或复制版本用于本发明的实践中（Odell 等人，1985，*Nature*，313：810-812；Rogers，美国专利第 5, 378, 619 号）。另外，可优选让异戊烯基转移酶基因在植物的特定组织（如叶、茎、根、块茎、种子、果实等）中表达，并且所选择的启动子应具有所需要的组织和发育特异性。

尤其感兴趣的是，使本发明的核酸序列从转录起始区开始表达，并优先在植物种子组织中表达。这种优先在种子中转录起始的序列例子包括那些源自编码植物储藏蛋白质基因序列或者油料种子中参与脂肪酸生物合成的基因的序列。这种启动子的例子包括来自诸如 napin（Kirdl 等，*Seed Sci. Res.*，1：209-219（1991））、菜豆蛋白、玉米醇溶蛋白、大豆胰蛋白酶抑制剂、ACP、硬脂酰-ACP 去饱和酶、大豆 β -conglycin 的 α' 亚单位（大豆 7s，（Chen 等 *Proc. Natl. Acad. Sci.*，83:8560-8564（1986）））和油质蛋白的基因的 5' 调节区。

指导具有异戊烯基转移酶的蛋白质定位于特定的亚细胞区室例如线粒体、内质网、液泡、叶绿体或其它质体区室中，也许是有益的。例如，当将本发明的目的基因靶向质体诸如叶绿体以表达时，构建体也要使用指导基因至质体的序列。这种序列在本文中称为叶绿体转运肽（CTP）或质体转运肽（PTP）。这样，在没有将目的基因直接插入到质体中的情况下，表达构建体将另外含有编码转运肽的基因，从而指导目的基因至质体。叶绿体转运肽可以是源自目的基因，也可以是源自具有 CTP 的异种序列。这种转运肽在本领域中是公知的。参见例如，Von Heijne 等（1991）*Plant Mol. Boil. Rep.* 9:104-126；Clark 等

(1989) J. Biol. Chem. 264:17544-17550; della-Cioppa 等(1987) Plant Physiol. 84:965-968; Romer 等(1993) Biochem. Biophys. Res Commun. 196:1414-1421; 和 Shah 等(1986) Science 233:478-481。

根据目的用途, 构建体可以含有编码整个异戊烯基转移酶蛋白质或其一个组成部分的核酸序列。例如, 当需要反义抑制给定的异戊烯基转移酶蛋白质时, 就不需要完整的异戊烯基转移酶序列。而且, 当用于构建体中的异戊烯基转移酶序列是旨在用作探针时, 则制备只含有异戊烯基转移酶编码序列的特定组成部分(例如, 被发现是编码高度保守的异戊烯基转移酶区的序列)的构建体可能是有益的。

技术人员会认识到, 用于抑制内源序列在宿主细胞中表达的方法有许多。这些方法包括但不限于反义抑制(Smith 等(1988) Nature, 334: 724-726)、共抑制(Napoli 等(1989) Plant Cell, 2: 279-289)、核酶(PCT 公开说明书, WO 97/10328)以及有义与反义的组合(Waterhouse 等(1998), Proc. Natl. Acad. Sci. USA 95:13959-13964)。在宿主细胞中抑制内源序列的方法一般使用了转录或转录及翻译至少一部分待抑制序列。这种序列可以与内源序列的编码区同源, 也可以与内源序列的非编码区同源。

本发明的植物表达构建体也可以提供调节转录本的终止区。转录本的终止区可以由编码异戊烯基转移酶的 DNA 序列提供, 或者由不同基因来源的方便的转录终止区提供, 例如, 与转录本起始区天然结合的转录本终止区。技术人员会认识到, 任何能够在植物细胞中终止转录的方便的转录本终止区, 都可以在本发明的构建体中予以采用。

或者, 也可以将构建体制备为可以指导异戊烯基转移酶序列直接从宿主植物细胞的质体表达。这种构建体和方法为本领域所公知, 并一般描述于, 例如 Svab 等(1990) Proc. Natl. Acad. Sci. USA 87:8526-8530 和 Svab 和 Maliga (1993) Proc. Natl. Acad. Sci. USA 90:913-917 以及美国专利第 5,693,507 号。

本发明中的异戊烯基转移酶构建体可以与另外的构建体一起在转化方法中使用, 该另外的构建体提供编码参与生产生育酚或生育酚前体(如同源酸和/或叶绿基焦磷酸)蛋白质的核苷酸序列的表达。参与生产尿黑酸的蛋白质的核苷酸序列在本领域中是公知的, 它包括但不限于

限于 4-羟苯基丙酮酸双加氧酶 (HPPD, EC1.13.11.27, 如在 Garcia 等 (1984) J. Mol. Biol. 180:1023-1051 中所述)、单一或双官能 tyrA (如 Xia 等 (1992) J. Gen Microbiol. 138:1309-1316, 以及 Hudson 等 (1984) J. Mol. Biol. 180:1023-1051 中所述)加氧酶、4-羟苯基丙酮酸二(9Cl), 4-羟苯基丙酮酸双加氧酶、对- 4-羟苯基丙酮酸双加氧酶、对- 4-羟苯基丙酮酸羟化酶、对- 4-羟苯基丙酮酸氧化酶、对- 4-羟苯基丙酮酸羟化酶、对- 4-羟苯基丙酮酸氧化酶、4-羟苯基乙酸, NAD(P)H : 氧化还原酶(1-羟化)、4-羟苯基乙酸 1-一加氧酶等。另外, 编码参与生产叶绿基焦磷酸蛋白质的核酸序列的表达构建体也可以与本发明中的异戊烯基转移酶构建体一起使用。编码参与生产叶绿基焦磷酸蛋白质的核酸序列在本领域中是公知的, 它包括但不限于牻牛儿苗焦磷酸合酶 (GGPPS), 牻牛儿苗焦磷酸还原酶 (GGH), 1-脱氧木酮糖 5-磷酸合酶, 1-脱氧-D-木糖-5-磷酸盐还原异构酶、4-二磷酸胞苷-2-C-甲基赤藓糖醇合酶、异戊基焦磷酸异构酶。

本发明中的异戊烯基转移酶序列可以用于制备转化构建体, 该构建体含有一个二级表达盒来表达参与生育酚生物合成的其它序列。本发明中令人感兴趣的另外的参与生育酚生物合成的序列包括但不限于 γ 生育酚甲基化转移酶 (Shintani 等 (1998) Science 282(5396):2098-2100)、生育酚环化酶和生育酚甲基化酶。

当将含有表达构建体的重组 DNA 构建体引入到植物的细胞、组织、器官中或植物中时, 我们就认为植物的细胞、组织、器官或植物已被转化、转染或转基因。转基因或转化的细胞或植物也包括细胞或植物的子代, 以及由育种项目而产生的子代, 该育种项目在杂交试验中采用这样的转基因植物作为亲代, 并且这些子代由于异戊烯基转移酶核酸序列的存在而表现出变化的表型。

具有异戊烯基转移酶并作为提高或减低其表达的目的 DNA 序列的植物表达或转录型构建体, 可用于广泛多种植物生命体中, 尤其是参与生成食用或工业用途的植物油的植物生命体。用于本发明方法的尤其优选的植物包括但不限于: 金合欢、苜蓿、茴香树、苹果树、杏树、洋蓟、芝麻菜、石刁柏、鳄梨树、芭蕉属植物、大麦、豆科植物、甜菜、黑莓、乌饭树、嫩茎花椰菜、球芽甘蓝、卷心菜、

canola、罗马甜瓜、胡萝卜、木薯、花椰菜、芹菜、樱桃树、菊苣属植物、芜荑叶、柑橘属植物、克莱门氏小柑橘、咖啡树、玉米、棉花、黄瓜、Douglas 冷杉、茄子、苣荬菜、escarole、桉树属植物、茴香、无花果属植物、大蒜、葫芦属植物、葡萄、葡萄柚、加甜味烟草、凉薯、kiwifruit、莴苣、韭葱、柠檬树、酸橙、火炬松、芒果、甜瓜、菌类植物、油桃、坚果、燕麦属植物、油棕榈、菜籽油油菜、秋葵、洋葱、柑桔、一种装饰性植物、番木瓜树、欧芹、豌豆、桃树、花生、梨树、胡椒属植物、柿树、松树、菠萝、车前草、李属植物、石榴树、杨属植物、马铃薯、南瓜、温柏树、放射状松树、菊苣、小萝卜、木莓、水稻、黑麦、高粱、美国长叶松、大豆、菠菜、南瓜、草莓、糖甜菜、甘蔗、向日葵、白薯、枫香属植物、红橘树、茶树、烟草、西红柿、黑小麦、草皮、芜菁、一种葡萄树、西瓜、小麦、薯蓣属植物和夏季产南瓜。

最优选的是温带油料作物。感兴趣的温带油料作物包括但不限于：油菜仔（Canula 和高芥子酸变种）、向日葵、红花、棉花、大豆、花生、椰子和油棕以及玉米。依据引入重组构建体至宿主细胞内所使用的方法，也许需要其它 DNA 序列。重要的是，本发明适用于双子叶种类的植物和单子叶种类的植物，并可以容易地适用于新的和/或改进的转化和调节技术。

尤其感兴趣的是，在植物中使用了异戊烯基转移酶构建体，以产生这样的植物或植物体部分（其中包括但不限于叶、茎、根、生殖性部分和种子），在具有转化植物细胞的植物体部分中，生育酚的含量被改变。

为了进行免疫学筛选，用纯化的蛋白质或其组成部分注射兔或小鼠就可以制备针对该蛋白质的抗体，这种制备抗体的方法为本领域人员所公知。可以制备产生单克隆抗体或多克隆抗体，尽管在一般情况下多克隆抗体对于基因分离更有用。可以进行 Western 分析，以确定在所希望的植物种类的粗提取物中有相关蛋白质的存在，例如采用针对被编码蛋白质的抗体通过交叉反应来确定。当观察到交叉反应时，通过筛选代表所希望植物种类的表达文库，将编码相关蛋白质的基因分离。表达文库可构建于多种可商购载体中，包括 λ gt11，正如在

Sambrook 等在 *Molecular Cloning: A Laboratory Manual* 第二版 (1989) Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 中所描述的。

为了证实被鉴定为异戊烯基转移酶的核酸序列所编码的蛋白质的活性和特异性, 采用杆状病毒表达系统在昆虫细胞培养物中进行了体外测定。这种杆状病毒表达系统为本领域所公知, 并且在 Lee 等人的美国专利 5, 348, 886 中有描述, 本文引用其全部作为参考。

另外, 可以制备其它表达构建体, 以使用不同的表达系统测定蛋白质活性。将这种表达构建体转化到酵母或原核细胞宿主内并测定异戊烯基转移酶活性。这种表达系统为本领域所公知, 并可容易地通过商业途径获得。

除了在本发明中所描述的序列外, 用于本发明的 DNA 编码序列也可源自藻、真菌、细菌、哺乳动物来源物、植物等。可以使用与异戊烯基转移酶的保守核苷酸或氨基酸序列相应的特征序列, 在已有的数据库中进行同源性搜索, 从而从其它来源物诸如植物和微生物中分离相当、相关基因。也可以在 EST 数据库中进行搜索。而且, 本发明还包含在功能上、酶本身方面与本文所公开的相当的 DNA 序列编码酶的用途, 其中该 DNA 序列是本文中所公开的与遗传密码的简并性相一致的核酸序列的简并性相当物。使用任何这些方法所鉴定的编码序列, 对其功能的证实, 可以通过缺乏特异性生化反应或者已经发生突变的合适生物体 (例如, 集胞蓝细菌属 (*Synechocystis*)、*Schewanella*、酵母、假单胞菌属 (*Pseudomonas*)、红细菌科 (*Rhodobacteria*) 等) 的突变体的互补来实现。DNA 编码区的序列可以依据密码子的使用选择性, 通过基因合成予以最优化, 从而在特定的宿主中实现表达的最大化。

为了改变宿主细胞中生育酚的生成, 依据本发明可以使用第二种表达构建体。例如, 异戊烯基转移酶表达构建体可以与第二种具有编码参与生育酚生物合成的蛋白的核苷酸序列的表达构建体一起引入宿主细胞。

在获得这种转基因植物中, 转化方法对于本发明并不是非常重要, 并且当前有多种植物转化方法可供使用。另外, 如果有更新可用

于转化作物的方法，此后可以直接应用这些新方法。例如，天然对农杆菌属（*Agrobacterium*）感染敏感的许多植物种类，可以通过农杆菌属介导转化的三分或二分载体方法予以成功地转化。在许多情况下，希望使构建体的一侧或两侧都与 T-DNA 相连，优选使左侧和右侧与 T-DNA 相连，更优选使右侧相连。这在构建体使用根癌农杆菌（*A. tumefaciens*）或毛根农杆菌（*A. rhizogenes*）作为转化方式时显得尤其有用，尽管 T-DNA 边缘区在其它转化方式中也有用途。另外，已经开发了显微注射、DNA 颗粒轰击和电穿孔技术，这些技术可以转化多种单子叶和双子叶植物种类。

在正常情况下，在 DNA 构建体中将包含这种结构基因，它具有在宿主中表达所必需的调节区并提供对转化细胞的选择。该基因可以提供对一种细胞毒试剂例如抗生素、重金属、毒素等的抗性，通过互补对营养缺陷型宿主提供原养，提供病毒免疫力等等。根据不同宿主种类的数目，引入表达构建体或其组分，可以使用一种或多种标记，从而对不同的宿主种类使用不同的选择条件。

在使用农杆菌属转化植物细胞时，可以使用这样一种载体，其被引入农杆菌属宿主后，可与存在于农杆菌属宿主中的 T-DNA 或 Ti-或 Ri-质粒进行同源重组。含有 T-DNA（用于重组）的 Ti-或 Ri-质粒可以被接臂（能引起瘿的形成）的，也可以被断臂（不能引起瘿的形成）的，在后一种情况下，只要 *vir* 基因存在于转化的农杆菌属宿主中即可。接臂的质粒能够产生正常植物细胞与瘿的混合物。

在将农杆菌属用作转化宿主植物细胞的媒介物的情况下，将通过 T-DNA 边缘区相连的表达或转录构建体插入到一种宽宿主范围的载体中，这种载体具有在大肠杆菌以及农杆菌属中复制的能力，在文献中有对多种宽宿主范围载体的描述。通常使用的是 pRK2 或其衍生物。参见例如，Ditta 等（*Proc.Nat. Acad. Sci. U.S.A.* (1980) 77:7347-7351）以及 EPA 0120515，此处引入作为参考。或者，也可以将准备在植物细胞中表达的序列插入到含有独立的复制序列的载体中，其中的一个复制序列在大肠杆菌中稳定载体，另一个在农杆菌属稳定载体。见例如，MacBride 等 *Plant Mol. Bol.* (1990) 14: 269-276，其中使用了

pRiHRI 复制起点 (Jouanin 等 *Mol. Gen. Genet.* (1985) 201: 370-374), 并使植物表达载体在宿主农杆菌属细胞中的稳定性增加。

包括表达构建体和 T-DNA 的可以是一种或多种标记, 它们允许对转化的农杆菌属以及转化的植物细胞进行选择。已经开发了若干种用于植物细胞的标记, 诸如对氯霉素、卡那霉素、氨基糖苷 G418、潮霉素等的抗性, 本发明并不要求必须使用某种标记, 而是根据特定的宿主以及构建的方式优选一种或另一种标记。

在使用农杆菌属转化植物细胞时, 可以将外植体与转化的农杆菌属合并在一起并温育足够的时间以使转化发生, 细菌被杀死, 然后将植物细胞培养于合适的选择培性培养基中。一旦愈伤组织形成, 通过使用合适的植物激素 (采用公知的方法) 可以催促茎干形成, 然后将茎干转移到生根培养基中以再生植物。接着可以让植物生长结种, 然后将种子用于建立重复生产以及用于分离植物油。

有几个可能的方法, 用来获得含有多个表达构建体的本发明的植物细胞。本发明包含了任何用于生成植物的方法, 该植物具有一种构建体, 其具有编码本发明的表达构建体的 DNA 序列, 以及至少一种其它构建体, 其具有另一种编码酶的 DNA 序列。例如, 可以通过将两种表达构建体合在单一种转化载体中, 或使用独立的载体将表达构建体与第二种构建体同时用于转化植物, 其中的每一种载体均表达所希望的基因。还可以将第二种构建体引入到已经用异戊烯转移酶表达构建体转化的植物中, 或者也可以将一个表达异戊烯转移酶构建体、另一个是表达第二种构建体的被转化植物, 互相进行杂交, 从而使两种构建体合于同一植物中。

可以将本发明的核酸序列用于构建体中, 以提供在多种宿主细胞 (真核细胞和原核细胞) 中该序列的表达。本发明的宿主细胞优选包括单子叶和双子叶植物细胞。

一般地, 技术人员对那些描述大分子 (例如, DNA 分子、质粒等) 的构建、操作和分离, 产生重组生物体以及筛选和分离克隆 (参见例如, Sambrook 等 *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press (1989); Maliga 等 *Methods in Plant Molecular Biology*, Cold Spring Harbor Press (1985)), 其全部内容在此处引入

作为参考；Birren 等 *Genome Analysis: Analysing DNA*, 1, Cold Spring Harbor, New York, 其全部内容在此处引入作为参考) 的特定条件及方法的标准资料都很熟悉。

在昆虫宿主细胞中表达序列的方法在本领域中是公知的。杆状病毒表达载体是重组的昆虫病毒，其中已将所选择的外源基因的编码序列插入到杆状病毒启动子的后面，以取代病毒基因，例如，多角体蛋白基因 (Smith 和 Summers, 美国专利 4, 745, 051, 其全部内容在此处引入作为参考)。杆状病毒表达载体为本领域所公知，并在例如 Doerfler 的 *Curr. Top. Microbiol. Immunol.*, 131: 51-68 (1968)；Luckrow 和 Summers 的 *Bio/Technology*, 6: 47-55 (1988a)；Miller 的 *Annual Review of Microbiology*, 42: 177-199 (1988)；Summers 的 *Curr. Comm. Molecular Biology*, Cold Spring Harbor Press, Cold Spring Harbor, N. Y. (1988)；Summers 和 Smith 的 *A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures*, Texas Ag. Exper. Station Bullentin No. 1555 (1988) 中对其进行了描述，其全部内容在此处引入作为参考。

在真菌宿主细胞中表达目的核酸序列的方法为本领域所公知。真菌宿主细胞可以是例如，酵母细胞或丝状真菌细胞。在酵母细胞中表达目的 DNA 序列的方法一般描述于“*Guide to Yeast Genetics and Molecular Biolgy*”，Guthrie 和 Fink 编辑，*Methods in Enzymology*, Academic Press, Inc., 第 194 卷 (1991 年)；以及“*Gene Expression Technology*”，Goeddel 编辑，*Methods in Enzymology*, Academic Press, Inc., 第 185 卷 (1991 年)。

可用作表达宿主的哺乳动物细胞是本领域所公知的，其中包括许多无限增殖化细胞系，可从美国典型培养物保藏中心 (ATCC, Manassas, VA) 得到它们，诸如 Hela 细胞、中国仓鼠卵巢 (CHO) 细胞、幼小仓鼠肾 (BHK) 细胞和若干其它细胞系。用于哺乳动物细胞的合适的启动子也是本领域所公知的，其中包括但不限于病毒性启动子诸如来自猴病毒 40 (SV40) (Fiers 等 *Nature*, 273: 113 (1978), 此处引入其全部内容作为参考)、劳斯肉瘤病毒 (RSV)、腺病毒 (ADV) 和牛乳头瘤病毒 (BPV) 的启动子。哺乳动物细胞也

需要终止子序列和 poly-A 添加序列。也可以将提高表达的增强子序列包括在内，并且也希望含有可促进基因扩增的序列（例如，甲氨蝶呤抗性基因）。

适于在哺乳动物细胞中复制的载体是本领域所公知的，其中可包括病毒性复制子，或者能够确保将编码表位的适当序列整合到宿主基因组中的序列。已经对那些极大地促进重组病毒构建的质粒载体进行了描述（参见例如，Mackett 等 J. Virol, 49: 857 (1984); Chakrabarti 等 Mol. Cell. Biol. 5: 3403 (1985); Moss, In: Gene Transfer Vectors For Mammalian Cells, Miller 和 Calos 编辑, Cold Spring Harbor Laboratory, N.Y. 第 10 页 (1987); 此处引入上述全部文献的全部内容作为参考）

到目前为止，我们已对本发明进行了综述，通过参考下面的实施例，将更容易了解本发明，所给出的这些实施例，仅用于说明，而不是构成对本发明的限制。

实施例

实施例 1：异戊二烯转移酶序列的鉴定

PSI-BLAST (Altschul, et al.(1997) Nuc Acid Res 25; 3389-3402) 的序列文件 (profiles) 是针对直链和芳香类的异戊烯转移酶产生的。为了得到直链的序列文件，作为对 NCBI 中非冗余的蛋白质数据库的查询,使用来源于 *Porphyra purpurea*(基因库, 收录号 1709766)的异戊烯转移酶。大肠杆菌的参与泛醌形成的酶 *ubiA*(genbank 收录号 1790473)作为起始序列来产生芳香类异戊烯转移酶的序列文件。这些序列文件用来检索公共和专利的 DNA 和蛋白质数据库。在拟南芥属中，鉴定了 7 种直链异戊烯转移酶：ATPT1 (SEQ ID NO: 9)、ATPT7 (SEQ ID NO: 10)、ATPT8 (SEQ ID NO: 11)、ATPT9 (SEQ ID NO: 13)、ATPT10 (SEQ ID NO: 14)、ATPT11 (SEQ ID NO: 15) 和 ATPT12 (SEQ ID NO: 16)；鉴定了 5 种芳香异戊烯转移酶：ATPT2 (SEQ ID NO: 1)、ATPT3 (SEQ ID NO: 3)、ATPT4 (SEQ ID NO: 5)、ATPT5 (SEQ ID NO: 7)、ATPT6 (SEQ ID NO: 8)。还鉴定了来源于其它植物与芳香

类异戊烯转移酶相关的芳香类异戊烯转移酶序列，如大豆（SEQ ID NOs: 19—23，在 SEQ ID NO: 24 中给出了 SEQ ID NO: 23 的推定氨基酸序列）和玉米（SEQ ID NOs:25—29 和 31）。由 ZMPT 的推定氨基酸序列（SEQ ID NO: 29）在 SEQ ID NO: 30 中给出。

使用附加的 Bioaccelerator 硬件和由 Compugen Ltd. 提供的 GenWeb 软件在 Silicon Graphics Unix 计算机上进行检索。该软件和硬件在使用序列文件作为查询工具检索 DNA 和蛋白质数据库中需要使用 Smith-Waterman 算法。用来查询蛋白质数据库的程序是序列文件索引。在该检索中，不是查询单一序列，而是查询基于氨基酸或核苷酸序列的多种排列的序列文件。这个序列文件是用来查询序列数据组，即序列数据库。该序列文件含有序列上每一个位点打分的相关信息，它有效地替代了用于标准查询的“打分矩阵”。用于以蛋白质序列文件查询核酸数据库的程序是 T 序列文件索引。T 序列文件索引使用氨基酸序列文件查询来检索核酸数据库。当检索进行时，数据库中的序列以 6 个读框为单位翻译成氨基酸序列。T 序列文件索引的结果文件与序列文件索引的只在一个方面有区别，即 T 序列文件索引中有一附加行指出最佳排列的读框。

Smith-Waterman 算法（Smith 和 Waterman (1981) *supra*）是用来查找一段序列与数据库中一组序列相似性的。E 分值以及其它序列信息，如保守的多肽序列，用来鉴定相关序列。

为了得到对应拟南芥属的异戊烯转移酶的全部编码区，设计了合成寡核苷酸引物来扩增含有异戊烯转移酶序列的部分 cDNA 克隆的 5' 末端和 3' 末端。引物按照相关拟南芥属异戊烯转移酶序列设计并用于 cDNA 末端快速扩增反应（Rapid Amplification of cDNA Ends, RACE）（Frohman 等 (1988) *Proc. Natl. Acad. Sci. USA* 85:8998-9002），该反应使用 Marathon cDNA 扩增试剂盒（Clontech Laboratories Inc, Palo Alto, CA）。

使用 ATPT2 序列进行其它 BLAST 检索，该序列是一种芳香异戊烯转移酶。其它序列在与 ATPT2 序列相似的大豆文库中鉴定。另外的大豆序列证实在氨基酸顺序上有 80% 相同性和 91% 相似性。

ATPT2 (SEQ ID NO:2)、ATPT3 (SEQ ID NO: 4)、 ATPT4 (SEQ ID NO:6)、 ATPT8 (SEQ ID NO: 12)和 ATPT12 (SEQ ID NO:17)之间的氨基酸序列排列使用 ClusfalW（图 1）进行，其相同和相似百分率在表 1 中给出。

表 1

		ATPT 2	ATPT 3	ATPT 4	ATPT 8	ATPT 12
ATPT 2	% 相同性		12	13	11	15
	% 相似性		25	25	22	32
	% 空位		17	20	20	9
ATPT 3	%相同性			12	6	22
	% 相似性			29	16	38
	% 空位			20	24	14
ATPT 4	%相同性				9	14
	% 相似性				18	29
	% 空位				26	19
ATPT 8	%相同性					7
	% 相似性					19
	%空位					20
ATPT 12	%相同性					
	% 相似性					
	% 空位					

实施例 2：表达构建体的制备

对一个由 pCGN3223 (描述于 USPN 5, 639, 790，其全部内容引入本文作为参考) 衍生而来的含 napin 盒的质粒进行修饰，使其更有助于克隆含有多限制性位点的大 DNA 片断，并且使多 napin 融合基因克隆至植物双表达载体。在使用限制性内切酶 BssH II 对构建载体 pCGN77645 消化以后，将序列为 CGCGATTTAAATGGCGCGCCCTGCAGGCGGCCGCCTGCAGGGC GCGCCATTTAAAT (SEQ ID NO:40) 的自身连接寡核苷酸组成的接头与克隆载体 pBC Sk+(Stratagen)连接在一起。得到的载体 pCGN7770 含有 pCGN3323 的 napin 种子特异表达盒的 pGCN7765 主链。

克隆盒 pCGN7787 (除 pCGN7770 的 napin 调节区以外与 pCGN7770 基本相同的调节单元)被双链 CAMV 35S 启动子和 tml 聚腺苷酸和转录终止区所代替。

植物转染双载体 pCGN5139 由 pCGN1558 构建 (McBride 和 Summerfelt (1990) Plant Molecular Biology,14:269-276)。pCGN1558 的多连接子被一个非多连接子的 Hind III/Asp718 片段所代替,后者含有独特的限制性内切酶位点 Asc I, Pac I, Xba I, Swa I, BamH I 和 Not I。Asp718 和 Hind III 限制性内切酶位点存在于 pCGN5139 内。

构建一系列快速双载体以使得 DNA 序列快速克隆到含有转录起始位点(起始子)和转录终止区的双载体中。

通过连接寡核苷酸 5'-TCGAGGATCCGCGGCCGCAAGCTTCCTGCAGG-3' (SEQ ID NO:41) 和 5'-TCGACCTGCAGGAAGCTTGCGGCCGCGGATCC-3' (SEQ ID NO:42) 到 Sal I /Xho I 消化的 pCGN7770 构建质粒 pCGN8618。含有 napin 启动子、多连接子(polylinker)和 napin 3'区的片段通过用 Asp718 I 消化而从 pCGN8618 切下;该片段被 Klenow 片段将 5'-突出端添入而变为平端,然后与 pCGN5139 连接,后者已被 Asp718 和 Hind III 消化,并由 Klenow 片段将 5'-突出端添入而变为平端。将含有为了使 napin 启动子与 pCGN5139 的 Asp718 I 的平端接近并且使 napin 的 3'-末端与 Hind III 位点平端接近而定向插入物的质粒进行序列测定,来确定插入的方向和克隆连接的完整性。得到的质粒称 pCGN8622。

通过连接寡核苷酸寡聚核苷酸 5'-TCGACCTGCAGGAAGCTTGCGGCCGCGGATCC-3' (SEQ ID NO:43) 和 5'-TCGAGGATCCGCGGCCGCAAGCTTCCTGCAGG-3' (SEQ ID NO:44)到 Sal I /Xho I 消化的 pCGN7770 构建质粒 pCGN8619。将含有 napin 启动子、多连接子(polylinker)和 napin 3'区的片段通过用 Asp718 I 消化而从 pCGN8619 切下;该片段被 Klenow 片段将 5'-突出端添入而变为平端,然后与 pCGN5139 连接,后者已被 Asp718 和 Hind III 消化,并由 Klenow 片段将 5'-突出端添入而变为平端。将含有为了使 napin 启动子与 pCGN5139 的 Asp718 I 的平端接近并且使 napin 的 3'-末端与 Hind III 位点平端接近而定向插入物的质粒进行序

列测定，来确定插入的方向和克隆连接的完整性。得到的质粒称 pCGN8623。

通过连接寡核苷酸寡聚核苷 5'-TCGAGGATCCGCGGCCGCAAGCTTCCTGCAGGAGCT-3'(SEQ ID NO:45)和 5'-CCTGCAGGAAGCTTGCGGCCGCGGATCC-3'(SEQ ID NO:46)到 Sal I / Sac I 消化的 pCGN7787 构建质粒 pCGN8620。将含有 d35S 启动子、多连接子(polylinker)和 trm 3'区的片段通过用 Asp718 I 完全和 Not I 部分消化而从 pCGN8620 切下；该片段被 Klenow 片段将 5'-突出端添入而变为平端，然后与 pCGN5139 连接，后者已被 Asp718 和 Hind III 消化，并由 Klenow 片段将 5'-突出端添入而变为平端。将含有为了使 d35S 启动子与 pCGN5139 的 Asp718 I 的平端接近并且使 tml 的 3'-末端与 Hind III 位点平端接近而定向插入物的质粒进行序列测定，来确定插入的方向和克隆连接的完整性。得到的质粒称 pCGN8624。

