

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公布说明书

[21] 申请号 200910166802.9

[51] Int. Cl.

C12N 15/11 (2006.01)

C12N 5/10 (2006.01)

C12N 15/82 (2006.01)

A01H 5/10 (2006.01)

A23L 1/20 (2006.01)

C11B 1/00 (2006.01)

[43] 公开日 2010年1月27日

[11] 公开号 CN 101633924A

[22] 申请日 2003.5.5

[21] 申请号 200910166802.9

分案原申请号 03815772.1

[30] 优先权

[32] 2002.5.3 [33] US [31] 60/377236

[32] 2003.5.5 [33] US [31] 10/429516

[71] 申请人 孟山都技术公司

地址 美国密苏里州

[72] 发明人 王琦 坦亚·法加利

罗纳德·巴萨纳 梁继红

蒂姆·N·奥尔马索夫

约翰·达布劳斯基

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 付磊

权利要求书1页 说明书57页 附图21页

[54] 发明名称

用于在植物中表达基因的种子特异性 USP 启动子

[57] 摘要

本发明涉及植物遗传工程领域。更特别地，本发明涉及种子特异性基因表达。本发明提供能够在种子中转录异源核酸序列的启动子，和改进、制备和使用该启动子的方法。

1. 一种核酸分子, 包含被可操作地连接到第二核酸的启动子, 其中该启动子在严紧条件下与选自 SEQ ID NO: 2 或 10 的核酸序列及其互补体杂交。
2. 含有权利要求 1 的核酸分子的转化植物细胞。
3. 权利要求 2 的转化植物细胞, 其中所述第二核酸是结构性核酸。
4. 权利要求 3 的转化植物细胞, 其中所述结构性核酸编码选自种子储存蛋白、脂肪酸途径酶、生育酚生物合成酶、氨基酸生物合成酶、类固醇途径酶或淀粉分支酶的蛋白质。
5. 权利要求 3 的转化植物, 其中所述结构性核酸被引导来表达反义 RNA 分子。
6. 一种制备转化植物的方法, 包括:
 - (a) 提供权利要求 1 的核酸分子; 和
 - (b) 用所述核酸分子转化植物。
7. 一种获得促进结构性核酸产物生成的种子的方法, 包括:
 - (a) 种植含有权利要求 1 的核酸分子的转化植物, 其中所述转化植物生成所述种子, 并且所述第二核酸在所述种子中被转录; 和
 - (b) 从所述被转化植物中分离所述种子。
8. 从包含权利要求 2-5 中任一项的转化植物细胞的转化植物制备的粗粉。
9. 从包含权利要求 2-5 中任一项的转化植物细胞的转化植物分离的油。
10. 含有权利要求 1 的核酸分子的转化大豆植株细胞。
11. 一种核酸分子, 包含被可操作地连接到外源可转录多核苷酸分子的启动子, 其中该启动子选自:
 - (1) 与 SEQ ID NO: 2 或 10 具有至少 95% 的序列同一性的核酸分子, 其中所述核酸分子具有启动子活性; 和
 - (2) 包含 SEQ ID NO: 2 或 10 的至少 95 个连续核苷酸的片段, 其中所述片段具有启动子活性。

用于在植物中表达基因的种子特异性 USP 启动子

本申请是以下申请的分案申请：申请日 2003 年 5 月 5 日，申请号 03815772.1 (PCT/US0313848)，标题“用于在植物中表达基因的种子特异性 USP 启动子”。

技术领域

本申请要求享有 2002 年 5 月 3 日申请的在先临时申请 US60/377,236 的优先权和利益。

本发明涉及植物遗传工程领域。更特别地，本发明涉及种子特异性基因表达。本发明提供能够在种子中转录异源核酸序列的启动子，和改进、制备和使用该启动子的方法。

背景技术

种子给人和家畜提供重要的食物蛋白来源。但是，种子中的蛋白含量往往是不完全的。例如，许多种子蛋白缺乏一种或多种必需氨基酸。克服这种缺陷可以通过遗传改进天然或非天然的蛋白质，使之具有营养更加全面的氨基酸组成(或某些其他的期望特性)，并且在转基因植物中过量表达改进的蛋白质。此外，一种或多种基因可以被导入到农作物中来控制代谢路径和改进游离氨基酸的含量。这些方法可以用于制备具有重要的农业(如产量)、营养学和药学特性的农作物。

尽管许多分子工具是有效的，种子的遗传修饰常常受到不能有效沉积被加工的蛋白的限制。许多细胞内过程可能影响蛋白质的总沉积，包括转录、翻译、蛋白质组装和折叠、转运和蛋白质水解。干扰这些过程的一种或多种能够增加在遗传工程种子中生产的蛋白质的含量。

基因导入会对植物生长和发育产生有害作用。在这种情况下，基因表达需要被限制在期望的目标组织中。例如，有必要以种子特异性或种子增强的方式表达氨基酸异常基因，以避免产生影响产量或其他农学性状的不期望的

表型。

基因的启动子部分在控制基因表达中起关键作用。在启动子区，转录元件被组装，转录被起始。相对于随后的基因表达阶段，该早期步骤通常是关键的调控步骤。启动子的转录起始可以以多种方式被调控。例如，启动子可以被存在特定化合物诱导，仅在特定组织中表达基因，或者组成型地表达编码序列。因此，改进编码序列的转录可以通过可操作地将编码序列连接到具有不同调控特性的启动子上。

发明内容

本发明包括和提供含有包含启动子的核酸分子的被转化的植物，该启动子包含在严紧条件下与选自 SEQ ID NO:1,2,3,4,9,10,或 11 的核酸序列及其互补体杂交的核酸序列。本发明包括和提供含有包含被可操作地连接到结构性核酸序列的启动子的核酸分子的被转化的植物，该启动子包含选自 SEQ ID NO:1,2,3,4,9,10,或 11 及其互补体的核酸序列。

本发明包括和提供含有包含启动子的核酸分子的被转化的植物，该启动子包含的核酸序列选自：与 SEQ ID NO:1 或其互补体具有超过约 85.5% 同一性的核酸序列，与 SEQ ID NO:2 或其互补体具有超过约 85.5% 同一性的核酸序列，与 SEQ ID NO:3 或其互补体具有超过约 97.1% 同一性的核酸序列，与 SEQ ID NO:4 或其互补体具有超过约 96.4% 同一性的核酸序列。

本发明包括和提供制备被转化的植物的方法，包括：提供包含被可操作地连接到结构性核酸序列的启动子的核酸分子，该启动子包含选自 SEQ ID NO:1,2,3,4,9,10,或 11 及其互补体的核酸序列；和用所述核酸分子转化植物。

本发明包括和提供在种子中表达结构核酸分子的方法，包括：种植含有包含被可操作地连接到结构性核酸序列的启动子的核酸分子的被转化植物，该启动子包含选自 SEQ ID NO:1,2,3,4,9,10,或 11 及其互补体的核酸序列，其中所述被转化的植物生产所述种子，并且所述结构核酸分子在所述分子中被转录；和分离所述种子。

本发明包括和提供一种获得结构核酸分子的产物生成升高的种子的方法，包括：种植含有包含被可操作地连接到所述结构性核酸分子的启动子的核酸分子的被转化植物，该启动子包含选自 SEQ ID NO:1,2,3,4,9,10,或 11 及其互补体的核酸序列，其中所述被转化的植物生产所述种子，并且所述结构

核酸分子在所述种子中被转录；和从所述转化植物中分离所述种子。

本发明包括和提供获得结构核酸分子的产物生成升高的食物(meal)的方法，包括：种植含有包含被可操作地连接到所述结构性核酸序列的启动子的核酸分子的被转化植物，该启动子包含选自 SEQ ID NO:1,2,3,4,9,10,或 11 及其互补体的核酸序列，其中所述被转化的植物生产所述种子，并且所述结构核酸分子在所述种子中被转录；和制备包含所述被转化植物或其部分的所述食物。

本发明包括和提供获得促进结构核酸分子的产物生成的原种的方法，包括：种植含有包含被可操作地连接到所述结构性核酸分子的启动子的核酸分子的被转化植物，该启动子包含选自 SEQ ID NO:1,2,3,4,9,10,或 11 及其互补体的核酸序列，其中所述被转化的植物生产所述种子，并且所述结构核酸分子在所述种子中被转录；和制备包含所述被转化植物或其部分的所述原种。

本发明包括和提供获得结构核酸分子的产物生成升高的油的方法，包括：种植含有包含被可操作地连接到所述结构性核酸序列的启动子的核酸分子的被转化植物，该启动子包含在严紧条件下与选自 SEQ ID NO:1,2,3,4,9,10,或 11 及其互补体的核酸序列杂交的核酸序列，其中所述被转化的植物生产种子，并且所述结构核酸分子在所述种子中被转录；和分离所述油。

本发明包括和提供一种含有载体的细胞，该载体包含选自 SEQ ID NO:1,2,3,4,9,10,或 11 及其互补体的核酸序列。

本发明包括和提供由一种或多种被转化的植物的种子制备的油，该被转化的植物含有包含启动子的核酸序列，该启动子包含选自 SEQ ID NO:1,2,3,4,9,10,或 11 及其互补体的核酸序列。

本发明包括和提供由一种或多种被转化的植物的种子制备的油，该被转化的植物含有包含可操作地连接到结构性核酸序列的启动子的核酸序列，该启动子包含选自 SEQ ID NO:1,2,3,4,9,10,或 11 及其互补体的核酸序列，其中该启动子相对于结构性核酸序列来说是异源的。

本发明包括和提供从含有核酸分子的被转化植物中生成的种子，该核酸分子包括：包含选自 SEQ ID NO:1,2,3,4,9,10,或 11 及其互补体的核酸序列的启动子。

本发明包括和提供包含含有核酸分子的被转化植物或其部分的原种，该核酸分子包含启动子，启动子包含选自 SEQ ID NO:1,2,3,4,9,10,或 11 及其互

补体的核酸序列。

本发明包括和提供包含来自含有核酸分子的被转化植物的植物原料的食物，该核酸分子包含启动子，启动子包含选自 SEQ ID NO:1,2,3,4,9,10,或 11 及其互补体的核酸序列。

本发明包括和提供种子容器(container)，其中至少约 25% 的所述种子包含被可操作地连接到结构性核酸序列的包含选自 SEQ ID NO:1,2,3,4,9,10,或 11 及其互补体的核酸序列的启动子，其中所述启动子相对于结构性核酸序列来说是异源的。

本发明包括和提供一种基本上纯的核酸分子，其包含选自 SEQ ID NO:1,2,3,4,9,10,或 11 及其互补体的核酸序列

本发明包括和提供一种基本上纯的核酸分子，该核酸分子具有的核酸序列选自：与 SEQ ID NO:1 或其互补体具有超过约 85.5% 同一性的核酸序列，与 SEQ ID NO:2 或其互补体具有超过约 85.5% 同一性的核酸序列，与 SEQ ID NO:3 或其互补体具有超过约 97.1% 同一性的核酸序列，与 SEQ ID NO:4 或其互补体具有超过约 96.4% 同一性的核酸序列。

本发明包括和提供含有包含启动子的核酸分子的被转化的大豆植物，该启动子包含在严紧条件下与选自 SEQ ID NO:1,2,3,4,9,10,或 11 及其互补体的核酸序列杂交的核酸序列。

本发明包括和提供含有的核酸分子的被转化的大豆植物，核酸分子包含被可操作地连接到结构性核酸序列的启动子，该启动子包含选自 SEQ ID NO:1,2,3,4,9,10,或 11 及其互补体的核酸序列。

附图说明

附图 1 是载体 pMON13773 的示意图。

附图 2 是载体 pMON58101 的示意图。

附图 3 是载体 pMON58102 的示意图。

附图 4 是载体 pMON58106 的示意图。

附图 5 是载体 pMON58110 的示意图。

附图 6 是载体 pMON58100 的示意图。

附图 7 是载体 pMON58107 的示意图。

附图 8 是载体 pMON58113 的示意图。

附图 9 是载体 pMON55526 的示意图。

附图 10 是载体 pMON58108 的示意图。

附图 11 是载体 pMON39319 的示意图。

附图 12 表示在多个启动子控制下在转基因大豆中表达的 GUS 的相对活性的图。

附图 13 是载体 pMON58130 的示意图。

附图 14 表示各种构建体的 GUS 的相对活性的图。

附图 15 是载体 pMON63604 的示意图。

附图 16 是载体 pMON63605 的示意图。

附图 17 是载体 pMON55542 的示意图。

附图 18 是载体 pMON63821 的示意图。

附图 19 是载体 pMON63819 的示意图。

附图 20 是载体 pMON63820 的示意图。

附图 21 是载体 pMON63654 的示意图。

序列简介

SEQ ID NO:1 是来自蚕豆(*Vicia faba*)的 USP88 启动子序列。

SEQ ID NO:2 是来自蚕豆的 eUSP88 启动子序列。

SEQ ID NO:3 是来自蚕豆的 USP99 启动子序列。

SEQ ID NO:4 是来自蚕豆的 USP91 启动子序列。

SEQ ID NO:5 是来自蚕豆的 USP 启动子序列。

SEQ ID NO:6 是用于扩增来自蚕豆的 USP 启动子的引物序列。

SEQ ID NO:7 是用于扩增来自蚕豆的 USP 启动子的引物序列。

SEQ ID NO:8 是用于扩增来自蚕豆的 USP 启动子的引物序列。

SEQ ID NO:9 是来自蚕豆的 USP99.5 启动子序列。

SEQ ID NO:10 是来自蚕豆的 USP95 启动子序列。

SEQ ID NO:11 是来自蚕豆的 USP68 启动子序列。

SEQ ID NO:12 是用于扩增来自蚕豆的 USP 启动子的引物序列。

SEQ ID NO:13 是用于扩增来自蚕豆的 USP 启动子的引物序列。

SEQ ID NO:14 是用于扩增来自蚕豆的 USP 启动子的引物序列。

SEQ ID NO:15 是用于扩增来自蚕豆的 USP 启动子的引物序列。

SEQ ID NO:16 是用于扩增是来自蚕豆的 USP 启动子的引物序列。

具体实施方式

定义

提供如下定义来帮助理解对本发明的详细描述。

短语“编码序列”、“结构序列”和“结构性核酸序列”是指包含核苷酸顺序排列的物理结构。核苷酸以一系列的三联体排列，每个三联体构成密码子。每个密码子编码特定氨基酸。因此，编码序列、结构序列和结构性核酸序列编码一系列氨基酸，构成蛋白质、多肽和肽序列。编码序列、结构序列和结构性核酸序列可以被包含在较大的核酸分子、载体等中。此外，这些序列中核苷酸的顺序排列可以以序列表、附图、表格、电子介质等形式表述。

短语“DNA 序列”、“核酸序列”和“核酸分子”是指包含核苷酸顺序排列的物理结构。DNA 序列或核苷酸序列可以被包含在较大的核苷酸分子、载体等中。此外，这些序列中核酸的顺序排列可以以序列表、附图、表格、电子介质等形式表述。

术语“表达”是指基因转录来制备相应的 mRNA，以及该 mRNA 被翻译来生产相应的基因产物(即肽、多肽和蛋白质)。

短语“反义 RNA 的表达”是指 DNA 转录来生成能够杂交到第二个 RNA 分子的第一个 RNA 分子

术语“同源性”是指按照位置相同的百分比，两条或多条核酸或氨基酸序列之间的相似水平(即序列相似性或同一性)。同源性还指不同核酸或蛋白质之间的相似功能特性的概念。

术语“异源的”是指来自不同来源的两条或多条核酸或蛋白质序列之间的关系。例如，如果启动子与编码序列的组合通常不是天然存在的，则启动子对于编码序列来说是异源的。此外，特定序列对于其所插入的细胞或生物体来说是“异源的”(即不是自然地存在于特定的细胞或生物体中)。

术语“杂交”是指当两条核酸链具有足够大的序列同一性时，第一条核酸链通过氢键碱基配对与第二条链结合的能力。当两条核酸分子在合适条件下相互退火时，发生杂交。

短语“被可操作地连接”是指两个或多个核酸区域或核酸序列的功能性的空间排列。例如，启动子区被置于相对于核酸序列的位置，使得核酸序列

的转录受该启动子区域引导。从而，启动子区域“被可操作地连接”到该核酸序列上。

术语或短语“启动子”或“启动子区域”是指核酸序列，通常存在于编码序列的上游(5')，能够引导核酸序列转录为 mRNA。一般地，启动子或启动子区域提供 RNA 聚合酶和正确起始转录所必需的其他因子的识别位点。如本文所预期，启动子或启动子区域包括启动子的变体，其通过插入或删除调控区域，进行随机或定点突变启动子等来获得。启动子的活性或强度可以根据其相对于转录活性预先已经被评价的启动子所产生的 RNA 的量、或者细胞或组织中蛋白质沉积的含量来测定。

短语“5'UTR”是指 DNA 上游的非翻译区，或者基因的编码区的 5'端的非翻译区。

短语“3'UTR”是指 DNA 下游的非翻译区，或者基因的编码区的 3'端的非翻译区。

短语“重组载体”是指任何介质，如质粒、粘粒、病毒、自主复制序列、噬菌体、线性单链的、环状单链的、线性双链的、或环状双链的 DNA 或 RNA 核苷酸序列。重组载体可以来自任何来源，能够进行基因组整合或自主复制。

短语“调控序列”是指位于编码序列的上游(5')、内部或下游(3')的核苷酸序列。编码序列的转录和表达通常受到调控序列存在与否的影响。

短语“基本上同源的”是指两条序列在序列上至少约 90% 相同，这通过本文描述的 BestFit 程序测定(版本 10; Genetics Computer Group, Inc., University of Wisconsin Biotechnology Center, Madison, WI)，采用默认的参数。

术语“转化”是指将核酸导入到受体宿主中。术语“宿主”是指细菌细胞、真菌、动物或动物细胞、植物或种子、或任何植物部分或组织，包括植物细胞、原生质体、愈伤组织、根、块茎、种子、茎、叶子、幼苗、胚和花粉。

如在此所用，短语“转基因植物”是指具有被导入的核酸的植物，该核酸被稳定地导入到植物的基因组中，例如，核或质体基因组。

如在此所用，短语“基本上被纯化的”是指基本上与通常在天然状态下与之相关的所有其他分子分开的分子。更优选地，基本上被纯化的分子是存在于制备物中的占优势种类。基本上被纯化的分子约 60%，优选约 75%，更加优选约 90% 和最优选约 95% 以上地不含天然混合物中存在的其他分子(不计溶剂)。短语“基本上被纯化的”不指包含在天然状态下存在的分子。

优选实施方案的详述

本发明提供能够在种子中转录异源结构核酸分子的启动子，以及改进、制备和使用这些启动子的方法。本发明还提供含有种子特异性启动子的组合物、被转化的宿主细胞和植物，以及制备和使用它们的方法。

核酸分子

本发明提供包含选自 SEQ ID NO:1,2,3,4,9,10,或 11 及其互补体的序列的核酸分子。SEQ ID NO:5 表示被报道的 USP 启动子。

核酸杂交是 DNA 操作领域的技术人员熟知的技术。特定的一对核酸的杂交特性指示它们的相似性或同一性。

低严谨条件可以被用于选择与靶核酸序列具有低序列同一性的核酸序列。可以采用条件如约 0.15 M 到约 0.9 M 氯化钠，温度范围在约 20°C 到约 55°C。

高严谨条件可以被用于选择与被公开的核酸序列具有较大同一性的核酸序列(Sambrook 等, 1989)。

高严谨条件通常包括在约 2X - 约 10X SSC (用蒸馏水稀释含有 3 M 氯化钠和 0.3 M 柠檬酸钠的 20X SSC 母液, pH 7.0 得到), 约 2.5X - 约 5X Denhardt 溶液(用蒸馏水稀释含有 1% (w/v) 牛血清白蛋白、1% (w/v) ficoll、和 1% (w/v) 聚乙烯吡咯烷酮的 50X 母液得到), 约 10 mg/mL - 约 100 mg/mL 鱼精子 DNA, 和约 0.02% (w/v) - 约 0.1% (w/v) SDS 的核酸杂交, 在约 50°C - 约 70°C 下温育数小时至过夜。高严谨条件优选通过 6X SSC, 5X Denhardt 液, 100 mg/mL 鱼精子 DNA 和 0.1% (w/v) SDS, 在 55°C 下温育数小时产生。

杂交后通常进行多个洗涤步骤。洗涤组合物通常包含 0.5X - 约 10X SSC, 和 0.01% (w/v) - 约 0.5% (w/v) SDS, 在约 20°C - 70°C 下温育 15min。优选地, 在 0.1X SSC、65°C 下至少洗涤一次后, 核酸片段仍然被杂交。

本发明的核酸分子优选在高严谨条件下与具有选自 SEQ ID NO:1,2,3,4,9,10,或 11 及其互补体的序列的核酸分子杂交。在一个优选实施方案中, 本发明的核酸分子在高严谨条件下与包含 SEQ ID NO:1 的核酸分子杂交。在一个优选实施方案中, 本发明的核酸分子在高严谨条件下与包含 SEQ ID NO:2 的核酸分子杂交。在一个优选实施方案中, 本发明的核酸分子在高严谨条件下与包含 SEQ ID NO:3 的核酸分子杂交。在一个优选实施方案中,

本发明的核酸分子在高严紧条件下与包含 SEQ ID NO:4 的核酸分子杂交。在一个优选实施方案中，本发明的核酸分子在高严紧条件下与包含 SEQ ID NO:9 的核酸分子杂交。在一个优选实施方案中，本发明的核酸分子在高严紧条件下与包含 SEQ ID NO:10 的核酸分子杂交。在一个优选实施方案中，本发明的核酸分子在高严紧条件下与包含 SEQ ID NO:11 的核酸分子杂交。

在一个优选实施方案中，本发明的核酸分子包含与 SEQ ID NO:1 的序列同一性大于约 85.5%，或大于约 87,88,89,90,91,92,93,94,95,96,97,98,或约 99% 的核酸序列。

在一个优选实施方案中，本发明的核酸分子包含与 SEQ ID NO:2 的序列同一性大于约 85.5%，或大于约 87,88,89,90,91,92,93,94,95,96,97,98,或约 99% 的核酸序列。

在一个优选实施方案中，本发明的核酸分子包含与 SEQ ID NO:3 的序列同一性大于约 97.1%，98,98.5,或大于约 99% 的核酸序列。

在一个优选实施方案中，本发明的核酸分子包含与 SEQ ID NO:4 的序列同一性大于约 96.4%,97,98,或约 99% 的核酸序列。

序列同一性的百分比优选采用如下方法确定。采用 Dnastar 软件包中的 EditSeq 应用程序(DNASTAR,Inc.,Madison,WI)将序列转化为 EditSeq DNA 序列文件。转化的文件被输入到 Dnastar 软件包的 Megalign 应用程序中。采用 Clustal 方法在默认设定(default setting)下将输入的序列与加权(weighted)的残基权重(weight)表比对。这种方法被用于确定本发明的启动子序列与其他序列以及相互之间的百分同一性。

用于确定百分同一性的另一种方法采用序列分析软件包的“BestFit”或“Gap”程序(版本 10;Genetics Computer Group,Inc.,University of Wisconsin Biotechnology Center, Madison,WI)。“Gap”采用 Needleman 和 Wunsch (1970)的算法来确定两条序列之间的排列，最大化匹配的数量并且最小化缺口数量。“BestFit”采用 Smith 和 Waterman 的局部同源性算法(Smith 和 Waterman,1981;Smith 等, 1983)，进行两条序列之间的最好相似性片段的最佳排列，并且插入缺口来最大化匹配数量。百分同一性最优选采用“BestFit”程序在默认参数下确定。

本发明还提供具有与任何 SEQ ID NO: 1,2,3,4,9,10,或 11,以及其互补体的百分同一性大于与 SEQ ID NO:5 的百分同一性的核酸分子片段。在一个优

选实施方案中，片段与任何 SEQ ID NO:1 - 4 的百分同一性至少比该片段与 SEQ ID NO:5 的百分同一性大至少约 1%，优选至少约 2,3,4,5,10,15,20,或大于至少约 30%。

在一个实施方案中，片段为本发明的核酸分子的约 50 个到约 600 个之间的连续核苷酸，约 50 个到约 550 个之间的连续核苷酸，约 50 个到约 500 个之间的连续核苷酸，约 50 个到约 450 个之间的连续核苷酸，约 50 个到约 400 个之间的连续核苷酸，约 50 个到约 350 个之间的连续核苷酸，约 50 个到约 300 个之间的连续核苷酸，约 50 个到约 250 个之间的连续核苷酸，约 50 个到约 200 个之间的连续核苷酸，约 50 个到约 150 个之间的连续核苷酸，约 50 个到约 100 个之间，约 15 个到约 100 个之间，约 15 个到约 50 个之间，约 15 个到约 25 个之间的连续核苷酸。

在另一个实施方案中，片段包含本发明的核酸序列的至少约 15,20,30,40,50,60,70,80,90,100,150,200,250,300,350,400,450,500,550,600,或约 650 个连续核苷酸。

本发明包括编码具有酶活性的多肽的核酸序列，该酶活性为类固醇途径酶鲨烯环氧酶、固醇甲基转移酶 I、固醇 C4 脱甲基酶、钝叶鼠曲草醇 (obtusifoliol)C14 α 脱甲基酶、固醇 C5 去饱和酶和固醇甲基转移酶 II。

鲨烯环氧酶(也称作鲨烯单加氧酶)催化鲨烯转化为环氧化鲨烯(2,3-氧化鲨烯)，环氧化鲨烯是植物甾醇类合成路径中的起始固醇分子的前体，环阿屯醇。这是该路径中首次报道的步骤，该路径是需氧的。环氧化鲨烯的形成也是最近经常报道的动物、真菌和植物的固醇生物合成中步骤。近来报道一些拟南芥属 (*Arabidopsis*) 和芸苔属 (*Brassica*) 环氧化鲨烯基因的同系物 (Schafer, U.A., Reed, D.W., Hunter, D.G., Yao, K., Weninger, A.M., Tsang, E.W., Reaney, M.J., MacKenzie, S. L., 和 Covello, P.S.(1999), *Plant Mol.Biol.*, 39(4):721-728)。这些作者还进行 PCT 申请，其中公开了使用环氧化鲨烯的反义技术来评价植物中的鲨烯水平的用途 (WO 97/34003)。

鲨烯环氧酶，也称作鲨烯单加氧酶，是参考号为 1.14.99.7 的酶，*Enzyme Nomenclature*, 1992, 146。

一些鲨烯环氧酶是本领域已知的。它们包括拟南芥鲨烯环氧酶蛋白质序列登录号 AC004786 拟南芥属 (*Arabidopsis*) 鲨烯环氧酶登录号 N64916，和拟南芥鲨烯环氧酶登录号 T44667。日本专利申请 07194381 A

公开了编码哺乳动物的鲨烯环氧化酶的 DNA。

本发明的另一个方面是包含编码鲨烯环氧化酶的核酸序列的重组构建体和载体，以及制备新的鲨烯环氧化酶方法，包括在使生成鲨烯环氧化酶的条件下培养用新的构建体或载体转化的宿主细胞一段时间，和回收由此生成的鲨烯环氧化酶。

S-腺苷 L-甲硫氨酸：固醇 C24 甲基转移酶(SMT1 和 SMT2)催化将甲基从辅助因子 S-腺苷 L-甲硫氨酸上转移到固醇侧链的 C24 中心(Bach, T.J.和 Benveniste, P. (1997), *Prog. Lipid Res.*, 36:197-226)。在高等植物细胞中，SMT 表示产生 24-甲基和 24-乙基固醇的能力(Schaffer, A., Bouvier-Nave, Benveniste, P., Schaller, H. (2000) *Lipids*, 35:263-269)。采用酵母 *erg6* 表达系统的 SMT 功能性特征清楚地证明，在特定植物种类中，SMT1 序列编码环阿屯醇-C24-甲基转移酶和 SMT2 序列编码 C24-亚甲基烯胆烷醇(methylenelophenol)-C24-甲基转移酶(Bouvier-Nave, P., Husselstein, T., 和 Benveniste, P. (1998), *Eur. J. Biochem.*, 246:518-529)。一些编码 SMT1 和 SMT2 的基因已经被报道和评论(Schaffer, A., Bouvier-Nave, Benveniste, P., Schaller, H. (2000) *Lipids*, 35:263-269)。表达 SMT1 或 SMT2 的同系物的转基因植物已经被试验(Schaffer, A., Bouvier-Nave, Benveniste, P., Schaller, H.(2000) *Lipids*, 35:263-269)。这些基因用于改进植物固醇组成的用途也被两个专利申请覆盖(WO 98/45457 和 WO 00/61771)。

本领域已知的固醇甲基转移酶 I 可被用于本发明。示例性的序列包括已知的拟南芥固醇甲基转移酶 I 蛋白序列登录号 U71400(以 SEQ ID NO:19 公开)、已知的烟草固醇甲基转移酶 I 蛋白序列登录号 U81312(以 SEQ ID NO:20 公开)和蓖麻(*Ricinus Communis*)固醇 C 甲基转移酶, *Eur. J. Biochem.*, 246(2):518-529(1997)(Complete cds, 登录号 g2246457)。

S-腺苷 L-甲硫氨酸固醇 C24 甲基转移酶——一种编码拟南芥 S-腺苷 L-甲硫氨酸固醇 C24 甲基转移酶的核酸序列已经被 Husselstein 等,(1996) *FEBS Letters* 381 : 87-92 发表。 Δ 24 固醇甲基转移酶是编号为 2.1.1.41 的酶, *Enzyme Nomenclature*, 1992,160。

固醇 C4 脱甲基酶催化个别脱甲基反应的前面反应，导致 C4 上的两个甲基基团去除。虽然在动物和真菌中，连续发生两个 C4 甲基基团的去除，但是已经报道在植物中，在第一次和第二次 C4 脱甲基之间还有其他步骤

(Bach, T. J.和 Benveniste, P.(1997), *Prog. Lipid Res.*, 36:197-226)。C4 脱甲基是由单加氧酶、NAD⁺依赖的固醇 4-脱羧酶和 NADPH 依赖的 3-类固醇还原酶组成的微粒体酶复合物催化的。

固醇 C14 脱甲基酶催化 C14 处的脱甲基，去除 C14 的甲基基团并且在该位置上产生双键。在真菌和动物中，这是固醇合成路径的第一步。但是，在高等植物中，14 α -甲基在一个 C4 甲基缺失后被去除。因此，虽然羊毛固醇是动物和真菌细胞中的 C14 脱甲基酶的底物，植物酶利用钝叶鼠曲草醇 (obtusifoliol) 作为底物。固醇 C14 脱甲基化通过细胞色素 P-450 复合物调节。去除 14 甲基的机制包括两个氧化步骤，在 C29 生成乙醇，然后转化为乙醛，以及另一个氧化步骤，涉及产生甲酸和典型的 8,14-双烯固醇的去甲酰基化 (Aoyama, Y., Yoshida, Y., Sonoda, Y., 和 Sato, Y. (1989) *J. Biol. Chem.*, 264:18502-18505)。来自 *Sorghum bicolor*(L) Moench 的钝叶鼠曲草醇 (obtusifoliol) C14 α -脱甲基酶已经采用由从内在 14 氨基酸序列设计的 PCR 引物生成的基因特异性探针克隆，并且在大肠杆菌中被功能性地表达 (Bak, S., Kahn, R. A., Olsen, C.E., 和 Halkier, B.A.(1997) *The Plant Journal*, 11(2):191-201)。此外，编码羊毛甾醇 14-脱甲基酶的酿酒酵母 *CYP51A1* 在烟草中被功能性地表达 (Grausem, B., Chaubet, N., Gigot, C., Loper, J. C., 和 Benveniste, P. (1995) *The Plant Journal*, 7(5):761-770)。

固醇 C14 脱甲基酶和序列是本领域已知的。例如，*Sorghum bicolor* 钝叶鼠曲草醇 C14 α 脱甲基酶 CYP51 mRNA 在 *Plant J.*, 11(2):191-201(1997) 中被描述 (全编码序列 (cds) 登录号 U74319)。

本发明的另一个方面在于包含编码新的钝叶鼠曲草醇 C14 α 脱甲基酶的核酸序列的重组构建体和载体，以及制备新的钝叶鼠曲草醇 C14 α 脱甲基酶的方法，包括在使生成钝叶鼠曲草醇 C14 α 脱甲基酶的条件下培养用新的构建体或载体转化的宿主细胞一段时间，和回收由此生成的钝叶鼠曲草醇 C14 α 脱甲基酶。

固醇 C5 去饱和酶催化 Δ^5 双键的插入，这通常在 Δ^7 -固醇水平上发生，从而形成 $\Delta^{5,7}$ 固醇 (Parks 等, *Lipids*, 30:227-230(1995))。该反应被报道包括立体特异地去除 5 α 和 6 α 氢原子，这两个氢原子分别由 (+) 和 (-) R-甲羟戊酸 4 pro-R 和 5 pro-S 氢生物合成地产生 (Goodwin, T.W.(1979) *Annu.Rev.Plant Physio.*, 30:369-404)。该反应显然是需要氧和 NADPH 或 NADH 的。去饱和

酶已经被报道是一种存在于微粒体中的多酶复合物。由去饱和酶本身、细胞色素 b_5 和吡啶核苷酸依赖的黄素蛋白组成。 Δ^5 -去饱和酶被报道是单加氧酶,其通过细胞色素 b_5 利用来自被还原的吡啶核苷酸的电子(Taton, M., 和 Rahier, A. (1996) Arch. Biochem. Biophys., 325:279-288)。编码固醇-C5 去饱和酶的拟南芥 cDNA 通过 *ERG3* 中缺乏 *erg3* 的酵母突变体的功能性互补物被克隆, *ERG3* 是编码麦角固醇生物合成必需的固醇 C5 去饱和酶的基因 (Gachotte D., Husselstein, T., Bard, M., Lacroute F., 和 Benveniste, P. (1996) *The Plant Journal*, 9(3):391-398)。已知的固醇 C5 去饱和酶可用于本发明, 包括拟南芥固醇 C5 去饱和酶蛋白质序列登录号 X90454, 以 SEQ ID NO:22 公开, 和在 *The Plant J.* 9(3):391-398(1996)中描述的固醇 C5 去饱和酶的拟南芥 mRNA (全编码序列登录号 g1061037)。

NCBI(国家生物技术信息中心)数据库显示了 37 条固醇去饱和酶的序列, 它们可以被用于本发明。如下是这种序列的示例。来自酵母: C5 固醇去饱和酶 NP_013157(酿酒酵母(*Saccharomyces cerevisiae*)); 推定的 C5 固醇去饱和酶裂殖体 (fission)T40027(粟酒裂殖酵母 (*Schizosaccharomyces pombe*)); C5 固醇去饱和酶裂殖体 T37759 (粟酒裂殖酵母); C5 固醇去饱和酶 JQ1146(酿酒酵母); C5 固醇去饱和酶 BAA21457 (粟酒裂殖酵母); C5 固醇去饱和酶 CAA22610(粟酒裂殖酵母); 推定的 C5 固醇去饱和酶 CAA16898(粟酒裂殖酵母); 可能的 C5 固醇去饱和酶 O13666 (*erg3_schpo*); C5 固醇去饱和酶 P50860(*Erg3_canga*); C5 固醇去饱和酶 P32353(*erg3_yeast*); C5,6 去饱和酶 AAC99343(白色假丝酵母(*Candida albicans*)); C5 固醇去饱和酶 BAA20292(酿酒酵母); C5 固醇去饱和酶 AAB39844(酿酒酵母); C5 固醇去饱和酶 AAB29844(酿酒酵母); C5 固醇去饱和酶 CAA64303(酿酒酵母); C5 固醇去饱和酶 AAA34595(酿酒酵母); C5 固醇去饱和酶 AAA34594 (酿酒酵母)。来自植物的: C5 固醇去饱和酶 S71251(拟南芥); 推定的固醇 C5 去饱和酶 AAF32466(拟南芥); 固醇 C5 去饱和酶 AAF32465(拟南芥); 推定的固醇去饱和酶 AAF22921(拟南芥(*Arabidopsis thaliana*)); Δ^7 -固醇 C5 去饱和酶(拟南芥); 固醇 C5,6 去饱和酶同系物 AAD20458(烟草); 固醇 C5 去饱和酶 AAD12944(拟南芥); 固醇 C5,6 去饱和酶 AAD04034 (烟草); 固醇 C5 去饱和酶 CAA62079 (拟南芥)。来自哺乳动物的: 固醇 C5 去饱和酶(*Mus musculus*) BAA33730; 固醇 C5 去饱和酶 BAA33729(*Homo sapiens*); 7-烯胆(甾)烷醇氧

化酶 CAB65928 (*Leishmania major*); 7-烯胆(甾)烷醇氧化酶(7-烯胆(甾)烷醇 C5 去饱和酶)088822(家鼠); 7-烯胆(甾)烷醇 C5 去饱和酶 075845 (*Homo sapiens*); Δ^7 -固醇 C5 去饱和酶 AAF00544 (*Homo sapiens*)。其他: 真菌固醇 C5 去饱和酶同系物 BAA18970 (*Homo sapiens*)。

对于可用于本发明的编码固醇 C5 去饱和酶的 DNA 序列, NCBI 核苷酸检索“固醇去饱和酶”获得 110 条序列。如下是这些序列的示例。NC_001139(酿酒酵母); NC_001145(酿酒酵母); NC_001144(酿酒酵母); AW700015(*Physcomitrella patens*); AB004539(粟酒裂殖酵母); 和 AW596303(大豆(*Glycine max*)); AC012188(拟南芥)。

结合导入 HMG-CoA 还原酶基因以及固醇甲基转移酶 II 基因到细胞中, 来除了减少 24-甲基固醇如菜油甾醇的沉积外, 还减少固醇路径中间的化合物沉积。

已知的固醇甲基转移酶 II 可以被用于本发明, 包括拟南芥固醇甲基转移酶 II 蛋白质序列(来自 FEBS Lett.381(12):87-92(1996)登录号 X89867 的完整 mRNA 编码序列), 以 SEQ ID NO:21 公开。

编码任何一种前述的影响固醇生物合成路径的酶的重组构建体可以被结合到包含重组构建体的重组载体中, 该重组构建体包含分离的 DNA 分子。这种载体可以是细菌的或植物的表达载体。

在优选实施方案中, 任何一种本发明的植物或生物体用本发明的核酸和编码选自鲨烯环氧酶、固醇甲基转移酶 I、固醇 C4 脱甲基酶、钝叶鼠曲草醇 (*obtusifoliol*)C14 α 脱甲基酶、固醇 C5 去饱和酶和固醇甲基转移酶 II 的一种成员的基因转化。在一个优选实施方案中, 本发明的植物或生物体用 SEQ ID NO:1-5 的一种或多种和编码选自鲨烯环氧酶、固醇甲基转移酶 I、固醇 C4 脱甲基酶、钝叶鼠曲草醇 C14 α 脱甲基酶、固醇 C5 去饱和酶和固醇甲基转移酶 II 的一种成员的基因转化。在另一个优选实施方案中, 本发明的植物或生物体用 SEQ ID NO:1-5 的一种或多种, 编码选自鲨烯环氧酶、固醇甲基转移酶 I、固醇 C4 脱甲基酶、钝叶鼠曲草醇 C14 α 脱甲基酶、固醇 C5 去饱和酶或固醇甲基转移酶 II 的一种成员的基因, 和在本文别处公开的编码生育酚途径酶的基因的一种或多种转化。在还有一个优选实施方案中, 本发明的植物或生物体用 SEQ ID NO:1-5 的一种或多种和编码选自鲨烯环氧酶、固醇甲基转移酶 I、固醇 C4 脱甲基酶、钝叶鼠曲草醇 C14 α 脱甲基酶、固醇 C5 去

饱和酶或固醇甲基转移酶 II 的一种成员的两种基因和在本文别处公开的编码生育酚途径酶的两种基因转化。任何上述的生育酚和固醇生物合成基因的结合可以在一种或多种本领域已知和在本文描述的构建体或载体中被导入到植物。

启动子

在一个实施方案中，任何被公开的核酸分子可以是启动子。在一个优选实施方案中，该启动子是组织或器官特异性的，并且优选是种子特异性的。在一个特别优选的实施方案中，启动子优先在胚乳或胚芽中表达相关的结构基因。在优选实施方案中，启动子是 USP 启动子。在特别优选的实施方案中，启动子是蚕豆(*vicia faba*)USP 启动子。

在一个方面，如果在该组织或器官中 mRNA 以比在其他组织或器官中至少高 10 倍，优选至少高 100 倍或至少高 1000 倍水平被表达，则该启动子被认为是组织或器官特异性的。可以在单个时间点上或多个时间点上测定 mRNA 水平，同样倍数增加可以是平均倍数增加或来自试验测定值的推定值。由于是进行水平的比较，可以采用任何测定 mRNA 水平的方法。在优选方面，被比较的组织或器官是种子或种子组织和叶片或叶片组织。在另一个优选方面，比较多个组织或器官。优选的多种比较是种子或种子组织与 2, 3, 4, 或多种组织或器官之间比较，这些组织或器官选自花组织、花原端、花粉、叶片、胚、苗、叶原基、苗端、根、根尖、维管组织和子叶。如在此所用，植物器官的例子是种子、叶片、根等。组织的例子是叶原基、根尖、维管组织等。

启动子的活性和强度可以根据 mRNA 或蛋白质沉积的含量来测定，其相对于 mRNA 或蛋白质沉积的总含量，特别地产生。启动子优选以超过 2.5% 的水平表达被可操作地连接的核酸序列；更加优选超过约 5, 6, 7, 8, 或约 9%；甚至更加优选超过约 10, 11, 12, 13, 14, 15, 16, 17, 18, 或约 19%，和最优选超过约 20% 的总 mRNA。

此外，启动子的活性或强度可以相对于被清楚确定的启动子(其转录活性预先被评价)来表达。例如，目标启动子可以被可操作地连接到报道序列(如 GUS)上，并被导入到特定细胞类型中。可以类似地制备已知的启动子，并导入到相同的细胞物质(context)中。然后通过比较相对于已知启动子报道基因的表达量来确定目标启动子的转录活性。细胞物质优选为大豆。

结构性核酸序列

本发明的启动子可以被可操作地连接到结构性核酸序列上,该核酸序列相对于启动子是异源的。结构性核酸序列一般可以是任何核酸序列,其转录水平有望被提高。结构性核酸序列优选编码适合于被结合到人类或动物的食物(diet)中的多肽,或该多肽产生某些其他农业上的重要特性。

合适的结构性核酸序列包括但是不限于编码种子贮存蛋白、脂肪酸途径酶、生育酚生物合成酶、氨基酸生物合成酶、固醇途径酶和淀粉分支酶的核酸序列。

优选的种子储存蛋白包括玉米蛋白(美国专利 4,886,878; 4,885,357; 5,215,912; 5,589,616; 5,508,468; 5,939,599; 5,633,436; 和 5,990,384; 专利申请: WO 90/01869, WO 91/13993, WO 92/14822, WO 93/08682, WO 94/20628, WO 97/28247, WO 98/26064,和 WO 99/40209), 7S 蛋白(美国专利 5,003,045 和 5,576,203), 巴西坚果(brazil nut)蛋白(美国专利 5,850,024), 不含苯丙氨酸的蛋白质(专利申请: WO 96/17064), 白蛋白(专利申请: WO 97/35023), β -伴球蛋白(conglycinin)(专利申请:WO 00/19839),11S(美国专利 6,107,051), α -hordothionin (美国专利 5,885,802 和 5885801), arcelin 种子储存蛋白(美国专利 5,270,200), 凝集素(美国专利 6,110,891),和麦谷蛋白(美国专利 5,990,389 和 5,914,450)。