通过连接寡核苷酸寡聚核苷 5'-TCGACCTGCAGGAAGCTTGCGGCCGCGGATCCAGCT-3'(SEQ ID NO:47)和 5'-GGATCCGCGGCCGCAAGCTTCCTGCAGG-3'(SEQ ID NO:48)到 Sal I / Sac I 消化的 pCGN7787 构建质粒 pCGN8621。将含有 d35S 启动子、多连接子(polylinker)和 trm 3'区的片段通过用 Asp718 I 完全和 Not I 部分消化而从 pCGN8621 切下；该片段被 Klenow 片段将 5'-突出端添入而变为平端，然后与 pCGN5139 连接，后者已被 Asp718 和 Hind III 消化，并由 Klenow 片段将 5'-突出端添入而变为平端。将含有为了使 d35S 启动子与 pCGN5139 的 Asp718 I 的平端接近并且使 tml 的 3'-末端与 Hind III 位点平端接近而定向插入物的质粒进行序列测定，来确定插入的方向和克隆连接的完整性。得到的质粒称 pCGN8625。

质粒构建体 pCGN8640 是上述 pCGN8624 的修饰。从编码细菌奇霉素和链霉素抗性 (Fling 等 (1985), Nucleic Acids Research 13(19):7095-7100) 的转位子分离的 938bp Pst I 的片段，是选择大肠杆菌和毛根农杆菌的决定因素，而且由 Pfu 聚合酶补为平端。该平端片段连接到被 Spe I 消化且被 Pfu 聚合酶补平的质粒 pCGN8624。对

含有 Pst I 片段的区域测序以确定插入物的方向和克隆连接的完整性。

按如下的方法将奇霉素抗性标记引入 pCGN8622 和 pCGN8623：
将从 pCGN8640 得到的 7.7kbp 的 Avr II -SnaB I 片段连接到从 pCGN8622 或 pCGN8623 得到的 10.9kbp 的 Avr II -SnaB I 片段。分别得到质粒 pCGN8641 和 pCGN8643。

通过连接寡核苷酸 5'-GATCACCTGCAGGAAGCTTGCGGCCGCGGATCCAATGCA-3' (SEQ ID NO:49) 和 5'-TTGGATCCGCGGCCGCAAGCTTCCTGCAGGT-3' (SEQ ID NO:50)到 Bam H I -Pst I 消化的 pCGN8641 构建质粒 pCGN8644。

设计将合成寡核苷酸用于聚合酶链反应 (PCR) 来扩增用于制备表达构建体的 ATPT2、ATPT3、ATPT4、ATPT8 和 ATPT12 的编码序列，见下面表 2。

表 2

名称	限制位点	序列	序列编号
ATPT2	5'Not I	GGATCCGCGGCCGCACAATGGAG TCTCTGCTCTCTAGTTCT	51
ATPT2	3'Sse I	GGATCCTGCAGGTCACCTCAAAA AAGGTAACAGCAAGT	52
ATPT3	5'Not I	GGATCCGCGGCCGCACAATGGCG TTTTTTGGGCTCTCCCGTGTTT	53
ATPT3	3'Sse I	GGATCCTGCAGGTTATTGAAAAC TCTTCCAAGTACAAC	54
ATPT4	5'Not I	GGATCCGCGGCCGCACAATGTGG CGAAGATCTGTTGTT	55
ATPT4	3'Sse I	GGATCCTGCAGGTCATGGAGAGT AGAAGGAAGGAGCT	56
ATPT8	5'Not I	GGATCCGCGGCCGCACAATGGTA CTTGCCGAGGTTCCAAAGCTTGCC TCT	57
ATPT8	3'Sse I	GGATCCTGCAGGTCACCTGTTTCT GGTGATGACTCTAT	58
ATPT12	5'Not I	GGATCCGCGGCCGCACAATGACT TCGATTCTCAACACT	59

ATPT12	3'Sse I	GGATCCTGCAGGTCAGTGTTCGA TGCTAATGCCGT	60
--------	---------	---	----

ATPT2、ATPT3、ATPT4、ATPT8 和 ATPT12 的编码序列均分别由上述表 2 中的 PCR 引物来扩增，并克隆到载体 TopoTA (Invitrogen) 中。含有相应异戊烯转移酶的构建体用 Not I 和 Sse8387 I 消化，并克隆到上述快速双载体中。

编码异戊烯转移酶 ATPT2 的序列按照有义方向克隆到 pCGN8640，以生成植物转化构建体 pCGN10800（图 2）。ATPT2 序列处于 35S 启动子的控制之下。

ATPT2 序列也按照反义方向克隆到构建体 pCGN8641 以生成 pCGN10801（图 3）。该构建体提供了 ATPT2 序列从 napin 启动子的反义表达。

ATPT2 编码序列也按照反义方向克隆到构建体 pCGN8643 以生成植物转化构建体 pCGN10802。

ATPT2 编码序列也按照反义方向克隆到载体 pCGN8644 以生成植物转化构建体 pCGN10803（图 4）。

ATPT4 编码序列克隆到载体 pCGN864，以生成植物转化构建体 pCGN10806（图 5）。ATPT2 序列克隆到载体 pCGN864 以生成植物转化构建体 pCGN10807（图 6）。ATPT3 编码序列克隆到载体 pCGN864 以生成植物转化构建体 pCGN10808（图 7）。ATPT3 编码序列按照有义方向克隆到载体 pCGN8640 以生成植物转化构建体 pCGN10809（图 8）。ATPT3 编码序列按照反义方向克隆到载体 pCGN8641 以生成植物转化构建体 pCGN10810（图 9）。ATPT3 编码序列克隆到载体 pCGN8643 以生成植物转化构建体 pCGN10811（图 10）。ATPT3 编码序列克隆到载体 pCGN8640 以生成植物转化构建体 pCGN10812（图 11）。ATPT4 编码序列克隆到载体 pCGN8640 以生成植物转化构建体 pCGN10813（图 12）。ATPT4 编码序列克隆到载体 pCGN8643 以生成植物转化构建体 pCGN10814（图 13）。ATPT4 编码序列克隆到载体 pCGN8641 以生成植物转化构建体 pCGN10815（图 14）。ATPT4 编码序列按照反义方向克隆到载体 pCGN8644 以

生成植物转化构建体 pCGN10816 (图 15)。ATPT2 编码序列克隆到载体 pCGN??? 以生成植物转化构建体 pCGN10817 (图 16)。ATPT8 编码序列按照有义方向克隆到载体 pCGN8643 以生成植物转化构建体 pCGN10819 (图 17)。ATPT12 编码序列克隆到载体 pCGN8644 以生成植物转化构建体 pCGN10824 (图 18)。ATPT12 编码序列克隆到载体 pCGN8641 以生成植物转化构建体 pCGN10825 (图 19)。ATPT8 编码序列克隆到载体 pCGN8644 以生成植物转化构建体 pCGN10826 (图 20)。

实施例 3 植物转化

转基因芸苔属植物是按照 Radke 等(Theor. Appl. Genet. (1988) 75: 685-694; Plant Cell Reports(1992) 11: 499-505)所述的方法通过毛根农杆菌介导的转化获得。转基因拟南芥植株可以按照 Valverbens 等 (Proc. Nat. Acad. Sci. (1988) 85: 5536-5540)、Bent 等 ((1994), Science 265: 1856-1860) 或 Bechtold 等 ((1993), C.R. Acad. Sci, Life Science 316: 1194-1199) 所述的方法通过毛根农杆菌介导的转化获得。其它种属的植物也可以利用相关技术进行类似的转化。

另外, 还可使用微粒轰击法 (如 Klein 等(Bio/Technology 10: 286-291)所述) 获得细胞核转化的植株。

实施例 4 异戊烯转移酶的鉴定

利用大肠杆菌 ubiA(Genebank accession 1790473)序列产生的 PSI-Blast 序列文件来分析集胞蓝细菌属基因组。该分析在集胞蓝细菌属基因组中分离了 5 个开放读框(ORFs), 它们都可能是异戊烯转移酶: slr0926 (注释为 ubiA (4-羟苯甲酸-八异戊烯转移酶), SEQ ID NO: 32)、slr1899 (注释为 ctaB (细胞色素 c 氧化酶折叠蛋白), SEQ ID NO: 33)、slr0056 (注释为 g4 (叶绿素合酶 33kd 亚单位), SEQ ID NO: 34)、slr1518(注释为 menA(甲基萘醌类生物合成蛋白), SEQ ID NO: 35)和 SLR1736 (注释为未知功能的推测蛋白, SEQ ID NO: 36)。

为了确定这些开放读框在生育酚生物合成中可能具有或其参与的作用，制备敲除构建体以阻断集胞蓝细菌属中破裂的 ORF。

设计寡核苷酸以扩增 slr1736ORF 的 5'末端区域和 3'末端区域，其中 5'末端为 5'-TAATGTGTACATTGTCGGCCTC (1736-5') (SEQ ID NO : 61) 和 5'-GCAATGTAACATCAGAGATTTTGAGACACAACGTGGCTTTCCACAATTCCCCGCACCGTC) (1736-kanpr1) (SEQ ID NO : 62) , 3' 端 为 5'-AGGCTAATAAGCACAAATGGGA(1736-3') (SEQ ID NO: 63)和 5'-GGTATGAGTCAGCAACACCTTCTTCACGAGGCAGACCTCAGCGGAATTGGTTTAGGTTATCCC(1736-kanpr2)(SEQ ID NO: 64)。1736-kanpr 1 和 1736-kanpr 2 寡核苷酸含有 20bp 与 slr1736 ORF 相同，另有 40bp 序列与卡那霉素抗性盒序列相同。完成这些寡核苷酸各自的 PCR 步骤，产物经凝胶纯化并与卡那霉素抗性基因连接，卡那霉素基因来源于 puc4K (Pharmacia)，其已被 Hinc II 消化，并从载体的主链中凝胶纯化出来。连接的片段在没有寡核苷酸时在下列条件下可以装配，该条件为在 100 μ l 反应体积中使用 pfu 聚合酶进行 40 个循环，每个循环为 94 $^{\circ}$ C 1 分钟、55 $^{\circ}$ C 1 分钟、72 $^{\circ}$ C 1 分钟零 5 秒 (Zhao, H 和 Arnold (1997) Nucleic Acid Res. 25(6): 1307-1308)。然后，在此装配反应体系中取 1 μ l 或者 5 μ l 进行扩增，扩增使用嵌套于 ORF 片段末端的 5'和 3'的寡核苷酸进行，这样得到的产物中含有 5'末端 100-200bp 碱基将被敲除的集胞蓝细菌属基因、卡那霉素抗性盒以及 3'末端 100-200 bp 碱基将被敲除的基因。该 PCR 产物随后被克隆到载体 pGem T easy (Promega)中，生成构建体 pMON21681，并用于集胞蓝细菌属的转化。

用下列引物，按照与上述同样的方法，还可合成其它序列用于制备集胞蓝细菌属敲除构建体的引物。使用引物 5'-GGATCCATGGTTGCCCAAACCCCATC(SEQ ID NO : 65) 和 5'-GCAATGTAACATCAGAGATTTTGAGACACAACGTGGCTTTGGGT AAGCA ACAATGACCGGC (SEQ ID NO: 66) 扩增 ubiA5'序列。使用合成寡核苷酸引物 5'-GAATTCTCAAAGCCAGCCCAGTAAC(SEQ ID NO : 67) 和 5'-

GGTATGAGTCAGCAACACCTTCTTCACGAGGCAGACCTCAGCGG
GTGCGAAAAGGGTTTTCCC(SEQ ID NO: 68)扩增 3'端区域。
扩增产物与卡那霉素抗性基因连接,卡那霉素基因来源于 puc4K,其
已被 Hinc II 消化,并从载体的主链中凝胶提纯出来。使用嵌套于
ORF 片段末端的 5'和 3'寡核苷酸(5'-
CCAGTGGTTTAGGCTGTGTGGTC(SEQ ID NO: 69)和 5'-CTGA
GTTGGATGTATTGGATC(SEQ ID NO: 70))扩增自身连接的片段,
这样得到的产物中含有 5'末端 100-200bp 碱基将被敲除的集胞蓝细菌
属基因、卡那霉素抗性盒以及 3'末端 100-200 bp 碱基将被敲除的基
因。该 PCR 产物随后被克隆到载体 pGem T easy (Promega)中,生成
构建体 pMON21682,并用于集胞蓝细菌属的转化。

用下列引物,按照与上述同样的方法,还可合成其它序列用于制
备集胞蓝细菌属敲除构建体的引物。使用引物 5'-
GGATCCATGGTTACTTCGACAAAAATCC(SEQ ID NO: 71)和 5'-
GCAATGTAAACATCAGAGATTTTGAGACACAACGTGGCTTTGCTA
GGCAACCGCTTAGTAC(SEQ ID NO: 72)扩增 sll 1899 5'序列。使
用合成寡核苷酸引物 5'-GA
ATTCTTAACCCAACAGTAAAGTTCCC(SEQ ID NO: 73)和 5'-
GGTATGAGTCAGCAACACCTTCTTCACGAGGCAGACCTCAGCGC
CGGCATTGTCTTTTACATG(SEQ ID NO: 74)扩增 3'端区域。
扩增产物与卡那霉素抗性基因连接,卡那霉素基因来源于 puc4K,其
已被 Hinc II 消化,并从载体的主链中凝胶提纯出来。使用嵌套于
ORF 片段末端的 5'和 3'寡核苷酸(5'-
GGAACCCTTGCAGCCGCTTC(SEQ ID NO: 75)和 5'-
GTATGCCCAACTGGTGCAGAGG(SEQ ID NO: 76))扩增自身连接的
片段,这样得到的产物中含有 5'末端 100-200bp 碱基将被敲除的集
胞蓝细菌属基因、卡那霉素抗性盒以及 3'末端 100-200 bp 碱基将被敲
除的基因。该 PCR 产物随后被克隆到载体 pGem T easy (Promega)
中,生成构建体 pMON21679,并用于集胞蓝细菌属的转化。

用下列引物,按照与上述同样的方法,还可合成其它序列用于制
备集胞蓝细菌属敲除构建体的引物。使用引物 5'-

GGATCCATGTCTGACACACAAAATACCG(SEQ ID NO: 77)和 5'-GCAATGTAACATCAGAGATTTTGAGACACAACGTGGCTTTCGCCAATACCAGCCACCAACAG (SEQ ID NO: 78) 扩增 slr 0056 5'序列。使用合成寡核苷酸引物 5'-GAATTCTCAAATCCCCGCATGGCCTAG (SEQ ID NO: 79)和 5'-GGTATGAGTCAGCAACACCTTCTTCACGAGGCAGACCTCAGCGGCCTACGGCTTGGACGTGTGGG (SEQ ID NO: 80) 扩增 3'端区域。扩增产物与卡那霉素抗性基因连接,卡那霉素基因来源于 puc4K,其已被 Hinc II 消化,并从载体的主链中凝胶提纯出来。使用嵌套于 ORF 片段末端的 5' 和 3' 寡核苷酸 (5'-CACTTGGATTCCCCTGATCTG(SEQ ID NO: 81) 和 5'-GCAATACCCGCTTGGAAAACG(SEQ ID NO: 82)) 扩增自身连接的片段,这样得到的产物中含有 5'末端 100-200bp 碱基将被敲除的集胞蓝细菌属基因、卡那霉素抗性盒以及 3'末端 100-200 bp 碱基将被敲除的基因。该 PCR 产物随后被克隆到载体 pGem T easy (Promega)中,生成构建体 pMON21677,并用于集胞蓝细菌属的转化。

用下列引物,按照与上述同样的方法,还可合成用于制备其它序列集胞蓝细菌属敲除构建体的引物。使用引物 5'-GGATCCATGACCGAATCTTCGCCCCTAGC(SEQ ID NO: 83)和 5'-GCAATGTAACATCAGAGATTTTGAGACACAACGTGGCTTTCATCC TAGGTAGCCGAGGCG(SEQ ID NO: 84) 扩增 slr 1518 5'序列。使用合成寡核苷酸引物 5'-GAATTCTTAGCCCAGGCCAGCCCAGCC(SEQ ID NO: 85)和 5'-GGTATGAGTCAGCAACACCTTCTTCACGAGGCAGACCTCAGCGGGGAATTGATTTGTTTAATTACC(SEQ ID NO: 86) 扩增 3'端区域。扩增产物与卡那霉素抗性基因连接,卡那霉素基因来源于 puc4K,其已被 Hinc II 消化,并从载体的主链中凝胶提纯出来。使用嵌套于 ORF 片段末端的 5' 和 3' 寡核苷酸 (5'-GCGATCGCCATTATCGCTTGG(SEQ ID NO: 87) 和 5'-GCAGACTGGCAATTATCAGTAACG(SEQ ID NO: 88)) 扩增自身连接的片段,这样得到的产物中含有 5'末端 100-200bp 碱基将被敲除的

集胞蓝细菌属基因、卡那霉素抗性盒以及 3'末端 100-200 bp 碱基将被敲除的基因。该 PCR 产物随后被克隆到载体 pGem T easy (Promega) 中，生成构建体 pMON21680，并用于集胞蓝细菌属的转化。

B. 集胞蓝细菌属的转化

将集胞蓝细菌属 6803 细胞培育至密度约为 2×10^8 细胞/ml 时离心收集，将细胞微粒重悬浮在新鲜的 BG-11 培养基(ATCC 培养基 616) 中，密度调整为 1×10^9 细胞/ml，马上用于转化。将 100μL 细胞用 5μL 微量制备的 DNA 混合。在光照条件下 30℃温育 4 小时。然后将混合物置于尼龙滤膜上，尼龙滤膜放置在 BG-11 琼脂上，琼脂中加有 pH 8 的 TES，使其培育 12-18 小时。随后将该滤膜转移至 BG-11+TES+5μg/ml 卡那霉素琼脂上培育直至 7-10 天内出现菌落(Packer 和 Glazer, 1988)。随后将菌落挑入含有 5μg/ml 卡那霉素的 BG-11 液体培养基中培育 5 天。然后将这些细胞转移至含有 10μg/ml 卡那霉素的 BG-11 液体培养基中培育 5 天。然后将这些细胞转移至含有 25 μg/ml 卡那霉素的 BG-11 液体培养基中培育 5 天。然后收集细胞，进行 PCR 分析以确定破裂的 ORF 的存在，同时进行 HPLC 分析以确定这种破裂的 ORF 对于生育酚水平是否有影响。

对为 slr1736 和 sll 1899 分离的集胞蓝细菌属中的 PCR 分析表明突变基因组中发生完全分离，表明在这些菌株中无法检测到任何野生型基因组的复制，这也表明天然基因的功能对细胞的功能并不重要。对这些分离株的 HPLC 分析表明，sll 1899 菌株生育酚水平没有可见的减少。但是，在承载 slr 1736 敲除的菌株中没有可检测水平的生育酚。

利用 Clustal W 对集胞蓝细菌属敲除株的氨基酸序列进行了比较，结果见表 3，表中列出了相同性百分比，类似性百分比和差异性百分比。序列的排列在图 21 中给出。

表 3:

		slr 1736	slr 0926	sll 1899	slr 0056	slr 1518
slr 1736	相同性%		14	12	18	11
	相似性%		29	30	34	26

slr 0926	差异性%	7	10	5
	相同性%	20	19	14
	相似性%	39	32	28
sll 1899	差异性%	7	9	4
	相同性%		17	13
	相似性%		29	29
slr 0056	差异性%		12	9
	相同性%			15
	相似性%			31
slr 1518	差异性%			8
	相同性%			
	相似性%			
	差异性%			

氨基酸序列的比较是在不同拟南芥异戊烯转移酶序列和集胞蓝细菌属序列间进行。比较结果见表 4。表中列出了相同性百分比，类似性百分比以及差异性百分比。序列对比结果见图 22。

表 4:

	ATPT2	slr 1736	ATPT3	slr 0926	ATPT4	sll 1899	ATPT12	slr 0056	ATPT8	slr 1518
ATPT2		29	9	9	8	8	12	9	7	9
slr1736		46	23	21	20	20	28	23	21	20
		27	13	28	23	29	11	24	25	24
			9	13	8	12	13	15	8	10
ATPT3			19	28	19	28	26	33	21	26
			34	12	34	15	26	10	12	10
				23	11	14	13	10	5	11
slr0926				36	26	26	26	21	14	22
				29	21	31	16	30	30	30
					12	20	17	20	11	14
ATPT4					24	37	28	33	24	29
					33	12	25	10	11	9
						18	11	8	6	7
sll1899						33	23	18	16	19
						28	19	32	32	33
							13	17	10	12
							24	30	23	26
							27	13	10	11

ATPT12	52	8	11
	66	19	26
	18	25	23
slr0056		9	13
		23	32
		10	8
ATPT8			7
			23
			7
slr1518			

4B. slr 1737 敲除株的制备

集胞蓝细菌属种 6803 slr 1737 敲除株按下列方法构建：利用 GPS™-1 Genome Priming System(New England Biolabs)，通过 Tn7 Transposase 系统，将卡那霉素抗性盒插入到 slr 1737 中。集胞蓝细菌属基因组文库克隆的含标靶的 orf 的 652 碱基对的质粒被用作靶 DNA(集胞蓝细菌属基因组碱基对 1324051-1324703；推定 orf 碱基对为 1323672-1324763，Cyanobase 注)。按照制造商的草案进行反应。然后将反应混合物转化到大肠杆菌 DH 10B electrocompetant 细胞中并摊铺。然后筛选转化形成的菌落，以使转位子通过用 M13 正向和反向通用引物扩增插入靶序列，结果产生了对于全长为约 2300 碱基对的片段 652 碱基对+约 1700 碱基对的产物，即转位子卡那霉素盒的大小。测定结束后需要确定标靶的 orf 内插入物大约的位点，因为据估计在集胞蓝细菌属中有效的同源重组需要 100 个碱基对的 orf 序列。这可以通过扩增反应进行，利用转位子末端的引物 Primer S(5'端)或 Primer N(3'端)，与 M13 的正向或反向引物结合使用。也就是说，将四种不同的引物结合用于描述可能的基因敲除株的图谱：引物 S-M13 正向、引物 S-M13 反向、引物 N-M13 正向、引物 N-M13 反向。用于转化集胞蓝细菌属和敲除株 slr 1737 的构建株被测定为由一个 150 个碱基对的 slr 1737 序列（位于插入转位子的 5'端）和约 500 bp 的序列（位于插入转位子的 3'端）组成，与 orf 和卡那霉素抗性盒的转录方

向一致。slr 1737 的核酸序列在 SEQ ID NO: 38 中给出, 推定的氨基酸序列在 SEQ ID NO: 39 中给出。

将集胞蓝细菌属 6803 细胞培育至密度约为 2×10^8 细胞/ml 时离心收集, 将细胞微粒重悬浮在新鲜的 BG-11 培养基中, 密度调整为 1×10^9 细胞/ml 并马上用于转化。将 100 μ L 细胞用 5 μ L 微量制备的 DNA 混合。在光照条件下 30 $^{\circ}$ C 温育 4 小时。然后将混合物置于尼龙滤膜上, 尼龙滤膜放置在 BG-11 琼脂上, 琼脂中添加有 pH 8 的 TES, 使其培育 12-18 小时。随后将该滤膜转移至 BG-11+TES+5 μ g/ml 卡那霉素琼脂上培育直至 7-10 天内出现菌落(Packer 和 Glazer, 1988)。随后将菌落挑入含有 5 μ g/ml 卡那霉素的 BG-11 液体培养基中培育 5 天。然后将这些细胞转移至含有 10 μ g/ml 卡那霉素的 BG-11 液体培养基中培育 5 天。然后将这些细胞转移至含有 25 μ g/ml 卡那霉素的 BG-11 液体培养基中培育 5 天。然后收集细胞, 进行 PCR 分析以确定破裂的 ORF 的存在, 同时进行 HPLC 分析以确定这种破裂的 ORF 对于生育酚水平是否有影响。

利用 slr 1737 orf 末端的引物对集胞蓝细菌属的分离株进行 PCR 分析, 结果显示突变基因组发生彻底分离, 表明在这些菌株中检测不到野生型基因组复制。这表明野生型基因对细胞的功能不重要。对基因敲除株 slr 1737 的 HPLC 分析表明检测不到生育酚产生。

4C. 植物异戊烯转移酶活性分析

[^{3} H]尿黑酸的 0.1% H_3PO_4 溶液(放射活性 40Ci/mmol)。按 Joo 等(1973) Can. J. Biochem. 51: 1527 所述方法合成叶绿基焦磷酸。按 Soll 等(1980) Phytochemistry 19: 215 的方法合成 2-甲基-6-叶绿基醌醇和 2, 3-二甲基-5-叶绿基醌醇。尿黑酸、 α , β , δ 和 γ -生育酚以及母育酚均为商购。

使用起泡空气在 30 $^{\circ}$ C、50 μ E $\cdot$ m $^{-2}$ $\cdot$ s $^{-1}$ 荧光和 70%相对湿度下, 在 BG-11 培养基上培育野生型集胞蓝细菌属种 Pcc6803。在该生物体的 slr 1736 敲除株(潜在的 PPT)的生长培养基上补充 25 μ g/ml 卡那霉素。收集 0.25 升到 1 升培养液中的细胞, 5000g 离心 10 分钟, 储存在-80 $^{\circ}$ C 下。

按照 Zak's 的操作方法 (Zak 等(1999) *Eur J. Biochem.* 261: 311), 稍作改动, 分离总膜。细胞在弗氏压碎器上破碎。在弗氏压碎处理前, 细胞用溶菌酶在 30℃下, 在含 7mM EDTA、5mM NaCl 和 10mM Hepes-NaOH 的 pH7.4 的培养基中温育 1 小时。5000g 离心 10 分钟, 收集去壁细菌细胞, 将其重悬浮在 pH 为 7.8 的 0.1-0.5mg 叶绿素 · mL⁻¹ 的 20mM 磷酸钾缓冲液中, 在溶液中加入适量的蛋白酶抑制剂混合物和 DNA 酶 I, DNA 酶 I 由 Boehringer Mannheim 获得。弗氏压碎处理在 100 MPa 下进行 2-3 次。破碎后, 细胞悬浮液在 5000g 离心 10 分钟以使未破碎细胞沉淀, 随后 100, 000g 离心 1 小时收集总膜。将最终所得的沉淀重悬浮在含 50mM Tris-HCl 和 4mM MgCl₂ 的缓冲液中。

叶绿体微粒是从 250g 市售的菠菜叶中分离的。将指定的叶片部分剪下, 放入研磨缓冲液中(2L/250g 叶片), 缓冲液含 2mM EDTA、1mM MgCl₂、1 mM MnCl₂, 0.33 M 山梨醇、0.1%抗坏血酸和 50 mM Hepes, pH 为 7.5。叶片在 1L 混合器中匀浆三次, 每次 3 秒, 然后用四层微孔布过滤。然后上清液在 5000g 离心 6 分钟。将叶绿体微粒重悬浮在少量研磨缓冲液中 (Douce 等 *Methods in Chloroplast Molecular Biology*, 239 (1982))。

微粒中的叶绿体可以用三种途径破碎。首先按每管 1mg 叶绿素将叶绿体分装于试管中, 在微量离心机中以 6000 rpm 的速度离心 2 分钟, 去除研磨缓冲液。每管中加入 200 μl Triton X-100 缓冲液(0.1% Triton X-100, 50 mM Tris-HCl, pH 7.6 和 4mM MgCl₂)或者是溶胀缓冲液(10 mM Tris pH 7.6, 4 mM MgCl₂), 并在 4℃下温育 0.5 小时, 随后破碎的叶绿体微粒马上用于分析。另外, 破碎的叶绿体微粒还可以通过在液 N₂ 中冷冻或者贮存在-80℃下 0.5 小时而获得, 随后进行分析。

在某些情形下, 叶绿体微粒可以进一步用 40%/80% percoll 梯度进行纯化, 以获得完整的叶绿体。完整的叶绿体用溶胀缓冲液破裂, 然后或者用于活性分析, 或者用于进一步用 20.5%/31.8%蔗糖密度梯度纯化被膜 (Sol 等, 1980, *supra*)。膜部分在 100, 000g 离心 40 分钟, 随后重悬浮在 50mM Tris-HCl pH 7.6 和 4mM MgCl₂ 中。

用适量的 1M Tris-NaOH pH10 混合不同数量的 [^3H] HGA 和 40-60 μM 未标记的 HGA (放射活性在 0.16-4Ci/mmol 范围内), 以将 pH 调至 7.6。用痕量的固态 NaBH_4 将 HGA 还原 4 分钟。除 HGA 外, 标准的温育混合物 (终体积 1ml) 含有 50 mM pH 为 7.6 的 Tris-HCl、3-5mM MgCl_2 和 100 μM 叶绿基焦磷酸。反应由加入集胞蓝细菌属总膜、菠菜叶绿体团微粒、菠菜破碎叶绿体或者菠菜被膜开始, 酶反应在 23 $^\circ\text{C}$ 或 30 $^\circ\text{C}$ 、黑暗或光照下进行 2 小时。反应通过加入液 N_2 冷冻来终止, 贮存在 -80 $^\circ\text{C}$ 或者直接进行萃取。

向每一分析混合液中加入恒定数量的母育酚, 用 2ml 氯仿/甲醇 (1: 2, v/v) 萃取反应产物, 以得到单相的溶液。其中加入 NaCl 溶液 (2ml; 0.9%), 剧烈振荡。重复萃取步骤 3 次。用 20 μm 过滤器过滤含有异戊二烯基醌的有机层, 在 N_2 环境下蒸发, 随后重悬浮在 100 μl 乙醇中。

样本主要通过正常相 HPLC 方法 (等度的 90% 己烷和 10% 甲基-叔丁基醚), 使用 Zorbax 硅胶柱 (4.6 $\times$ 250mm) 进行分析。样品也可以用反相 HPLC 方法 (等度的 0.1% H_3PO_4 的甲醇溶液), 使用 Vydac 201 HS54 C18 柱 (4.6 $\times$ 250mm) 并同时使用 All-tech C18 防护柱进行分析。产物的量基于底物的放射活性进行计算, 基于内标的量, 按照回收百分比进行调整。

叶绿素的量按照 Arnon (1949) Plant Physiol. 24: 1 描述的方法进行测定。蛋白质的含量用 Bradford 法, 用 γ 球蛋白作为标准物进行测定 (Bradford, (1976) Anal. Biochem. 72: 248)。

分析结果证明, 在制备集胞蓝细菌属 slr 1736 敲除株中生成了 2-甲基-6-叶绿基质体醌。slr 1736 敲除株的叶绿基异戊烯转移酶活性的分析结果在图 23 中给出。

4D. ATPT₂ 对 slr 1736 基因敲除株的功能补偿

为检测 ATPT₂ 能否补偿集胞蓝细菌属 6803 基因敲除株 slr 1736 的功能, 构建了表达 ATPT₂ 序列的质粒, 其启动子为 TAC。载体质粒 psl 1211 来自于 Washington 大学 Dr. Himadri Pakrasi 实验室, 此质粒是在质粒 RSF1010 基础上改进的, RSF1010 具有宽范围的宿主 (Ng

W.O., Zentella R., Wang. Y., Taylor J-S.A., Pakrasi, H.B. 2000. PhrA, 集胞蓝细菌属种 PCC 6803 植株中主要的光反应因子编码环丁烷-嘧啶二聚体特异性 DNA 光裂解酶(待发表)。通过 PCR 技术, 利用下述引物可以将 ATPT₂ 基因从载体 pCGN10817 中分离出来: ATPT2 nco.pr. 5'-CCATGGATTCGAGTAAAGTTGTCGC(SEQ ID NO : 89) ; ATPT2 ri.pr 5'-GAATTCACTTCAAAAAAGGTAACAG(SEQ ID NO: 90)。这些引物除去了 ATPT₂ 序列 5'端大约 112 bp, 这段 DNA 被认为是叶绿体的转运肽。这些引物还在 5'末端引入了 NcoI 位点, 3'末端引入了 EcoR I 位点, 以便能使其亚克隆至另一载体。将使用这些引物所得的 PCR 产物和 pCGN10817 连接到 pGEM T easy 中, 得到载体 pMON21689, 用 M13 正向和反向引物进行测序, 验证序列正确与否。然后将 pMON21689 的 NcoI/EcoRI 片段与 psl 1211 的 EagI/EcoRI 和 EagI/NcoI 片段连接。通过接合作用将质粒 pMON21690 引入 slr 1736 集胞蓝细菌属 6803 KO 植株。slr 1736 细胞(一种辅助菌株)和含 pMON21690 的 DH 10B 细胞培育至对数期(O.D.600=0.4), 通过离心收集 1ml 细胞。用无菌 BG-11 溶液将细胞微粒冲洗 2 次, 然后重悬浮在 200 μ l BG-11 中。在灭菌 Eppendorf 管中混合下列溶液: 50 μ l SL906、50 μ l 含 pMON21690 的 DH10B 和 100 μ L 新鲜的 slr 1736 集胞蓝细菌属 6803 KO 菌株(O.D.730=0.2-0.4) 培养物。马上将细胞混合物转移至置于 BG-11 上的硝酸纤维素膜, 在 30 $^{\circ}$ C、2500 LUX(50 μ e)光照下温育 24 小时。然后将滤液转移至添加 10 μ g/ml 庆大霉素的 BG-11 上, 如上述方法温育约 5 天。当菌落出现时, 挑取菌落, 培养在 BG-11+10 μ g/ml 庆大霉素液体培养基中(Elhai, J. 和 Wolk, P. 1988. Conjugal tranfer of DNA to Cyanobacteria. Methods in Enzymology 167: 747-54)。然后离心收集 1ml 液体培养物, 用乙醇/邻苯三酚萃取, HPLC 分离, 分析生育酚含量。Slr 1736 集胞蓝细菌属 6803 KO 菌株不含有可检测的生育酚, 而用 pMON 21690 转化的 Slr 1736 集胞蓝细菌属 6803 KO 菌株含有可检测的 α -生育酚。用 psl 1211(载体对照)转化的集胞蓝细菌属 6803 株也同样产生 α -生育酚。

实施例 5：转基因植物分析

用高压液相色谱（HPLC）分析经 ATPT 蛋白有义或反义表达构建体转化的拟南芥属植物，以检测总生育酚及特定的生育酚（ α -、 β -、 γ -及 δ -生育酚）水平的改变。

按照下述方法制备用于 HPLC 的叶子与种子提取物：对于种子提取物，在一消毒微量离心管中加入 10mg 种子、1g 微珠（Biospec）和 500 μ l 1% 邻苯三酚（Sigma Chem）的乙醇溶液。混合物于一小珠搅拌器（Beadbeater, Biospec）内快速震荡 3 分钟。然后将提取物经 0.2 μ m 滤膜过滤至自动样品收集管中。过滤的提取物按下述方法用于 HPLC 分析。

叶子萃取物是这样制备的：用在萃取前将 30-50mg 叶子组织与 1g 微珠混合后在液氮中冷冻。萃取时将 500 μ l 1% 邻苯三酚的乙醇溶液加入叶子/微珠混合物中，并在珠搅拌器（Biospec）上快速摇动 1 分钟。最终形成的混合物于 14000 rpm 离心 4 分钟，并在 HPLC 分析之前按上述方法过滤。

HPLC 的操作中使用的是 4.6 mm $\times$ 250mm 的 HPLC Zorbax 硅胶柱，其具有荧光检测器、带通滤波器和狭缝光栅，激发波长为 290nm，发射波长为 336nm。溶剂 A 是己烷，溶剂 B 是甲基叔丁基醚。进样体积为 20 μ l，流速为 1.5ml/min，采用梯度洗脱（表 5），洗脱时间是 12min（40 $^{\circ}$ C）。

表 5:

时间	溶剂 A	溶剂 B
0 分钟	90 %	10 %
10 分钟	90 %	10 %
11 分钟	25 %	75 %
12 分钟	90 %	10 %

溶于 1% 邻苯三酚/乙醇溶液的生育酚标准品也同时进样作为对照（包括 α -、 γ -、 β -、 δ -生育酚及生育酚（母育酚））（均来自于 Matreya）。

α -、 β -、 δ -、 γ -生育酚的标准曲线采用 Chemstation 软件计算。组分 X 的绝对含量如下计算：X 的绝对量=响应值 $\times$ RF_X $\times$ 稀释

因子，响应值 x 是组分 X 峰下面积， RF_x 是组分 X 的响应因子（含量 x /响应值 x ），稀释因子是 $500 \mu l$ 。而 ng/mg 组织是指总 ng 组分/ mg 植物组织。

含有 pMON10822 质粒从 napin 启动子开始表达 ATAT2 序列的转基因拟南芥属种子提取物的 HPLC 分析结果见图 24。

从 napin 启动子（pMON10822 质粒）开始表达 ATAT2 序列的拟南芥属种子组织的 HPLC 分析结果说明种子中生育酚水平增加了。其总生育酚水平比未转化（野生型）拟南芥属植物的总生育酚水平增加了 50-60%之多（图 24）。

另外，特定生育酚的水平在从 napin 启动子开始表达 ATAT2 核酸序列的转基因拟南芥属植物中也是增高的。这些植株中的 δ -生育酚水平比野生型拟南芥属植株种子中的 δ -生育酚水平增加 3 倍还多。在表达 ATAT2 核酸序列的转基因拟南芥属植株中 γ -生育酚水平比非转基因对照植株中的 γ -生育酚水平增加了近 60%之多。另外， α -生育酚水平比非转基因对照植株中的 α -生育酚水平增加了 3 倍之多。

含有 pMON10803 质粒从增强的 35S 启动子开始表达 ATAT2 序列的转基因拟南芥属植株种子提取物的 HPLC 分析结果见图 25。

本专利说明书中提到的所有出版物和专利申请都代表了与本发明有关的领域的技术人员的技术水平。所有出版物及专利申请在此以相同的程度被引用作为参考，每个出版物或专利申请都被具体地且独立地引用作为参考。

尽管为了能够更清晰地理解，已经通过图示说明和例证的方法对上述发明的细节加以描述，但显然在所附的权利要求范围内还可以做某些改变和修改。

序列表

<110> M·拉斯纳
M·波斯特-比特恩米勒
B·萨维季
J·韦斯

<120> 编码参与生育酚合成的蛋白质的核酸序列

<130> 17133/00/WO

<150> 60/129,899

<151> 1999-04-15

<150> 60/146,461

<151> 1999-07-30

<150> PCT/US00/10368

<151> 2000-04-14

<160> 94

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 1182

<212> DNA

<213> 拟南芥属种 (*Arabidopsis* sp)

<400> 1

```
atggagtctc tgctctctag ttcttctctt gtttccgctg ctggtgggtt ttgttggaag    60
aagcagaatc taaagctcca ctctttatca gaaatccgag ttctgcgttg tgattcgagt    120
aaagtgtctg caaaaccgaa gtttaggaac aatcttggtt ggctgatgg tcaaggatct    180
tcattgttgt tgatccaaa acataagtcg agatttcggg ttaatgccac tgcgggtcag    240
cctgaggctt tcgactcgaa tagcaaacag aagtctttta gagactcgtt agatgcgttt    300
tacaggtttt ctaggcctca tacagttatt ggcacagtgc ttagcatttt atctgtatct    360
ttcttagcag tagagaaggt ttctgatata tctcctttac ttttactgg catcttgag    420
gctgttggtg cagctctcat gatgaacatt tacatagttg ggctaaatca gttgtctgat    480
gttgaaatag ataaggttaa caagccctat ctccattgg catcaggaga atattctggt    540
aacaccggca ttgcaatagt agcttccttc tccatcatga gtttctggct tgggtggatt    600
gttggttcat ggccattggt ctgggctctt tttgtgagtt tcatgctcgg tactgcatac    660
tctatcaatt tgccactttt acggtggaag agatttgcag tggttgcagc aatgtgtatc    720
ctcgctgtcc gagctattat tgttcaaacc gccttttata tacatattca gacacatgtg    780
tttgaagac caatctgtt cactaggcct cttattttcg ccactgcgtt tatgagcttt    840
ttctctgtcg ttattgcatt gtttaaggat atacctgata tcgaagggga taagatatc    900
ggaatccgat cattctctgt aactctgggt cagaaacggg tgttttgac atgtgttaca    960
ctacttcaaa tggcttacgc tgttgcaatt ctagttggag ccacatctcc attcatatgg   1020
agcaaagtca tctcggttgt gggcatgtt atactcgaa caactttgtg ggctcgagct   1080
aagtcggtt atctgagtag caaaaccgaa ataacttcat gttatatgtt catatggaag   1140
ctcttttatg cagagtactt gctgttacct ttttgaagt ga                        1182
```

<210> 2

<211> 393

<212> PRT

<213> 拟南芥属种 (*Arabidopsis* sp)

<400> 2

```

Met Glu Ser Leu Leu Ser Ser Ser Ser Leu Val Ser Ala Ala Gly Gly
 1           5           10           15
Phe Cys Trp Lys Lys Gln Asn Leu Lys Leu His Ser Leu Ser Glu Ile
 20           25           30
Arg Val Leu Arg Cys Asp Ser Ser Lys Val Val Ala Lys Pro Lys Phe
 35           40           45
Arg Asn Asn Leu Val Arg Pro Asp Gly Gln Gly Ser Ser Leu Leu Leu
 50           55           60
Tyr Pro Lys His Lys Ser Arg Phe Arg Val Asn Ala Thr Ala Gly Gln
 65           70           75           80
Pro Glu Ala Phe Asp Ser Asn Ser Lys Gln Lys Ser Phe Arg Asp Ser
 85           90           95
Leu Asp Ala Phe Tyr Arg Phe Ser Arg Pro His Thr Val Ile Gly Thr
 100          105          110
Val Leu Ser Ile Leu Ser Val Ser Phe Leu Ala Val Glu Lys Val Ser
 115          120          125
Asp Ile Ser Pro Leu Leu Phe Thr Gly Ile Leu Glu Ala Val Val Ala
 130          135          140
Ala Leu Met Met Asn Ile Tyr Ile Val Gly Leu Asn Gln Leu Ser Asp
 145          150          155          160
Val Glu Ile Asp Lys Val Asn Lys Pro Tyr Leu Pro Leu Ala Ser Gly
 165          170          175
Glu Tyr Ser Val Asn Thr Gly Ile Ala Ile Val Ala Ser Phe Ser Ile
 180          185          190
Met Ser Phe Trp Leu Gly Trp Ile Val Gly Ser Trp Pro Leu Phe Trp
 195          200          205
Ala Leu Phe Val Ser Phe Met Leu Gly Thr Ala Tyr Ser Ile Asn Leu
 210          215          220
Pro Leu Leu Arg Trp Lys Arg Phe Ala Leu Val Ala Ala Met Cys Ile
 225          230          235          240
Leu Ala Val Arg Ala Ile Ile Val Gln Ile Ala Phe Tyr Leu His Ile
 245          250          255
Gln Thr His Val Phe Gly Arg Pro Ile Leu Phe Thr Arg Pro Leu Ile
 260          265          270
Phe Ala Thr Ala Phe Met Ser Phe Phe Ser Val Val Ile Ala Leu Phe
 275          280          285
Lys Asp Ile Pro Asp Ile Glu Gly Asp Lys Ile Phe Gly Ile Arg Ser
 290          295          300
Phe Ser Val Thr Leu Gly Gln Lys Arg Val Phe Trp Thr Cys Val Thr
 305          310          315          320
Leu Leu Gln Met Ala Tyr Ala Val Ala Ile Leu Val Gly Ala Thr Ser
 325          330          335
Pro Phe Ile Trp Ser Lys Val Ile Ser Val Val Gly His Val Ile Leu
 340          345          350
Ala Thr Thr Leu Trp Ala Arg Ala Lys Ser Val Asp Leu Ser Ser Lys
 355          360          365
Thr Glu Ile Thr Ser Cys Tyr Met Phe Ile Trp Lys Leu Phe Tyr Ala
 370          375          380
Glu Tyr Leu Leu Leu Pro Phe Leu Lys
 385          390

```

<210> 3

<211> 1224

<212> DNA

<213> 拟南芥属种 (*Arabidopsis* sp)

<400> 3

```

atggcggtttt ttgggtcttc ccgtgtttca agacggttgt tgaaatcttc cgtctccgta    60
actccatctt cttctctgc tcttttgcaa tcacaacata aatccttgtc caatcctgtg    120
actaccatt acacaaatcc tttactaag tggtatcctt catggaatga taattaccaa    180
gtatggagta aaggagaga attgcatcag gagaagtttt ttggtgttg ttgaattac    240
agattaatth gtggaatgtc gtcgtcttct tcggttttgg agggaaagcc gaagaaagat    300
gataaggaga agagtgatgg tggtgtgtt aagaaagctt cttggataga ttgtattta    360
ccagaagaag ttagaggtta tgctaagctt gctcgattgg ataaacccat tggaacttgg    420
ttgcttgctg gcccttgat gtggtcgatt gcgttggctg ctgacctgg aagcctcca    480
agttttaaat atatggtttt atttggttgc ggagcattac ttcttagagg tgctggtgt    540
actataaatg atctgcttga tcaggacata gatacaaagg ttgatcgtac aaaactaaga    600
cctatcgcca gtggtctttt gacaccattt caagggattg gatttctcgg gctgcagttg    660
cttttaggct tagggattct tctccaactt aacaattaca gccgtgtttt aggggcttca    720
tctttgttac ttgtcttttc ctaccactt atgaagaggt ttacattttg gcctcaagcc    780
tttttaggtt tgaccataaa ctggggagca ttgttaggat ggactgcagt taaaggaagc    840
atagcaccat ctattgtact ccctctctat ctctccggag tctgctggac cttgtttat    900
gatactatth atgcacatca ggacaaagaa gatgatgtaa aagttggtgt taagtcaaca    960
gcccttagat tcggtgataa tacaagctt tggtaactg gatttggcac agcatccata   1020
ggttttcttg cactttctgg attcagtca gatctcgggt ggcaatatta cgcactactg   1080
gccgtgcat caggacagtt aggatggcaa atagggacag ctgactatc atctggtgct   1140
gactgcagta gaaaatttgt gtcgaacaag tggtttggtg ctattatatt tagtgagtt   1200
gtacttgaa gaagttttca ataa                                     1224

```

<210> 4

<211> 407

<212> PRT

<213> 拟南芥属种 (*Arabidopsis* sp)

<400> 4

```

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

```

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

<210> 5

<211> 1296

<212> DNA

<213> 拟南芥属种 (*Arabidopsis* sp)

<400> 5

atgtggcgaa gatctgttgt ttctcgttta tcttcaagaa tctctgtttc ttcttcgtta 60
 ccaaacccta gactgattcc ttggtcccg gcgaattatgt ccgttaatag cttctcccag 120
 cctccggtct cgacggaatc aactgctaag ttagggatca ctggtgttag atctgatgcc 180
 aatcgagttt ttgccactgc tactgccgcc gctacagcta cagctaccac cggtgagatt 240
 tcgtctagag ttgcggcttt ggctggatta gggcatcact acgctcgttg ttattgggag 300
 ctttctaaag ctaaaacttag tatgcttggt gttgcaactt ctggaactgg gtatattctg 360
 ggtacgggaa atgctgcaat tagcttcccg gggcttttgt acacatgtgc aggaaccatg 420
 atgattgctg catctgctaa ttccttgaat cagatttttg agataagcaa tgattctaag 480
 atgaaaagaa cgatgctaag gccattgcct tcaggacgta ttagtgttcc acacgctgtt 540
 gcatgggcta ctattgctgg tgcttctggt gcttgtttgt tggccagcaa gactaatatg 600
 ttggctgctg gacttgcatc tgccaatctt gtactttatg cgtttgttta tactccgttg 660
 aagcaacttc accctatcaa tacatgggtt ggcgctgttg ttgggtgctat cccacccttg 720
 cttgggtggg cggcagcgtc tggtcagatt tcatacaatt cgatgattct tccagctgct 780
 ctttactttt ggcagatacc tcattttatg gcccttgcac atctctgccg caatgattat 840

```

gcagctggag gttacaagat gttgtcactc tttgatccgt caggaagag aatagcagca    900
gtggctctaa ggaactgctt ttacatgata cctctcggtt tcatcgcta tgactggggg    960
ttaacctcaa gttggttttg cctcgaatca acatttctca cactagcaat cgctgcaaca   1020
gcattttcat tctaccgaga cggaccatg cataaagcaa ggaaatggt ccatgccagt    1080
cttctcttcc ttctgtttt catgtctggt cttcttctac accgtgtctc taatgataat    1140
cagcaacaac tcgtagaaga agccggatta acaaatctg tatctggtga agtcaaaact    1200
cagaggcgaa agaaacgtgt ggctcaacct ccggtggctt atgcctctgc tgcaccgttt    1260
cttttctccc cagctccttc cttctactct ccatga                                1296

```

<210> 6

<211> 431

<212> PRT

<213> 拟南芥属种 (*Arabidopsis* sp)

<400> 6

```

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

```

```

305          310          315          320
Leu Thr Ser Ser Trp Phe Cys Leu Glu Ser Thr Leu Leu Thr Leu Ala
          325          330          335
Ile Ala Ala Thr Ala Phe Ser Phe Tyr Arg Asp Arg Thr Met His Lys
          340          345          350
Ala Arg Lys Met Phe His Ala Ser Leu Leu Phe Leu Pro Val Phe Met
          355          360          365
Ser Gly Leu Leu Leu His Arg Val Ser Asn Asp Asn Gln Gln Gln Leu
          370          375          380
Val Glu Glu Ala Gly Leu Thr Asn Ser Val Ser Gly Glu Val Lys Thr
385          390          395          400
Gln Arg Arg Lys Lys Arg Val Ala Gln Pro Pro Val Ala Tyr Ala Ser
          405          410          415
Ala Ala Pro Phe Pro Phe Leu Pro Ala Pro Ser Phe Tyr Ser Pro
          420          425          430

```

<210> 7

<211> 479

<212> DNA

<213> 拟南芥属种 (*Arabidopsis* sp)

<400> 7

```

ggaaactccc ggagcacctg tttgcaggta ccgctaacct taatcgataa tttatttctc      60
ttgtcaggaa ttatgtaagt ctggtggaag gctcgcatac catttttgca ttgcctttcg      120
ctatgatcgg gtttactttg ggtgtgatga gaccaggcgt ggctttatgg tatggcgaaa      180
acccatTTTT atccaatgct gcattccctc ccgatgattc gttctttcat tcctatacag      240
gtatcatgct gataaaactg ttactgggtac tggtttgat ggtatcagca agaagcgcgg      300
cgatggcggt taaccggtat ctgcacaggc attttgacgc gaagaaccgc cgtactgcca      360
tccgtgaaat acctgcgggc gtcatatctg ccaacagtgc gctggtgttt acgataggct      420
gctgcgtggt attctgggtg gcctgttatt tcattaacac gatctgtttt tacctggcg      479

```

<210> 8

<211> 551

<212> DNA

<213> 拟南芥属种 (*Arabidopsis* sp)

<220>

<221> misc_feature

<222> (1)... (551)

<223> n = A, T, C or G

<400> 8

```

ttgtggctta caccttaatg agcatacgcc agnccattac ggctcgtaa tcggcgccat      60
ngccgngct gntgcaccgg tagtgggcta ctgcgccgtg accaatcagc ttgatctagc      120
ggctcttatt ctgttttta ttttactgtt ctggcaaag cgcattttt acgcgatttc      180
cattttcagg ctaaaagact tttcagcggc ctgtattccg gtgctgcca tcattaaaga      240
cctgcgctat accaaaatca gcatgctggt ttactggtggc ttatttacac tggctgctat      300
catgccggcc ctcttagggt atgccggttg gatttatggg atagcggcct taattttagg      360
cttgatttgg ctttatattg ccatacaagg attcaagacc gccgatgatc aaaaatggtc      420
tcgtaagatg tttggatctt cgattttaat cattaccctc ttgtcggtaa tgatgcttgt      480
ttaaacttac tgccctctga agtttatata tcgataattt cagcttaagg aggcttagtg      540
gttaattcaa t                                     551

```

<210> 9

<211> 297

<212> PRT

<213> 拟南芥属种 (*Arabidopsis* sp)

<400> 9

```

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

```

<210> 10

<211> 561

<212> DNA

<213> 拟南芥属种 (*Arabidopsis* sp)

<400> 10

```

aagcgcatcc gtcctcttct acgattgccg ccagccgcat gtatggctgc ataaccgacc      60
gccctatcc gctcgggcc gcggtcgaat tcattcacac cgcgacgtg ctgcatgacg      120
acgtcgtcga tgaaagcgat ttgcgccg gccgcgaaag cgcgcataag gttttcgga      180
atcaggcgag cgtgctcgtc ggcgatttcc ttttctcccg cgccttcag ctgatggtgg      240

```

aagacggctc gctcgacgcg ctgcgcattc tctcggatgc ctccgccgtg atcgcgcagg 300
 gcgaagtgat gcagctcggc accgcgcgca atcttgaaac caatatgagc cagtatctcg 360
 atgtgatcag cgcaagacc gccgcgctct ttgccgccgc ctgcgaaatc ggcccgggtga 420
 tggcgaacgc gaaggcggaa gatgctgccg cgatgtgcga atacggcatg aatctcggta 480
 tcgccttcca gatcatcgac gaccttctcg attacggcac cggcggccac gccgagcttg 540
 gcaagaacac gggcgacgat t 561

<210> 11

<211> 966

<212> DNA

<213> 拟南芥属种 (*Arabidopsis* sp)

<400> 11

atgtacttg ccgaggttcc aaagcttgcc tctgctgctg agtacttctt caaaaggggt 60
 gtgcaaggaa aacagtttcg ttcaactatt ttgctgctga tggcgacagc tctgaatgta 120
 cgcgttccag aagcattgat tggggaatca acagatatag tcacatcaga attacgcgta 180
 aggcaacggg gtattgctga aatcactgaa atgatacacg tcgcaagtct actgcacgat 240
 gatgtcttgg atgatgccga tacaaggcgt ggtgttggtt ccttaaatgt tgtaatgggt 300
 aacaagatgt cgtatttagc aggagacttc ttgctctccc gggcttgtgg ggctctcgct 360
 gctttaaaga acacagaggt ttagcatta cttgcaactg ctgtagaaca tcttgttacc 420
 ggtgaaacca tggaaataac tagttcaacc gagcagcgtt atagtatgga ctactacatg 480
 cagaagacat attataagac agcatcgcta atctctaaca gctgcaaagc tgttgccgtt 540
 ctactggac aaacagcaga agttgccgtg ttagcttttg agtatgggag gaatctgggt 600
 ttagcattcc aattaataga cgacattctt gatttcacgg gcacatctgc ctctctcgga 660
 aagggatcgt tgcagatat tcgccatgga gtcataacag cccaatcct ctttgccatg 720
 gaagagtttc ctcaactacg cgaagttgtt gatcaagttg aaaaagatcc taggaatgtt 780
 gacattgctt tagagtatct tgggaagagc aagggaatac agagggcaag agaattagcc 840
 atggaacatg cgaatctagc agcagctgca atcggtctc tacctgaaac agacaatgaa 900
 gatgtcaaaa gatcgaggcg ggcacttatt gacttgaccc atagagtcac caccagaaac 960
 aagtga 966

<210> 12

<211> 321

<212> PRT

<213> 拟南芥属种 (*Arabidopsis* sp)

<400> 12

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

130 135 140
 Glu Ile Thr Ser Ser Thr Glu Gln Arg Tyr Ser Met Asp Tyr Tyr Met
 145 150 155 160
 Gln Lys Thr Tyr Tyr Lys Thr Ala Ser Leu Ile Ser Asn Ser Cys Lys
 165 170 175
 Ala Val Ala Val Leu Thr Gly Gln Thr Ala Glu Val Ala Val Leu Ala
 180 185 190
 Phe Glu Tyr Gly Arg Asn Leu Gly Leu Ala Phe Gln Leu Ile Asp Asp
 195 200 205
 Ile Leu Asp Phe Thr Gly Thr Ser Ala Ser Leu Gly Lys Gly Ser Leu
 210 215 220
 Ser Asp Ile Arg His Gly Val Ile Thr Ala Pro Ile Leu Phe Ala Met
 225 230 235 240
 Glu Glu Phe Pro Gln Leu Arg Glu Val Val Asp Gln Val Glu Lys Asp
 245 250 255
 Pro Arg Asn Val Asp Ile Ala Leu Glu Tyr Leu Gly Lys Ser Lys Gly
 260 265 270
 Ile Gln Arg Ala Arg Glu Leu Ala Met Glu His Ala Asn Leu Ala Ala
 275 280 285
 Ala Ala Ile Gly Ser Leu Pro Glu Thr Asp Asn Glu Asp Val Lys Arg
 290 295 300
 Ser Arg Arg Ala Leu Ile Asp Leu Thr His Arg Val Ile Thr Arg Asn
 305 310 315 320
 Lys

<210> 13

<211> 621

<212> DNA

<213> 拟南芥属种 (*Arabidopsis* sp)

<400> 13

gctttctcct ttgctaattc ttgagctttc ttgatccac cgcgatttct aactatttca 60
 atcgcttctt caagcgatcc aggtcacaa aactcagact caatgatctc tcttagcett 120
 ggctcattct ctagecgcaa gatcactggc gccgttatgt tacctttggc taagtcatta 180
 gctgcaggct tacctaactg ctctgtggac tgagtgaagt ccagaatgtc atcaactact 240
 tgaaaagata aaccgagatt ctccccgaac tgatacattt gctctgagac cttgctttcg 300
 actttactga aaattgctgc tcctttggtg cttgcagcta ctaatgaagc tgtctttag 360
 taactcttta gcatgtagtc atcaagcttg acatcacaa cgaataaact cgatgcttgc 420
 tttatctcac cgcttgcaaa atctttgatc acctgcaaaa agataaatca agatcagac 480
 caaatgttct ttgtattgag tagcttcac taatctcaga aaggaatatt acctgactta 540
 tgagcttaat gacttcaagg ttttcgagat ttgtaagtac catgatgctt gagcaacatg 600
 aatccccag ctaatacagc t 621

<210> 14

<211> 741

<212> DNA

<213> 拟南芥属种 (*Arabidopsis* sp)