优选脂肪酸途径酶包括硫酯酶(美国专利 5,512,482; 5,530,186; 5,945,585; 5,639,790; 5,807,893; 5,955,650; 5,955,329; 5,759,829; 5,147,792; 5,304,481; 5,298,421; 5,344,771; 和 5,760,206),和去饱和酶(美国专利 5,689,050; 5,663,068; 5,614,393; 5,856,157; 6,117,677; 6,043,411; 6,194,167; 5,705,391; 5,663,068; 5,552,306; 6,075,183; 6,051,754; 5,689,050; 5,789,220; 5,057,419; 5,654,402; 5,659,645; 6,100,091; 5,760,206; 6,172,106; 5,952,544; 5,866,789; 5,443,974; 和 5,093,249)。优选生育酚生物合成途径酶包括 *tyrA*, *slr1736*, *ATPT2*, *dxs*, *dxr*, *GGPPS*, *HPPD*, *GMT*, *MT1*, *tMT2*, *AANT1*, *slr1737*, 和尿黑酸的双加氧酶的反义构建体(Kridl 等, *Seed Sci. Res.*, 1:209:219(1991); Keegstra,*Cell*,56(2): 247-53(1989); Nawrath,等., *Proc. Natl. Acad. Sci. (U.S.A.)*, 91:12760-12264 (1994); Xia 等 *J. Gen. Microbiol.*, 138:1309-1316 (1992); cyanobase <http://www.kazusa.or.jp/cyanobase>; Lois etal., *Proc. Natl. Acad. Sci. (U.S.A.)*, 95 (5):2105-2110 (1998); Takahashi 等 *Proc. Natl.*

Acad. Sci. (U.S.A.), 95(17), 9879-9884 (1998); Norris 等, *Plant Physiol.*, 117: 1317-1323 (1998); Bartley 和 Scolnik, *Plant Physiol.*, 104: 1469-1470 (1994); Smith 等, *Plant J.*, 11:83-92(1997); WO 00/32757; WO 00/10380; Saint Guily, 等, *Plant Physiol.*, 100(2):1069-1071(1992); Sato 等, *J.DNA Res.*, 7(1):31-63 (2000))。

各种基因及其编码的参与生育酚生物合成的蛋白质被列举在下面的表 1 中。

表 1: 基因及其编码的参与生育酚生物合成的蛋白质

基因 ID	酶的名称
tyrA	预苯酸脱氢酶(Prephenate dehydrogenases)
Slr1736	来自 Synechocystis 的叶绿基异戊二烯基转移酶
ATPT2	来自拟南芥的叶绿基异戊二烯基转移酶
DXS	1-脱氧木酮糖-5-磷酸合成酶
DXR	1-脱氧木酮糖-5-磷酸还原酮异构酶(reductoisomerase)
GGPPS	香叶基香叶基焦磷酸合成酶
HPPD	p-羟基苯基丙酮酸双加氧酶
AANTI	酰苷酰转运蛋白
slr1737	生育酚环化酶
IDI	异戊烯二磷酸异构酶
GGH	香叶基香叶基还原酶
GMT	γ 甲基转移酶

上表给出的“基因 ID”确定与所列酶相关的基因。任何列于表 1 中的出现在本公开中的基因 ID 是指编码在表 1 中与基因 ID 相关的酶的基因。

优选的氨基酸生物合成酶包括邻氨基苯甲酸合成酶(美国专利 5,965,727; 专利申请: WO 97/26366, WO99/11800, 和 WO 99/49058), 色氨酸脱羧酶(专利申请: WO 99/06581), 苏氨酸脱羧酶(美国专利 5,534,421 和 5,942,660, 专利申请: WO 95/19442), 苏氨酸脱氨酶(专利申请: WO 99/02656, 和 WO 98/55601), 和天门冬氨酸激酶(美国专利 5,367,110; 5,858,749; 和 6,040,160)。

优选的淀粉分支酶包括那些列举在美国专利 6,232,122 和 6,147,279; 和专利申请 WO 97/22703 中的酶。

此外, 启动子和结构性核酸序列可以设计成下调特定核酸序列。这一般是通过将启动子连接到结构性核酸序列上来实现, 该序列以反义方向被引导。本领域的普通技术人员熟悉这种反义技术。任何核酸序列可以以这种方式被反面调节。

这种调控的靶点可能包括具有低必需氨基酸含量的多肽, 但是这种多肽

在特定组织中以相对较高水平被表达。例如， β -伴球蛋白和大豆球蛋白在种子被丰富地表达，但是从必需氨基酸角度考虑是缺乏营养的。这种反义方法还可以被用于有效地从植物来源的饲料中去除不期望的蛋白质，如拒食素(如凝集素)、白蛋白和过敏原，或者下调参与降解期望的化合物如必需氨基酸的分解代谢酶。

改进的结构性核酸序列

本发明的启动子还可以被可操作地连接到被改进的结构性核酸序列上，该核酸序列相对于启动子是异源的。这种结构性核酸序列可以被改进来产生各种期望的特性。例如，结构性核酸序列可以被改进来增加必需氨基酸的含量、提高氨基酸序列的翻译、改变翻译后的修饰(如磷酸化位点)、将翻译产物转运到细胞内外的腔室、改善蛋白质稳定性、插入或删除细胞信号基序等。

在优选实施方案中，结构性核酸序列被增强来编码至少一种和更加优选2,3,或4种必需氨基酸含量升高的多肽，这些必需氨基酸选自组氨酸、赖氨酸、甲硫氨酸或苯丙氨酸。如果需要，还可以添加非必需氨基酸，从结构和营养上提高该多肽。特别适合于这种升高的结构性核酸序列包括那些编码天然多肽的序列，这些多肽以相对高的水平被表达，具有特别低的必需氨基酸含量或包含这两种情况。这种多肽的例子是种子贮存蛋白，如大豆球蛋白和 β -伴球蛋白。其他合适的靶包括 arcelin、菜豆蛋白、凝集素、玉米蛋白和白蛋白。

结构性核酸序列中的密码子使用

由于遗传密码的简并性，不同核苷酸密码子可以被用于编码特定氨基酸。宿主细胞通常具有密码子使用的选择模式。结构性核酸序列优选被构建来利用特定宿主细胞的密码子选择模式。这通常在被转化的宿主细胞中增强结构性核酸序列的表达。任何上述核酸和氨基酸序列可以被改进来体现含有它们的宿主细胞或生物体的优选密码子选择。在植物中进行最佳密码子使用的对结构性核酸序列的修饰在美国专利 5,689,052 中被描述。

结构性核酸序列的其他修饰

上述结构性核酸序列中的其他变体可以编码当与用于加工得到它们的蛋白质比较时具有相当或更好特性的蛋白质。突变包括删除、插入、截断、取代、融合、基序序列改组等。

对结构性核酸序列的突变可以以特定或随机的方式被导入，这两种方法

是分子生物学领域的技术人员熟知的。还有许多定点突变技术,一般采用寡核苷酸来在结构性核酸序列的特定位点上导入突变。例子包括单链拯救(Kunkel等,1985),特定位点消除(Deng和Nickloff,1992),缺口保护(Vandeyar等,1988),和PCR(Costa等,1996)。随机或非特异性突变可以化学试剂产生(综述,参见Singer和Kusmierck,1982)如亚硝基胍(Cerda-Olmedo等,1968;Guerola等,1971),和2-氨基嘌呤(Rogan和Bessman,1970);或通过生物学方法如通过增变株传代(Greener等,1997)。用于制备上面描述的改变的其他方法被描述在Ausubel等(1995);Bauer等(1985);Craig(1985);Frits Eckstein等(1982);Sambrook等(1989);Smith等(1981);和Osuna等(1994)。

这种修饰可能在氨基酸序列上产生保守或非保守的变化。保守变化是不改变最终的蛋白质的氨基酸序列的变化。在优选实施方案中,蛋白质具有5个和500个之间的保守变化,更加优选10个和300个之间的保守变化,甚至更加优选25个和150个之间的保守变化和最优选在5个和25个之间的保守变化或在1个和5个之间的保守变化。

非保守变化包括添加、删除和取代,这产生被改变的氨基酸序列。在优选的实施方案中,蛋白质具有5个和500个之间的非保守变化,更加优选10个和300个之间的非保守变化,甚至更加优选25个和150个之间的非保守变化和最优选在5个和25个之间的非保守变化或在1个和5个之间的非保守变化。

可以对本文公开的蛋白质序列和编码它们的核酸序列进行修饰,保留这些分子的期望特性。如下是基于改变蛋白质的氨基酸序列来产生相当的或可能是改进的第二代分子的讨论。氨基酸变化可以根据表2中给出的密码子来改变结构性核酸序列的密码子来实现。

表 2: 氨基酸的密码子简并性

氨基酸	单字母	三字母	密码子
丙氨酸	A	Ala	GCA GCC GCG GCT
半胱氨酸	C	Cys	TGC TGT
天冬氨酸	D	Asp	GAC GAT
谷氨酸	E	Glu	GAA GAG
苯丙氨酸	F	Phe	TTC TTT
甘氨酸	G	Gly	GGA GGC GGG GGT
组氨酸	H	His	CAC CAT
异亮氨酸	I	Ile	ATA ATC ATT
赖氨酸	K	Lys	AAA AAG
亮氨酸	L	Leu	TTA TTG CTA CTC CTG CTT
甲硫氨酸	M	Met	ATG
天冬酰胺	N	Asn	AAC AAT
脯氨酸	P	Pro	CCA CCC CCG CCT
谷氨酰胺	Q	Gln	CAA CAG
精氨酸	R	Arg	AGA AGG CGA CGC CGG CGT
丝氨酸	S	Ser	AGC AGT TCA TCC TCG TCT
苏氨酸	T	Thr	ACA ACC ACG ACT
缬氨酸	V	Val	GTA GTC GTG GTT
色氨酸	W	Trp	TGG
酪氨酸	Y	Tyr	TAC TAT

一些氨基酸可以被用于取代蛋白质序列中的其他氨基酸，而不使期望活性产生可评估的损失。因此，可预期地在肽序列或蛋白质序列或者它们的相应核酸序列中进行各种改变，而不使生物活性产生可检测的损失。

在进行这种改变时，考虑氨基酸的亲水指数。亲水氨基酸指标在赋予一种蛋白质交互式生物学功能中的重要性是本领域认可的(Kyte 和 Doolittle,1982)。一般认为，氨基酸的相对亲水性特性产生生成的蛋白质的二级结构，该二级结构随后限定该蛋白质与其他分子的互动，例如酶、底物、

受体、DNA、抗体、抗原等。

可以根据每种氨基酸的疏水性和电荷特征赋予其亲水指数。即：异亮氨酸(+4.5)；缬氨酸(+4.2)；亮氨酸(+3.8)；苯丙氨酸(+2.8)；半胱氨酸/半胱氨酸(+2.5)；甲硫氨酸(+1.9)；丙氨酸(+1.8)；甘氨酸(-0.4)；苏氨酸(-0.7)；丝氨酸(-0.8)；色氨酸(-0.9)；酪氨酸(-1.3)；脯氨酸(-1.6)；组氨酸(-3.2)；谷氨酸/谷氨酰胺/天冬氨酸/天冬酰胺(-3.5)；赖氨酸(-3.9)；和精氨酸(-4.5)。

本领域已知，一些氨基酸可以被其他具有相似亲水指数或值的氨基酸取代，并且仍然能够生成具有相似生物学活性的蛋白质，即仍然获得具有生物学功能的蛋白质。在进行这种改变中，亲水指数在 ± 2 内的氨基酸取代是优选的，亲水指数在 ± 1 内的氨基酸取代是更加优选的，亲水指数在 ± 0.5 内的氨基酸取代是最优选的。

本领域还理解，类似氨基酸的取代可以根据亲水性有效地进行。美国专利 4,554,101 指出蛋白质的最大局部平均亲水性受其邻近的氨基酸的控制，与蛋白质的生物学性质相关。如下亲水性值被赋予氨基酸：精氨酸/赖氨酸(+3.0)；天冬氨酸/谷氨酸(+3.0 \pm 1)；丝氨酸(+0.3)；天冬酰胺/谷氨酰胺(+0.2)；甘氨酸(0)；苏氨酸(-0.4)；脯氨酸(-0.5 \pm 1)；丙氨酸/组氨酸(-0.5)；半胱氨酸(-1.0)；甲硫氨酸(-1.3)；缬氨酸(-1.5)；亮氨酸/异亮氨酸(-1.8)；酪氨酸(-2.3)；苯丙氨酸(-2.5)和色氨酸(-3.4)。

一般认为，一些氨基酸可以被其他具有相似亲水性值的氨基酸取代，并且仍然能够生成具有相似生物学活性的蛋白质，即仍然获得具有生物学功能的蛋白质。在进行这种改变中，亲水指数在 ± 2 内的氨基酸取代是优选的，亲水指数在 ± 1 内的氨基酸取代是更加优选的，亲水指数在 ± 0.5 内的氨基酸取代是最优选的。

如上简要描述，氨基酸取代是基于氨基酸侧链取代的相对相似性，如其疏水性、亲水性、电荷、大小等。考虑各种前述特性的示例性取代是本领域技术人员已知的，包括：精氨酸和赖氨酸；谷氨酸和天冬氨酸；丝氨酸和苏氨酸；谷氨酰胺和天冬酰胺；和缬氨酸、亮氨酸和异亮氨酸。如果相对于加工获得它们的未改进的多肽，生成的蛋白质具有改进的对瘤胃的抗性，增强对蛋白质水解降解的抵抗，或者同时具有改进的对瘤胃的抗性并增强对蛋白质水解降解的抵抗，被认为是不期望的改变也可以被使用。或者，进行改变以改善代谢酶的动力学。

在一个优选方面,被改进的蛋白质选自种子贮存蛋白质、脂肪酸途径酶、生育酚生物合成酶、氨基酸生物合成酶或淀粉分支酶。

重组载体

任何一种上面描述的启动子和结构性核酸序列可以在重组载体中被提供。重组载体一般包括,从5'到3'方向:引导结构性核酸序列转录的启动子和结构性核酸序列。合适的启动子和结构性核酸序列包括本文描述的那些。如果需要,重组载体还可以包括3'转录终止子、3'多聚腺苷酸化信号、其他非翻译核酸序列、转运和靶向核酸序列、可选择标记、增强子和操作子。

用于制备重组载体的方法是本领域熟知的。用于制备特别适合于植物转化的重组载体的方法被描述在美国专利4,971,908;4,940,835;4,769,061;和4,757,011中。这些类型的载体也已经被评述(Rodriguez等,1988;Glick等,1993)。

用于在高等植物中表达核酸的典型载体是本领域熟知的,包括来自根癌农杆菌的肿瘤诱导(Ti)质粒的载体(Rogers等,1987)。其他可以用于植物转化的重组多肽包括pCaMVCN转化控制载体,也已经被描述(Fromm等,1985)。

在一个实施方案中,多个USP启动子在单构建体中被可操作地连接到任何结构基因的组合上。在优选实施方案中,任何1,2,3,4,5,或6或多种包含SEQ ID NO: 1,2,3,4,9,10,或11的核酸分子的组合在单个构建体中被可操作地连接到任何结构基因的组合上。在优选实施方案的另一个方面,核酸分子被修饰。这种修饰包括例如去除或添加一种或多种结构性或功能性的元件。

重组载体中的其他启动子

还可以在重组载体中提供一种或多种其他启动子。这些启动子可以被可操作地连接到,例如但不限于,任何上述的结构性核酸序列上。可选择地,启动子可以被可操作地连接到其他核酸序列上,如那些编码转运肽、可选择的标记蛋白的序列或反义序列。

这些其他的启动子可以根据将插入载体的细胞的类型来选择。此外,在细菌、酵母和植物中起作用的启动子都在本领域中被充分地教导。其他启动子还可以根据其调控特性来选择。这种特性的例子包括增强转录活性、可诱导性、组织特异性和发育阶段特异性。在植物中,可诱导的、病毒的或合成来源的、组成型活性的、时间调控的和空间调控的启动子都已经被描述(Poszkowski等,1989;Odell等,1985;Chau等,1989)。

常用的组成型启动子包括 CaMV 35S 启动子(Odell 等, 1985)、增强的 CaMV 35S 启动子、玄参(Figwort)花叶病毒(FMV)启动子(Richins 等,1987)、甘露碱合成酶(mas)启动子, 胭脂碱合成酶(nos)启动子和章鱼碱合成酶(ocs)启动子。

有用的可诱导的启动子包括被水杨酸或聚丙烯酸诱导的启动子(PR-1; Williams 等,1992),由应用安全剂诱导的启动子(替代苯磺酰胺除草剂; Hershey 和 Stoner,1991)、热休克启动子(Ou-Lee 等,1986; Ainley 等,1990)、来自菠菜亚硝酸盐还原酶结构性核酸序列的硝酸盐诱导的启动子(Back 等,1991)、激素诱导的启动子(Yamaguchi-Shinozaki 等,1990; Kares 等,1990)、和与 RuBP 羧化酶小亚基和 LHCP 家族相关的光诱导的启动子(Kuhlemeier 等,1989; Feinbaum 等, 1991; Weisshaar 等, 1991; Lam 和 Chua,1990; Castresana 等,1988; Schulze-Lefert 等,1989)。

有用的组织或器官特异性启动子的例子包括 β -伴球蛋白, (Doyle 等, 1986; Slighton 和 Beachy, 1987)和其他种子特异性启动子(Knutzon 等,1992; Bustos 等,1991; Lam 和 Chua, 1991)。用于在种子中优先表达的植物功能性启动子包括来自植物储存蛋白和来自参与油料种子中的脂肪酸生物合成的启动子。这种启动子的例子包括来自这种的结构性核酸序列的 5'调控区, 如 napin(Kridl 等,1991)、菜豆蛋白、玉米蛋白、大豆胰岛素抑制剂、ACP、硬脂酰-ACP 去饱和酶和 oleosin。种子特异性调控还在 EP 0255378 中被讨论。

另一种示例性的种子特异性启动子是凝集素启动子。大豆种子中的凝集素启动子由单一的结构性核酸序列(Lel)编码, 该序列仅在种子发育过程中被表达。凝集素结构性核酸序列和种子特异性启动子已经被表征, 并被用于在转基因烟草植物中引导种子特异性表达(Vodkin 等, 1983; Lindstrom 等,1990)。

重组载体中的其他特别优选的启动子包括胭脂碱合成酶(nos)、甘露碱合成酶(mas)和章鱼碱合成酶(ocs)启动子, 在根癌农杆菌的肿瘤诱导质粒中携带这些启动子; 花椰菜花叶病毒(CaMV)19S 和 35S 启动子; 增强的 CaMV 35S 启动子; 玄参花叶表达病毒(FMV) 35S 启动子; 来自核酮糖-1,5-二磷酸羧化酶的小亚基的光诱导启动子(ssRUBISCO); 烟草 EIF-4A 启动子(Mandel 等,1995); 玉米蔗糖合成酶 1(Yang 和 Russell,1990); 玉米乙醇脱氢酶 1(Vogel 等,1989); 玉米集光(light harvesting)复合物(Simpson,1986); 玉米热休克蛋白(Odell 等,1985); 来自拟南芥的几丁质酶启动子(Samac 等,1991); 来自椰菜的

LTP(脂肪转运蛋白)启动子(Pyee 等,1995); 矮牵牛查耳酮异构酶(Van Tunen 等, 1988); 豆富含甘氨酸的蛋白 1(Keller 等, 1989); 马铃薯 patatin (Wenzler 等, 1989); 来自玉米的泛素启动子(Christensen 等, 1992); 和来自稻的肌动蛋白启动子(McElroy 等,1990)。

其他启动子优选是种子选择性的、组织选择性的、组成型的或可诱导的。启动子最优选为胭脂碱合成酶(nos)、章鱼碱合成酶的(ocs)、甘露氨酸合成酶(mas)的花椰菜花叶病毒 19S 和 35S (CaMV19S,CaMV35S)、增强的 CaMV(eCaMV)、核酮糖 1,5-二磷酸羧化酶(ssRUBISCO)、玄参花叶病毒(FMV)、CaMV 来源的 AS4、烟草 RB7、小麦 POX1、烟草 EIF-4、凝集素蛋白(Lel)或稻 RC2 启动子。

具有其他结构性核酸序列的重组载体

重组载体可能还含有一种或多种其他结构性核酸序列。这些其他结构性核酸序列一般是适合于用在重组载体中的任何序列。这种结构性核酸序列包括但是不限于上述序列。其他结构性核酸序列还可以被可操作地连接到任何上述启动子上。一种或多种结构性核酸序列可以各自地被连接到各个启动子上。可选择地, 结构性核酸序列可以被可操作地连接到单一启动子上(即单一操纵子)。

其他结构性核酸序列包括但是不限于那些编码种子储存蛋白、脂肪酸途径酶、生育酚生物合成酶、氨基酸生物合成酶和淀粉分支酶的核酸。

优选的种子储存蛋白质包括玉米蛋白(美国专利 4,886,878; 4,885,357;5,215,912; 5,589,616; 5,508,468; 5,939,599; 5,633,436;和 5,990,384; 专利申请: WO 90/01869, WO 91/13993, WO 92/14822, WO 93/08682, WO 94/20628, WO 97/28247, WO 98/26064,和 WO 99/40209), 7S 蛋白质(美国专利 5,003,045 和 5,576,203), 巴西坚果蛋白(美国专利 5,850,024), 无苯丙氨酸蛋白(专利申请:WO 96/17064),白蛋白(专利申请: WO 97/35023), β -伴球蛋白(专利申请: WO 00/19839),11S (美国专利 6,107, 051), α -hordothionin(美国专利 5,885,802 和 5,88,5801) arcelin 种子储存蛋白(美国专利 5,270,200)凝集素(美国专利 6,110,891)和麦谷蛋白(美国专利 5,990,389 和 5,914,450)。

优选的脂肪酸途径酶包括硫酯酶(美国专利 5,512,482; 5,530,186; 5,945,585; 5,639,790; 5,807,893; 5,955,650; 5,955,329; 5,759,829; 5,147,792; 5,304,481; 5,298,421; 5,344,771; 和 5,760,206) 和去饱和酶(美国专利

5,689,050;5,663,068; 5,614,393; 5,856,157; 6,117,677; 6,043,411; 6,194,167; 5,705,391; 5,663,068; 5,552,306; 6,075,183; 6,051,754; 5,689,050; 5,789,220; 5,057,419; 5,654,402; 5,659,645; 6,100,091; 5,760,206; 6,172,106; 5,952,544; 5,866,789; 5,443,974;和 5,093,249)。