<400> 14

ggtgagtttt gttaatatgt atgagattca tctatttttg tcataaaatt gtttggtttg 60
 gtttaaactc tgtgtataat tgcaggaaag gaaacagttc atgagctttt cggcacaaga 120
 gtagcgggtc tagctggaga ttcatgttt gctcaagcgt catggtactt agcaaactc 180
 gagaatcttg aagttattaa gtcacatcag caggctacta gttactctta cattgttttt 240

```

ctatgaggtt gagctatgaa tctcatttcg ttgaataatg ctgtgcctca aacttttttt 300
catgttttca ggtgatcaaa gacttttcaa gcggagagat aaagcaggcg tccagcttat 360
ttgactgcga caccaagctc gacgagtact tactcaaaag tttctacaag acagcctctt 420
tagtggctgc gaggacacaaa ggagctgcca ttttcagcag agttgagcct gatgtgacag 480
aacaaatgta cgagtttggg aagaatctcg gtctctcttt ccagatagtt gatgatattt 540
tggatttcac tcagtcgaca gacgagctcg ggaagccagc agggagtgat ttggctaaag 600
gtaacttaac agcacctgtg attttcgctc tggagaggga gccaaaggcta agagagatca 660
ttgagtcaaa gttctgtgag gcgggttctc tggagaagc gattgaagcg gtgacaaaag 720
gtggggggat taagagagca c 741

```

<210> 15

<211> 1087

<212> DNA

<213> 拟南芥属种 (*Arabidopsis* sp)

<400> 15

```

cctcttcagc caatccagag gaagaagaga caacttttta tctttcgtca agagtctccg 60
aaaacgcagc gttttatgct ctctcttctg ccctcacctc acaagacgca gggcacatga 120
ttcaaccaga gggaaaaagc aacgataaca actctgcttt tgatttcaag ctgtatatga 180
tccgcaaagc cgagtctgta aatgcggctc tcgacgtttc cgtaccgctt ctgaaacccc 240
ttacgatcca agaagcggtc aggtactctt tgctagccgg cggaaaacgt gtgaggcctc 300
tgctctgcat tgccgcttgt gagcttgtgg ggggcgacga ggctactgcc atgtcagccg 360
cttgccggt cgagatgata cacacaagct ctctcattca tgacgatctt ccgtgcatgg 420
acaatgccga cctccgtaga ggcaagccca ccaatcaca ggtagttgt ttaattatat 480
gaaggctcag agataatgct gaactagtgt tgaaccaatt tttgctcaa caaggtatat 540
ggagaagaca tggcggtttt ggcaggtgat gcactccttg cattggcgtt tgagcacatg 600
acggttgtgt cgagtgggtt ggtcgtccc gagaagatga ttcgcgccgt ggttgagctg 660
gccagggcca tagggactac agggctagtt gctggacaaa tgatagacct agccagcgaa 720
agactgaatc cagacaaggt tggattggag catctagagt tcatccatct ccacaaaacg 780
gcggcattgt tggaggcagc ggcagtttta ggggttataa tgggaggtgg aacagaggaa 840
gaaatcgaaa agcttagaaa gtatgctagg tgtattggac tactgtttca ggttgttgat 900
gacattctcg acgtaacaaa atctactgag gaattgggta agacagccgg aaaagacgta 960
atggccggaa agctgacgta tccaaggctg ataggtttg agggatccag ggaagttgca 1020
gagcacctga ggagagagc agaggaaaag cttaaagggt ttgatccaag tcaggcggcg 1080
cctctgg 1087

```

<210> 16

<211> 1164

<212> DNA

<213> 拟南芥属种 (*Arabidopsis* sp)

<400> 16

```

atgacttcga ttctcaacac tgtctccacc atccactctt ccagagttac ctccgtcgat 60
cgagtcggag tcctctctct tcggaattcg gattccgttg agttcactcg ccggcgttct 120
ggtttctcga cgttgatcta cgaatcacc ggcgggagat ttgttgtcg tgcggcggag 180
actgatactg ataaagttaa atctcagaca cctgacaagg caccagccgg tggttcaagc 240
attaaccagc ttctcggtat caaaggagca tctcaagaaa ctaataaatg gaagattcgt 300
cttcagctta caaaaccagt cacttggcct cacttggtt ggggagtcgt ctgtggtgct 360
gctgcttcag ggaacttca ttggaccca gaggatgttg ctaagtcgat tctttgcatg 420
atgatgtctg gtccttgtct tactggctat acacagacaa tcaacgactg gtatgataga 480
gatatcgacg caattaatga gccatatcgt ccaattccat ctggagcaat atcagagcca 540
gaggttatta cacaagctg ggtgctatta ttgggaggtc ttggtattgc tggaatatta 600
gatgtgtggg cagggcatac cactccact gtcttctatc ttgctttggg aggatcattg 660
ctatcttata tatactctgc tccactctt aagctaaaac aaaatggatg ggttggaat 720

```

```

tttgacttg gagcaageta tattagtttg ccatggtggg ctggccaagc attgtttggc 780
actcttacgc cagatgttgt tgttctaaca ctcttgtaca gcatagctgg gttaggaata 840
gccattgta acgacttcaa aagtgttgaa ggagatagag cattaggact tcagtctctc 900
ccagtagctt ttggcaccga aactgcaaaa tggatatgcg ttggtgctat agacattact 960
cagctttctg ttgccggata tctattagca tctgggaaac cttattatgc gttggcggtg 1020
gttgctttga tcattcctca gattgtgttc cagtttaaact actttctcaa ggaccctgtc 1080
aaatacgacg tcaagtacca ggcaagcgcg cagccattct tgggtgctcg aatatttgta 1140
acggcattag catcgcaaca ctga 1164

```

<210> 17

<211> 387

<212> PRT

<213> 拟南芥属种 (*Arabidopsis* sp)

<400> 17

```

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

```



```

attggctttc caagatcatt gggttttctt gttgcattca tgaccttcta ctctttgggt    60
ttggcattgt ccaaggatat acctgacgtt gaaggagata aagagcacgg cattgattct    120
tttcagtac gtctaggtca gaaacgggca ttttgatttt gcgtttcctt ttttgaaatg    180
gctttcggag ttggtatcct ggccggagca tcatgctcac acttttgac taaaattttc    240
acgggtatgg gaa                                         253

```

<210> 21

<211> 275

<212> DNA

<213> 甘氨酸种 (Glycine sp)

<400> 21

```

tgatcttcta ctctctgggt atggcattgt ccaaggatat atctgacgtt aaaggagata    60
aagcatacgg catcgatact ttagcgatac gtttgggtca aaaatgggta ttttgatttt    120
gcattatcct ttttgaaatg gcttttgag ttgccctctt ggcaggagca acatcttctt    180
acctttggat taaaattgtc acgggtcttg gacatgctat tcttgcttca attctcttgt    240
accaagccaa atctatatac ttgagcaaca aagtt                                         275

```

<210> 22

<211> 299

<212> DNA

<213> 甘氨酸种 (Glycine sp)

<220>

<221> misc_feature

<222> (1)... (299)

<223> n = A, T, C or G

<400> 22

```

ccanaatang tncatcttng aaagacaatt ggctcttcca acacacaagt ctgcatgtga    60
agaagaggcc aattgtcttt ccaagatcac ttatngtggc tattgtaatc atgaacttct    120
tctttgtggg tatggcattg gcaaaggata tacctanctg ttgaaggaga taaaatatat    180
ggcattgata cttttgcaat acgtataggt caaaaacaag tattttggat ttgtattttc    240
ctttttgaaa ggctttcgga gtttccttag tggcaggagc aacatcttct agccttgggt    299

```

<210> 23

<211> 767

<212> DNA

<213> 甘氨酸种 (Glycine sp)

<400> 23

```

gtggaggctg tggttgctgc cctgtttatg aatatattata ttgttggttt gaatcaattg    60
tctgatgttg aaatagacaa gataaacaag ccgtatcttc cattagcatc tggggaatat    120
tcctttgaaa ctggtgtcac tattgttgca tctttttcaa ttctgagttt ttggcttggc    180
tgggtttag gttcatggcc attattttgg gccctttttg taagctttgt gctaggaact    240
gcttattcaa tcaatgtgcc tctgttgaga ttgaagaggt ttgcagtgtc tgcagcgatg    300
tgcattctag ctgttcgggc agtaatagtt caacttgcatt ttttcttca catgcagact    360
catgtgtaca agaggccacc tgtcttttca agaccattga ttttgctac tgcattcatg    420
agcttcttct ctgtagtatt agcactgttt aaggatatac ctgacattga aggagataaa    480
gtattttgca tccaatcttt ttcagtgtgt ttaggtcaga agccggtgtt ctggacttgt    540
gttacccttc ttgaaatagc ttatggagtc gccctcctgg tgggagctgc atctccttgt    600
ctttggagca aaattttcac gggctctggga cacgctgtgc tggcttcaat tctctggttt    660
catgccaaat ctgtagattt gaaaagcaaa gcttcgataa catccttcta tatgtttatt    720
tggaagctat tttatgcaga atacttactc attccttttg ttagatg                                         767

```

tggtgtatga tactatatat gcgcatacagg tgtttcgcta tccctacttt catattaatc	240
cttgatgaag tggccatttc atgttgctgc ggtggctcta tacttgcata tctccatgca	300
tctcaggaca aagangatga cctgaaagta ggagtccaag tccacagctt aagatttggg	360

<210> 26

<211> 299

<212> DNA

<213> 玉蜀黍属种 (*Zea sp*)

<220>

<221> misc_feature

<222> (1)... (299)

<223> n = A, T, C or G

<400> 26

gatggttgca gcacttgcaa ataccctcaa ccagggtgtt gngataaaaa atgatgctaa	60
aatgaaaagg acaatgcgtg cccctgccca tctggtcgca ttagtcctgc acatgctgcg	120
atgtgggcta caagtgttgg agttgcagga acagctttgt tggcctggaa ggctaattggc	180
ttggcagctg ggcttgacgc ttctaacttt gttctgtatg catttgtgta tacgccgttg	240
aagcaataac accctgttaa tacatgggtt ggggcagtcg ttggtgcat cccaccact	299

<210> 27

<211> 255

<212> DNA

<213> 玉蜀黍属种 (*Zea sp*)

<220>

<221> misc_feature

<222> (1)... (255)

<223> n = A, T, C or G

<400> 27

anacttgcat atctccatgc ntctcaggac aaangangatg acctgaaagt aggtgtcaag	60
tccacagcat taagatttgg agatttgacc nnatactgna tcagtggctt tggcgcggca	120
tgcttcggca gcttagcact cagtgtgtac aatgctgacc ttggttggtg tttagtgtga	180
tgcttgagcg aagaatggta tngtttttac ttgatattga ctccagacct gaaatcatgt	240
tggacagggt ggccc	255

<210> 28

<211> 257

<212> DNA

<213> 玉蜀黍属种 (*Zea sp*)

<400> 28

attgaagggg ataggactct ggggcttcag tcacttcctg ttgcttttgg gatggaaact	60
gcaaaatgga tttgtgttgg agcaattgat atactcaat tatctgttgc aggttaccta	120
ttgagcaccg gtaagctgta ttatgccctg gtgttgcttg ggctaacaat tcctcaggtg	180
ttctttcagt tccagtactt cctgaaggac cctgtgaagt atgatgtcaa atatcaggca	240
agcgcacacac cattctt	257

<210> 29

<211> 368

<212> DNA

<213> 玉蜀黍属种 (*Zea sp*)

<210> 24
 <211> 255
 <212> PRT
 <213> 甘氨酸种 (Glycine sp)

<400> 24
 Val Glu Ala Val Val Ala Ala Leu Phe Met Asn Ile Tyr Ile Val Gly
 1 5 10 15
 Leu Asn Gln Leu Ser Asp Val Glu Ile Asp Lys Ile Asn Lys Pro Tyr
 20 25 30
 Leu Pro Leu Ala Ser Gly Glu Tyr Ser Phe Glu Thr Gly Val Thr Ile
 35 40 45
 Val Ala Ser Phe Ser Ile Leu Ser Phe Trp Leu Gly Trp Val Val Gly
 50 55 60
 Ser Trp Pro Leu Phe Trp Ala Leu Phe Val Ser Phe Val Leu Gly Thr
 65 70 75 80
 Ala Tyr Ser Ile Asn Val Pro Leu Leu Arg Trp Lys Arg Phe Ala Val
 85 90 95
 Leu Ala Ala Met Cys Ile Leu Ala Val Arg Ala Val Ile Val Gln Leu
 100 105 110
 Ala Phe Phe Leu His Met Gln Thr His Val Tyr Lys Arg Pro Pro Val
 115 120 125
 Phe Ser Arg Pro Leu Ile Phe Ala Thr Ala Phe Met Ser Phe Phe Ser
 130 135 140
 Val Val Ile Ala Leu Phe Lys Asp Ile Pro Asp Ile Glu Gly Asp Lys
 145 150 155 160
 Val Phe Gly Ile Gln Ser Phe Ser Val Cys Leu Gly Gln Lys Pro Val
 165 170 175
 Phe Trp Thr Cys Val Thr Leu Leu Glu Ile Ala Tyr Gly Val Ala Leu
 180 185 190
 Leu Val Gly Ala Ala Ser Pro Cys Leu Trp Ser Lys Ile Phe Thr Gly
 195 200 205
 Leu Gly His Ala Val Leu Ala Ser Ile Leu Trp Phe His Ala Lys Ser
 210 215 220
 Val Asp Leu Lys Ser Lys Ala Ser Ile Thr Ser Phe Tyr Met Phe Ile
 225 230 235 240
 Trp Lys Leu Phe Tyr Ala Glu Tyr Leu Leu Ile Pro Phe Val Arg
 245 250 255

<210> 25
 <211> 360
 <212> DNA
 <213> 玉蜀黍属种 (Zea sp)

<220>
 <221> misc_feature
 <222> (1)... (360)
 <223> n = A, T, C or G

<400> 25
 ggcgtcttca cttgttcttg tcttctcgta tcccctgatg aagaggttca cattttggcc 60
 tcaggcttat cttggcctga cattcaactg gggagcttta ctagggtggg ctgctattaa 120
 ggaaagcata gaccctgcaa atcatccttc cattgtatac agctggtatt tgttgacgc 180

<400> 29

```

atccagttgc aaataataat ggcgttcttc tctgttgtaa tagcactatt caaggatata      60
cctgacatcg aaggggaccg catattcggg atccgatacct tcagcgtccg gttagggcaa     120
aagaaggctct tttagatctg cgttggcttg cttagatgg cctacagcgt tgcgatactg     180
atgggagcta cctcttcctg ttgtggagc aaaacagcaa ccatcgctgg ccattccata     240
cttgccgcga tcctatggag ctgcgcgcga tcggtggact tgacgagcaa agccgcaata     300
acgtccttct acatgttcat ctggaagctg ttctacgcgg agtacctgct catccctctg     360
gtgcgggtg

```

<210> 30

<211> 122

<212> PRT

<213> 玉蜀黍属种 (*Zea* sp)

<400> 30

```

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

```

<210> 31

<211> 278

<212> DNA

<213> 玉蜀黍属种 (*Zea* sp)

<400> 31

```

tattcagcac cacctctcaa gctcaagcag aatggatgga ttgggaactt cgctctgggt      60
gcgagttaca tcagcttgcc ctggtgggct ggccaggcgt tatttgaac tcttacacca     120
gatatcattg tcttgactac ttgtacagc atagctgggc tagggattgc tattgtaaat     180
gatttcaaga gtattgaagg ggataggact ctggggcttc agtcacttcc tgttgctttt     240
gggatggaaa ctgcaaaatg gatttgtgtt ggagcaat

```

<210> 32

<211> 292

<212> PRT

<213> 集胞蓝细菌属种 (*Synechocystis* sp)

<400> 32

```

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

```

```

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

```

<210> 33

<211> 316

<212> PRT

<213> 集胞蓝细菌属种 (*Synechocystis* sp)

<400> 33

```

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

```

Gln Pro Arg His Ala Leu Ile Phe Ala Leu Ala Leu Gly Val Leu Ser
 100 105 110
 Phe Ala Leu Leu Ala Thr Phe Val Asn Val Leu Ser Gly Cys Leu Ala
 115 120 125
 Leu Ser Gly Ile Val Phe Tyr Met Leu Val Tyr Thr His Trp Leu Lys
 130 135 140
 Arg His Thr Ala Gln Asn Ile Val Ile Gly Gly Ala Ala Gly Ser Ile
 145 150 155 160
 Pro Pro Leu Val Gly Trp Ala Ala Val Thr Gly Asp Leu Ser Trp Thr
 165 170 175
 Pro Trp Val Leu Phe Ala Leu Ile Phe Leu Trp Thr Pro Pro His Phe
 180 185 190
 Trp Ala Leu Ala Leu Met Ile Lys Asp Asp Tyr Ala Gln Val Asn Val
 195 200 205
 Pro Met Leu Pro Val Ile Ala Gly Glu Glu Lys Thr Val Ser Gln Ile
 210 215 220
 Trp Tyr Tyr Ser Leu Leu Val Val Pro Phe Ser Leu Leu Leu Val Tyr
 225 230 235 240
 Pro Leu His Gln Leu Gly Ile Leu Tyr Leu Ala Ile Ala Ile Ile Leu
 245 250 255
 Gly Gly Gln Phe Leu Val Lys Ala Trp Gln Leu Lys Gln Ala Pro Gly
 260 265 270
 Asp Arg Asp Leu Ala Arg Gly Leu Phe Lys Phe Ser Ile Phe Tyr Leu
 275 280 285
 Met Leu Leu Cys Leu Ala Met Val Ile Asp Ser Leu Pro Val Thr His
 290 295 300
 Gln Leu Val Ala Gln Met Gly Thr Leu Leu Leu Gly
 305 310 315

<210> 34

<211> 324

<212> PRT

<213> 集胞蓝细菌属种 (*Synechocystis* sp)

<400> 34

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

```

145          150          155          160
Ala Pro Pro Leu Lys Leu Lys Gln Asn Gly Trp Leu Gly Asn Tyr Ala
          165          170          175
Leu Gly Ala Ser Tyr Ile Ala Leu Pro Trp Trp Ala Gly His Ala Leu
          180          185          190
Phe Gly Thr Leu Asn Pro Thr Ile Met Val Leu Thr Leu Ile Tyr Ser
          195          200          205
Leu Ala Gly Leu Gly Ile Ala Val Val Asn Asp Phe Lys Ser Val Glu
          210          215          220
Gly Asp Arg Gln Leu Gly Leu Lys Ser Leu Pro Val Met Phe Gly Ile
225          230          235          240
Gly Thr Ala Ala Trp Ile Cys Val Ile Met Ile Asp Val Phe Gln Ala
          245          250          255
Gly Ile Ala Gly Tyr Leu Ile Tyr Val His Gln Gln Leu Tyr Ala Thr
          260          265          270
Ile Val Leu Leu Leu Leu Ile Pro Gln Ile Thr Phe Gln Asp Met Tyr
          275          280          285
Phe Leu Arg Asn Pro Leu Glu Asn Asp Val Lys Tyr Gln Ala Ser Ala
          290          295          300
Gln Pro Phe Leu Val Phe Gly Met Leu Ala Thr Gly Leu Ala Leu Gly
305          310          315          320
His Ala Gly Ile

```

<210> 35

<211> 307

<212> PRT

<213> 集胞蓝细菌属种 (*Synechocystis* sp)

<400> 35

```

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

```

Ala Ile Ile Leu Phe Cys Ser His Phe His Gln Val Glu Asp Asp Leu
 195 200 205
 Ala Ala Gly Lys Lys Ser Pro Ile Val Arg Leu Gly Thr Lys Leu Gly
 210 215 220
 Ser Gln Val Leu Thr Leu Ser Val Val Ser Leu Tyr Leu Ile Thr Ala
 225 230 235 240
 Ile Gly Val Leu Cys His Gln Ala Pro Trp Gln Thr Leu Leu Ile Ile
 245 250 255
 Ala Ser Leu Pro Trp Ala Val Gln Leu Ile Arg His Val Gly Gln Tyr
 260 265 270
 His Asp Gln Pro Glu Gln Val Ser Asn Cys Lys Phe Ile Ala Val Asn
 275 280 285
 Leu His Phe Phe Ser Gly Met Leu Met Ala Ala Gly Tyr Gly Trp Ala
 290 295 300
 Gly Leu Gly
 305

<210> 36

<211> 927

<212> DNA

<213> 集胞蓝细菌属种 (*Synechocystis* sp)

<400> 36

```

atggcaacta tccaagcttt ttggcgcttc tcccgcctcc ataccatcat tggtaacaact    60
ctgagcgtct gggctgtgta tctgttaact attctcgggg atggaaactc agttaactcc    120
cctgcttccc tggatttagt gttcggcgct tggctggcct gcctgttggg taatgtgtac    180
atttgcggcc tcaaccaatt gtgggatgtg gacattgacc gcatcaataa gccgaatttg    240
ccctagctca acggagattt ttctatcgcc caggcgccgtt ggattgtggg actttgtggc    300
gttgcttccct tggcgatcgc ctggggatta gggctatggc tggggctaac ggtgggcatt    360
agtttgatta ttggcacggc ctattcgggt cgcgcagtga ggtaaagcg cttttccctg    420
ctggcgcccc tgtgtattct gacggtgcgg ggaattgtgg ttaacttggg cttattttta    480
tttttttaga ttggtttagg ttatcccccc actttaataa ccccatctg ggttttgact    540
ttatttatct tagttttcac cgtggcgatc gccattttta aagatgtgcc agatatggaa    600
ggcgatcggc aatttaagat tcaaaacttta actttgcaaa tcggcaaaca aaacgttttt    660
cggggaacct taattttact cactggttgt tatttagcca tggcaatctg gggcttatgg    720
gcggctatgc ctttaatac tgctttcttg attgttccc atttgtgctt attagcetta    780
ctctggtggc ggagtcgaga tgtacactta gaaagcaaaa ccgaaattgc tagtttttat    840
cagtttattt ggaagctatt tttcttagag tacttgcgtg atcccttggc tctgtggtta    900
cctaattttt ctaatactat tttttag                                     927

```

<210> 37

<211> 308

<212> PRT

<213> 集胞蓝细菌属种 (*Synechocystis* sp)

<400> 37

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

Asn Gln Leu Trp Asp Val Asp Ile Asp Arg Ile Asn Lys Pro Asn Leu
 65 70 75 80
 Pro Leu Ala Asn Gly Asp Phe Ser Ile Ala Gln Gly Arg Trp Ile Val
 85 90 95
 Gly Leu Cys Gly Val Ala Ser Leu Ala Ile Ala Trp Gly Leu Gly Leu
 100 105 110
 Trp Leu Gly Leu Thr Val Gly Ile Ser Leu Ile Ile Gly Thr Ala Tyr
 115 120 125
 Ser Val Pro Pro Val Arg Leu Lys Arg Phe Ser Leu Leu Ala Ala Leu
 130 135 140
 Cys Ile Leu Thr Val Arg Gly Ile Val Val Asn Leu Gly Leu Phe Leu
 145 150 155 160
 Phe Phe Arg Ile Gly Leu Gly Tyr Pro Pro Thr Leu Ile Thr Pro Ile
 165 170 175
 Trp Val Leu Thr Leu Phe Ile Leu Val Phe Thr Val Ala Ile Ala Ile
 180 185 190
 Phe Lys Asp Val Pro Asp Met Glu Gly Asp Arg Gln Phe Lys Ile Gln
 195 200 205
 Thr Leu Thr Leu Gln Ile Gly Lys Gln Asn Val Phe Arg Gly Thr Leu
 210 215 220
 Ile Leu Leu Thr Gly Cys Tyr Leu Ala Met Ala Ile Trp Gly Leu Trp
 225 230 235 240
 Ala Ala Met Pro Leu Asn Thr Ala Phe Leu Ile Val Ser His Leu Cys
 245 250 255
 Leu Leu Ala Leu Leu Trp Trp Arg Ser Arg Asp Val His Leu Glu Ser
 260 265 270
 Lys Thr Glu Ile Ala Ser Phe Tyr Gln Phe Ile Trp Lys Leu Phe Phe
 275 280 285
 Leu Glu Tyr Leu Leu Tyr Pro Leu Ala Leu Trp Leu Pro Asn Phe Ser
 290 295 300
 Asn Thr Ile Phe
 305

<210> 38

<211> 1092

<212> DNA

<213> 集胞蓝细菌属种 (*Synechocystis* sp)

<400> 38

atgaaatttc cgccccacag tggttaccat tggcaaggtc aatcaccttt ctttgaaggt 60
 tggtacgtgc gcctgctttt gccccaatcc ggggaaagtt ttgcttttat gtactccatc 120
 gaaaatcctg ctagcgatca tcattacggc ggcgggtgctg tgcaaatttt agggccggct 180
 acgaaaaaac aagaaaatca ggaagaccaa cttgtttggc ggacatttcc ctcggtaaaa 240
 aaattttggg ccagtcctcg ccagtttgcc ctagggcatt ggggaaaatg tagggataac 300
 aggcaggcga aaccctact ctccgaagaa ttttttgcca cggtaagga aggttatcaa 360
 atccatcaa atcagcacca aggacaaatc attcatggcg atcgccattg tcgttggcag 420
 ttcaccgtag aaccggaagt aacttggggg agtcctaacc gatttcctcg ggctacagcg 480
 ggttggcttt cttttttacc cttgtttgat cccggttggc aaattctttt agcccaaggt 540
 agagcgcacg gctggctgaa atggcagagg gaacagtatg aatttgacca cgccctagtt 600
 tatgccgaaa aaaattgggg tctactcctt cctccccgtt ggttttggct ccaagcaaat 660
 tattttcttg accatccagg actgagcgtc actgccgtg gcggggaacg gattgttctt 720
 ggtcgcctcg aagaggtagc ttaattggc ttacatcacc aagtaattt ttacgaattt 780
 ggccccggcc atggcacagt cacttggcaa gtagctccct ggggccgttg gcaattaaaa 840
 gccagcaatg ataggtattg ggtcaagttg tccggaaaaa cagataaaaa aggcagttta 900

```

gtccacactc ccaccgccca gggcttacaa ctcaactgcc gagataccac taggggctat    960
ttgtatttgc aattgggatc tgtgggtcac ggctgatag tgcaagggga aacggacacc    1020
gcggggctag aagttggagg tgattgggtt ttaacagagg aaaatttgag caaaaaaaca    1080
gtgccattct ga                                     1092

```

<210> 39

<211> 363

<212> PRT

<213> 集胞蓝细菌属种 (*Synechocystis* sp)

<400> 39

```

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

```

	340	345	350
Glu Glu Asn Leu Ser Lys Lys Thr Val Pro Phe			
355	360		

<210> 40
 <211> 56
 <212> DNA
 <213> 人工序列

<220>
 <223> 人工序列的说明: 寡核苷酸

<400> 40
 cgcgatttaa atggcgcgcc ctgcaggcgg ccgcctgcag ggcgcgccat ttaaat 56

<210> 41
 <211> 32
 <212> DNA
 <213> 人工序列

<220>
 <223> 人工序列的说明: 寡核苷酸

<400> 41
 tcgaggatcc gcggccgcaa gcttcctgca gg 32

<210> 42
 <211> 32
 <212> DNA
 <213> 人工序列

<220>
 <223> 人工序列的说明: 寡核苷酸

<400> 42
 tcgacctgca ggaagcttgc ggccgcggat cc 32

<210> 43
 <211> 32
 <212> DNA
 <213> 人工序列

<220>
 <223> 人工序列的说明: 寡核苷酸

<400> 43
 tcgacctgca ggaagcttgc ggccgcggat cc 32

<210> 44
 <211> 32
 <212> DNA
 <213> 人工序列

<220>

<223> 人工序列的说明： 寡核苷酸

<400> 44

tcgaggatcc gcggccgcaa gcttcctgca gg 32

<210> 45

<211> 36

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明： 寡核苷酸

<400> 45

tcgaggatcc gcggccgcaa gcttcctgca ggagct 36

<210> 46

<211> 28

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明： 寡核苷酸

<400> 46

cctgcaggaa gcttgccggcc gcggatcc 28

<210> 47

<211> 36

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明： 寡核苷酸

<400> 47

tcgacctgca ggaagcttgc ggccgcggat ccagct 36

<210> 48

<211> 28

<212> DNA

<213>人工序列

<220>

<223> 人工序列的说明： 寡核苷酸

<400> 48

ggatccgcgg ccgcaagctt cctgcagg 28

<210> 49

<211> 39

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明: 寡核苷酸

<400> 49

gatcacctgc aggaagcttg cggccgcgga tccaatgca

39

<210> 50

<211> 31

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明: 寡核苷酸

<400> 50

ttggatccgc ggccgcaagc ttctgcagg t

31

<210> 51

<211> 41

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明: 寡核苷酸

<400> 51

ggatccgcgg ccgcacaatg gagtctctgc tctctagttc t

41

<210> 52

<211> 38

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明: 寡核苷酸

<400> 52

ggatcctgca ggtcacttca aaaaaggtaa cagcaagt

38

<210> 53

<211> 45

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明: 寡核苷酸

<400> 53

ggatccgcgg ccgcacaatg gcgttttttg ggctctcccg tgttt

45

<210> 54

<211> 40

<212> DNA

<213> 人工序列

<220>

<223>人工序列的说明： 寡核苷酸

<400> 54

ggatcctgca ggttattgaa aacttcttcc aagtacaact

40

<210> 55

<211> 38

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明： 寡核苷酸

<400> 55

ggatccgagg ccgcacaatg tggcgaagat ctgttggt

38

<210> 56

<211> 37

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明： 寡核苷酸

<400> 56

ggatcctgca ggtcatggag agtagaagga aggagct

37

<210> 57

<211> 50

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明： 寡核苷酸

<400> 57

ggatccgagg ccgcacaatg gtacttgccg aggttccaaa gcttgccctct

50

<210> 58

<211> 38

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明： 寡核苷酸

<400> 58

ggatcctgca ggtcacttgt ttctggtgat gactctat

38

<210> 59

<211> 38

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明： 寡核苷酸

<400> 59

ggatccgcgg ccgcacaatg acttcgattc tcaacact

38

<210> 60

<211> 36

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明： 寡核苷酸

<400> 60

ggatcctgca ggtcagtggt gcgatgctaa tgccgt

36

<210> 61

<211> 22

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明： 寡核苷酸

<400> 61

taatgtgtac attgtcggcc tc

22

<210> 62

<211> 60

<212> DNA

<213>人工序列

<220>

<223> 人工序列的说明： 寡核苷酸

<400> 62

gcaatgtaac atcagagatt ttgagacaca acgtggcttt ccacaattcc ccgcaccgtc

60

<210> 63

<211> 22

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明： 寡核苷酸

<400> 63

aggctaataa gcacaaatgg ga

22

<210> 64

<211> 63

<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的说明: 寡核苷酸	
<400> 64	
ggtatgagtc agcaacacct tcttcacgag gcagacctca gcggaattgg tttaggttat	60
ccc	63
<210> 65	
<211> 26	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的说明: 寡核苷酸	
<400> 65	
ggatccatgg ttgcccaaac cccatc	26
<210> 66	
<211> 61	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的说明: 寡核苷酸	
<400> 66	
gcaatgtaac atcagagatt ttgagacaca acgtggcttt gggtaagcaa caatgaccgg	60
c	61
<210> 67	
<211> 25	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的说明: 寡核苷酸	
<400> 67	
gaattctcaa agccagccca gtaac	25
<210> 68	
<211> 63	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的说明: 寡核苷酸	
<400> 68	
ggtatgagtc agcaacacct tcttcacgag gcagacctca gcgggtgcga aaagggtttt	60

ccc	63
<210> 69	
<211> 23	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的说明: 寡核苷酸	
<400> 69	
ccagtggttt aggctgtgtg gtc	23
<210> 70	
<211> 21	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的说明: 寡核苷酸	
<400> 70	
ctgagttgga tgtattggat c	21
<210> 71	
<211> 28	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的说明: 寡核苷酸	
<400> 71	
ggatccatgg ttacttcgac aaaaatcc	28
<210> 72	
<211> 60	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的说明: 寡核苷酸	
<400> 72	
gcaatgtaac atcagagatt ttgagacaca acgtggcttt gctaggcaac cgcttagtac	60
<210> 73	
<211> 28	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的说明: 寡核苷酸	