优选的生育酚生物合成酶包括 *tyrA*,*slr1736*, *ATPT2*, *dxs*, *dwr*, *GGPPS*, *HPPD*, *GMT*, *MT1*, *tMT2*, *AANT1*, *slr1737* 和尿黑酸双加氧酶的反义构建体 (Kridl 等, *Seed Sci. Res.*, 1:209:219 (1991); Keegstra, *Cell*,56(2):247-53 (1989); Nawrath,等, *Proc. Natl. Acad. Sci.(U.S.A.)*, 91:12760-12764(1994); Xia 等, *J. Gen. Microbiol.*, 138:1309-1316 (1992); Cyanobase <http://www.kazusa.or.jp/cyanobase> ; Lois 等, *Proc. Natl. Acad. Sci. (U.S.A.)*, 95(5): 2105-2110 (1998); Takahashi 等 *Proc. Natl. Acad. Sci. (U.S.A.)*, 95 (17): 9879-9884 (1998); Norris 等, *Plant Physiol.*, 117:1317-1323 (1998); Bartley 和 Scolnik, *Plant Physiol*, 104:1469-1470 (1994); Smith 等, *Plant J.*, 11:83-92(1997); WO 00/32757; WO 00/10380; Saint Guily 等, *Plant Physiol.*, 100(2):1069-1071(1992); Sato 等., *J. DNA Res.*,7(1): 31-63(2000)。

优选的氨基酸生物合成酶包括邻氨基苯甲酸合成酶(美国专利 5,965,727, 专利申请: WO 97/26366,WO 99/11800 和 WO 99/49058),色氨酸脱羧酶(专利申请: WO 99/06581),苏氨酸脱羧酶(美国专利 5,534,421 和 5,942,660;专利申请: WO 95/19442),苏氨酸脱氨酶(专利申请 WO 99/02656 和 WO 98/55601), 和天冬氨酸激酶(美国专利 5,367,110;5,858,749 和 6,040,160)。

优选的淀粉分支酶包括在美国专利 6,232,122 和 6,147,279 以及专利申请 WO 97/22703 中列举的那些。

可选择地, 第二条结构性核酸序列可能被设计来下调特定核酸序列。这一般通过将该第二条结构性氨基酸以反义方向可操作地与启动子连接来实现。本领域技术人员熟悉这种反义技术。任何核酸序列以这种方式被负调控。优选的靶核酸序列含有低含量的必需氨基酸, 但是在特定组织中以较高水平被表达。例如, β -伴球蛋白和大豆球蛋白在种子中被丰富地表达, 但是从必需氨基酸来考虑是缺乏营养的。这种反义方法也可以被用于有效地从植物来源的 foodstuff 去除其他不期望的蛋白质, 如拒食素(如凝集素)、白蛋白和过敏原, 或者下调参与降解期望的化合物如必需氨基酸的分解代谢酶。

可选择的标记

载体或构建体还可以包括可选择的标记。可选择的标记还可以被用于选择含有外源遗传物质的植物或植物细胞。这种例子包括但不限于: neo 基因(Potrykus 等,1985),其编码卡那霉素抗性,可以用卡那霉素、RptII、G418、hpt 等选择; bar 基因,其编码 bialaphos 抗性;突变 EPSP 合成酶基因(Hinchee 等,1988; Reynaerts 等,1988); aadA (Jones 等,1987),其编码草甘膦抗性;腈水解酶基因,其赋予对 bromoxynil 的抗性(Stalker 等,1988);突变的乙酰乙酸合成酶基因(ALS),其赋予咪唑酮和磺脲抗性(欧洲专利申请 154,204 (1985)), ALS(D'Halluin 等,1992),和甲氧蝶呤抗性 DHFR 基因(Thillet 等,1988)。可选择的标记优选是 GUS、绿色荧光蛋白(GFP)、新霉素磷酸转移酶 II (*npt II*),荧光素酶(LUX),抗生素抗性编码序列或除草剂(例如,草甘膦)抗性编码序列。可选择的标记最优选为卡那霉素、潮霉素或除草剂抗性标记。

载体或构建体还可以包括筛选标记。筛选标记可以被用以监控表达。示例性的筛选标记包括: β -葡萄糖苷酸酶或 *uidA* 基因(GUS),其编码其各种发光底物已知的酶(Jefferson, 1987); R-基因座基因,其编码在植物组织中调节花色苷色素(红色)生成的产物(Dellaporta 等,1988); β -内酰胺酶基因(Sutcliffe 等,1978),编码一种各种发光底物已知的酶的基因(例如 PADAC,发光头孢菌素(cephalosporin)); 荧光素酶基因(Ow 等,1986); *xyIE* 基因(Zukowsky 等,1983),其编码能够转化发光的儿茶酚的儿茶酚双加氧酶; α -淀粉酶基因(Ikatu 等,1990); 酪氨酸酶基因(Katz 等,1983),其编码能够将酪氨酸氧化为 DOPA 和多巴醌的酶,随后多巴醌浓缩为黑色素; α -半乳糖苷酶,其转化发光的 α 半乳糖底物。

术语或短语“可选择的或筛选的标记基因”中包括的还有编码可选择标记的基因,这些可选择标记的分泌可以被检测来作为鉴定或选择被转化的细胞的手段。例子包括编码可以被抗体作用确定的可分泌抗原标记,或者甚至是可以酶促地检测的可分泌酶。可分泌蛋白分为许多种类,包括可检测的小的扩散性蛋白(如通过 ELISA),在细胞外溶液中可检测的小的活性酶(如 α -淀粉酶, β -内酰胺酶, 磷丝菌素转移酶)或者被插入或捕获在细胞壁上的蛋白质(例如包括先导序列的蛋白质如存在于延展的表达单位中的蛋白或烟草 PR-S)。其他可能的可选择的和/或可筛选的标记基因对于本领域技术人员来说是显而易见的。

重组载体中的其他元件

各种顺式作用非翻译的 5'和 3'调控序列可以被包括在重组核酸载体中。任何一种这样调控序列可以与其他调控序列一起被提供在重组载体中。这种组合可以被设计和改进来产生期望的调控特性。

3'非翻译区通常具有转录终止信号和聚腺苷酸化信号，聚腺苷酸化信号在植物中起作用来将腺苷酸添加到 mRNA 的 3'端。这些元件可以从胭脂碱合成酶(nos)编码序列、大豆 7S α '储存蛋白编码序列、arcelin-5 编码序列、白蛋白编码序列和豌豆 ssRUBISCO E9 编码序列的 3'区获得。一般地，位于多聚腺苷酸化位点下游的数百个碱基对的核酸序列起终止转录的作用。这些区域是被转录 mRNA 有效多聚腺苷酸化所必需的。

转录增强子也可以被结合作为重组载体的一部分。因此，重组载体优选地含有一个或多个 5'非翻译前导序列，其起增强核酸序列表达的作用。这种增强子序列可能有望增加或改变生成的 mRNA 的转录效率。优选的 5'核酸序列包括 dSSU 5'、PetHSP70 5'和 GmHSP17.9 5'(美国专利 5,362,865)。

重组载体可能还包含编码转运肽的核酸序列。这种肽可以被用于引导蛋白质到细胞外区域、质体或到细胞内外的一些其他腔室(参见，如 EP 0218571、美国专利 4,940,835; 5,88,624; 5,610,041; 5,618,988 和 6,107,060)。

重组载体中的结构核酸序列可能包含内含子。该内含子对于结合性核酸序列来说可能是异源的。优选的内含子包括水稻肌动蛋白内含子和玉米 HSP70 内含子。

融合蛋白

任何上述的结构性核酸序列及其改进形式可以与其他核酸序列连接来编码融合蛋白。其他核酸序列优选编码至少 1 种氨基酸、肽或蛋白质。存在许多可能的融合组合。

例如，融合蛋白可能提供一种“标记”的表位来方便融合蛋白的检测，例如，GST,GFP,FLAG,或多聚 HIS。这种融合优选编码 1 - 50 个氨基酸，更加优选 5 - 30 个其他的氨基酸，甚至最优选在 5 - 20 个氨基酸。

可选择地，融合可能具有调控、酶促、细胞信号或细胞内转运的功能。例如，编码质粒转运肽的序列可以被添加来引导融合蛋白到种子中的叶绿体中。这种融合配偶体优选编码 1 - 1000 个其他氨基酸，更加优选 5 - 500 个其他氨基酸，和甚至更加优选 10 - 250 个氨基酸。

序列分析

在本发明中，序列相似性或同一性优选采用序列分析软件包™的“Best Fit”或“Gap”程序来确定(版本 10; Genetics Computer Group, Inc., University of Wisconsin Biotechnology Center, Madison, WI)。“Gap”采用 Needleman 和 Wunsch(1970)的算法来寻找两条序列的排列，这种排列使匹配数量最大化和缺口数量最小化。“BestFit”进行两条序列之间的相似性的最好片段的最佳排列。最佳排列采用 Smith 和 Waterman (Smith and Waterman, 1981 ; Smith 等, 1983)的局部同源性算法通过插入缺口来最小化匹配数量来发现。

上面描述的序列分析软件包含有大量其他有用的序列分析工具来确定本公开的核苷酸和氨基酸序列的同系物。例如，“BLAST”程序在特定数据库(例如由 Bethesda, MD 的国家生物信息中心(NCBI)维护的序列数据库)中检索与被检索的序列(肽或核酸)相似的序列；“FastA” (Lipman 和 Pearson, 1985; 参见还有 Pearson 和 Lipman, 1988; Pearson, 1990)进行 Pearson 和 Lipman 检索来确定被检索序列与一组相同类型(核酸或蛋白质)的序列之间的相似性。“TfastA” 进行 Pearson 和 Lipman 检索来确定一种蛋白质的被查询序列与任何一组核苷酸序列(在进行比较之前，以所有 6 种阅读框翻译核苷酸序列)之间的相似性；“FastX”进行 Pearson 和 Lipman 检索来确定被检索的核酸序列与一组蛋白质序列之间的相似性，考虑移码。“TfastX”进行 Pearson 和 Lipman 检索来确定被检索的蛋白质序列与任何一种核苷酸序列之间的相似性，考虑移码(在进行比较之前，翻译核酸序列的两条链)。

探针和引物

能够特异性地杂交到互补性核酸序列上的短核酸序列可以被制备，并被用于本发明。这种短的核酸分子可以被用作探针来鉴定在特定样本中是否存在互补的核酸序列。因此，通过构建与特定核酸序列的小部分互补的核酸探针，核酸序列的存在可以被检测和评价。

可选择地，短的核酸序列可以被用作寡核苷酸引物，采用 PCR 技术来扩增或突变互补的核酸序列。这些引物可能还方便扩增相关互补的核酸序列(例如，来自其他物种的相关核酸序列)。

短的核酸序列可以被用作引物和特异地作为 PCR 引物。PCR 探针是能够在双链结构中与其他核酸一起起始聚合酶活性的核酸分子。本领域存在各种用于确定 PCR 引物结构的方法和 PCR 技术。采用程序如 Primer3 问(www.genome.wi.mit.edu/cgi-bin/primer/primer3.cgi), STSPipeline (www.genome.wi).

mit.edu/cgi-bin/www.STSPipeline)或者 GeneUp(Pesole 等, 1998)在计算机中进行的检索可以被用于确定潜在的 PCR 引物。

任何本文公开的核酸序列可以被用作引物或探针。这些探针或引物的使用极大地方便确定转基因植物, 该转基因植物含有本发明公开的启动子和结构性核酸序列。这种探针或引物还可以被用于筛选与本申请公开的启动子和结构性核酸序列相关或具有同源性的其他核酸序列的 cDNA 或基因组文库。

引物或探针通常与被鉴定、扩增或突变的核酸序列的一部分互补, 该序列足够长以与其互补体形成稳定的和序列特异性的双链核酸分子。引物或探针优选为长约 10 到约 200 个核苷酸, 更加优选为长约 10 到约 100 个核苷酸, 甚至更加优选为长约 10 到约 50 个核苷酸, 和最优选为长约 14 到约 30 个核苷酸。

引物或探针例如但是不限于通过直接化学合成、通过 PCR(美国专利 4,683,195 和 4,683,202)或通过从较大的核酸分子上切除核酸特异性片段来制备。

转基因植物和转化的植物宿主细胞

本发明还涉及转基因植物和转化的宿主细胞, 其包含被可操作地连接到异源的结构性核酸序列的启动子。其他核酸序列也可能与启动子和结构性核酸序列一起被导入到植物或宿主细胞中。这些其他序列可能包括 3'转录终止子、3'多聚腺苷酸信号、其他的非翻译核酸序列、转运或靶向序列、可选择标记、增强子和操纵子。上面描述了本发明的优选核酸序列, 包括重组载体、结构性核酸序列、启动子和其他调控元件。

在优选实施方案中, 转基因植物和被转化的宿主细胞包含来自蚕豆的启动子。在更加优选的实施方案中, 转基因植物和被转化的宿主细胞包含任何本文描述的本发明的核酸分子, 包括选自 SEQ ID NO:1,2,3,4,9,10,或 11 和其互补体的核酸序列。

在特别优选的实施方案中, 本发明的转基因植物是大豆植物。在优选实施方案中, 本发明的大豆植物包括一种或多种被导入的本发明的核酸分子。在优选实施方案中, 本发明的被转化大豆植物包含选自 SEQ ID NO:1,2,3,4,5,9,10,或 11 的核酸分子。在优选实施方案中, 本发明的转化大豆植物包含由 SEQ ID NO: 4 组成的核酸分子。在优选实施方案中, 本发明的转化大豆植物包含由 SEQ ID NO: 5 组成的核酸分子。

在本发明的一些实施方案中，一种植物、细胞或生物体的一种或多种成分与具有“相似遗传背景”的植物、细胞或生物体比较。在优选方面，“相似遗传背景”是被比较的生物体共同具有 50% 或更多的其核酸遗传物质的背景。在更加优选方面，相似遗传背景是被比较的生物体共同具有 75% 或更多，甚至更加优选约 90% 或更多的其核酸遗传物质的背景。在另一个甚至更加优选方面，相似的遗传背景是被比较的生物体是植物背景，除了采用植物转化技术导入最先导入的任何遗传物质外，该植物是同基因的。

用于制备这种重组载体的方法在本领域是熟知的。例如，用于制备突变特别适合于植物转化的重组载体的方法被描述在美国专利 4,971,908; 4,940,835; 4,769,061; 和 4,757,011 中。这些载体也被评述(Rodriguez 等, 1988; Glick 等, 1993)。

用于在细胞和高等植物中表达核酸的典型载体在本领域是熟知的，包括来自根癌农杆菌的肿瘤诱导(Ti)质粒的载体(Rogers 等, 1987)。可用于植物转化的其他重组载体也已经被描述(Fromm 等, 1985)。这种重组载体的元件包括但是不限于上面讨论的那些。

转化的宿主细胞通常是任何与本发明相容的细胞。转化的宿主植物或细胞可以是或来源于单子叶植物或双子叶植物，包括但是不限于，canola、海甘蓝(crambe)、玉米、芥菜、蓖麻子、芝麻、棉籽、亚麻子、大豆、Arabidopsis phaseolus、花生、紫花苜蓿、小麦、稻、燕麦、高粱、油菜籽、黑麦、tritordeum、粟类(millet)、酥油草(fescue)、多年生黑麦草(perennial ryegrass)、甘蔗、酸果(cranberry)、番木瓜(papaya)、香蕉、红花、油椰子(oil palms)、亚麻(flax)、香瓜(muskmelon)、苹果、黄瓜、石斛兰(dendrobium)、唐菖蒲(gladiolus)、菊花(chrysanthemum)、百合(liliacea)、棉花、桉树(eucalyptus)、向日葵、芸苔(Brassica campestris)、欧洲油菜(Brassica napus)、草坪草(turfgrass)、甜菜(sugarbeet)、咖啡树(coffee)和薯蓣属(dioscorea)(Christou, Particle Bombardment for Genetic Engineering of Plants, Biotechnology Intelligence Unit, 学院出版社, 圣地亚哥, 加利福尼亚(1996)), canola、玉米、芸苔、欧洲油菜、油菜籽、大豆、红花、小麦，水稻和向日葵是优选的，而 canola、油菜籽、玉米、芸苔，欧洲油菜、大豆、向日葵、红花、油椰子和花生是更加优选的。在特别优选的实施方案中，植物或细胞是或者来源于 canola。在另一个特别优选实施方案中，植物或细胞是或来源于欧洲油菜。在另一个特别优选的实施方案中，植物或细胞

是或者来源于大豆。

大豆细胞或植物优选是原种(elite)大豆品系。“原种品系”是任何从繁育和从优良的农学操作选择中获得的品系。原种品系的例子是农场主和大豆育种工作者能够购买到的品系，如 HARTZ™ 品种 H4994, HARTZ™ 品种 H5218, HARTZ™ 品种 H5350, HARTZ™ 品种 H5545, HARTZ™ 品种 H5050, HARTZ™ 品种 H5454, HARTZ™ 品种 H5233, HARTZ™ 品种 H5488, HARTZ™ 品种 HLA572, HARTZ™ 品种 H6200, HARTZ™ 品种 H6104, HARTZ™ 品种 H6255, HARTZ™ 品种 H6586, HARTZ™ 品种 H6191, HARTZ™ 品种 H7440, HARTZ™ 品种 H4452 Roundup Ready™, HARTZ™ 品种 H4994 Roundup Ready™, HARTZ™ 品种 H4988 Roundup Ready™, HARTZ™ 品种 H5000 Roundup Ready™, HARTZ™ 品种 H5147 Roundup Ready™, HARTZ™ 品种 H5247 Roundup Ready™, HARTZ™ 品种 H5350 RoundupReady™, HARTZ™ 品种 H5545 Roundup Ready™, HARTZ™ 品种 H5855 Roundup Ready™, HARTZ™ 品种 H5088 RoundupReady™, HARTZ™ 品种 H5164 RoundupReady™, HARTZ™ 品种 H5361 RoundupReady™, HARTZ™ 品种 H5566 RoundupReady™, HARTZ™ 品种 H5181RoundupReady™, HARTZ™ 品种 H5889RoundupReady™, HARTZ™ 品种 H5999 Roundup Ready™, HARTZ™ 品种 H6013 Roundup Ready™, HARTZ™ 品种 H6255 Roundup Ready™, HARTZ™ 品种 H6454 Roundup Ready™, HARTZ™ 品种 H6686 Roundup Ready™, HARTZ™ 品种 H7152 Roundup Ready™, HARTZ™ 品种 H7550 Roundup Ready™, HARTZ™ 品种 H8001 Roundup Ready™(HARTZ SEED, Stuttgart, AR); A0868, AG0901, A1553, A1900, AG1901, A1923, A2069, AG2101, AG2201, A2247, AG2301, A2304, A2396, AG2401, AG2501, A2506, A2553, AG2701, AG2702, A2704, A2833, A2869, AG2901, AG2902, AG3001, AG3002, A3204, A3237, A3244, AG3301, AG3302, A3404, A3469, AG3502, A3559, AG3601, AG3701, AG3704, AG3750, A3834, AG3901, A3904, A4045 AG4301, A4341, AG4401, AG4501, AG4601, AG4602, A4604, AG4702, AG4901, A4922, AG5401, A5547, AG5602, A5704, AG5801, AG5901, A5944, A5959, AG6101, QR4459, and QP4544 (Asgrow Seeds, Des Moines, IA); DeKalb 品种 CX445 (DeKalb, IL)。

本发明还涉及一种制备转化植物的方法，该植物从 5'到 3'方向包括被可

操作地连接到异源结构性核酸序列的启动子。其他序列也可以与启动子和结构性核酸来一起被导入到植物中。这些其他可能包括 3'转录终止子、3'多聚腺苷酸化信号、其他非翻译序列、转运或靶向序列、可选择标记、增强子和操作子。优选的重组载体、结构性核酸序列、启动子和其他调控序列包括但不限于本文描述的那些。

这种方法通常包括选择合适的植物、用重组载体转化植物，获得转化的宿主细胞的步骤。

有许多用于将核酸导入植物的方法。合适方法包括细菌感染(如农杆菌)、双元人工染色体载体、直接呈递核酸(如通过 PEG 介导的转化、干燥/抑制介导的核酸摄取、电穿孔、用碳化硅纤维激发，和加速核酸包被的粒子等(在 Potrykus 等, 1991 中被评述))。

用于将核酸导入细胞中的技术是本领域技术人员熟知的。一般可以将方法分为四类：(1) 化学方法(Graham 和 van der Eb, 1973; Zatloukal 等, 1992); (2) 物理方法如微注射(Capecci, 1980)、电穿孔(Wong 和 Neumann, 1982; Fromm 等, 1985; 美国专利 5,384,253)和粒子加速(Johnston 和 Tang, 1994; Fynan 等, 1993); (3) 病毒载体(Clapp, 1993; Lu 等, 1993; Eglitis 和 Anderson, 1988); 和(4)受体介导的机制(Curiel 等, 1992; Wagner 等, 1992)。可选择地，可以通过直接注射植物生殖器官将核酸直接导入到花粉中(Zhou 等, 1983; Hess, 1987; Luo 等, 1988; Pena 等, 1987)。在另一个方面，核酸还可以被注射到未成熟的胚中(Neuhaus 等, 1987)。

本领域已经教导了从被转化的植物原生质体或外植体再生、发育和培养植物 (Weissbach 和 Weissbach, 1988; Horsch 等, 1985)。转化株通常在存在选择性培养液时被培养，该培养液选择被成功转化的细胞，并且包括再生植物幼苗(Fraley 等, 1983)。这种幼苗通常在 2-4 个月之内获得。

然后幼苗被转移到合适的诱导根的培养液中，该培养液含有选择试剂和防止细菌生长的抗生素。许多幼苗将会生长出根来。然后这些幼苗被转移到土壤或者其他介质中，使根继续发育。所概述的方法通常根据所选用的特定植物而改变。

优选地，再生的转基因植物是自花授粉的，产生纯合的转基因植物。可选择地，从再生转基因植物中获得的花粉可以与非转基因植物杂交，优选与农业上重要的品种的近交系杂交。相反，来自非转基因植物的花粉可以被用

于对再生的转基因植物授粉。

转基因植物可以将编码增强基因表达的核酸序列传递给其后代。转基因植物优选是编码增强基因表达的核酸纯合的，并且通过有性繁殖将该序列传递给所有其后代。后代可能从由转基因植物生成的种子中生长得到。然后，这些额外的植物被自花授粉来产生该植物的真正繁殖品系。

评价来自这些植物的后代的基因表达。基因表达可以通过许多常规方法如 western 印迹、northern 印迹、免疫沉淀和 ELISA 来检测。

本发明的植物或试剂可以被用于方法中，例如但是不限于，来获得在其中表达结合核酸分子的种子，获得结构基因产生升高的种子，获得结构基因产生升高的食物，获得结构基因产生升高的原种和获得结构基因产生升高的油。

被用于这种方法中的植物可以被加工。植物或植物部分可以被与其他植物部分分开或分离。用于该目的的优选植物部分是种子。一般认为，甚至在与其它植物部分分开或分离后，被分离或分开的植物部分可能被其他植物部分污染。在优选方面，被分开的植物部分超过被分开的物质的约 50%(w/w)，更加优选地，超过被分开的物质的约 75%(w/w)，和甚至更加优选超过被分开的物质的约 90%(w/w)。通过这种方法生成的本发明的植物或植物部分可以用已知技术加工成产品。优选的产品是食物、原种(feedstock)和油。

饲料、食物、蛋白质和油的制备物

任何本发明的植物或其部分可以被加工来制备饲料、食物、蛋白质或油制备物。用于该目的的特别优选的植物部分是种子。在优选实施方案中，饲料、食物、蛋白质或油制备物是给反刍动物设计的。制备饲料、食物、蛋白质或油制备物的方法在本领域是熟知的。参见，如美国专利 4,957,748; 5,100,679; 5,219,596; 5,936,069; 6,005,076; 6,146,669;和 6,156,227。在优选实施方案中，蛋白质制备物是高蛋白制备物。这种高蛋白制备物优选具有超过约 5% w/v 的蛋白质含量，更加优选超过约 10% w/v，和甚至更加优选超过约 15% w/v。在优选的油制备物中，油制备物是高油含量的制备物，具有来自本发明的植物或其部分的油成分，超过约 5% w/v，更加优选超过约 10% w/v，和甚至更加优选超过约 15% w/v。在一个优选实施方案中，油制备物是液体，体积超过 1,5,10,或 50L。本发明提供从本发明的植物制备或通过本发明的方法生成的油。这种油可以是任何生成的产物的次要或主要成分。而且，

这种油可以与其他油混合。在更加优选的实施方案中，从本发明的植物制备或由本发明的方法生成的油在体积或重量上组成任何产物的油成分的约0.5%，1%，5%，10%，25%，50%，75%，或约90%以上。在另一个优选实施方案中，油制备物被混合，在体积上占该混合物的约10%，25%，35%，50%，或约75%以上。从本发明的植物中制备的油可以与一种或多种有机溶剂或石油馏出物混合。

在另一个实施方案中，本发明的食物与其他食物混合。在优选实施方案中，从本发明的植物中制备的或通过本发明方法生成的食物在体积或重量上占任何产物的食物成分的约0.5%，1%，5%，10%，25%，50%，75%，或约90%以上。在另一个实施方案中，食物制备物可以被混合，并且在体积上占该混合物的约10%，25%，35%，50%，或约75%以上。

种子容器

植物的种子被放置在容器中。如在此所用，容器是任何能够装载这种种子的物体。容器优选含有超过约500，1000，5000，或约25000粒种子，其中至少约10%，25%，50%，75%，或约100%的种子是来自本发明的植物。

育种程序

本发明的植物可以是育种程序的一部分或者来自育种程序。育种方法的选择取决于植物繁殖模式、被改进的特性的遗传性和商业上所用的栽培种的类型(如F₁杂合栽培种、纯系栽培种等)。被选择的非限制性的用繁殖本发明的植物的方法被列举在下文中。育种程序可以采用标记辅助性选择任何杂交后代来加强。还应当理解，任何商业性的和非商业性的栽培种可以被用于育种程序。因素如，emergence vigor、营养势(vegetative vigor)、应激耐受、抗病性、分枝、开花、结果、种子大小、种子密度、持久性和脱粒性等通常会决定该选择。

对于高度遗传的特性，选择在单个区域中被评价的优良个体植物是有效的，而对于低遗传力的特性，应当在从重复评价相关植物家族获得的平均值的基础上进行选择。常用的选择方法一般包括家谱选择、改进的家谱选择、群体选择和轮回选择。在优选实施方案中，进行回交或轮回育种程序。

遗传的复杂性影响育种方法的选择。回交育种可以被用于将一种或少数高度遗传的特性优良基因转移给期望的栽培种。这种方法已经被广泛地用于繁育抗病的栽培种。各种轮回选择技术被用于改进由大量基因控制的数量遗

传特性。轮回选择在自花授粉的农作物中的用途取决于授粉的难易程度、每次自花授粉的成功杂交的频率和每次成功杂交的杂交后代的数量。

繁育品系可以被检测，在代表商业目标地区的环境下与合适的标准物比较两代或多代。最好的品系是新的商业栽培中的候选；那些仍然缺乏特性的品系可以被用作亲本来制备作为进一步选择的新群体。

一种确定较好的植物的方法是观察其相对于其他试验植物和相对于广泛种植的标准栽培种的性能。如果一次观察不是决定性的，重复观察可以提供对其遗传价值的较好评价。育种人员能够选择和杂交两种或多种亲本品系，接着通过重复自花授粉和选择，产生许多新的遗传组合。

开发新的栽培种需要开发和选择品种，杂交这些品种，和选择较好的杂合杂交。杂合的种子可以通过手工杂交被选择的雄性繁育亲本或者通过采用雄性不育体系来制备。根据某些单基因特性来选择杂种，如荚颜色、花的颜色、种子产量、软毛(pubescence)颜色或除草剂抗性，该特性表明种子是真正杂种。关于亲本品系的其他数据，以及杂种的表型，影响育种人员的决定是否继续特定的杂合杂交。

家谱育种和轮回选择育种方法可以被用于从繁育群体中开发栽培种。育种程序将来自两种或多种栽培种或各种广泛来源的期望特性结合到育种库(breeding pool)中，通过自花授粉和选择期望表型从育种库中开发栽培种。新的栽培种可以被评价来确定哪种具有商业潜能。

家谱育种通常被用来改进自花授粉的农作物。两个具有期望、互补特性的亲本进行杂交来制备 F_1 。通过自花授粉一株或多株 F_1 来制备 F_2 群体。从最好的家族中选择最佳个体。重复检测家族可以在 F_4 代开始进行来提高低遗传力的特性的选择效率。在近交的高级阶段(即 F_6 和 F_7)，最好的品系或表型相似品系被检测作为新的栽培种的潜在释放。

回交育种已经被用于将简单遗传、高遗传特性的基因转移到期望的纯合栽培种或作为轮回亲本的近交品系中。被转移的特性的来源被称为供体亲本。生成的植物被期望具有轮回亲本(如栽培种)的特征和从供体亲本转移来的期望特性。在开始杂交后，具有供体亲本表型的个体被选择和重复与轮回亲本杂交(回交)。生成的亲本被期望具有轮回亲本(如栽培种)的特征和从供体亲本转移来的期望特性。

在严格意义上，单一种子传代程序是指种植一种分离群体，每棵植物收

获一粒种子样本，并且使用该种子样本来种植下一代。当该群体已经从 F_2 发展到期望水平的近交程度时，产生品系的植物各自追溯到不同的 F_2 个体。由于一些种子未能够发芽或者一些植物不能生产至少一粒种子，群体中植物的数量逐代减少。结果，当世代进展完成时，并非所有最初被取样在群体中的 F_2 植物都被后代表现。

在多种子程序中，育种人员通常从群体的每棵植物中收获一个或多个荚，并将它们在一起脱粒来形成批量。批量的一部分被用来种植下一代，一部分被保存。该程序已经被称作为改进的单一种子传代或荚-批量技术。

多种子程序已经被用于在收获时节约劳力。用机器脱粒荚显著快于单一种子程序中通过手工从每个荚中取出种子。多种子程序还使得可能在每代近交中种植相同数量种子群体。

对于通常用于不同特性和农作物的其他育种方法的描述可以在一些参考书籍之一找到(如 Fehr, Principles of Cultivar Development, 1, 2-3(1987))。

本发明的转基因植物还可以采用孤雌生殖方法繁殖。孤雌生殖是植物中的繁殖的遗传控制方法，其中胚乳不是通过卵子和精子的结合形成。有三种基本的孤雌生殖繁殖类型：1)其中胚在从核来源的胚囊中由染色体不减数的卵子发育得到的孤雌生殖，2)其中胚在从大孢子母细胞来源的胚囊中由染色体不减数的卵子发育得到的二倍数孢子形成(diplospory)，和3)胚直接从体细胞发育得到的不定胚生殖。在大多数孤雌生殖形式中，假受精或极体核受精来生成胚乳对于种子发育能力是必需的。在无孢子生殖中，保育栽培种可以被用作种子中胚乳形成的花粉来源。由于栽培种的不减数卵子孤雌发育，保育栽培种不影响无孢子生殖的单性生殖的栽培种的遗传学，而是可能生成胚乳。孤雌生殖在经济上是重要的，特别是在转基因植物中，因为无论多大的杂合程度，其使任何的基因型进行纯育。因此，通过孤雌生殖繁殖，杂合的转基因植物能够在整个重复的生活周期中保持其遗传的忠实性。用于制备孤雌生殖的植物的方法在本领域是已知的。参见，美国专利 5,811,636。

其他生物体

本发明的核酸可以被导入任何细胞或生物体中，如哺乳动物细胞、哺乳动物、鱼细胞、鱼、鸟细胞、鸟、藻类细胞、藻类、真菌细胞、真菌或细菌细胞。优选的宿主和转化体包括：真菌细胞如曲霉菌(Aspergillus)、酵母、哺乳动物、特别是牛和猪、昆虫、细菌和藻类。优选的细菌是大肠杆菌和根癌

农杆菌。

转化这种细胞或生物体的方法在本领域是已知的(EP0238023;Yelton等,1984; Malardier 等,1989;Becker 和 Guarente;Ito 等,1983;Hinnen 等,1978;以及 Bennett 和 LaSure,1991)。制备来自这种生物体的蛋白质的方法也是已知的(Kudla 等, 1990; Jarai 和 Buxton, 1994; Verdier, 1990; MacKenzie 等, 1993; Hartl 等, 1994; Bergeron 等, 1994; Demolder 等, 1994; Craig, 1993; Gething 和 Sambrook, 1992; Puig 和 Gilbert, 1994; Wang 和 Tsou,1993; Robinson 等, 1994; Enderlin 和 Ogrydziak, 1994; Fuller 等, 1989; Julius 等, 1984; 以及 Julius 等, 1983)。

实施例

提供如下实施例, 并且这些实施例无论如何不被解释为限制本发明的保护范围。

实施例 1-生成来自蚕豆的 USP 启动子的克隆

USP 启动子通过 PCR 扩增(Expand High Fidelity PRC System,产品目录号 1732641, Roche Molecular Biochemicals, Indianapolis,IN)从蚕豆基因组 DNA 中获得, 使用按照公开的序列(GenBank Accession X56240)设计的引物。用于扩增 USP 启动子的引物为: 5'-AAACTGCAGCAAATTTACACATTG-3'(SEQ ID NO: 6);和 5'-AAACCATGGTTGACTGGCTATG-3'(SEQ ID NO:7)。然后, 分离的扩增产物被亚克隆到载体 pMON13773 中(附图 1), 生成克隆 pMON58101(USP99)(附图 2)、pMON58102(USP91)(附图 3), 和 pMON58106(USP88)(附图 4)。

实施例 2-生成嵌合的 eUSP88 启动子

采用 pMON58106 作为模板进行 PCR 反应(Expand High Fidelity PRC System, 产品目录号 1732641, Roche Molecular Biochemicals,Indianapolis,IN)。如下引物被用于扩增: 5'-AAACTGCAGCAAATTTACACATTG-3'(SEQ ID NO:6); 和 5'-AAACTGCAGGACTACATGCATAAC-3'(SEQ ID NO:8)。被扩增的启动子片段用 Pst I 限制性酶消化, 并连接到 pMON58106 DNA 上, pMON58106 DNA 通过 Pst I 消化和 CIP 碱性磷酸酶处理来线性化。选择具有期望方向的质粒, 称作为 pMON58110(eUSP88) (附图 5)。

实施例 3-采用大豆子叶瞬时转化来评价 USP 启动子

在开花后 25-28d 收获来自大豆植物的种子(Asgrow A3244), 并且在 GAMBORG 培养基中(G5893, Sigma Company,St.Louis,MO), 在 25°C 下在黑暗中渗透处理过夜, GAMBORG 培养液中添加了 50 mM 谷氨酰胺、111 mM 麦芽糖、125mM 棉子糖、125 mM 甘露糖醇和 3g/l 纯化的琼脂,pH 5.6。得到的 125 片子叶对半剪切, 并采用基因枪技术用纯化的 pMON13773 (7Sα'), pMON58101(USP99), pMON58102(USP91), pMON58106 (USP88) 和 pMON58110 (Minimum 35S)的超螺旋 DNA 轰击(Maliga 等,1995, "Methods in Plant Molecular Biology, A Laboratory Course Manual, "Cold Spring Harbor Laboratory Press,47)。包括与各种启动子构建体为 1: 1 摩尔比的分离的 e35S 起始的荧光素酶构建体作为表达对照。被轰击的组织在 25°C 下培育 48h。

从六种被轰击的大豆子叶中提取蛋白质, 采用 1ml 含有 0.1 M 磷酸钾(pH 7.8), 10mM DTT, 1mM EDTA, 5% 甘油和蛋白酶抑制剂的提取缓冲液(1 片 /50ml, Roche Molecular Biochemicals,产品目录号 1697498, Indianapolis,IN)。100μl 蛋白质提取物的等分试样被用于荧光素酶分析, 通过 Promega 进行 "Steady-Glo"程序 (产品目录号 E2510,Promega Corporation,Madison,WI)。50μl 蛋白质提取物的等分试样被用于稍作改进的标准 GUS 分析操作(Maliga 等, 1995, "Methods in Plant Molecular Biology, A Laboratory Course Manual", Cold Spring Harbor Laboratory Press, page 29)。每种样本分析两次, 平均值用于数据分析。GUS 活性用荧光素酶活性规格化, 相对启动子强度通过任意地设定基准启动子 7Sα'(pMON13773)(附图 1) 为 100%来表示。独立地重复试验三次。结果(附图 14)表明, 当与 7Sα' 启动子、通常用于在大豆种子中高水平表达的基准启动子比较时, USP 启动子显著地提高 GUS 表达。最小的启动子/GUS 构建体(pMON58100)是采用 35 最小启动子起始 GUS 的低表达对照, 以约相当于 7Sα'构建体(pMON13773)表达水平的 20%被表达(附图 1)。

实施例 4-制备含有 USP 启动子的转基因大豆植物

为了在转基因大豆中检测 USP 启动子的强度, 通过如下稍加改进的标准分子克隆操作构建 pMON55526(Arc5/GUS/NOS), pMON55542 (T-Arc5/GUS/NOS), pMON63605(USP91/GUS/NOS), pMON58107 (USP88/GUS/NOS), pMON63604 (USP99/GUS/NOS)和 pMON58113 (eUSP88/GUS/NOS) (Sambrook 等, Molecular Cloning: A laboratory manual,1989,Cold Spring Harbor Laboratory Press; Maliga 等, Methods in Plant Molecular Biology: A laboratory course manual, 1995, Cold

Spring Harbor Laboratory Press)。由 FMV 启动子、转运肽序列、CP4 编码基因和 E9 3'UTR 组成的表达盒被包含在所有转化载体中，作为可选择标记。

对于粒子轰击转化方法，大豆种子(Asgrow A3244,A4922)发芽过夜(18-24h)，并切除分生组织外植体。去除初生(primary)叶来暴露分生组织。制备的外植体在 OR 培养基(参见，美国专利 5,914,451 来获得对 OR 培养液的描述)4℃下在黑暗中保存 2d，在 15℃下黑暗中保存 1d。在即将轰击之前，立即将制备的外植体放置在靶向培养液中(参见美国专利 5,914,451 来获得对靶向培养液的描述)，分生组织位于垂直于粒子呈递的方向。含有 DNA 的 pMON55526, pMON58107,或 pMON55542 用 CaCl₂ 和亚精胺被沉淀在细微的金颗粒上，随后重悬在乙醇中。将悬浮液覆盖到聚酯薄片上，然后将该薄片放置在放电装置上。通过释放大约 60%电容将颗粒加速到植物组织中。一般地，轰击目标一次。在轰击后，将外植体放到选择培养液中(WPM+75μM 草甘膦,参见美国专利 5,914,451 中的描述)5-7 周，选择和生长转基因的幼苗。约在轰击后 5-7 周收获表型阳性的幼苗，放置到选择性生根培养液中(BRM+25mM 草甘膦,参见美国专利 5,914,451 中的描述)2-3 周。生根的幼苗被移植到温室中，栽培在土壤中。在选择中仍然健康生长的但是不生根的幼苗被转移到非选择性生根培养液中(BRM,如上述)再培养 2 周。在被转移到温室并栽培在土壤之前，来自任何在脱离选择时生根的幼苗的根被检测植物可选择标记的表达。植物在标准的温室条件下被维持直到收获 RI 种子。

对于农杆菌转化方法，商业上可获得的大豆种子(Asgrow A3244, A4922)发芽过夜(大约 10-12h)，切除分生组织外植体。初生叶被去除或不被去除，暴露分生组织，外植体被放置在创伤容器(wounding vessel)中。含有 pMON58113(附图 8), pMON63605,或 pMON63604 的农杆菌菌株 ABI 生长到对数期。通过离心来收集细胞，并重悬在含有诱导物的培养液中。在不晚于种子萌发开始的 14h 内将大豆外植体和被诱导的农杆菌培养物被混合，并用超声波创伤。

在创伤后，外植物体在农杆菌中培育约 1h。在该培育步骤之后，通过吸取去除农杆菌，将外植体放置共培养 2-4d。将外植体转移到选择培养基中(WPM+0.075 mM 草甘膦+控制农杆菌过度生长的抗生素)5-7 周来进行选择，并使转基因幼苗生长。在轰击后约 5-7 周收获表型阳性的幼苗，并放置到选择性生根培养液中(BRM+25mM 草甘膦)2-3 周。生根的幼苗被移植到温

室中，栽培在土壤中。在选择中仍然健康生长的但是不生根的幼苗被转移到非选择性生根培养液中(无草甘膦的 BRM)再培养 2 周。在被转移到温室并栽培在土壤之前，来自任何在脱离选择时生根的幼苗的根被检测植物可选择的标记草甘膦抗性的表达。植物在标准的温室条件下被维持直到收获 RI 种子。

分析来自被选择的植物的成熟种子的 GUS 活性。为了分析 GUS 活性，分别碾磨 8 粒来自各个转基因事件(品系)的种子。约 20mg 被碾磨的种子组织用 200 μ l 提取缓冲液提取，该缓冲液中含有 0.1 M 磷酸钾(pH 7.8), 10mM DTT, 1mM EDTA, 5%甘油和蛋白酶抑制剂(1 片/50ml, 产品目录号 1697498, Roche Molecular Biochemicals, Indianapolis, IN)。提取物的蛋白质含量采用 Bio-Rad Protein Assay (产品目录号 61234A, Bio-Rad Laboratories, Hercules, California)确定，采用稍加改进的 GUS 标准分析方法测定 GUS 活性(Maliga 等, 1995, "Methods in Plant Molecular Biology, A Laboratory Course Manual", Cold Spring Harbor Laboratory Press, 29)。GUS 活性被相对于蛋白质浓度规格化。每种样本被分析两次，并且平均值被用数据分析。

如果 8 粒种子都没有可检测的 GUS 活性，则排除该事件(品系)。如果特定事件的至少一粒种子具有可检测的 GUS，该情形被认为是阳性转基因。这种情况下具有最高 GUS 活性的种子样本被选择作为这种情形的代表，因为其更有可能在纯合种子中反映 GUS 活性。对于各个构建体，通常确定 10 - 20 个阳性事件。阳性情况下的平均被比较以证明启动子的相对强度，如附图 12 所示。数据表明，在转基因大豆种子中 USP88 比 T-Arc5 启动子强约 3 倍，并且比 Arc5 启动子强约 6 倍(附图 12)。该结果还表明，在转基因大豆种子中，eUSP88 启动子比 T-Arc5 启动子强约 10 倍(附图 12)。

实施例 5-制备种子中含有高水平游离色氨酸的转基因大豆植物

构建农杆菌转化载体 pMON58130(附图 13)来证明，USP99 启动子在转基因大豆中起始对色氨酸不敏感的 C28 玉米邻氨基苯甲酸合成酶 α 亚基的效率(美国专利 6,118,047)。通过如下稍加改进的标准分子克隆操作构建载体 pMON58130(附图 13)(Sambrook 等, Molecular Cloning: A laboratory manual, 1989, Cold Spring Harbor Laboratory Press; Maliga 等, Methods in Plant Molecular Biology: A laboratory course manual, 1995, Cold Spring Harbor Laboratory Press)。由 FMV 启动子、转运肽序列、CP4 编码基因和 E9 3'UTR

组成的表达盒被包括在转化载体中，作为可选择的标记。

构建其他农杆菌载体 pMON63654 来证明，USP99 启动子在转基因大豆中引导对色氨酸反馈不敏感的根癌农杆菌邻氨基苯甲酸合成酶突变体 (F298W) 的效率(美国专利申请 60/288,904)。通过如下稍加改进的标准分子克隆操作构建载体 pMON63654 (附图 21) (Sambrook 等, *Molecular Cloning: A laboratory manual*, 1989, Cold Spring Harbor Laboratory Press; Maliga 等, *Methods in Plant Molecular Biology: A laboratory course manual*, 1995, Cold Spring Harbor Laboratory Press)。由 FMV 启动子、HSP70 5'UTR、CTP2、CP4 编码序列和 E9 3'UTR 组成的表达盒被包括在转化载体中作为可选择的标记。在 pMON63654 中(附图 21)，USP99 启动子被连接到由叶绿体转运肽序列 CTP1 和农杆菌邻氨基苯甲酸合成酶(F298W)突变体组成的基因的上游。NOS 3'UTR 被用于发出转录终止和腺苷酸化的信号。