<400> 73	
gaattcttaa cccaacagta aagttccc	28
<210> 74	
<211> 63	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的说明: 寡核苷酸	
<400> 74	
ggtatgagtc agcaacacct tcttcacgag gcagacctca gcgccggcat tgtcttttac	60
atg	63
<210> 75	
<211> 20	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的说明: 寡核苷酸	
<400> 75	
ggaacccttg cagccgcttc	20
<210> 76	
<211> 22	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的说明: 寡核苷酸	
<400> 76	
gtatgcccaa ctggtgcaga gg	22
<210> 77	
<211> 28	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的说明: 寡核苷酸	
<400> 77	
ggatccatgt ctgacacaca aaataccg	28
<210> 78	
<211> 62	
<212> DNA	
<213> 人工序列	
<220>	

<223> 人工序列的说明： 寡核苷酸

<400> 78

gcaatgtaac atcagagatt ttgagacaca acgtggcttt cgccaatacc agccaccaac 60
ag 62

<210> 79

<211> 27

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明： 寡核苷酸

<400> 79

gaattctcaa atccccgcat ggcctag 27

<210> 80

<211> 65

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明： 寡核苷酸

<400> 80

ggtatgagtc agcaacacct tcttcacgag gcagacctca gcggcctacg gcttggacgt 60
gtggg 65

<210> 81

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明： 寡核苷酸

<400> 81

cacttggatt cccctgatct g 21

<210> 82

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明： 寡核苷酸

<400> 82

gcaataccgg cttggaaaac g 21

<210> 83

<211> 29

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明： 寡核苷酸

<400> 83

ggatccatga ccgaatcttc gccccctagc

29

<210> 84

<211> 61

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明： 寡核苷酸

<400> 84

gcaatgtaac atcagagatt ttgagacaca acgtggcttt caatcctagg tagccgaggc
g

60

61

<210> 85

<211> 27

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明： 寡核苷酸

<400> 85

gaattcttag ccagggccag ccagacc

27

<210> 86

<211> 66

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明： 寡核苷酸

<400> 86

ggtatgagtc agcaacacct tcttcacgag gcagacctca gcggggaatt gatttgttta
attacc

60

66

<210> 87

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的说明： 寡核苷酸

<400> 87

gcgatcgcca ttatcgcttg g

21

<210> 88
 <211> 24
 <212> DNA
 <213> 人工序列

<220>
 <223> 人工序列的说明: 寡核苷酸

<400> 88
 gcagactggc aattatcagt aacg 24

<210> 89
 <211> 25
 <212> DNA
 <213> 人工序列

<220>
 <223> 人工序列的说明: 寡核苷酸

<400> 89
 ccatggattc gagtaaagtt gtcgc 25

<210> 90
 <211> 25
 <212> DNA
 <213> 人工序列

<220>
 <223>人工序列的说明: 寡核苷酸

<400> 90
 gaattcactt caaaaaaggt aacag 25

<210> 91
 <211> 4550
 <212> DNA
 <213> 拟南芥属种 (*Arabidopsis* sp)

<400> 91
 attttacacc aatttgatca cttaactaaa ttaattaaat tagatgatta tcccaccata 60
 tttttgagca ttaaaccata aaaccatagt tataagtaac tgttttaatc gaatatgact 120
 cgattaagat taggaaaaat ttataaccgg taattaagaa aacattaacc gtagtaaccg 180
 taaatgccga ttcctccctt gtctaaaaga cagaaaacat atattttatt ttgcccata 240
 tgtttcactc tatttaattt caggcacaat acttttggtt ggtaacaaaa ctaaaaagga 300
 caacacgtga tacttttcct cgtccgtcag tcagattttt tttaactag aaacaagtgg 360
 caaatctaca ccacattttt tgcttaatct attaacttgt aagttttaaa ttcttaaaaa 420
 agtctaacta attcttctaa tataagtaca ttccctaaat ttcccacaaa gtcaaattaa 480
 taattttcaa aatctaactt aaatatctaa taattcaaaa tcattaaaaa gacacgcaac 540
 aatgacacca attaatcatc ctgacccac acaattctac agttctcatg cttaaccata 600
 ttttttgctc tctgttcctt caaaatcatt tctttctctt ctttgattcc caaagatcac 660
 ttctttgtct ttgatttttg attttttttc tctctggcgt gaaggaagaa gctttatttc 720
 atggagtctc tgctctctag ttcttctctt gtttcgctg gtaaatctcg tccttttctg 780
 gtttcagggt ttatttggtg ttttaggttc gtttttgtga ttcagaacca tacaanaagt 840
 ttgaactttt ctgaatataa aataaggaaa aagtttcgat ttttataatg aattgtttac 900

tagatcgaag	taggtgacaa	aggttattgt	gtggagaagc	ataatttctg	ggcttgactt	960
tgaattttgt	ttctcatgca	tgcaacttat	caatcagctg	gtgggttttg	ttggaagaag	1020
cagaatctaa	agctccactc	tttatcaggt	tcgttagggg	tttatgggtt	tttgaaatta	1080
aatactcaat	catcttagtc	tcattattct	attgggtgaa	tcacattttc	taatttgga	1140
tttatgagac	aatgtatgtt	ggacttagtt	gaagttcttc	tctttgggta	tagttgaagt	1200
gttactgatg	ttgttttagct	ctttacacca	atatatacac	ccaattttgc	agaaatccga	1260
gttctgcgtt	gtgattcgag	taaagttgtc	gcaaaaccga	agtttaggaa	caatcttgtt	1320
aggcctgatg	gtcaaggatc	ttcattgttg	ttgtatccaa	aacataagtc	gagatttcgg	1380
gttaatgcca	ctgcgggtca	gcctgaggct	ttcgactcga	atagcaaaca	gaagtccttt	1440
agagactcgt	tagatgcgtt	ttacaggttt	tctaggcctc	atacagttat	tggcacagtt	1500
aagtttctct	ttaaaaatgt	aactctttta	aaacgcaatc	tttcagggtt	ttcaaggaga	1560
taacattagc	tctgtgattg	gatttgcagg	tgcttagcat	tttatctgta	tctttcttag	1620
cagtagagaa	ggtttctgat	atatctcctt	tacttttcac	tggcatcttg	gaggtaatga	1680
atatataaca	cataatgacc	gatgaagaag	atacattttt	ttcgtctctc	tgtttaaaca	1740
attgggtttt	gttttcaggc	tgttgttgca	gctctcatga	tgaacattta	catagttagg	1800
ctaaatcagt	tgtctgatgt	tgaatatagat	aaggtaacat	gcaaattttc	ttcatatgag	1860
ttcgagagac	tgatgagatt	aatagcagct	agtgccctaga	tcattctctat	gtgggttttt	1920
gcaggttaac	aagccctatc	ttccattggc	atcaggagaa	tattctgtta	acaccggcat	1980
tgcaatagta	gcttccttct	ccatcatggg	atgggtccat	tttcacaaaa	ttcaacttt	2040
tagaattcta	taagttactg	aaatagtttg	ttataaatcg	ttatagagtt	tctggccttg	2100
gtggattgtt	ggttcatggc	cattgttctg	ggctcttttt	gtgagtttca	tgcctcggtac	2160
tgcatactct	atcaatgtaa	gtaagtttct	caatactaga	atttggtcga	aatcaaaatc	2220
tgcagtttct	agtttttaggt	taatgaggtt	ttaataactt	acttctacta	caaacagttg	2280
ccacttttac	gggtgaaaag	atttgcattg	gttgacgcaa	tgtgtatcct	cgtgtgccga	2340
gctattattg	ttcaaatcgc	cttttatcta	catattcagg	tactaaacca	ttttccttat	2400
gttttgtagt	tgttttcatc	aaaatcacct	ttatattact	aaagctgtga	aactttgttg	2460
cagacacatg	tgtttggaag	accaatcttg	ttacttaggc	ctcttatttt	cgccactgcg	2520
tttatgagct	ttttctctgt	cgttattgca	ttgtttaagg	taaacaaaga	tggaaaaaga	2580
ttaaatctat	gtatacttaa	agtaaagcat	tctactgtta	ttgatgagaa	gttttctttt	2640
ttgggtggat	gcaggatata	cctgatatcg	aaggggataa	gatattcgga	atccgatcat	2700
tctctgtaac	tctgggtcag	aaacgggtac	gatatctaaa	ctaaagaaat	tgttttgact	2760
caagtgttgg	attaagatta	cagaagaaag	aaaactgttt	ttgtttcttg	caaaattcag	2820
gtgttttgga	catgtgttac	actacttcaa	atggcttacg	ctgttgcaat	tctagtggga	2880
gccacatctc	cattcatatg	gagcaaagtc	atctcggtaa	caatctttct	ttacccatcg	2940
aaaactcgct	aattcatcgt	ttgagtggta	ctggtttcat	tttgttccgt	tctgttgatt	3000
ttttttcagg	ttgtgggtca	tgttatactc	gcaacaactt	tgtgggctcg	agctaagtcc	3060
gttgatctga	gtagcaaaac	cgaataaact	tcattgttata	tgttcatatg	gaaggttaga	3120
ttcgtttata	aatagagtct	ttactgcctt	tttatgcgct	ccaatttgga	attaaaatag	3180
cctttcagtt	tcacgaatc	accattatac	tgataaattc	tcatttctgc	atcagctctt	3240
ttatgcagag	tacttgctgt	tacctttttt	gaagtgactg	acattagaag	agaagaagat	3300
ggagataaaa	gaataagtca	tcactatgct	tctgttttta	ttacaagttc	atgaaattag	3360
gtagtgaact	agtgaattag	agttttattc	tgaacatggg	cagactgcaa	aaatatgtca	3420
aagatatgaa	tttctgttgg	gtaaagaagt	ctctgcttgg	gcaaaatctt	aaggttcggg	3480
gtgttgatat	aatgctaagc	gaagaaatcg	attctatgta	gaaatttccg	aaactatgtg	3540
taaacatgtc	agaacatctc	cattctatat	cttcttctgc	aagaaagctc	tgttttttatc	3600
acctaaactc	tttatctctg	tgtagttaag	atatgtatat	gtacgtgact	acattttttt	3660
gttgatgtaa	tttgcagaac	gtatggattt	ttgttagaaa	gcatgagttc	gaaagtatat	3720
gtttatatat	atggataatt	cagacctaac	gtcgaagctc	acaagcataa	attcactact	3780
atagtttgct	ctgtaataga	tagttccatt	gatgtcttga	aactgtacgt	aactgcctgg	3840
gcgtttttgtg	gttgatactg	actactgagt	gttctttgtg	agtgttgtaa	gtatacaaga	3900
agaagaatat	aggctcacgg	gaacgactgt	gggtgaagat	gaaatggaga	tcatcacgta	3960
gcggctttgc	caagaccgga	gtcacgatcg	agtctatgaa	gtctttacag	ctgctgatta	4020
tgattgacca	ttgcttagag	acgcattgga	atcttactag	ggacttgctt	gggagtttct	4080
tcaagtacgt	gtcagatcat	acgatgtagg	agatttcacg	gctttgatgt	gtttgtttgg	4140

agtcacaatg	cttaatgggc	ttattggccc	aataatagct	agctcttttg	ctttagccgt	4200
ttcgtttgtc	ccctgggtgt	gagtattatt	agggtatggt	gtgaccaaag	tcaccagacc	4260
tagagtgaat	ctagtagagt	cctagaccat	ggctccatggc	ttttatttgt	aatttgaaaa	4320
atgaacaatt	ctttttgtaa	ggaaaacttt	tatatagtag	acgtttacta	tatagaaact	4380
agttgaacta	acttcgtgca	attgcataat	aatgggtgga	aatagagggt	gcaaaactca	4440
ataaacattt	cgacgtacca	agagttcgaa	acaataagca	aaatagattt	ttttgcttca	4500
gactaatttg	tacaatgaat	ggttaataaa	ccattgaagc	ttttattaat		4550

<210> 92

<211> 4450

<212> DNA

<213> 拟南芥属种 (*Arabidopsis* sp)

<400> 92

tttaggttac	aaaatcaatg	atattgcgta	tgtcaactat	aaaagccaaa	agtaaagcct	60
cttgtttgac	cagaagggtca	tgatcattgt	atacatacag	ccaaactacc	tcctggaaga	120
aaagacatgg	atcccaaaca	acaacaatag	cttcttttac	aagaaccagt	agtaactagt	180
cactaatcta	aaagagttaa	gtttcagctt	ttctggcaat	ggctccttga	tcatttcaat	240
cctgaaggag	acccactttg	tagcaagacc	atgtcctctg	tttcacttac	agtgtgtctc	300
aaaagtctac	ttcaattctt	catatatagg	ttcctcacac	tacagcttca	tcctcattcg	360
ttgacagaga	gagagtcttt	attgaaaact	tcttccaagt	acaactccac	taaatataat	420
agcaccaaac	cacttgttcg	acacaaatct	gtacagatat	aaaaacacta	ttaggttttc	480
caaggcaaat	cacataattg	gattgtgaaa	gagtacaaaa	gataaaccca	aattttcata	540
ctttctactg	cagtcagcac	cagatgataa	gtcagctgtc	cctatttgcc	atcctaactg	600
tcctgatgca	gcggccagtg	atgcgtaata	ttgccaccct	taatcattag	agcgagaaac	660
aaaaagaatc	aaaagacagt	aaatggaatt	aggaatcaca	aatgagtcct	tgtaaagttt	720
attgagtacc	gagatctgca	ctgaatccag	aaagtgaag	aaaacctatg	gatgctgtgc	780
caaatccagt	taaccaaagc	tttgtattat	caccgaatct	aagggtgtgt	gacttaacac	840
caacttttac	atcatcttct	ttgtcctgga	gacacaatat	attagacatt	agtccatgga	900
aaaaaaatga	tttaacctag	aatatctcaa	aattacttgc	ataaaaactg	aacttgagct	960
gaaatttttg	gttcgtagct	tgtggcatat	actatttcat	tttcaatggg	ccacaaggt	1020
aactttcttt	tctcacttct	gttgcaaacg	ggaagacttt	tatggggcta	actcttcact	1080
taaagtatag	aaatcagatg	gaaaagggtg	gagatcaggg	taattttctt	ctttatgatt	1140
gacaaaagtc	gaacatcgaa	atggatgcat	ttgcatgaga	catgaaacaa	aagctgaaaa	1200
agaaatctgt	ggtgggtgaag	ctagaaaaag	aaaacaaagc	aagcaatatg	cacacattga	1260
gattaactac	tttgctactg	gtcataatca	aatagatttt	gaagctaaaa	aataaaaagt	1320
gaatatacct	gatgtgcata	aatagtatca	taaacaaggg	tccagcagac	tccggagaga	1380
tagagaggga	gtacaataga	tggtgctatg	cttcctttta	ctgcagtcca	tcctaacaat	1440
gtccccagct	ttatgtgtcaa	acctaataag	gcttgaggct	gcaattataa	aaacgaatca	1500
atcataagaa	aatcagaaaa	tatataatgt	ctaactttga	gaagccagaa	tagattttaa	1560
ttacccaaaa	tgtaaacctc	ttcataagtg	ggtaggaaaa	gacaagtaac	aaagatgaag	1620
ccctataaac	acggctgcag	aatatataca	ctgaaatgag	ctcaagtaga	aaagaatttg	1680
atcacaaaaac	taaagacaag	acctgagaac	atatcttcag	aatttggggc	aactacataa	1740
gggtgaacca	tatgtgtatg	tgaattttta	aacaacact	tgcaaatagc	cgactttagg	1800
gcaagtaaaa	aatccaaaca	aacctgtaat	tgtaagtgtg	gagaagaatc	cctaagccta	1860
aaagcaactg	cagcccagaga	aatccaatcc	cttgaaatgg	tgtcaaaaga	ccactggcga	1920
taggtcttag	ttttgtacga	tcaacctgga	tataaaagaa	atttgtaaga	caacataatc	1980
taaaacaaaa	caaccataca	aaatcttgag	ctttacatac	aagcaaccca	tctttgttta	2040
tggaagaatg	aatccagtta	catgaatgct	gtgtatctac	cctaactact	aaacacatat	2100
ttcaatcgaa	aaacatattc	caccttcacc	atatctaaca	cctgaagtct	ttcacttttt	2160
gaacgaagtc	atcagaacat	gcagataagc	tattacccaa	aacagagata	tgactggaaa	2220
tgttgtcgta	aattgatcca	acatagaaaa	atcaagacca	gttccagatg	tcaaagcaat	2280
aacactttcc	caccatggtt	acagaaacca	tagttacaca	aaacatgttt	cctaaccaca	2340
catactaaag	ggatatataa	atttgacatc	actttatcac	cataccataa	gatagcttaa	2400

```

aaacaaactg acctttgtat ctatgtcctg atcaagcaga tcatttatag tacaaccagc 2460
acctctaaga agtaatgctc cgcaacccaa taaagccata tatttaaaac ttggaaggct 2520
tccaggatca gcagccaacg caatcgacct atacaacaat gatggagatt cagagtatcg 2580
atctatttac atagctctgg aactagatcc atgacgaaac atggaacatc gttataatat 2640
ctaaagactt ccaaacagat tcttgagtaa gaaaccaggt ggaactatag tactgtaaca 2700
tatataaaat caaagaaaac tcaggtttat agcattatcc aatcctgatt tctgccaatc 2760
cttaaccact ctcccatgct atcaaaaacc tcagctcaag atcactactac ctaattgcct 2820
atgagctctt gggaagatca ttatggattt gataactgaa aaaagtaaca gagaaatagc 2880
agactgcaag aactactcca aacttctcca ctgatatgta tgtagtctaa caataataaa 2940
cagacataaa ttcttttatc aagcttcaag agcaagttag tcagaaaaca tcacagccaa 3000
accaaccagg aaaacacata actttatcac ataaaactaa atttaagtga atctgactta 3060
acataaacca tcctttggga cgaaaggaaa ctatataaac atgcagctctt tctttccctc 3120
agctattctt tcggatggat tataatgaat ctcaaaaagt aaatgtcttg attctcagct 3180
acattactca aaggcgaaga taaacttacc acatacaagg ccacgcaagc aaccaagttc 3240
caatgggttt atccaatcga gcaagcttag cataacctct aacttcttct ggtaaataca 3300
aatctatcca agaagcttcc ttaacaacaa caccatcact ctctcctta tcattcttct 3360
tcggctttcc ctccaaaacc gaagaagacg acgacattcc acaaatatct ctgtaattcc 3420
aaccaacacc aaaaaacttc tcttgatgca attctcttcc ttactccat acttggtaat 3480
tatcattcca tgaaggataa cacttagtga aaggatttgt gtaatgggta gtcacaggat 3540
tggaacagga tttatgttgt gattgcaaaa gagcagagga agaagatgga gttacggaga 3600
cggaagattt caacaaccgt ctggaacac gggagagccc aaaaaacgcc atctttgaga 3660
gaaattgttg cctggaagaa acaaagactt gagatttcaa acgtaagtga attcttacga 3720
acgaaagcta acttctcaag agaatcagat tagtgattcc tcaaaaacaa acaaaactat 3780
ctaatttcag ttctcagatga tgaagcctta agaacttaga acctccatgg cgtttctaat 3840
ctctcagaga taatcgaatt ccttaaacaa tcaaagctta gaaagagaag aacaacaaca 3900
acaacaaaaa aatcagatt aacaaccgac cagagagcaa cgacgacgcc ggcgagaaag 3960
agcacgtcgt ctccgagcaa gacttcttct ccagtaaccc ggatggatcg ttaatgggcc 4020
tgtagattat tatatttggg ccgaaacaat tgggtcagca aaaacttggg ggataatgaa 4080
gaaacacgta cagtatgcat ttaggtcca aattaattgg ccatataatt cgaatcagat 4140
aaactaatca accctacct tacttatttc tactgtttt tatttctacc ttagtagttg 4200
aagaaacact ttattttatc ttttcgggac ccaaatttga taggatcggg ccattactca 4260
tgagcgtcag acacatatta gccttatcag attagtgggg taaggttttt ttaattcggg 4320
aagaagcaac aatcaatgtc ggagaaatta aagaatctgc atgggcgtgg cgtgatgata 4380
tgtgcatatg gagtcagttg ccgatcatat ataactattt ataaactaca tataaagact 4440
actaatagat 4450

```

<210> 93

<211> 2850

<212> DNA

<213> 拟南芥属种 (*Arabidopsis* sp)

<400> 93

```

aattaaaatt tgagcggctt aaaccattag accgtttaga gatccctcca acccaaaata 60
gtcgattttc acgtcttgaa catatattgg gccttaattct gtgtggttag taaagacttt 120
tattggtcaa agaaaaacaa ccatggccca acatgttgat acttttattt aattatacaa 180
gtacccctga attctctgaa atatatttga ttgaccaga tattaatttt aattatcatt 240
tctgttaaaa gtgaaggagt caccgtgact cgtcgtaatc tgaaccaat ctgttcatat 300
gatgaagaag tttctctcgt tctctccaa cgcgtagaaa attctgacgg cttacgatg 360
tgcggaagat ctgttgttta tcgtttctct tcaagaatct ctgtttcttc ttcgttacca 420
aaccttagac tgattccttg gtcccgcgaa ttatgtgccg ttaatagctt ctccagcct 480
ccggtctcga cggaatcaac tgctaagtta gggatcactg gtgttagatc tgatgccaat 540
cgagtttttg cactgctac tgccgcgct acagctacag ctaccacgg tgagatttcg 600
tctagagttg cggttttggc tggattaggg catcactacg ctctgtgtta ttgggagctt 660
tctaaagcta aacttaggta tgtgtttact tttcttttct catgaaaaat ctgaaaattt 720

```

```

ccaattgttg gattcttaaa ttctcatttg ttttatgggt gtagtatgct tgtggttgca 780
acttctggaa ctgggtatat tctgggtacg gaaatgctg caattagctt cccggggctt 840
tgttacacat gtgcaggaaac catgatgatt gctgcatctg ctaattcctt gaatcaggtc 900
attgaaatgt tgagaagttc ataaatttcg aatccttggt gtgtttatgt agttgatctt 960
gcttgcttat gtttatgtag ttgaaaagtt taaaaatttc taatccttgg tagttgatct 1020
cgcttgcttg ttttttcatt ttctagattt ttgagataag caatgattct aagatgaaaa 1080
gaacgatgct aaggccattg ccttcaggac gtattagtggt tccacacgct gttgcatggg 1140
ctactattgc tgggtcttct ggtgcttggt tggtgccag caaggtgaat gtttgttttt 1200
ttatatgtga tttctttggt ttatgaatgg gtgattgaga gattatggat ctaaactttt 1260
gcttcacga caaggttatt gcagactaat atgttggtg ctggacttgc atctgccaat 1320
cttgactttt atgcgttggt ttatactccg ttgaagcaac ttcacctat caatacatgg 1380
gttggcgctg ttgttggtgc tatccaccc ttgcttggtt aaatttttgc tccttttctt 1440
ctttatttta gcagattctg tttgttgga tactgctttt aattcaaaat gtagtcattg 1500
ttcaccaatt ctatgcttat ctattttgtg tgtgtcagg tgggcggcag cgtctggtca 1560
gatttcatac aattcgatga ttcttcacg tgcctttac ttttggcaga tacctcattt 1620
tatggccctt gcacatctct gccgaatga ttatgcagct ggagggtga accatatggt 1680
gtcatatgag attagaatgt ctcttccat gtagtggtga tcttgaacta gttcaatttc 1740
gtggaatgat cagagtgtcc tagatagtg cacagcagtc gacattttag tggctagata 1800
atgagttctt tccgttagag ataaacattc gcgaacattg tttccagctt ccgcgacca 1860
acttctgatt ttgtttcttg gtaccttggt ttcagttaca agatgttgct actctttgat 1920
ccgtcaggga agagaatagc agcagtggtc ctaaggaaact gcttttcat gatccctctc 1980
ggtttcatcg cctatgactg tgagtcttgt agattcatct tttttttgta gtttattgac 2040
tgcattgctg tatctgattt ttgctgttcc ttccaatttt tgtgacaggg gggttaacct 2100
caagttggtt ttgcctcgaa tcaacacttc tcacactagc aatcgctgca acagcatttt 2160
cattctaccg agaccggacc atgcataaag caaggaaaat gttccatgcc agtcttctct 2220
tccttctgtt ttcatgtct ggtcttcttc tacaccgtgt ctctaattgat aatcagcaac 2280
aactcgtaga agaagccgga ttaacaaatt ctgtatctgg tgaagtcaaa actcagaggc 2340
gaaagaaacg tgtggctcaa cctccggttg cttatgcctc tgcctgaccg tttcctttcc 2400
tcccagctcc ttccttctac tctccatgat aacctttaag caagctattg aatttttgga 2460
aacagaaatt aaaaaaaaaa tctgaaaagt tcttaagttt aatctttggt taataatgaa 2520
gtggagaacg catacaagtt tatgtatttt ttctcatctc cacataattg tattttttct 2580
ctaagtatgt ttcaaatgat acaaaataca tactttatca attatctgat caaattgatg 2640
aatttttgag ctttgacgtg ttaggtctat ctaataaacg tagtaacgaa tttggttttg 2700
gaaatgaaat ccgataaccg atgatggtgt agagttaaac gattaaaccg ggttggttaa 2760
aggtctcgag tctcgacggc tgcggaaatc ggaaaatcac gattgaggac tttgagctgc 2820
cacgaagatg gcgatgaggt tgaaatcaat 2850

```

<210> 94

<211> 3660

<212> DNA

<213> 拟南芥属种 (*Arabidopsis* sp)

<400> 94

```

tattttgtatt tttattgtta aattttatga tttcaccggt tatatatcat cccatattaa 60
tatttagattt attttttggg ctttatttgg gttttcgatt taaactgggc ccattctgct 120
tcaatgaaac cctaattgggt tttgtttggg ctttgattt aaaccgggcc cattctgctt 180
caatgaaggt cctttgtcca acaaaactaa catccgacac aactagtatt gccaagagga 240
tcgtgccaca tggcagttat tgaatcaaag gccgccaaaa ctgtaacgta gacattactt 300
atctccgta acggacaacc actcgtttcc cgaacagca actcacagac tcacaccact 360
ccagtctccg gcttaactac caccagagac gattctctct tccgtcggtt ctatgacttc 420
gattctcaac actgtctcca ccatccactc ttccagagtt acctccgtcg atcgagtcgg 480
agtcctctct ctccggaatt cggattccgt tgagttcact cgcggcggtt ctggtttctc 540
gacgttgatc tacgaatcac ccggtagtta gcattctggt ggatagattg atgaatgttt 600
tcttcgattt tttttttact gatcttggtg tggatctctc gtagggcgga gatttgttgt 660