对于农杆菌转化方法，将商业上可获得的大豆种子(AsgrowA3244, A4922)过夜发芽(大约 10-12h)发芽，并切除分生组织外植体。初生叶被去除或不被去除，暴露分生组织，外植体被放置在创伤容器(wounding vessel)中。含有 pMON58130, 或 pMON63654 的农杆菌菌株 ABI 生长到对数期。通过离心来收集细胞，并重悬在含有诱导物的培养基中。在不晚于种子萌发开始的 14h 内将大豆外植体和被诱导的农杆菌培养物混合，并采用超声波创伤。

在创伤后，外植物体在农杆菌中培育约 1h。在该培育步骤之后，通过吸取去除农杆菌，将外植体放置共培养 2-4d。此时，将外植体转移到选择培养液中(WPM+0.075 mM 草甘膦+控制农杆菌过度生长的抗生素)5-7 周来进行选择，并使转基因幼苗生长。在轰击后约 5-7 周收获表型阳性的幼苗，并放置到选择性生根培养液中(BRM+25mM 草甘膦)2-3 周。生根的幼苗被移植到温室中，栽培在土壤中。在选择中仍然健康生长的但是不生根的幼苗被转移到非选择性生根培养液中(无草甘膦的 BRM)再培养 2 周。在被转移到温室并栽培在土壤之前，来自任何在脱离选择时生根的幼苗的根被检测植物可选择的标记草甘膦抗性的表达。植物在标准的温室条件下被维持直到收获 RI 种子。

为了分析游离的色氨酸，10 粒来自各个转基因事件(品系)的成熟 RI 种子分别被碾碎。约 50mg 来自各个种子的被碾碎物质被放置到各个离心管中并称重。将 1ml 5%三氯乙酸加入到各个样本中。摇晃样本，在室温下混合

15min。然后以 14,000 rpm 离心 15min。取出一些上清液，装入 HPLC 管中并密封。采用 Zorbax Eclipse-AAA 柱和 Agilent 1100 HPLC 分析游离的氨基酸 (Agilent Technical Publication,"Amino Acid Analysis Using Zorbax Eclipse-AAA Columns and the Agilent 1100 HPLC." 2000 年 3 月 17 日)。因为各种情况的 RI 种子最可能由一群相互分离的种子构成，在各种事件的 10 粒种子中，具有最高色氨酸读数的种子被选择作为该组中最有可能是纯合的代表。数据总结在下面的表 3 和 4 中。

还分析了 10 粒随机选择的非转基因的 Asgrow A3244 种子，具有最高色氨酸含量的被包括在表中作为阴性对照。从 pMON58130-1 到 pMON58130-23 表示采用载体 pMON58130 产生的不同情形。与非转基因 A3244 比较，在大部分转基因事件中具有高水平的 Trp 沉积。在采用 pMON58130 和 pMON63654 的多个转基因事件中检测到 Trp 沉积升高。

表 3: pMON58130 事件中的色氨酸浓度

事件编号	Trp (ppm)
A3244	306
pMON58130-1	484
pMON58130-2	3104
pMON58130-3	8237
pMON58130-4	7734
pMON58130-5	432
pMON58130-6	4540
pMON58130-7	4698
pMON58130-8	361
pMON58130-9	344
pMON58130-10	6435
pMON58130-11	5310
pMON58130-12	283
pMON58130-13	200
pMON58130-14	90
pMON58130-15	5479
pMON58130-16	6316
pMON58130-17	1516
pMON58130-18	3714
pMON58130-19	4480
pMON58130-20	636
pMON58130-21	534
pMON58130-22	872
pMON58130-23	4986

表 4: pMON63654 事件的色氨酸浓度

pMON 号	描述	事件编号	TRP 平均值 (ppm)	TRP 最大值 (ppm)
63654	USP99-F298W-AS	27581	10, 591	19, 630
63654	USP99-F298W-AS	27654	1, 802	17, 796
63654	USP99-F298W-AS	28034	7, 186	21, 278

实施例 6-生成来自蚕豆的 USP 启动子的克隆

通过 PCR 扩增(Expand High Fidelity PCR System,产品目录号 1732641, Roche Molecular Biochemicals,Indianapolis,IN)和 Universal GenomeWalker® 试剂盒(产品目录号 K1807-1 BD Biosciences, Palo Alto, 加利福尼亚)从蚕豆基因组 DNA 中获得延伸序列 USP 启动子, 采用按照公开的序列(GenBank Accession X56240)设计的 3'引物。5'引物按照 Genome Walker 方法设计。用于第一轮扩增 USP 启动子的引物为: GATAAAACAGTGAGATGTGCAAACCTCC (uspGW-P-down)(SEQID NO:12)和 GTAATACGACTCACTATAGGGC(AP1,连接引物 1,由试剂盒提供) (SEQ ID NO:13)。

采用嵌套引物 CCATGGAGATCTGACTGGCTATGAAGAAATTATAATCG (uspGW-N-downBgl2NcoI) (SEQ ID NO: 14)和 ACTATAGGGCACGCGTGGT (AP2,嵌套连接引物 2) (SEQID NO:15)从这些初步的 PCR 产物中扩增启动子。

1 微升 PCR 片段洗脱物被用作模板用于第三轮 PCR, 采用 uspGW-N-down Bgl2NcoI 和 CTGCAGGTCGACGGCCCGGGCTGGT (AP6-PstI/SrfI) (SEQ ID NO: 16)来将便利的 5'限制性位点加入到推定的启动子片段中。然后将分离的扩增产物亚克隆到载体 pMON8677 中, 生成克隆 pMON63821(USP99.5) (附图 18), pMON63819(USP95)(附图 19),和 pMON63820 (USP68) (附图 20)。

实施例 7-采用大豆子叶瞬时转化来评价 USP 启动子

在开花后 25-28d 收获大豆植物的种子(Asgrow A3244), 并且在 GAMBORG 培养液(G5893, Sigma Company,St.Louis,MO)中在 25°C 下黑暗中渗透处理过夜, 该培养液中加入 50 mM 谷氨酰胺,111 mM 麦芽糖,125 mM 棉子糖,125 mM 甘露醇和 3g/l 纯化的琼脂,pH 5.6。得到的 125 片子叶被对半

剪切，并且采用粒子枪技术(Maliga 等, 1995, "Methods in Plant Molecular Biology, A Laboratory Course Manual," Cold Spring Harbor Laboratory Press,47) 用纯化的 pMON63819(USP95),pMON63820(USP68),pMON63621 (USP99.5), 和 pMON58101(USP99)的超螺旋 DNA 轰击。包括以与各种启动子构建体 1:1 摩尔比的分离的 e35S 引导的荧光素每构建体作为表达的 内部标准。被轰击的组织在 25°C 下培养 48h。

采用 1ml 提取缓冲液从 6 片被轰击的大豆子叶中提取蛋白质，该缓冲液含有 0.1 M 磷酸钾(pH 7.8), 10mM DTT, 1mM EDTA, 5% 甘油和蛋白酶抑制剂 (1 片 /50ml, Roche Molecular Biochemicals, 产品目录号 1697498, Indianapolis,IN)。100 μ l 蛋白质提取物的等分试样被用于采用 Promega(产品目录号 E2510, Promega Corporation Madison, WI)通过“Steady-Glo”操作分析荧光素酶。50 μ l 蛋白质提取物的等分试样被用于稍加改进的标准 GUS 分析操作 (Maliga 等, 1995, "Methods in Plant Molecular Biology, A Laboratory Course Manual", Cold Spring Harbor Laboratory Press,29)。各种样本被分析两次，平均值被用于数据分析。GUS 活性使用荧光素酶活性规格化，相对启动子强度通过比较 pMON58101(USP99)的表达与 pMON63819 (USP95), pMON63820 (USP68), pMON63621 (USP99.5)表达来表示(表 5)。pMON63621(USP99.5)具有显著高于 USP99 的表达，pMON63820 (USP68)具有显著低于 USP99 的表达，而 pMON63819 (USP95)具有与 USP99 相似的表达。在该瞬时表达系统中，所有这三种构建体都被证实含有活性启动子。

表 5: 转基因表达系统中的启动子活性

构建体	启动子	相对 GUS 活性	标准方差	启动子大小(bp)
pMON58101	USP99	1.59	0.10	682
pMON63819	USP95	1.47	0.21	1464
pMON63820	USP68	0.52	0.17	1301
pMON63821	USP99.5	2.81	0.82	1748

参考文献

- Ainley *et al.*, *Plant Mol. Biol.*, 14:949, 1990.
- Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley and Sons, Inc., 1995.
- Bartley and Scolnik, *Plant Physiol.*, 104:1469-1470, 1994.
- Back *et al.*, *Plant Mol. Biol.*, 17:9, 1991.
- Bauer *et al.*, *Gene*, 37:73, 1985.
- Becker and Guarente, In: Abelson and Simon (eds.), *Guide to Yeast Genetics and Molecular Biology, Methods Enzymol.*, 194:182-187, Academic Press, Inc., New York.
- Bennett and LaSure (eds.), *More Gene Manipulations in Fungi*, Academic Press, California, 1991.
- Bergeron *et al.*, *TIBS*, 19:124-128, 1994.
- Bustos *et al.*, *EMBO J.*, 10:1469-1479, 1991.
- Castresana *et al.*, *EMBO J.*, 7:1929-1936, 1988.
- Capecchi, *Cell*, 22(2):479-488, 1980.
- Cerda-Olmedo *et al.*, *J. Mol. Biol.*, 33:705-719, 1968.
- Chau *et al.*, *Science*, 244:174-181, 1989.
- Christensen *et al.*, *Plant Mol. Biol.*, 18:675-689, 1992.
- Clapp, *Clin. Perinatol.*, 20(1):155-168, 1993.
- Costa *et al.*, *Methods Mol. Biol.*, 57:31-44, 1996.
- Craik, *BioTechniques*, 3:12-19, 1985.
- Craig, *Science*, 260:1902-1903, 1993.
- Curiel *et al.*, *Hum. Gen. Ther.*, 3(2):147-154, 1992.
- Cyanobase, <http://www.kazusa.or.jp/cyanobase>.
- Dellaporta *et al.*, *Stadler Symposium*, 11:263-282, 1988.
- Demolder *et al.*, *J. Biotechnology*, 32:179-189, 1994.
- Deng and Nickloff, *Anal. Biochem.*, 200:81, 1992.
- D'Halluin *et al.*, *Bio/Technology*, 10:309-314, 1992.
- Doyle *et al.*, *J. Biol. Chem.*, 261:9228-9238, 1986.
- Eglitis and Anderson, *Biotechniques*, 6(7):608-614, 1988.
- Enderlin and Ogrydziak, *Yeast*, 10:67-79, 1994.
- Feinbaum *et al.*, *Mol. Gen. Genet.*, 226:449-456, 1991.
- Fraley *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)*, 80:4803, 1983.
- Frits Eckstein *et al.*, *Nucleic Acids Research*, 10:6487-6497, 1982.
- Fromm *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)*, 82(17):5824-5828, 1985.

- Fromm *et al.*, *Bio/Technology*, 8:833, 1990.
- Fuller *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)*, 86:1434-1438, 1989.
- Fynan *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)*, 90(24):11478-11482, 1993.
- Gething and Sambrook, *Nature*, 355:33-45, 1992.
- Glick *et al.*, *Methods in Plant Molecular Biology and Biotechnology*, CRC Press, Boca Raton, FL., 1993.
- Graham and Van der Eb, *Virology*, 54(2):536-539, 1973.
- Greener *et al.*, *Mol. Biotechnol.*, 7:189-195, 1997.
- Hartl *et al.*, *TIBS*, 19:20-25, 1994.
- Hershey and Stoner, *Plant Mol. Biol.*, 17:679-690, 1991.
- Hess, *Intern Rev. Cytol.*, 107:367, 1987.
- Hinchee *et al.*, *Bio/Technology*, 6:915-922, 1988.
- Hinnen *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)*, 75:1920, 1978.
- Horsch *et al.*, *Science*, 227:1229-1231, 1985.
- Ikatu *et al.*, *Bio/Technol.*, 8:241-242, 1990.
- Jarai and Buxton, *Current Genetics*, 26:2238-2244, 1994.
- Jefferson (I), *Plant Mol. Biol. Rep.*, 5:387-405, 1987.
- Jefferson (II) *et al.*, *EMBO J.*, 6:3901-3907, 1987.
- Johnston and Tang, *Methods Cell Biol.*, 43(A):353-365, 1994.
- Jones *et al.*, *Science*, 266:789-793, 1994.
- Jones *et al.*, *Mol. Gen. Genet.*, 1987.
- Julius *et al.*, *Cell*, 32:839-852, 1983.
- Julius *et al.*, *Cell*, 37:1075-1089, 1984.
- Kares *et al.*, *Plant Mol. Biol.*, 15:905, 1990.
- Katz *et al.*, *J. Gen. Microbiol.*, 129:2703-2714, 1983.
- Keegstra, *Cell*, 56(2):247-253, 1989.
- Keller *et al.*, *EMBO L.*, 8:1309-1314, 1989.
- Knutzon *et al.*, *Proc. Natl. Acad. Sci (U.S.A.)*, 89:2624-2628, 1992.
- Kridl *et al.*, *Seed Sci. Res.*, 1:209, 1991.
- Kudla *et al.*, *EMBO*, 9:1355-1364, 1990.
- Kuhlemeier *et al.*, *Seeds*, 1:471, 1989.
- Kunkel, *Proc. Natl. Acad. Sci. (U.S.A.)*, 82:488-492, 1985.
- Kyte and Doolittle, *J. Mol. Biol.*, 157:105-132, 1982.
- Lam and Chua, *J. Biol. Chem.*, 266:17131-17135, 1990.

- Lam and Chua, *Science*, 248:471, 1991.
- Lipman and Pearson, *Science*, 227:1435-1441, 1985.
- Lindstrom *et al.*, *Developmental Genetics*, 11:160, 1990.
- Lois *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)*, 95(5):2105-2110, 1998.
- Lu *et al.*, *J. Exp. Med.*, 178(6):2089-2096, 1993.
- Luo *et al.*, *Plant Mol Biol. Reporter*, 6:165, 1988.
- MacKenzie *et al.*, *Journal of Gen. Microbiol.*, 139:2295-2307, 1993.
- Malardier *et al.*, *Gene*, 78:147-156, 1989.
- Mandel *et al.*, *Plant Mol. Biol.*, 29:995-1004, 1995.
- McElroy *et al.*, *Seeds*, 2:163-171, 1990.
- Nawrath *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)*, 91:12760-12764, 1994.
- Needleman and Wunsch, *Journal of Molecular Biology*, 48:443-453, 1970.
- Neuhaus *et al.*, *Theor. Appl. Genet.*, 75:30, 1987.
- Odell *et al.*, *Nature*, 313:810, 1985.
- Osuna *et al.*, *Critical Reviews In Microbiology*, 20:107-116, 1994.
- Ou-Lee *et al.*, *Proc. Natl. Acad. Sci (U.S.A.)*, 83:6815, 1986.
- Ow *et al.*, *Science*, 234:856-859, 1986.
- Pearson and Lipman, *Proc. Natl. Acad. Sci. (U.S.A.)*, 85:2444-2448, 1988.
- Pearson, "Rapid and Sensitive Sequence Comparison with FASTP and FASTA." In *Methods in Enzymology*, (R. Doolittle, ed.), 183:63-98, Academic Press, San Diego, California, 1990.
- Pearson, *Protein Science*, 4:1145-1160, 1995.
- Pena *et al.*, *Nature*, 325:274, 1987.
- Poszkowski *et al.*, *EMBO J.*, 3:2719, 1989.
- Potrykus *et al.*, *Ann. Rev. Plant Physiol. Plant Mol. Biol.*, 42:205, 1991.
- Potrykus *et al.*, *Mol. Gen. Genet.*, 199:183-188, 1985.
- Puig and Gilbert, *J. Biol. Chem.*, 269:7764-7771, 1994.
- Pyee *et al.*, *Plant J.*, 7:49-59, 1995.
- Reynaerts *et al.*, Selectable and Screenable Markers. In Gelvin and Schilperoort. *Plant Molecular Biology Manual*, Kluwer, Dordrecht, 1988.
- Richins *et al.*, *Nucleic Acids Res.*, 20:8451, 1987.
- Robinson *et al.*, *Bio/Technology*, 1:381-384, 1994.
- Rodriguez *et al.* Vectors: A Survey of Molecular Cloning Vectors and Their Uses, Butterworths, Boston, 1988.
- Rogan and Bessman, *J. Bacteriol.*, 103:622-633, 1970.

- Rogers *et al.*, *Meth. In Enzymol.*, 153:253-277, 1987.
- Saint Guily *et al.*, *Plant Physiol.*, 100(2):1069-1071, 1992.
- Samac *et al.*, *Seeds*, 3:1063-1072, 1991.
- Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.
- Sato *et al.*, *J. DNA Res.*, 7(1):31-63, 2000.
- Schulze-Lefert *et al.*, *EMBO J.*, 8:651, 1989.
- Simpson, *Science*, 233:34, 1986.
- Singer and Kusmierk, *Ann. Rev. Biochem.*, 52:655-693, 1982.
- Slighton and Beachy, *Planta*, 172:356, 1987.
- Smith *et al.*, In: *Genetic Engineering: Principles and Methods*, Setlow *et al.*, Eds., Plenum Press, NY, 1-32, 1981.
- Smith and Waterman, *Advances in Applied Mathematics*, 2:482-489, 1981.
- Smith *et al.*, *Nucleic Acids Research*, 11:2205-2220, 1983.
- Smith *et al.*, *Plant J.*, 11:83-92, 1997.
- Stalker *et al.*, *J. Biol. Chem.*, 263:6310-6314, 1988.
- Sutcliffe *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)*, 75:3737-3741, 1978.
- Takahashi *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)*, 95(17):9879-9884, 1998.
- Thillet *et al.*, *J. Biol. Chem.*, 263:12500-12508, 1988.
- Vandeyar *et al.* *Gene*, 65:129-133, 1988
- Van Tunen *et al.*, *EMBO J.*, 7:1257, 1988.
- Verdier, *Yeast*, 6:271-297, 1990.
- Vodkin *et al.*, *Cell*, 34:1023, 1983.
- Vogel *et al.*, *J. Cell Biochem., (Suppl)* 13D:312, 1989.
- Wagner *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)*, 89(13):6099-6103, 1992.
- Wang and Tsou, *FASEB Journal*, 7:1515-1517, 1993.
- Weissbach and Weissbach, *Methods for Plant Molecular Biology*, (Eds.), Academic Press, Inc., San Diego, California, 1988.
- Weisshaar *et al.*, *EMBO J.*, 10:1777-1786, 1991.
- Wenzler *et al.*, *Plant Mol. Biol.*, 12:41-50, 1989.
- Williams, *et al.*, *Biotechnology*, 10:540-543, 1992.
- Wong and Neumann, *Biochim. Biophys. Res. Commun.*, 107(2):584-587, 1982.
- Xia *et al.*, *J. Gen. Microbiol.*, 138:1309-1316, 1992.
- Yang *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)*, 87:4144-48, 1990.
- Yamaguchi-Shinozaki *et al.*, *Plant Mol. Biol.*, 15:905, 1990.

Yelton *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)*, 81:1470-1474, 1984.

Zhou *et al.*, *Methods in Enzymology*, 101:433, 1983.

Zukowsky *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)*, 80:1101-1105, 1983.

美国专利 4,683,195; 4,683,202; 4,757,011; 4,769,061; 4,885,357, 4,886,878, 4,940,835;
4,957,748; 4,971,908; 5,057,419; 5,093,249; 5,100,679; 5,147,792; 5,215,912;
5,219,596; 5,270,200; 5,298,421; 5,304,481; 5,344,771; 5,362,865; 5,576,203;
5,508,468; 5,003,045; 5,955,329; 5,367,110; 5,858,749; 6,040,160; 5,610,041;
5,618,988; 6,107,060; 5,811,636; 4,766,072; 5,003,045; 5,576,203; 5,384,253;
5,443,974; 5,512,482; 5,530,186; 5,534,421; 5,552,306; 5,589,616; 5,508,468;
5,614,393; 5,663,068; 5,663,068; 5,633,436; 5,639,790; 5,654,402; 5,659,645;
5,689,050; 5,689,050; 5,689,052; 5,705,391; 5,760,206; 5,759,829; 5,789,220;
5,807,893; 5,850,024; 5,856,157; 5,866,789; 5,885,802; 5,885,801; 5,914,450;
5,942,660; 5,945,585; 5,952,544; 5,955,650; 5,965,727; 5,995,329; 5,990,384;
5,990,389; 5,936,069; 5,939,599; 6,005,076; 6,051,754; 6,075,183; 6,043,411;
6,100,091; 6,107,051; 6,110,891; 6,117,677; 6,194,167; 6,146,669; 6,147,279;
6,156,227; 6,172,106; and 6,232,122.

欧洲专利: 0 154 204; 0 238 023; 和 0 255 378.

专利申请: WO 90/01869, WO 91/13993, WO 92/14822, WO 93/08682,
WO 94/20628, WO 95/19442, WO 97/26366, WO 97/28247, WO 97/22703,
WO 98/55601, WO 98/26064, WO 96/17064, WO 97/35023, WO 00/19839,
WO 99/06581, WO 99/02656, WO 99/40209, WO 99/11800, WO 99/49058,
WO 00/32757, WO 00/10380.