```

gcgtgcggcg	gagactgata	ctgataaagg	tatgattttt	tagttgtttt	tattttctct	720
ctcttcaaaa	tctctttttc	aaacactgtg	gcgtttgaat	ttccgacggc	agttaaactc	780
cagacacctg	acaaggcacc	agccgggtgt	tcaagcatta	accagcttct	cggtatcaaa	840
ggagcatctc	aagaaactgt	aattttgttc	atctcctcag	aatcttttaa	attatcatac	900
ttgtggataa	tgatgtgtta	gttttaggaat	tttctacta	aaggtaatct	cttttgagga	960
caagcttctg	ttttagctta	gaaatgatgt	gaaaatgttg	tttgtagct	aaaaagagtt	1020
tgttgttata	tctgtattc	agaataaatg	gaagattcgt	cttcagctta	caaaaccagt	1080
cacttggcct	ccactggttt	ggggagtcgt	ctgtgggtgt	gctgcttcag	gtaatcatac	1140
gaacctcttt	tggatcatgc	aatactgtac	agaaagtttt	ttcattttcc	ttccaattgt	1200
ttcttctggc	agggaaactt	catttgaccc	cagaggatgt	tgctaagtcg	attctttgca	1260
tgatgatgtc	tggtccttgt	cttactggct	atacacaggt	ctggttttac	acaacaaaaa	1320
gctgacttgt	tcttattcta	gtgcatttgc	ttggtgctac	aataacctag	acttgctgat	1380
ttccagacaa	tcaacgactg	gtatgataga	gatatcgacg	caattaatga	gccatctcgt	1440
ccaattccat	ctggagcaat	atcagagcca	gaggtaactg	agacagaaca	ttgtgagctt	1500
ttatctcttt	tgtgattctg	atttctcctt	actccttaaa	atgcagggtta	ttacacaagt	1560
ctgggtgcta	ttattgggag	gtcttggtat	tgctggaata	ttagatgtgt	gggtaagttg	1620
gcccttctga	cattaactag	tacagttaaa	gggcacatca	gatttgctaa	aatcttccct	1680
tatcaggcag	ggcataccac	tcccactgtc	ttctatcttg	ctttgggagg	atcattgcta	1740
tcttatatat	actctgctcc	acctcttaag	gtaagtttta	ttcctaactt	ccactctcta	1800
gtgataagac	actccatcca	agttttggag	ttttgaatat	cgatatctga	actgatctca	1860
ttgcagctaa	aacaaaatgg	atgggttggg	aattttgcac	ttggagcaag	ctatattagt	1920
ttgccatggt	aagatatctc	gtgtatcaat	aatatatggc	gttggtctca	tctcattgat	1980
ttgtttcttg	ctcacttgac	tgataggtgg	gctggccaag	cattgttttg	cactcttacg	2040
ccagatgttg	ttgttctaac	actcttgtag	agcatagctg	gggtactctt	ttggcaaacc	2100
ttttatgttg	cttttttctg	tatctgttgt	aatatgctct	tgcttcatgt	tgtacctttg	2160
tgataatgca	gttaggaata	gccattgtta	acgacttcaa	aagtgttgaa	ggagatagag	2220
cattaggact	tcagtctctc	ccagtagctt	ttggcaccga	aactgcaaaa	tggatatgcg	2280
ttggtgctat	agacattact	cagctttctg	ttgccggtat	gtactatcca	ctgtttttgt	2340
gcagctgtgg	cttctatttc	tttcccttga	tcttatcaac	tggaatttca	ccaatggtaa	2400
agcacaatt	aatgaagctg	aatcaacaaa	ggcaaaacat	aaaagtacat	tctaatagaaa	2460
tgagctaata	aagaggagcg	atctactttt	atgtttcatt	agtgtagattg	atggattttc	2520
atttcatgct	tctaaaacaa	gtattttcaa	cagtgtcatg	aaataacaga	acttatatct	2580
tcatattgtac	ttttactagt	ggatgagtta	cacaatcatt	gttatagaac	caaatcaaa	2640
gtagagatca	tcattagtat	atgtctattt	tggttgacag	atatctatta	gcacttgga	2700
aaccttatta	tgctgtggcg	ttggttgctt	tgatcattcc	tcagattgtg	ttccaggtaa	2760
agacgttaac	agtctcacat	tataattaat	caaattcttg	tactcgtct	gattgctaca	2820
ctcgttcta	taaaactgcag	tttaataact	ttctcaagga	ccctgtcaaa	tacgacgtca	2880
agtaccaggt	aagtcaactt	agtacacatg	tttgtgttct	tttgaaatat	ctttgagagg	2940
tctcttaate	agaagttgct	tgaaacactc	atcttgatta	caggcaagcg	cgcagccatt	3000
cttgggtgctc	ggaatatttg	taacggcatt	agcatcgcaa	cactgaaaaa	ggcgtatttt	3060
gatgggggtt	tgtcgaaaagc	agaggtgttg	acacatcaaa	tgtgggcaag	tgatggcatc	3120
aactagttta	aaagattttg	taaaatgtat	gtaccgttat	tactagaaac	aactcctgtt	3180
gtatcaattt	agcaaaacgg	ctgagaaatt	gtaattgatg	ttaccgtatt	tgcgtccat	3240
ttttgcattt	cctgctcata	tcgaggattg	gggtttatgt	tagttctgtc	acttctctgc	3300
tttcagaatg	tttttgtttt	ctgtagtgga	ttttaactat	tttcatcact	ttttgtattg	3360
attctaaaca	tgtatccaca	taaaaacagt	aataataaaa	aatgataact	cctcaaaact	3420
tttataatct	aaatctaaca	actagctagt	aacccaacta	acttcataca	attaatttga	3480
gaaactacaa	agactagact	atacatatgt	tatttaacaa	cttgaaactg	tgttattact	3540
acctgatttt	tttctattct	acagccattt	gatatgctgc	aatcttaaca	tatcaagtct	3600
cacgttggtg	gacacaacat	actatcacia	gtaagacacg	aagtaaaacc	aaccggcaac	3660

[illegible]