<110> Wang, Qi
Fagaly, Tanya
Bassuner, Ronald Liang, Jihon
oulmassov, Tim Dabrowski, John

<120> 用于在植物中表达基因的种子特异性 USP 启动子

<130> REN-00-122

<160> 16

<170> PatentIn version 3.1

<210> 1

<211> 634

<212> DNA

<213> 蚕豆(*Vicia faba*)

<400> 1

```
ctgcagcaaa tttacacatt gtcactaac gtctaaatca ttgtaatttg tttttgtttt 60
aatatgtgtg ttatgaactt gattttcaat aatTTTTaaa tttggtacca gtattataac 120
atcttttgtg ctaacggttg ccaacactta gcaatttgta agttgattaa ttgattctaa 180
acttttattg tcttcttaat tcatgctgat aaatatatgc tgataaaaat taaagtgaat 240
atggtaccac aagTTTTtg agactgttgc catatacacc aaacattcaa taattcttga 300
ggataataat ggtaccacac aagctttgag gtgcatgaac gtcacgtgga caaaaggttt 360
agtaattttt caagacaaca atggtaccac acacaagttt tgaggtgcat gcatggatgc 420
cctgtggaaa gtttaaaaat attttggaaa tgatttgc at ggaagccatg tgtaaaacca 480
tgacatccac ttggaggaag caataatgaa gaaaactaca aatttacatg caactagtta 540
tgcatgtagt ctatataatg aggattttgc aatactttca ttcataaaca ctcactaagt 600
tttacacgat ttcatttct tcatagccag tcaa 634
```

<210> 2

<211> 1185

<212> DNA

<213> 蚕豆(*Vicia faba*)

<400> 2

```
ctgcagcaaa tttacacatt gtcactaac gtctaaatca ttgtaatttg tttttgtttt 60
aatatgtgtg ttatgaactt gattttcaat aatTTTTaaa tttggtacca gtattataac 120
atcttttgtg ctaacggttg ccaacactta gcaatttgta agttgattaa ttgattctaa 180
acttttattg tcttcttaat tcatgctgat aaatatatgc tgataaaaat taaagtgaat 240
atggtaccac aagTTTTtg agactgttgc catatacacc aaacattcaa taattcttga 300
ggataataat ggtaccacac aagctttgag gtgcatgaac gtcacgtgga caaaaggttt 360
agtaattttt caagacaaca atggtaccac acacaagttt tgaggtgcat gcatggatgc 420
cctgtggaaa gtttaaaaat attttggaaa tgatttgc at ggaagCcatg tgtaaaacca 480
```

tgacatccac	ttggaggaag	caataatgaa	gaaaactaca	aatttacatg	caactagtta	540
tgcatgtagt	cctgcagcaa	atttacacat	tgctactaaa	cgtctaaatc	attgtaattt	600
gtttttgttt	taatattgtg	gttatgaact	tgattttcaa	taatttttaa	atttgggtacc	660
agtattataa	catcttttgt	gctaacgggt	gccaacactt	agcaatttgt	aagttgatta	720
attgattcta	aactttttatt	gtctttcttaa	ttcatgctga	taaataatg	ctgataaaaa	780
ttaaagttaa	tatggtacca	caagtttttg	gagactgttg	ccatatacac	caaacattca	840
ataattcttg	aggataataa	tgggtaccaca	caagctttga	ggtgcatgaa	cgtcacgtgg	900
acaaaagggt	tagtaatttt	tcaagacaac	aatgttacca	cacacaagtt	ttgaggtgca	960
tgcatggatg	ccctgtggaa	agtttaaaaa	tattttggaa	atgatttgca	tggagccat	1020
gtgtaaaacc	atgacatcca	cttggaggaa	gcaataatga	agaaaactac	aaatttacat	1080
gcaactagtt	atgcatgtag	tctatataat	gaggattttg	caatactttc	attcataaac	1140
actactaag	ttttacacga	ttatcatttc	ttcatagcca	gtcaa		1185

<210> 3

<211> 688

<212> DNA

<213> 蚕豆(*Vicia faba*)

<400> 3

ctgcagcaaa	tttacacatt	gccactaaac	gtctaaacce	ttgtaatttg	tttttatttc	60
gctatgtgtg	ttatgtattt	aatttgcgat	aaatttttat	atttgggtact	aaatttataa	120
caccttttat	gctaacgttt	gccaacactt	agcaatttgc	aagttgatta	atcgattcta	180
aattattttt	gtctttctaaa	tacatatact	aatcaactgg	aaatgtaa	atttggctaat	240
atttctacta	taggagaatt	aaagtgagtg	aatatggtac	cacaaggttt	ggagatttaa	300
ttgttgcaat	gcttgcatgg	atggcatata	caccaaacat	tcaataattc	ttgaggataa	360
taatgggtacc	acacaagctt	tgaggtgcat	gaacgtcacg	tggacaaaag	gtttagtaat	420
ttttcaagac	aacaatgtta	ccacacacaa	gttttgaggt	gcatgcatgg	atgccctgtg	480
gaaagtttaa	aaatattttg	gaaatgattt	gcatggaagc	catgtgtaaa	accatgacat	540
ccacttggag	gaagcaataa	tgaagaaaac	tacaaattta	catgcaacta	gttatgcatg	600
tagtctatat	aatgaggatt	ttgcaatact	ttcattcata	aacactcact	aagttttaca	660
cgattatcat	ttcttcatag	ccagttcaa				688

<210> 4

<211> 744

<212> DNA

<213> 蚕豆(*Vicia faba*)

<400> 4

ctgcagcaaa	tttacacatt	gccactaaac	gtctaaacce	ttgtaatttg	tttttatttc	60
actatgtgtg	ttacgtattt	aatttgcgat	aaatttttat	atttgggtact	aaatttataa	120
caccttttat	gctaacgttt	gccaacacct	agcaatttgc	aagttgaaat	ttataacacc	180
ttttatgcta	acgtttgcca	acacttagca	atttgcaagt	tgattaatcg	attctaaatt	240
atttttgtct	tctaaataca	tatactaatac	aactggaaat	gtaaataattt	gctaataattt	300
ctactatagg	agaattaaag	tgagtgaata	tgggtaccaca	aggtttggag	atttaattgt	360
tgcaatgctt	gcatggatgg	catatacacc	aaacattcaa	taattcttga	ggataataat	420
ggtaccacac	aagctttgag	gtgcatgaac	gtcacgtgga	caaaaaggtt	agtaattttt	480
caagacaaca	atgttaccac	acacaagttt	tgaggtgcat	gcatggatgc	cctgtggaaa	540

gtttaaaaat attttggaaa tgatttgcac ggaagccatg tgtaaaacca tgacatccac 600
 ttggaggaag caataatgaa gaaaactaca aatttacatg caactagtta tgcacgtatg 660
 ctatataatg aggattttgc aatactttca ttcataaaca ctactaagt tttacacgat 720
 tatcatttct tcatagccag tcaa 744

<210> 5

<211> 688

<212> DNA

<213> 蚕豆(*Vicia faba*)

<400> 5

ctgcagcaaa tttacacatt gccactaaac gtctaaacc ttgtaatttg tttttgtttt 60
 actatgtgtg ttatgtattt gatttgcgat aaatttttat atttggact aaatttataa 120
 caccttttat gctaacgttt gccaacactt agcaatttgc aagttgatta attgattcta 180
 aattattttt gtcttctaaa tacatatact aatcaactgg aatgtaaatt atttgctaat 240
 atttctacta taggagaatt aaagtgagt aatatggtac cacaaggttt ggagatttaa 300
 ttgttgcaat gctgcatgga tggcatatac accaaacatt caataattct tgaggataat 360
 aatggtacca cacaagattt gaggtgcatg aacgtcacgt ggacaaaagg tttagtaatt 420
 tttcaagaca acaatgttac cacacacaag ttttgaggtg catgcatgga tgcccctgtg 480
 gaaagttaa aatattttg gaaatgattt gcatggaagc catgtgtaaa accatgacat 540
 ccacttggag gatgcaataa tgaagaaaac taaaattta catgcaacta gttatgcatg 600
 tagtctatat aatgaggatt ttgcaatact ttattcata cacactcact aagttttaca 660
 cgattataat ttcttcatag ccagtcaa 688

<210> 6

<211> 24

<212> DNA

<213> 合成的

<400> 6

aaactgcagc aaatttacac attg 24

<210> 7

<211> 22

<212> DNA

<213> 合成的

<400> 7

aaacctgggt tgactggcta tg 22

<210> 8

<211> 24

<212> DNA

<213> 合成的

<400> 8

aaactcjcagg actacatgca taac 24

<210> 9

<211> 1749

<212> DNA

<213> 蚕豆(*Vicia faba*)

<400> 9

```

acttcgacca actagagttc gatccaaaat cccacatacg caataccaaa cattaactga 60
agtttttcaa ggaggctcaa aactatcgag aatgtctcat catcattgac atcatcgtct 120
cgaagctccg tgaaagaaca aacacattat gccaaacatt atctttatga atatattacc 180
acagaccatc actccgacct tgcaatcaag aatgggtgtca ttgtcgttac cctaatggta 240
gggtggatca aacaagaggt ctatgttagt tttatcggac ccctaaaaag tacaatgaga 300
taaccctatg ttttttagtgt tgttttatgt atttagaggg tttgctcatg catgcgtagt 360
ggtagaagt aaggttgtct agatcaaaaat tttacataag atccggagag aataacaaga 420
ttgtaaaaata taaaatcgta gttaagatta gaagatttta ctcaattatt tttgtaagat 480
tgacagcgaa taacatcgta aaacataaga taataaaaaa aaagtaagat cctactaaaa 540
tgaaaaaaaa atagtatata tttaaagtat ttttacctga tataatgatg tatgtaagta 600
ggcaagtgta tttgtgagaa aaaaatattc tttttcatat tctttaaaca tttacgattg 660
aggttttatt aaatatttgt taatgtttag acattagaga catatatagt caattactaa 720
gattatcata agtctactta aaacaaatct tatatgttaa aagtttattc tctaaatcct 780
aaatctaaaa tttcatttca aaaggtgaaa attcatctcc gctgaacatc attgtgcttc 840
cacatcttgt catcattctc acattttaat ggtgggtggtc gtaaaacggt gaatcatagt 900
caaagctggc aaagatcgca caaaatcaat aattttaaaa gaatttttat tcaatttagc 960
tacaattcgc gatttactt gaaatctgac aaccatacat atttatatac cgataaaata 1020
taggctatat tatacatttg ctttaaaaaa aaactgaaaa actcctgcag caaatttaca 1080
cattgccact aaacgtctaa acccttgtaa tttgtttttg ttttactatg tgtgttatgt 1140
at ttgatttg cgataaattt ttatatttgg tactaaattt ataacacctt ttatgctaac 1200
g tttgccaac acttagcaat ttgcaagttg attaattgat tctaaattat ttttgtcttc 1260
taaatacata tactaatcaa ctggaaatgt aatatattgc taatatttct actataggag 1320
aattaaagt agtgaatatg gtaccacaag gtttgagatg ttaattgttg caatgctgca 1380
tggatggcat atacacaaa cattcaataa ttcttgagga taataatggt accacacaag 1440
at ttgaggtg cargaacgtc acgtggacaa aaggttttagt aatttttcaa gacagcaatg 1500
ttaccacaca caagttttga ggtgcatgca tggatgcctt gtggaaagtt taaaaatatt 1560
ttggaaatga tttgcatgga agccatgtgt aaaacctatg catccacttg gaggatgcaa 1620
taatgaagaa aactacaaat ttacatgcaa ctagtattgc atgtagtcta tataatgagg 1680
at ttgcaat actttcattc atacacactc actaagtttt acacgattat aattttctca 1740
tagccagtc 1749

```

<210> 10

<211> 1465

<212> DNA

<213> 蚕豆(*Vicia faba*)

<400> 10

```

cctatgtagt ttttatcgga cacctaaaaa gtacaatgag ataaccttat gtttttagtg 60
ttgttttatg tgttttagagt gtttgctcat gcatagtggt tagaagtaag gttgtcaaga 120
tcgaaacttt acataagatc tggagagaat aataagattg taaaatataa aatttttagtt 180
aagataagaa gattttactc aattatTTTT gtaagattga cagggataaa cataatcgta 240
aatcataaga tagtaacaaa aaaagtaaga tctactaaa ataaaaaaat taatagtata 300
tatttaaagt atttttacat gatataatga tgtatgtaag tatgcaagtg tattttgtgag 360
aaaaaaaaata ctctttttca tattctttta acattttatga ttgagatttt attaaatatt 420

```

tgттаатгтт	tagacattag	agacatatat	ggtcaattat	taggattatc	ataagtttag	480
ttaaaacaaa	tcttatatgt	taaaagtta	ttctctaac	cttaaacta	aaactttatt	540
tcacaaagcg	aaaattcatc	cccgcigaat	cgtaaattat	gtgcttccac	atcttgatc	600
cattctcaca	ttttaatagt	ggtggtcgta	aaacggtaa	tcatggccaa	agctgacaaa	660
gatcgcacaa	aatcaataat	ttatagattt	ttattcaat	ttagctacaa	tacacgagtc	720
tacttgaaat	cgagattttg	acaaccatat	atattcatat	acagataaaa	tgtaggctat	780
gttatacatt	tgcccttaaa	aaaactgaaa	aactcctgca	gcaaatttac	acattgccac	840
taaacgtcta	aacccttgta	atttgttttt	gtttactat	gtgtgttatg	aacttgattt	900
tcaataattt	ttaaatttgg	taccagtatt	ataacatctt	ttgtgctaac	ggttgccaac	960
acagcaattt	gtaagttgat	taattgattc	tatactttta	ttgtctcctt	aattcatgct	1020
gataaatata	tgctgataaa	aattaaagt	aatatggtac	cacaaggttt	ggagatttaa	1080
ttgttgcaat	gctgcatgga	tggcatatac	accaaacatt	caataattct	tgaggataat	1140
aatggtacca	cacaagcttt	gagtgcatg	aacatcgcgt	ggacaaaagg	tttagtaatt	1200
tttcaagaca	acaatgttac	cacacacaag	ttttgaggtg	catgcatgga	tgccctgtgg	1260
aaagtttaaa	aatattttgg	aatgatttg	catggaagcc	atgtgtaaaa	ccatgacatc	1320
cacttgagg	aagcaataat	gaagaaaact	acaaatttac	atgcaactag	ttatgcatgt	1380
agtctatata	atgaggattt	tgcaataact	tcattcataa	acactcacta	agttttacac	1440
gattataatt	tcttcatagc	cagtc				1465

<210> 11

<211> 1302

<212> DNA

<213> 蚕豆(*Vicia faba*)

<400> 11

cctatgtag	ttttatcgga	cccctaaaa	gtacaatgag	ataaccttat	gttttttagtg	60
ttgttatatg	tgtttagagg	gtttgctcat	gcatagtgga	tagaagtaag	gttgccaaga	120
tcaaaatttt	acataagatc	aggagagaat	aacaagattg	taaaatataa	aattctagtt	180
aagataagaa	gattttactc	aattattttt	gtaagattga	cagggaataa	cataatagta	240
aattataata	tagtaacaaa	aaaagtaaga	tcctactaaa	ataaaaaaaaa	taatagtata	300
tatttaaaagt	attttaacat	gatataatga	tgtatgtaag	tatgcaagtg	tatttgtgag	360
aaaaaaaaaa	aactcttttt	catattcttt	aacattttat	gattgaggtt	ttattaaata	420
tttgtaatg	tttagacatt	agagacatat	atggtcaatt	attaggatta	tcataagttt	480
agttaaaata	aatcttata	gttaaaagt	tattctctaa	accttatatc	taaaatttca	540
tttcaaaagg	ccaaaattta	tctccgctga	accgtaatta	ttgtgcttcc	gcatcttgtc	600
atcattctca	tattttaata	gtgggtgctg	taaaacggtg	aatcatggtc	aaagctgaca	660
aagatcgcac	aaaatccata	attttataga	ttttttattc	aatttagcta	caatacgcga	720
ttctacttga	aatcgagatt	ttgacaacta	tacatattca	tatacagata	aatatagtc	780
tatgttatac	atttgctttt	aaaaaaaaaa	actgaaaaac	tctgcagca	aatttacaca	840
ttgccactaa	acgtctaaac	ccttgtaatt	tgtttttggt	ttaatattgtg	tgttatgaac	900
ttgatttga	ataattttta	aatttggtag	tagtattata	acaccttttg	tgctaacggg	960
tgccaacact	tagtaatttg	taagttgatt	aattgattct	aaactattat	tgtcttctta	1020
aatcatatcc	taaataatcg	aaaatgtaa	tatatgctga	taaaatttaa	agtgaatatg	1080
gtaccacaag	tttttgaaa	gtttaaaaat	attttgaaa	tgatttgcac	ggaagccatg	1140
tgtaaadcca	tgacatccac	ttggaggaag	caataatgaa	gaaaactaca	aatttacatg	1200
caactagtta	tgcatgtagt	ctatataatg	aggattttgc	aatactttca	ttcatacaca	1260