一
五

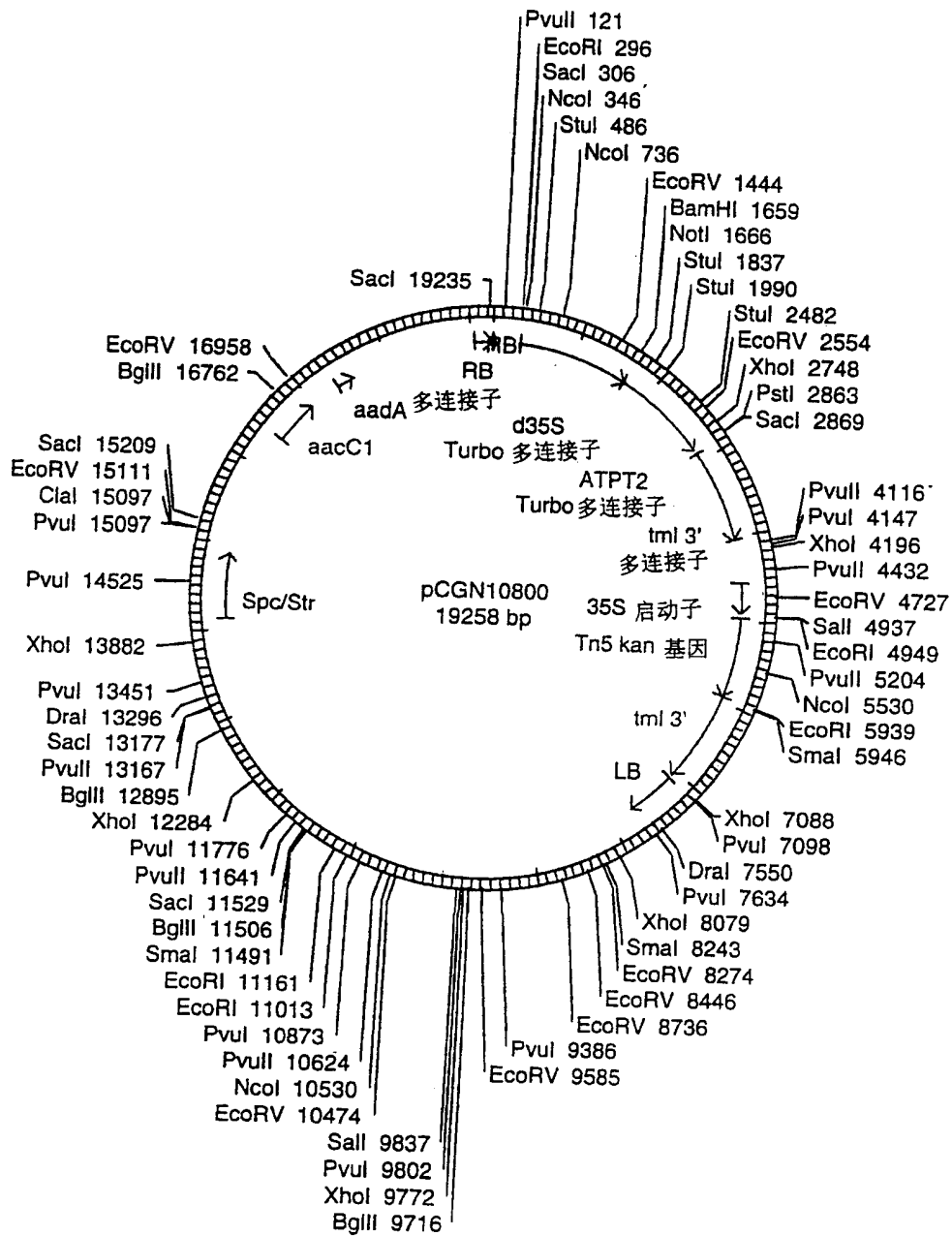


图 2

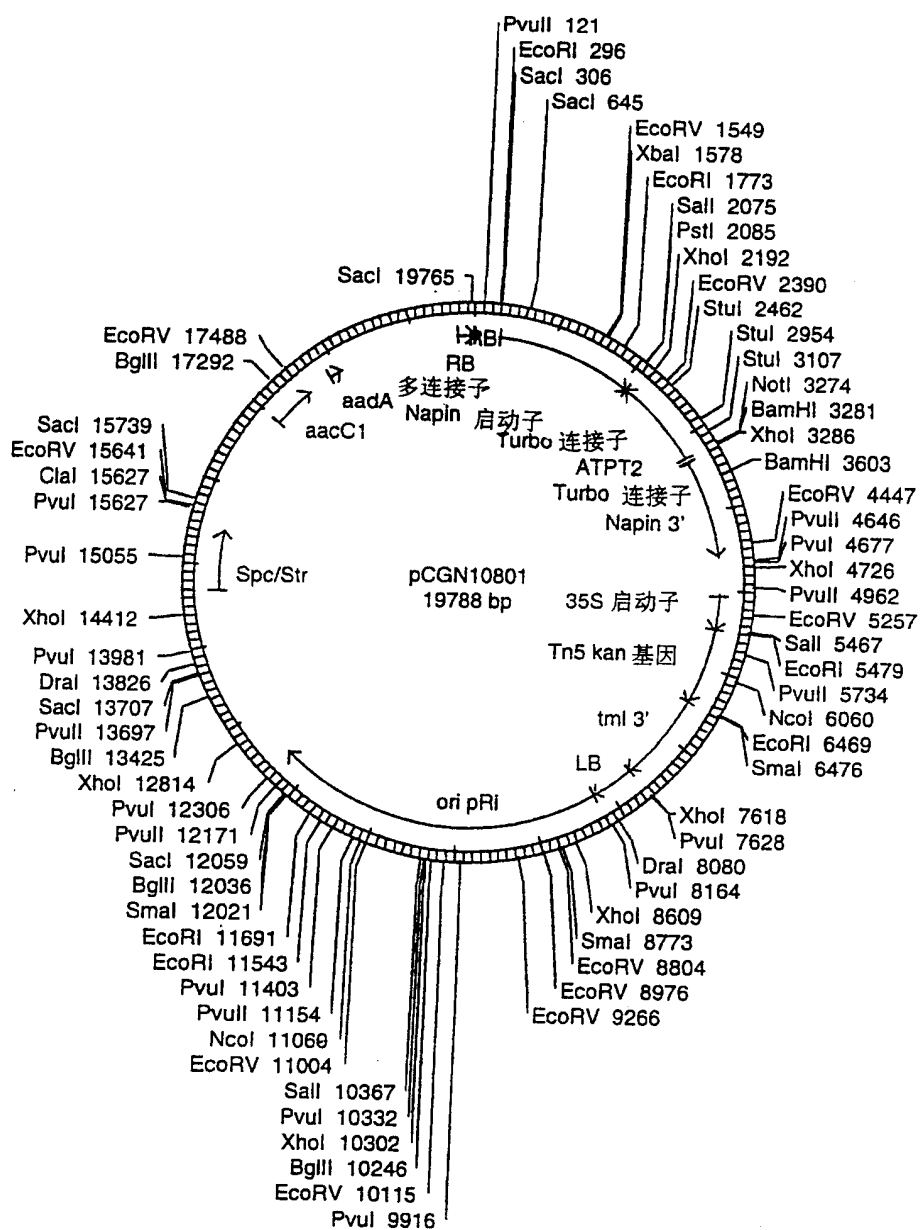


图 3

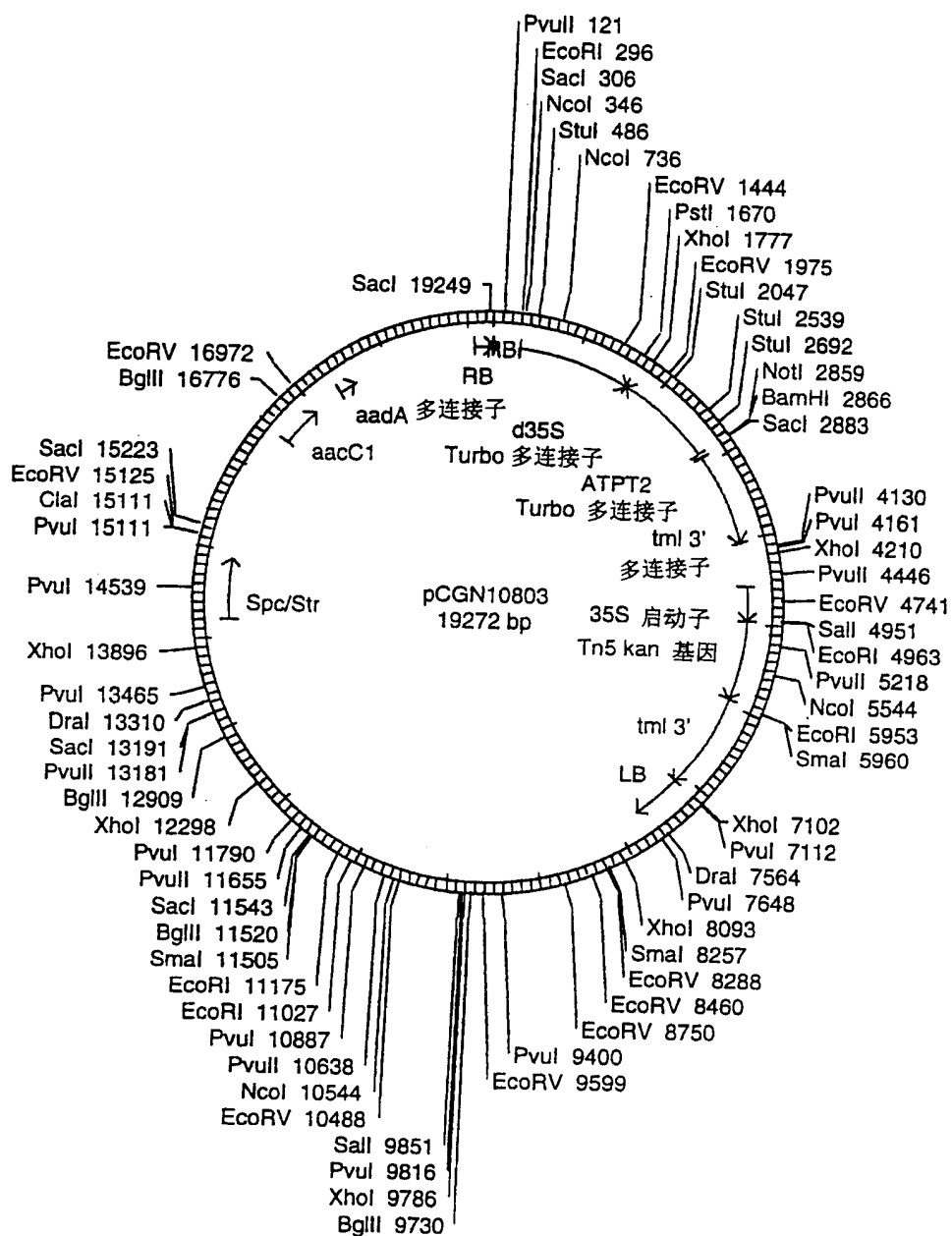


图 4

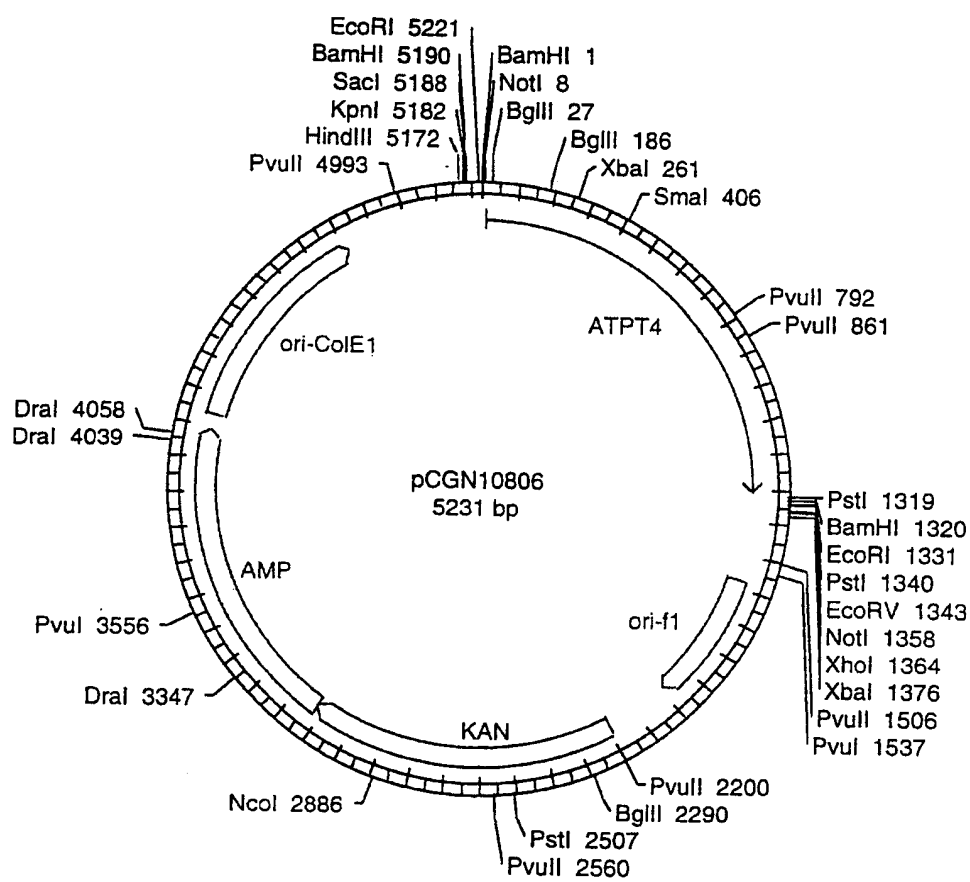


图 5

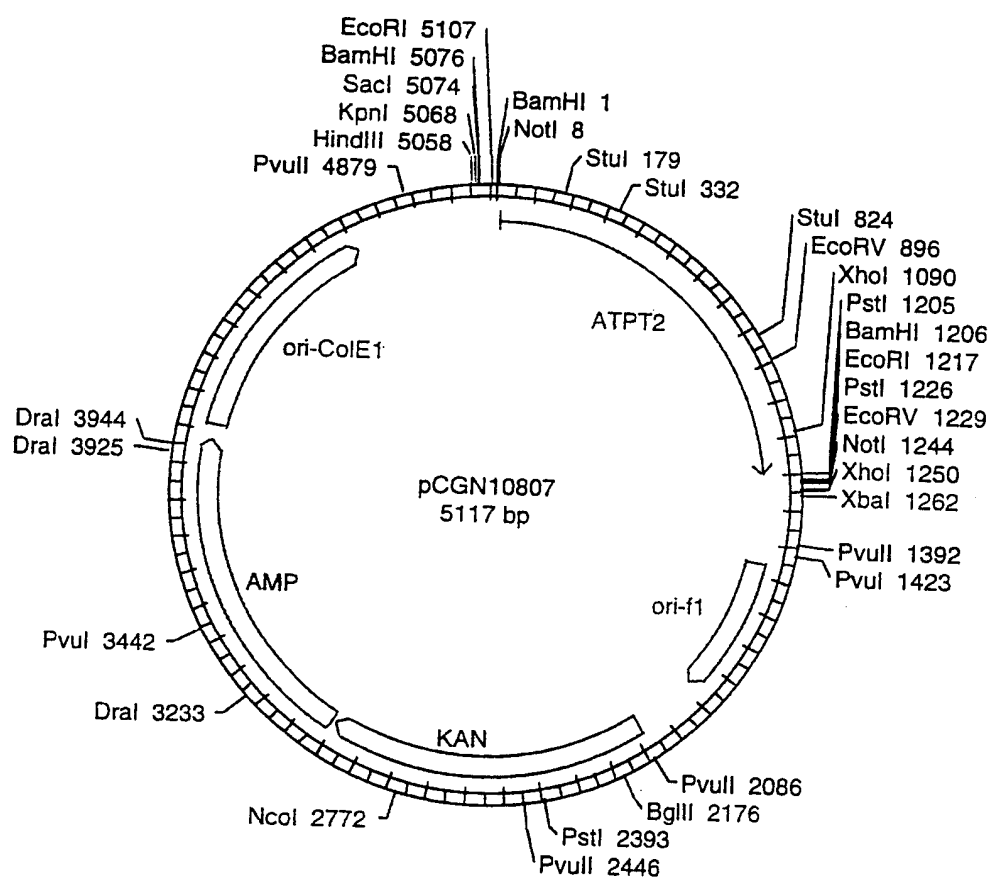


图 6

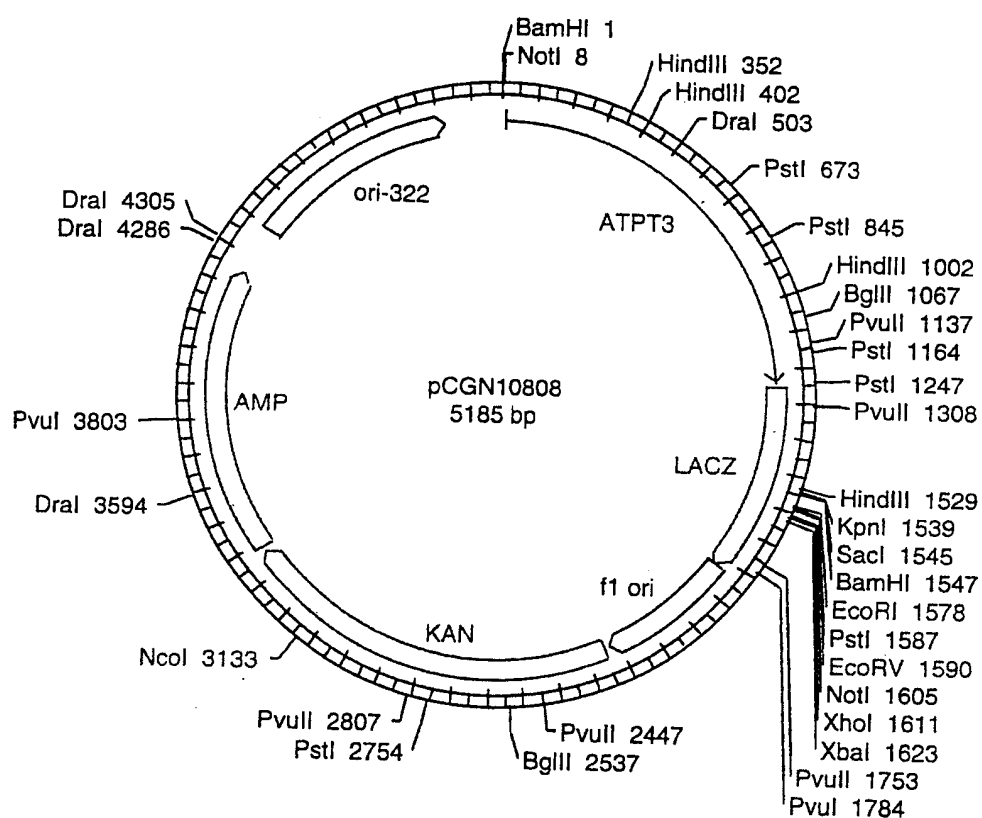


图 7

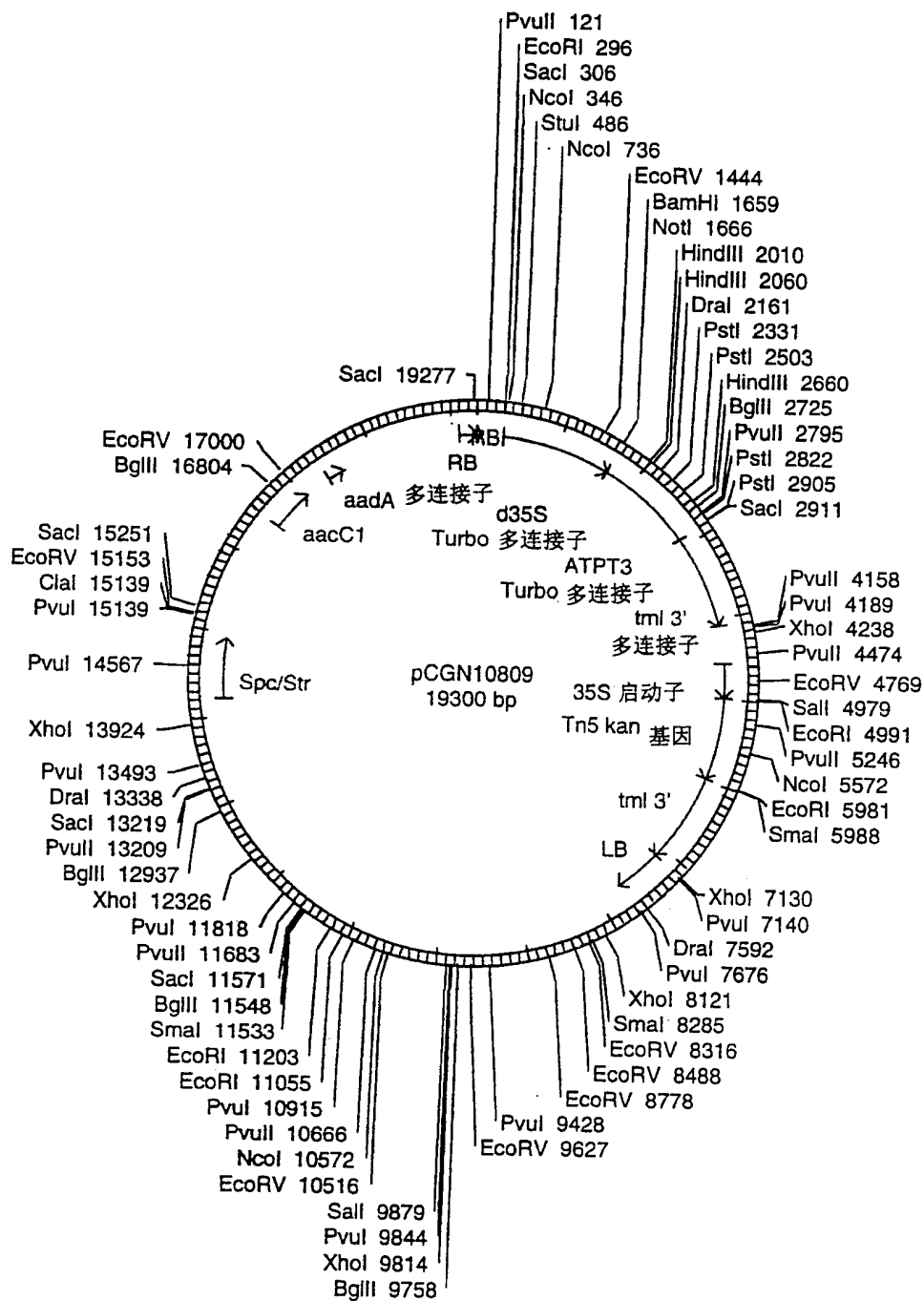


图 8

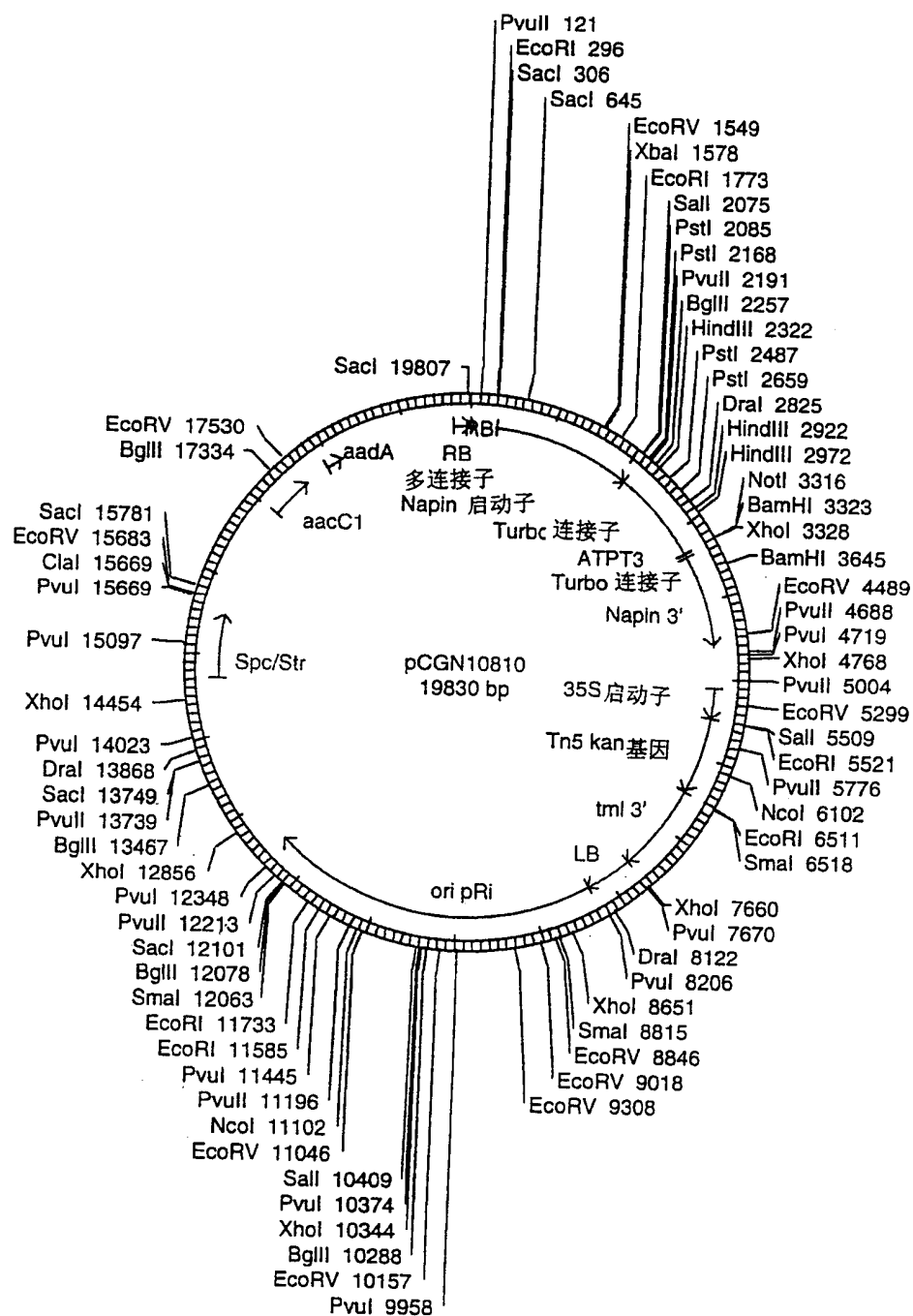


图 9

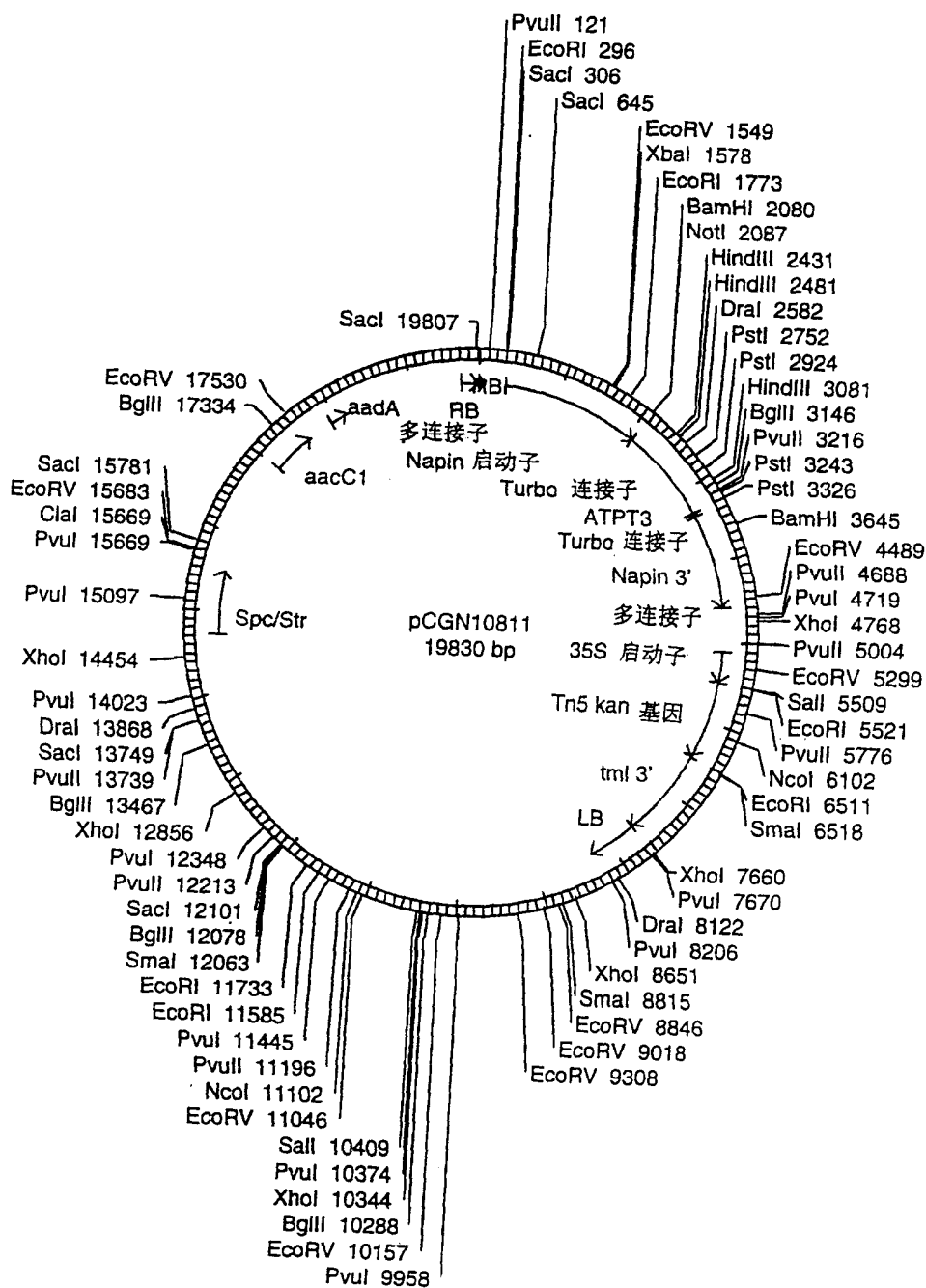


图 10

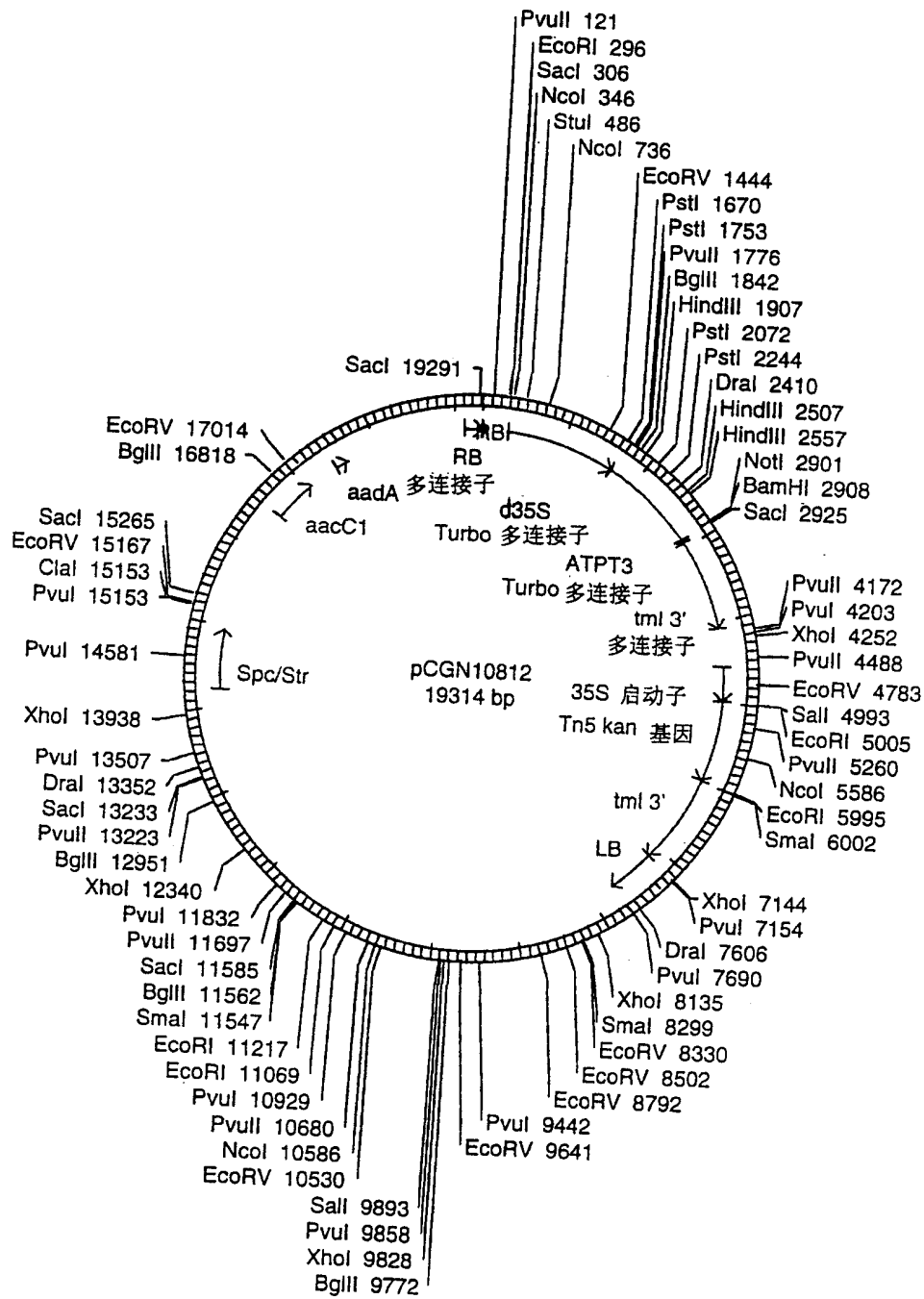


图 11

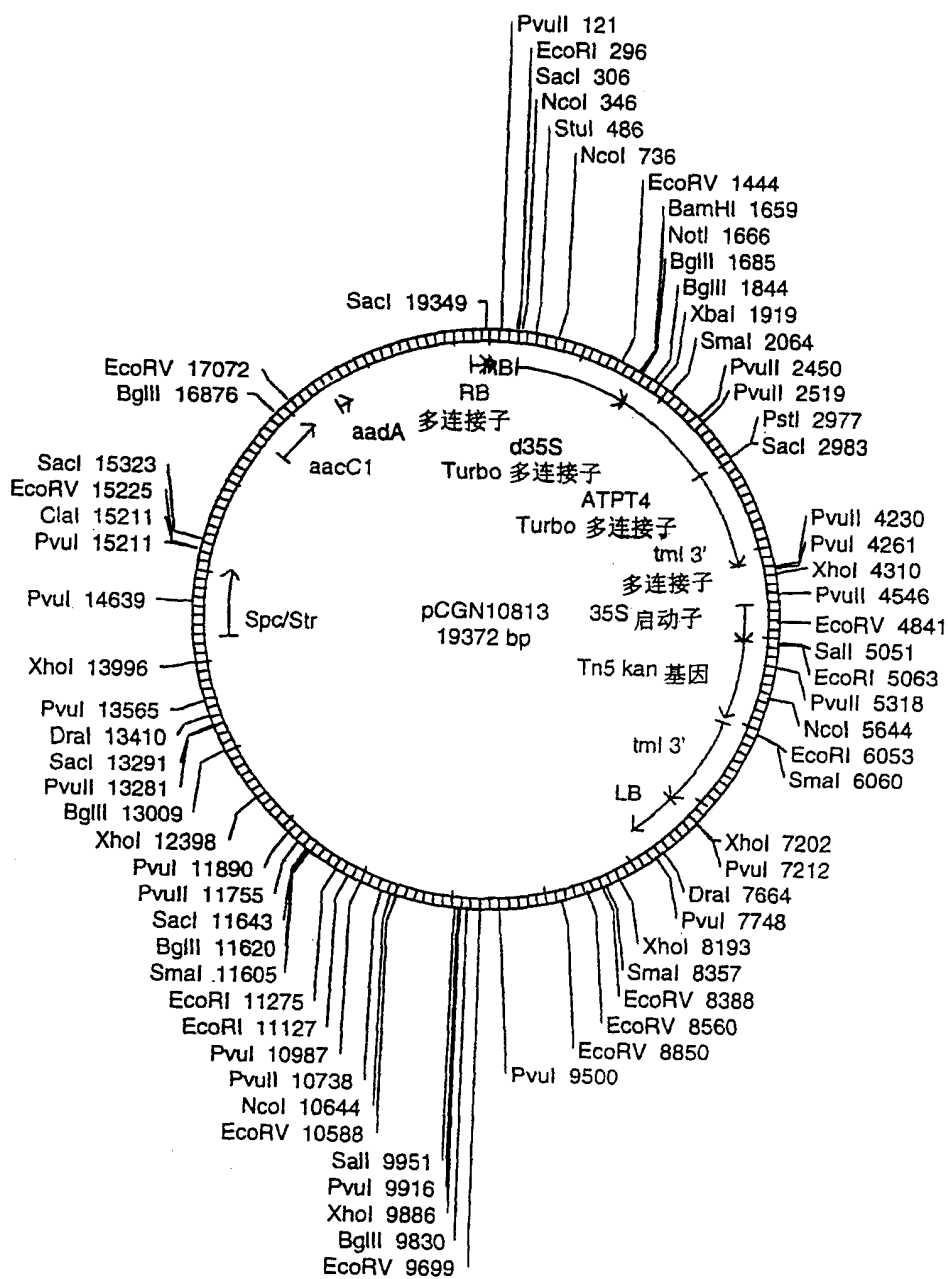


图 12

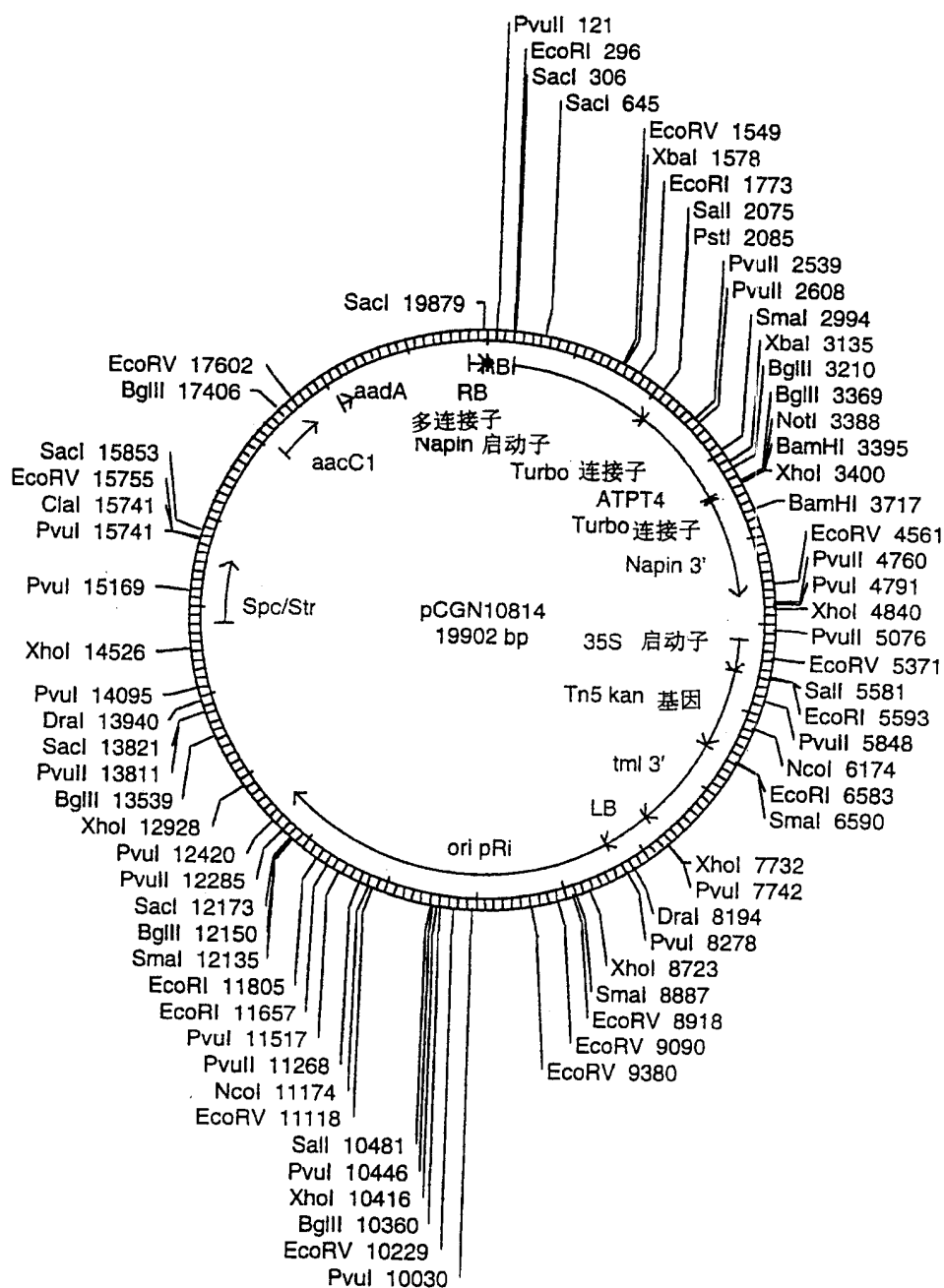


图 13

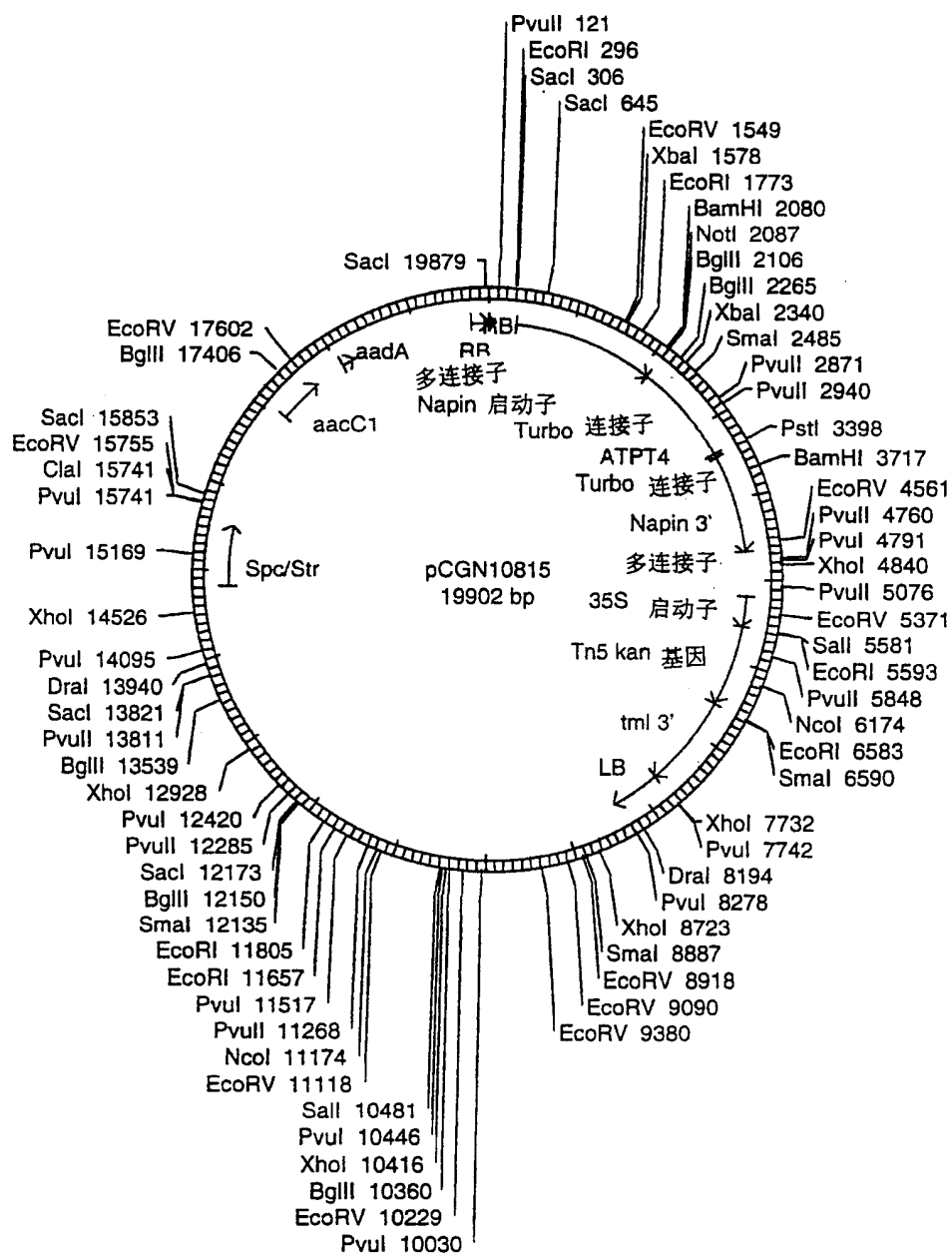


图 14

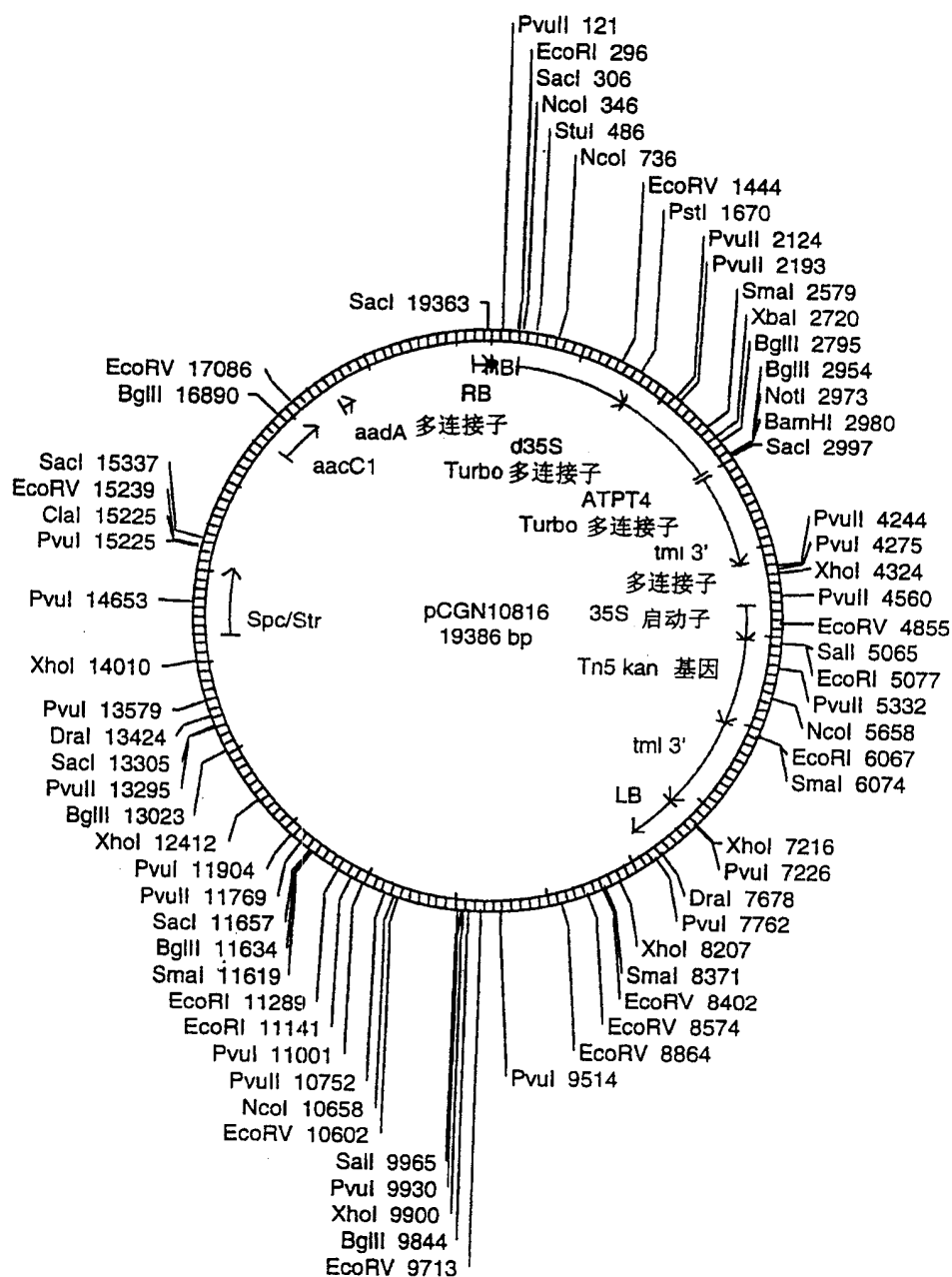


图 15

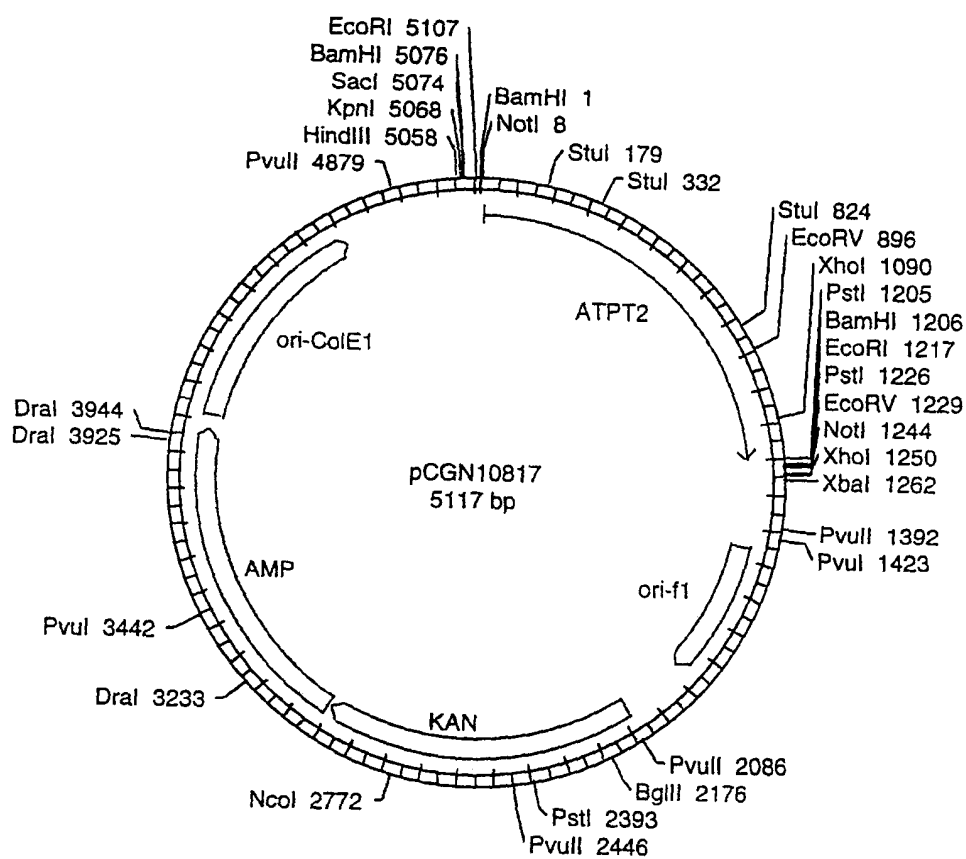


图 16

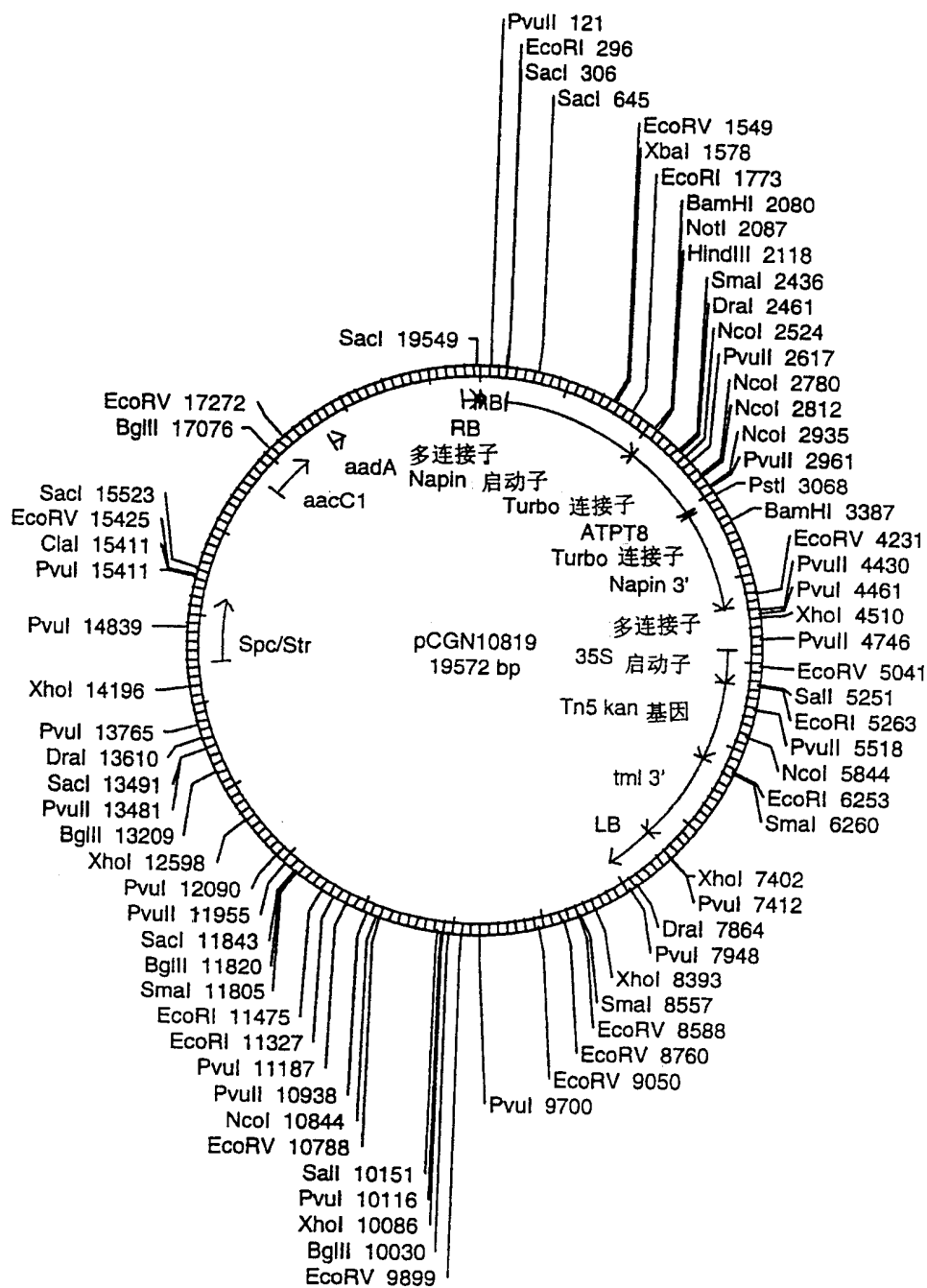


图 17

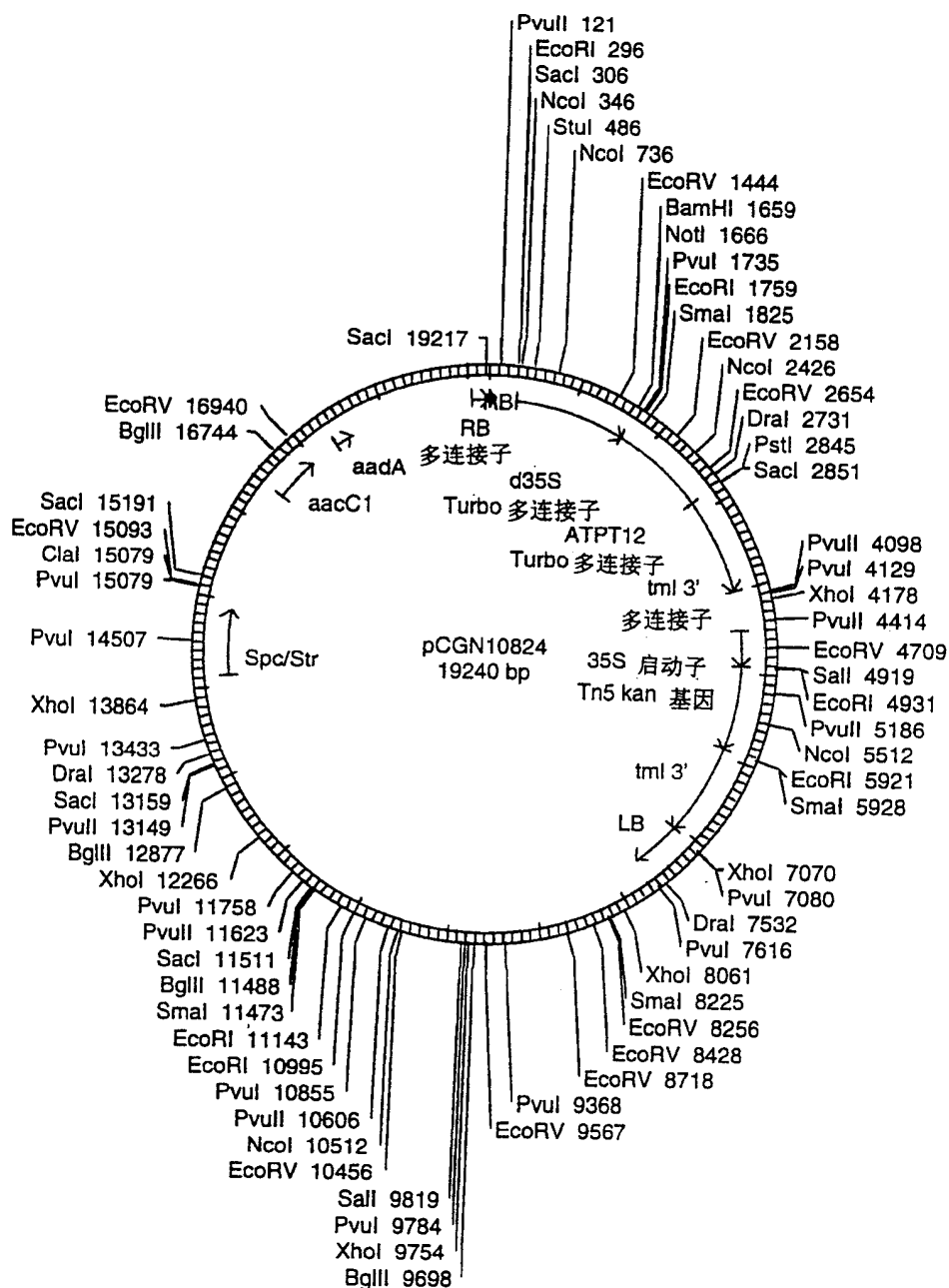


图 18

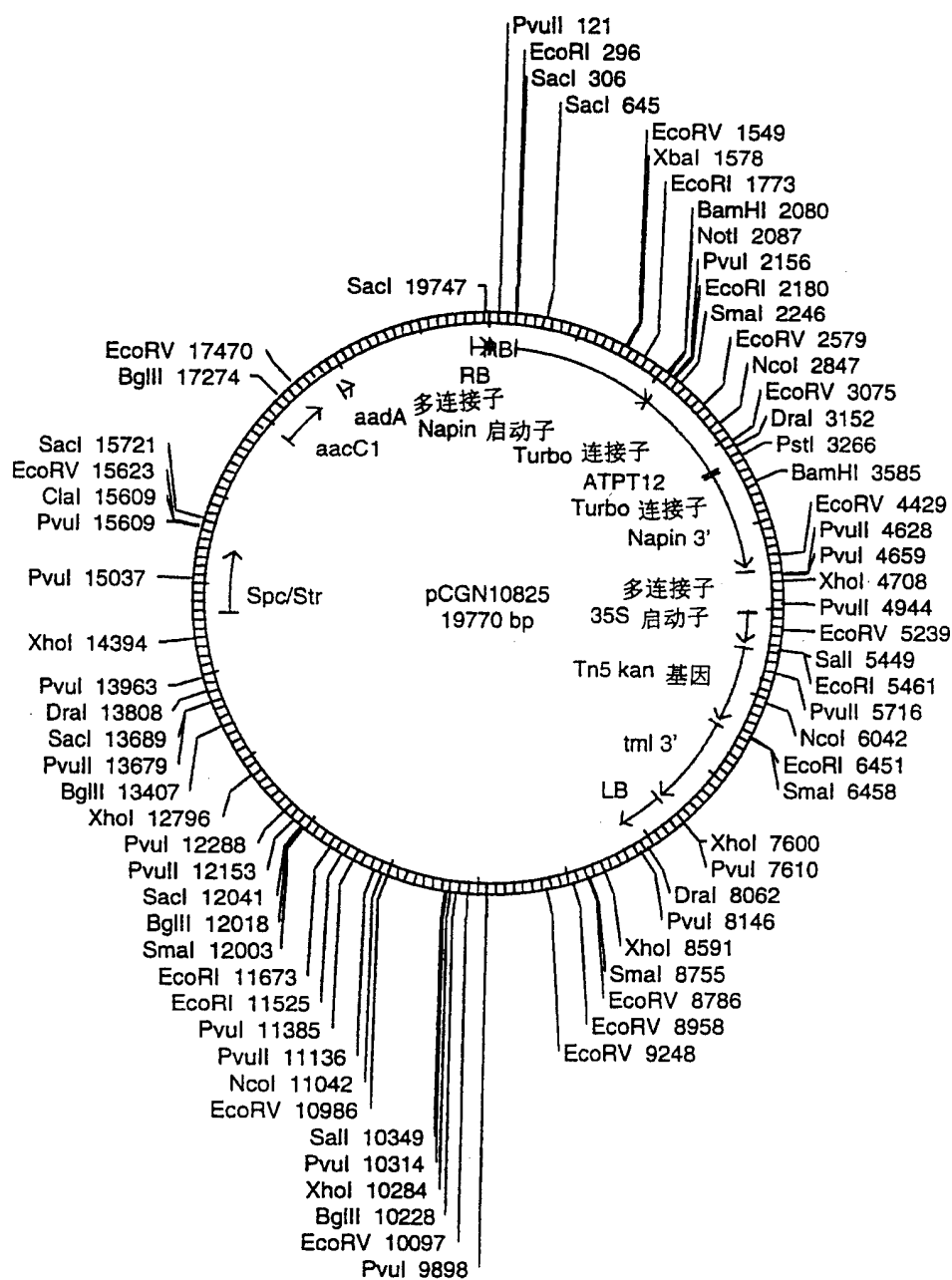


图 19

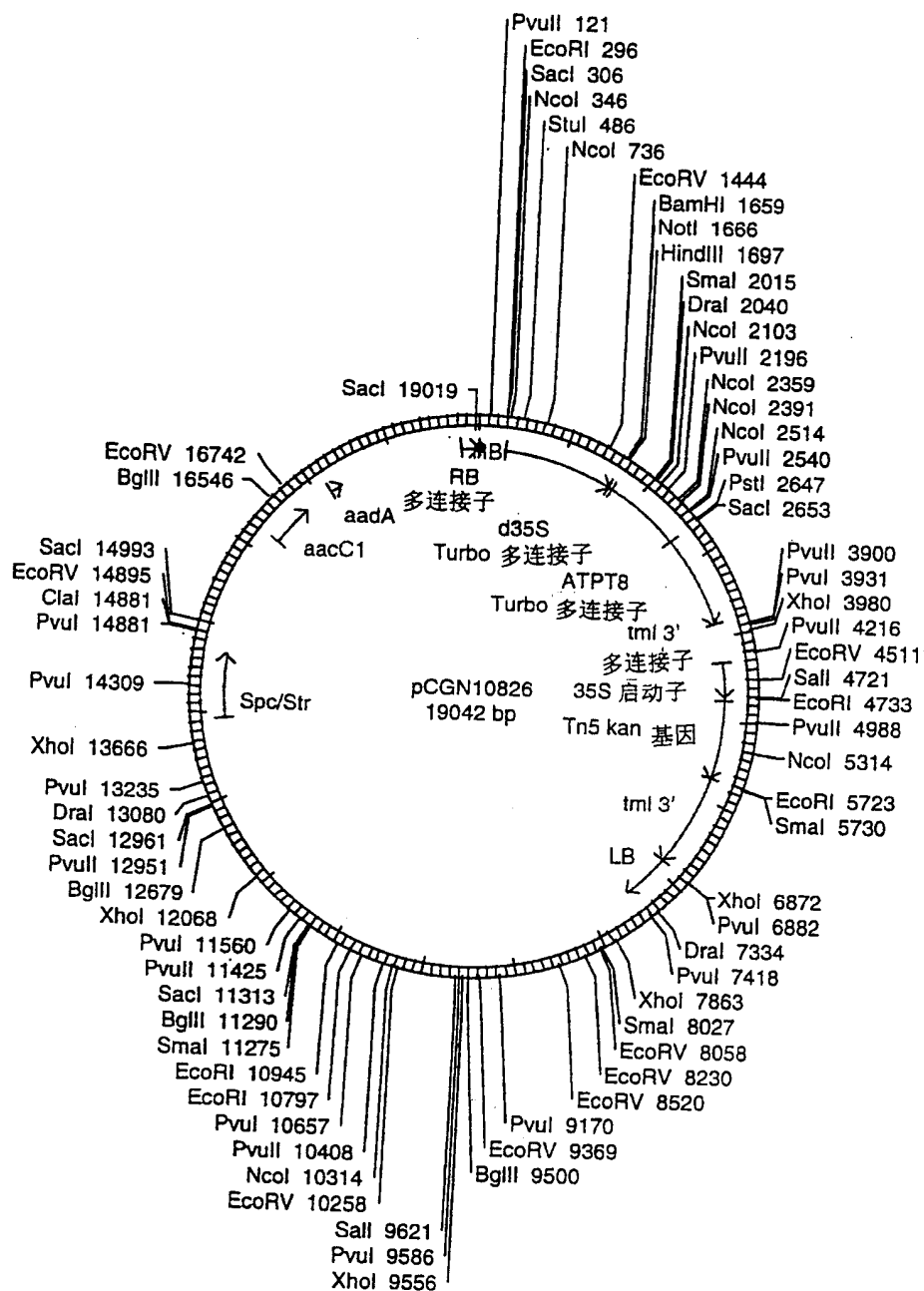


图 20


```

ATPT2 : * 20 * 40 * 60 * 80
SLR1736 : --MESLLSSSLVSAAGGFCWKQKQLKHLHSLSEIRVLRCDSSKVVAKPKFRNNLVRPDGQSSLLYPKHKSRRFVNATAGQ : 80
ATPT3 : MAFFGLSRVSRLLKSSVTPSPSSSALLQSQHKSLSNPVTHYTNPTFKCYPWNNDYQVWSKGRHLHQEKFFGVGWNRYRLICMSSS : 89
SLR0926 : -----MWRRSVYRFSSRISVSSSLPNRPLPWSRELCAVNSFSQP-----PVSTESTAKLGITGVRS DANRVFATA : 67
SLL1899 : -----MTSILNTVSTIHSRVTSVDRGVLSLRNSDSVEFT-----RRRSGFSTLIYESPGRRFVVRAAETDT : 4
ATPT12 : -----MVAOTATATG-----EISSRVAALAGLGHVYAR-----CYWELSKAKLSMLVATSG-----TGYYLGTNAALSPFGCYTCAG : 63
SLR0056 : -----TKIHRQHDSMG-----AVCKSYQLTGP-----RIIPLLTATTAASMI-----ASER--VDLPKLTILG : 4
ATPT8 : -----DKVKSQTPDKAP-----AGGSSINQLLGGKAS-----QETNKKIRLOLTKPVT--PPLVGVVCG--AASNFHWTPEDAKSLC : 4
SLR1518 : -----QNT-GONQAKA-----RQLLGMKGAAP-----GESSIWKIRLOLMKPIT--IPLINGVVCG--AASSGYIWSVEDFKAATC : 73
ATPT8 : -----EVPKLASAAEY-----FFKRGVQKQF-----RSTILLATALNRVP-----EALIGEST--DIVTSE--RVQR : 63
SLR1518 : -----SPLAPSTAPAT-----RKLWLAALKP-----PMVTAVVPIITVG-----SAVYGLTG--QWHDGFTIFLL : 59

```



```

ATPT2 : * 100 * 120 * 140 * 160 * 1
SLR1736 : PEAFDSNSKQK-----SFRDSIDAFYR-----FSRPHTVGTVLSLS-----VSFLVEKVS--DISPLFTGLE : 140
ATPT3 : -----MATIQAFWR-----FSRPHTEGTTLSNAVY-----LTLILGDN-SVNSPAS--DLVFG : 49
SLR0926 : SSVLEGPKKDDKEKSDGVVVKKASWIDLYPEEVRGYAKLARLDKPIGTWLAWPCSS-----IALADPS--LPSPFYALFGC : 170
ATPT4 : -----MVAOTATATG-----PLWETIIVL-----LRWHKPAGRILMIPALMA-----VCLAQ--G--LPPLP--GTAL : 56
SLL1899 : TAAATATATG-----EISSRVAALAGLGHVYAR-----CYWELSKAKLSMLVATSG-----TGYYLGTNAALSPFGCYTCAG : 138
SLR1899 : TKIHRQHDSMG-----AVCKSYQLTGP-----RIIPLLTATTAASMI-----ASER--VDLPKLTILG : 60
ATPT12 : DKVKSQTPDKAP-----AGGSSINQLLGGKAS-----QETNKKIRLOLTKPVT--PPLVGVVCG--AASNFHWTPEDAKSLC : 139
SLR0056 : QNT-GONQAKA-----RQLLGMKGAAP-----GESSIWKIRLOLMKPIT--IPLINGVVCG--AASSGYIWSVEDFKAATC : 73
ATPT8 : EVPKLASAAEY-----FFKRGVQKQF-----RSTILLATALNRVP-----EALIGEST--DIVTSE--RVQR : 63
SLR1518 : SPLAPSTAPAT-----RKLWLAALKP-----PMVTAVVPIITVG-----SAVYGLTG--QWHDGFTIFLL : 59

```



```

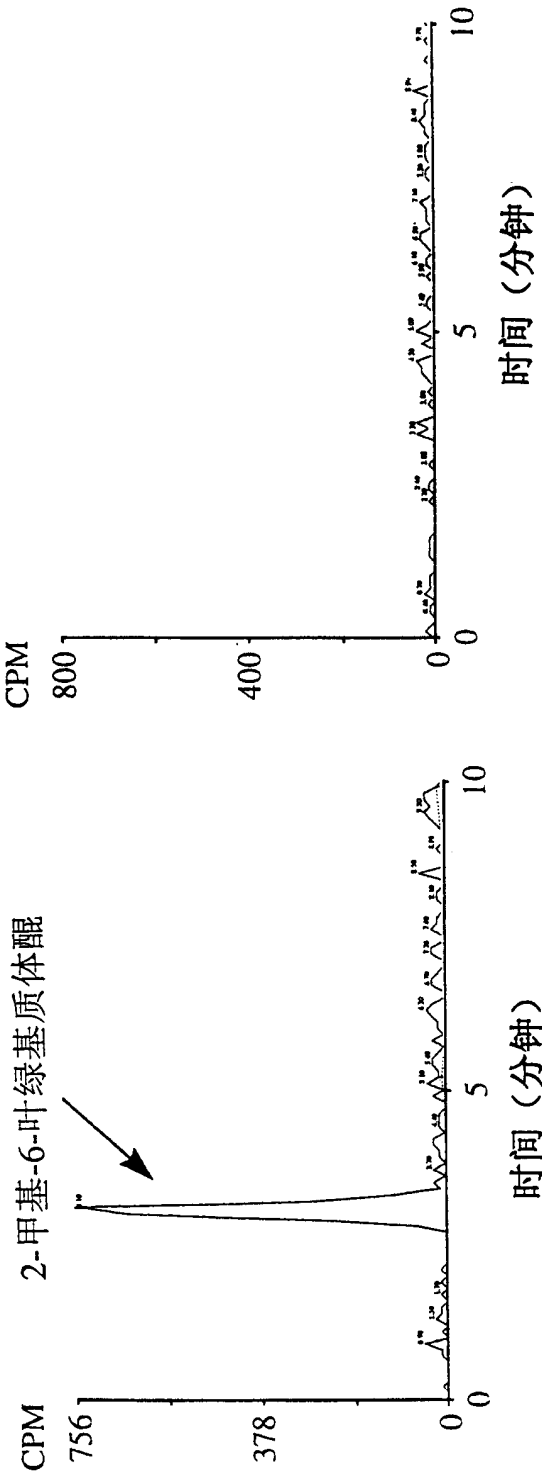
ATPT2 : * 80 * 200 * 220 * 240 * 260
SLR1736 : AVAAALMNIYIVGLNQSEVEIKVKNKPYLPLASGEYSNTGIAIASFSIMSFWLGVIVGWSWPLFWALF--SFGTYS--IN--PLR : 228
ATPT3 : AWLACLLGNVYVGLNQSEVEIKVKNKPYLPLANGDFS--AQGRWGLCGVAS--A--WGLG--LWLGLTVG--SL--GT--Y--S--PPVR : 134
SLR0926 : GAIL--IRGAGCTINDLEQLITTKVDITKLR--IASGLT--PFOG--GFLG--OL--LGLG--ILQLNNYSR--GAS--S--LVF : 246
ATPT4 : GTLA--TSGLGCVMND--ER--IN--POVE--TKQRLAARA--S--VQVG--GVA--VAL--CAG--L--FYLTP--SFW--CV--A--P--IV : 132
SLL1899 : TMTI--AASANSLOH--EISN--SKMK--TMLR--LPSGR--SVPH--AVAT--AGASGACL--L--SKTNM--AAG--AS--N--L--LYAF : 215
ATPT12 : GTLA--AASAOQLNCL--DQ--LE--YEML--TRAR--IPAG--OPRHA--FALA--GV--SFL--L--L--TFVNV--SGC--ALSG--I--F--YML : 137
SLR0056 : MMSGPCITGYTQTLNDEMER--IDAINEPYRPI--SGAISEPEVITO--VWL--LGG--GILG--VW--GHT--TPTV--F--ALG--G--SY : 223
ATPT8 : MMSGPLTGYTQTLNDEMER--IDAINEPYRPI--SGAISEPEVITO--VWL--LGG--GILG--VW--GHT--TPTV--F--ALG--G--SY : 157
SLR1518 : GIAE--ITEMIHVASL--HDDVLELDADT--RGVGSNVMGNKMSV--AGDF--SRACGL--A--L--KNTE--VA--LAT--VEHLV--TGETM : 144
ATPT8 : SADA--IAWINLS--D--S--TGIDVR--AHSVVNLTG--NRNLVFL--SNFF--LAG--LGLMSMS--WRAQDWT--LE--L--GV--F--F--GY : 138

```

图 22 1/2

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

图 22 2/2



集胞蓝细菌属 6083 野生型 集胞蓝细菌属 slr1736 敲除株

图 23

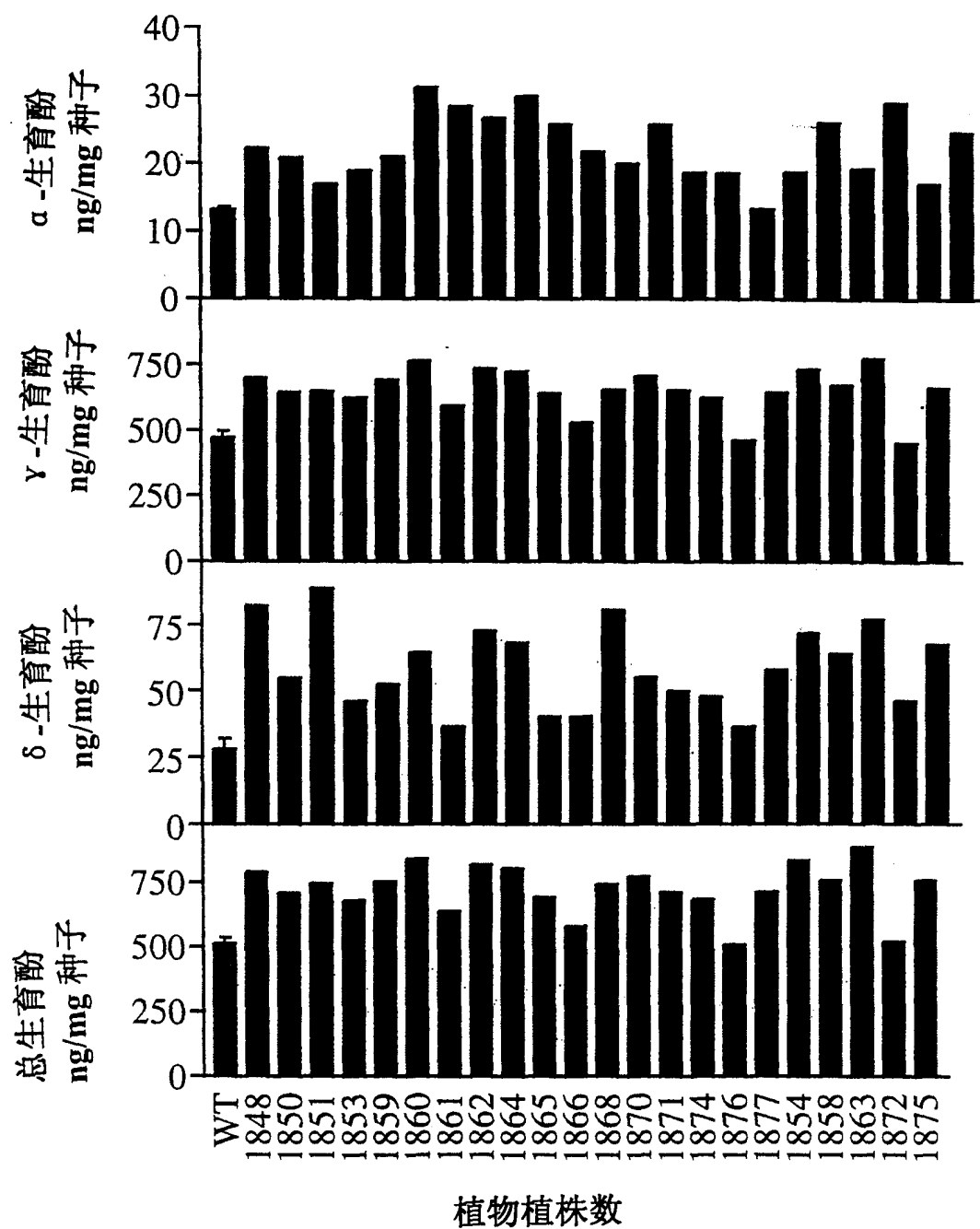


图 24

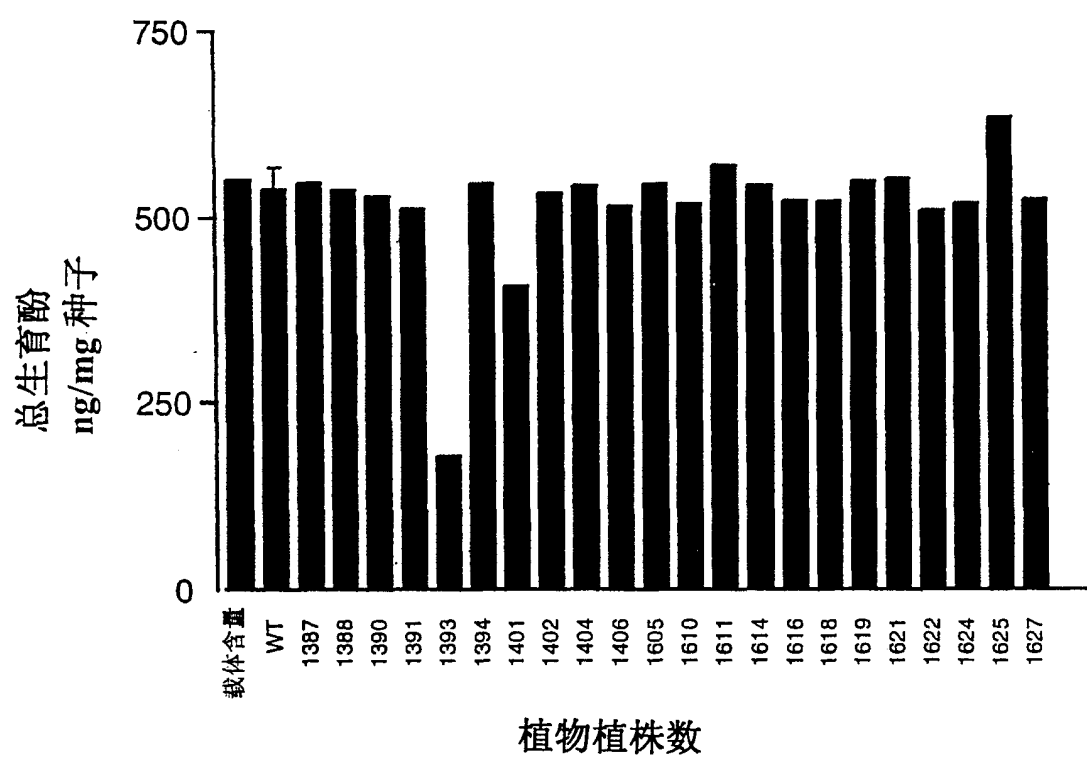


图 25