ctcactaagt ttacacgat tataãtttct tcatagccag tc

1302

<210> 12

<211> 27

<212> DNA

<213> 合成的

<400> 12

gataaaacag tgagatgtgc aaactcc 27

<210> 13

<211> 22

<212> DNA

<213> 合成的

<400> 13

gtaatacgac tcactatagg gc 22

<210> 14

<211> 38

<212> DNA

<213> 合成的

<400> 14

ccatggagat ctgactggct atgaagaaat tataatcg 38

<210> 15

<211> 19

<212> DNA

<213> 合成的

<400> 15

actatagggc acgcgtggt 19

<210> 16

<211> 25 <212> DNA

<213> 合成的

<400> 16

ctgcaggtcg acggcccggg ctggt 25

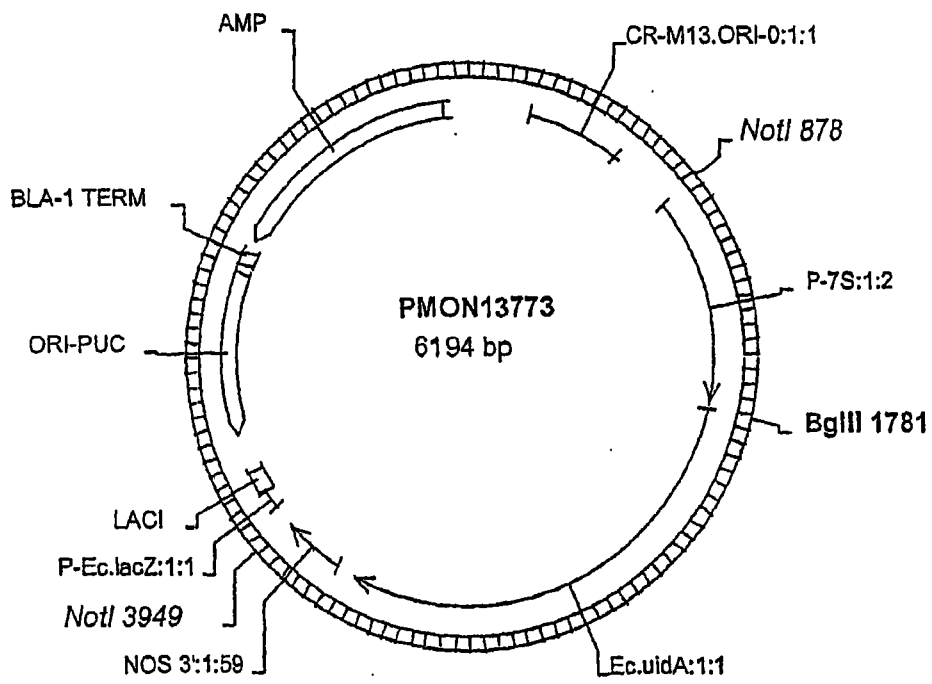


图 1

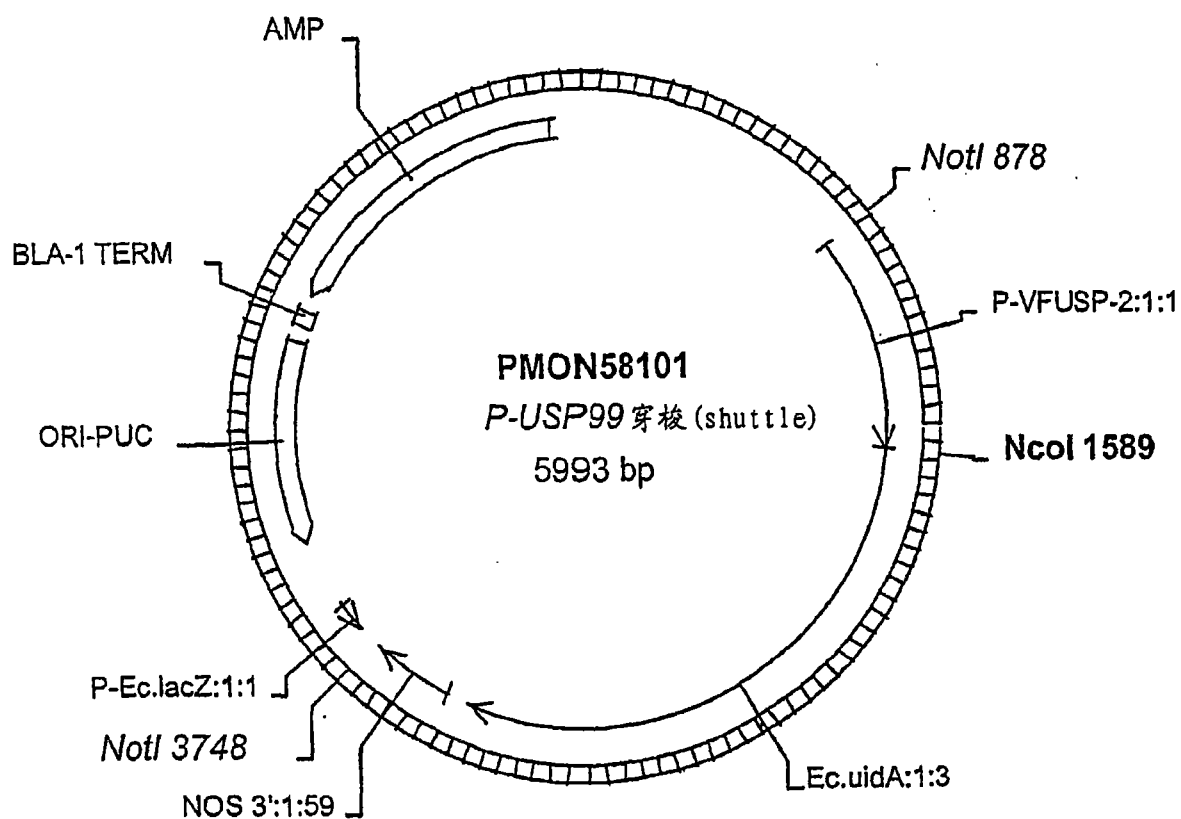


图 2

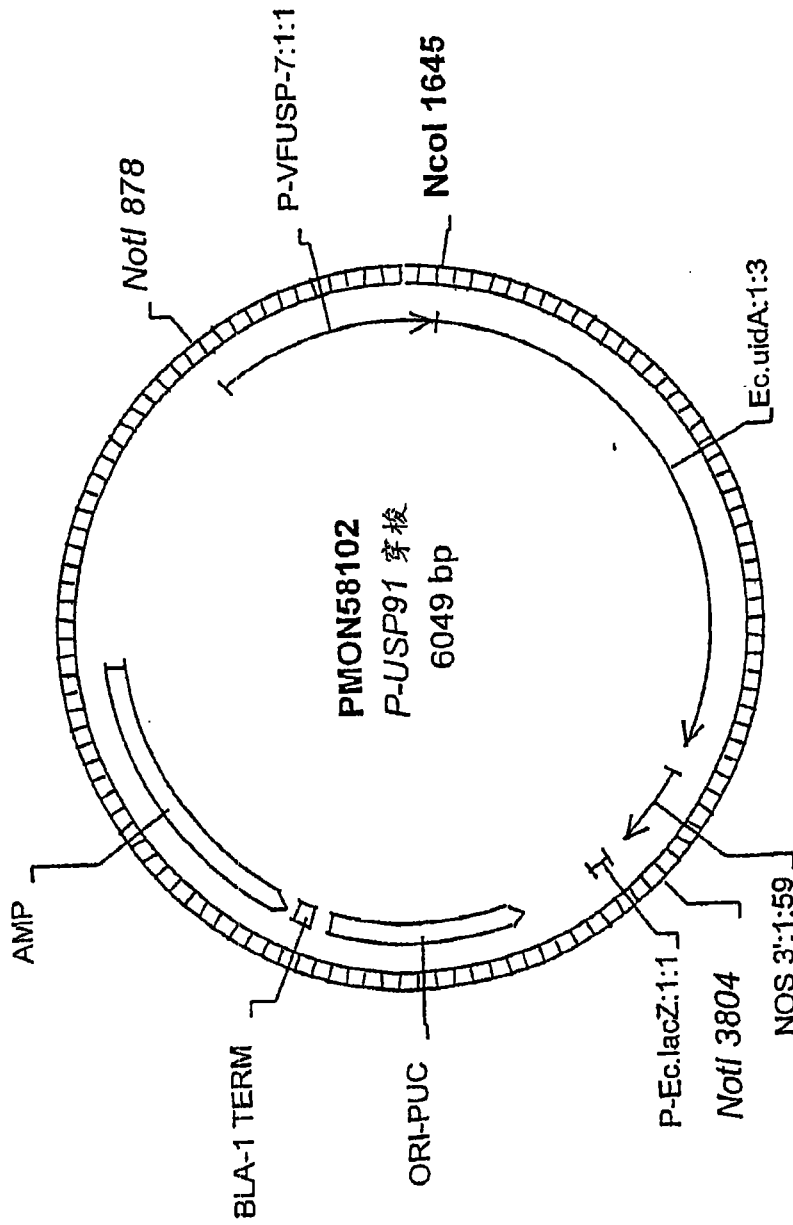


图 3

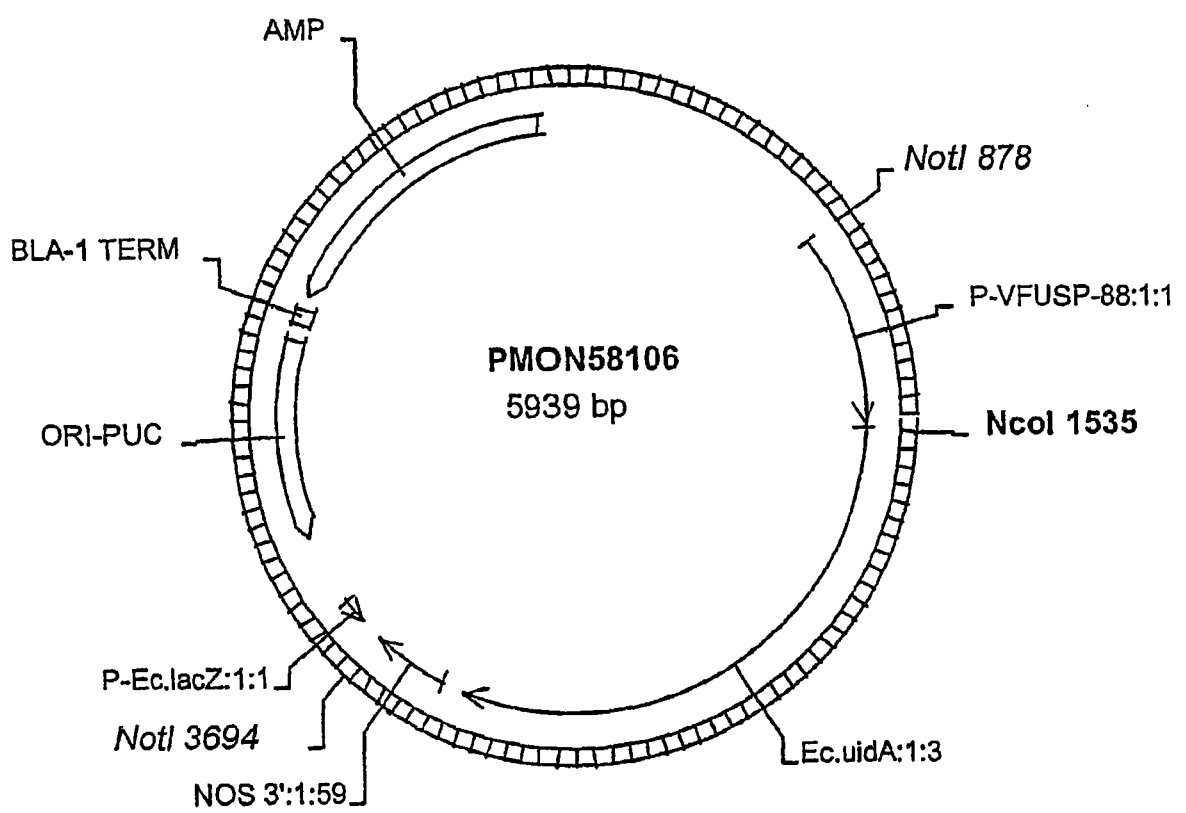


图 4

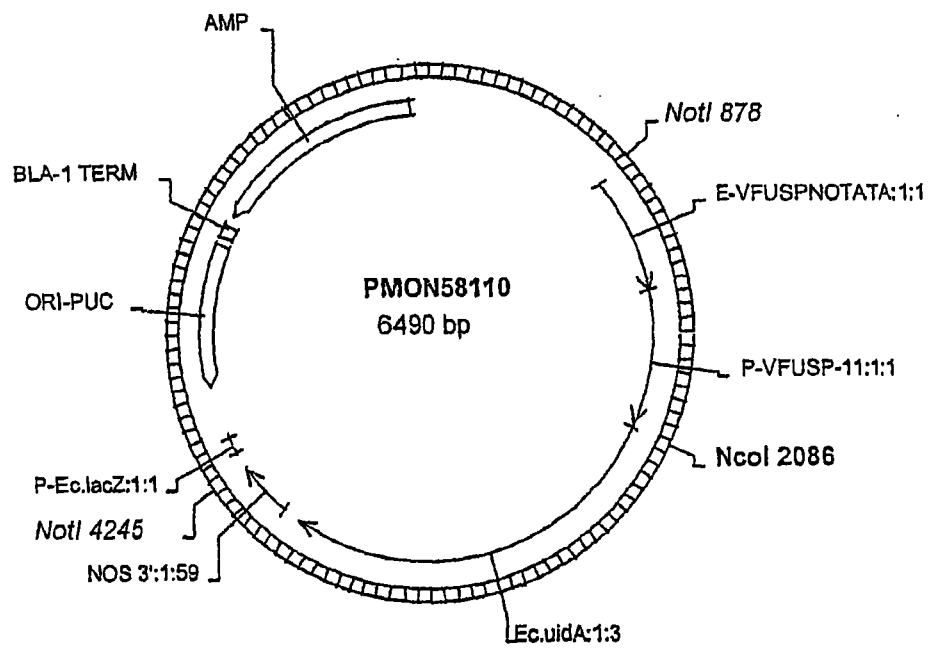


图 5

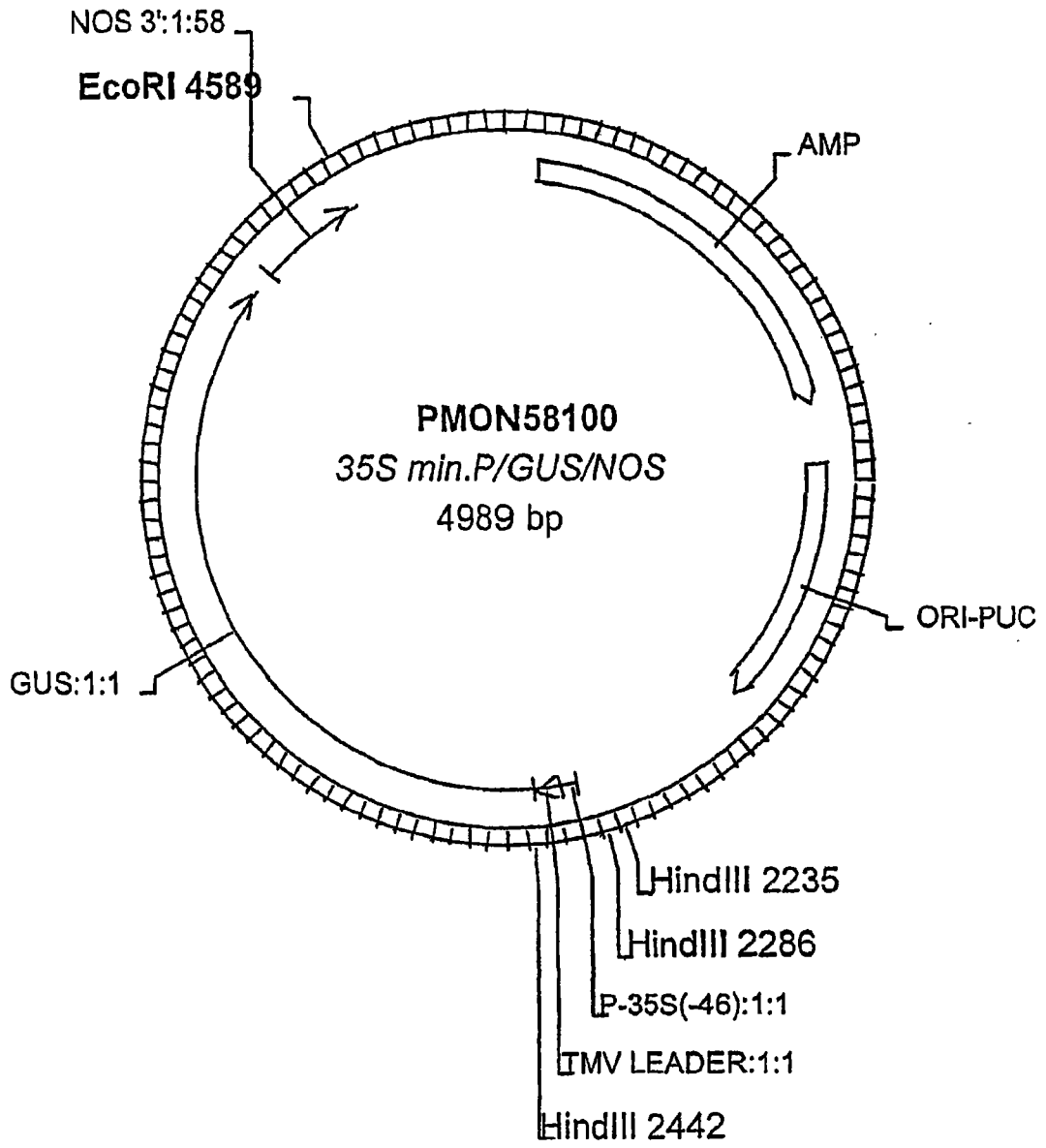


图 6

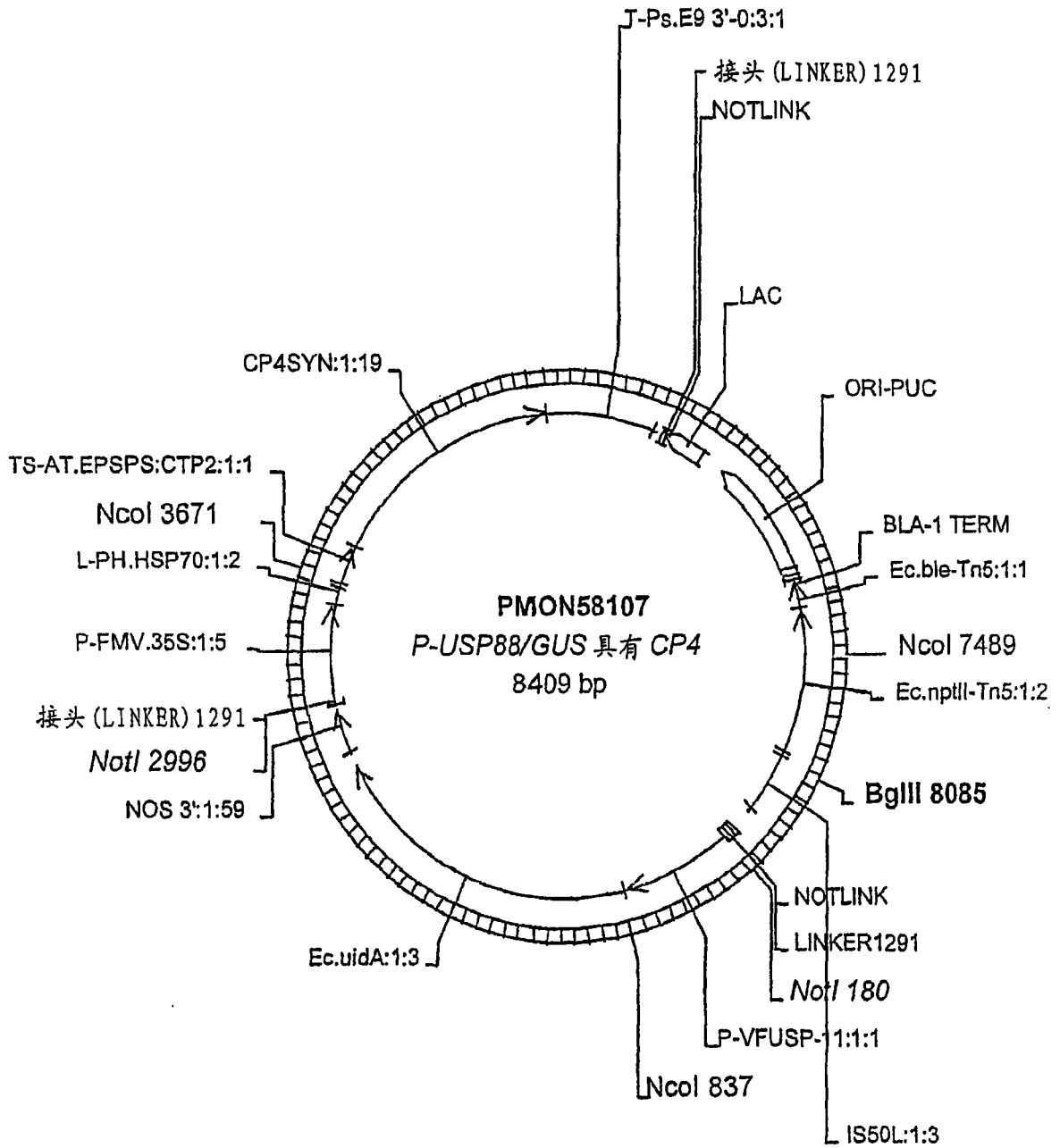


图 7

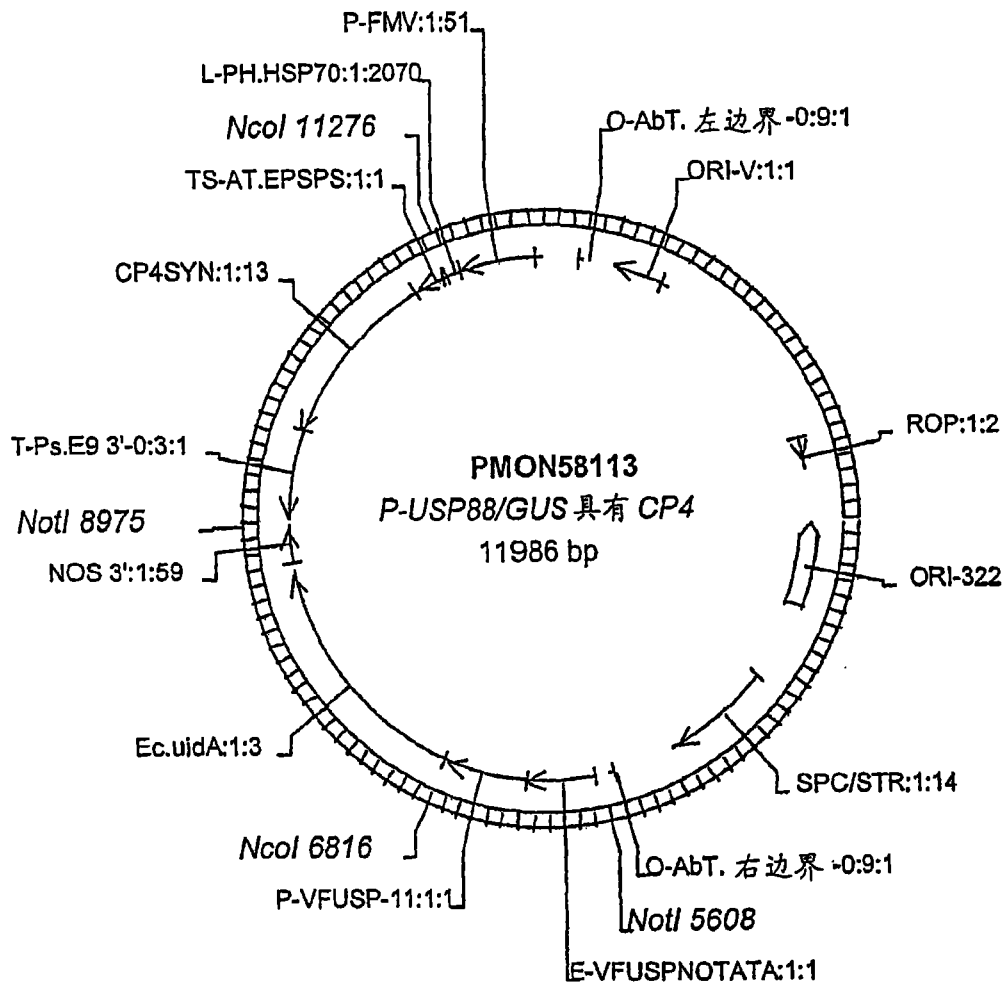


图 8

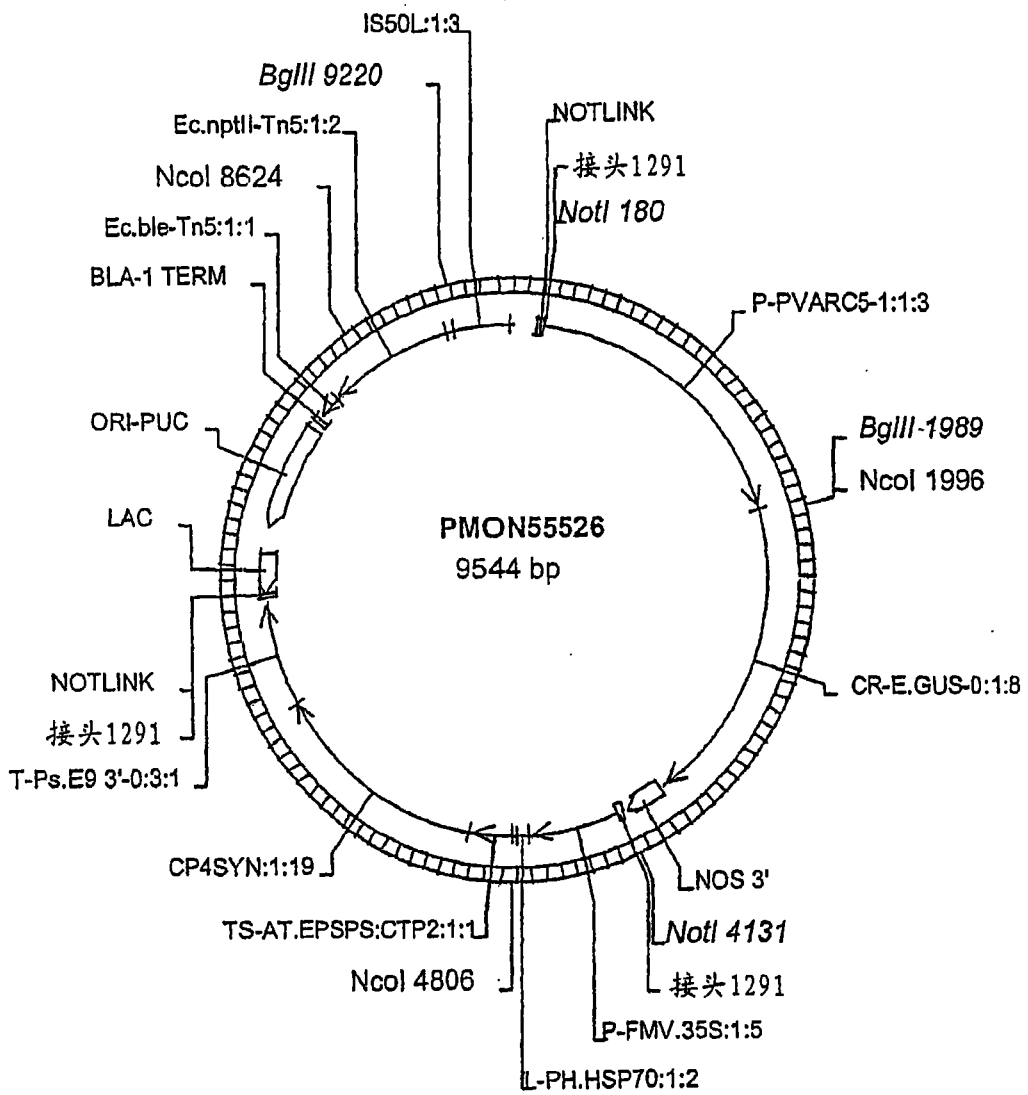


图 9

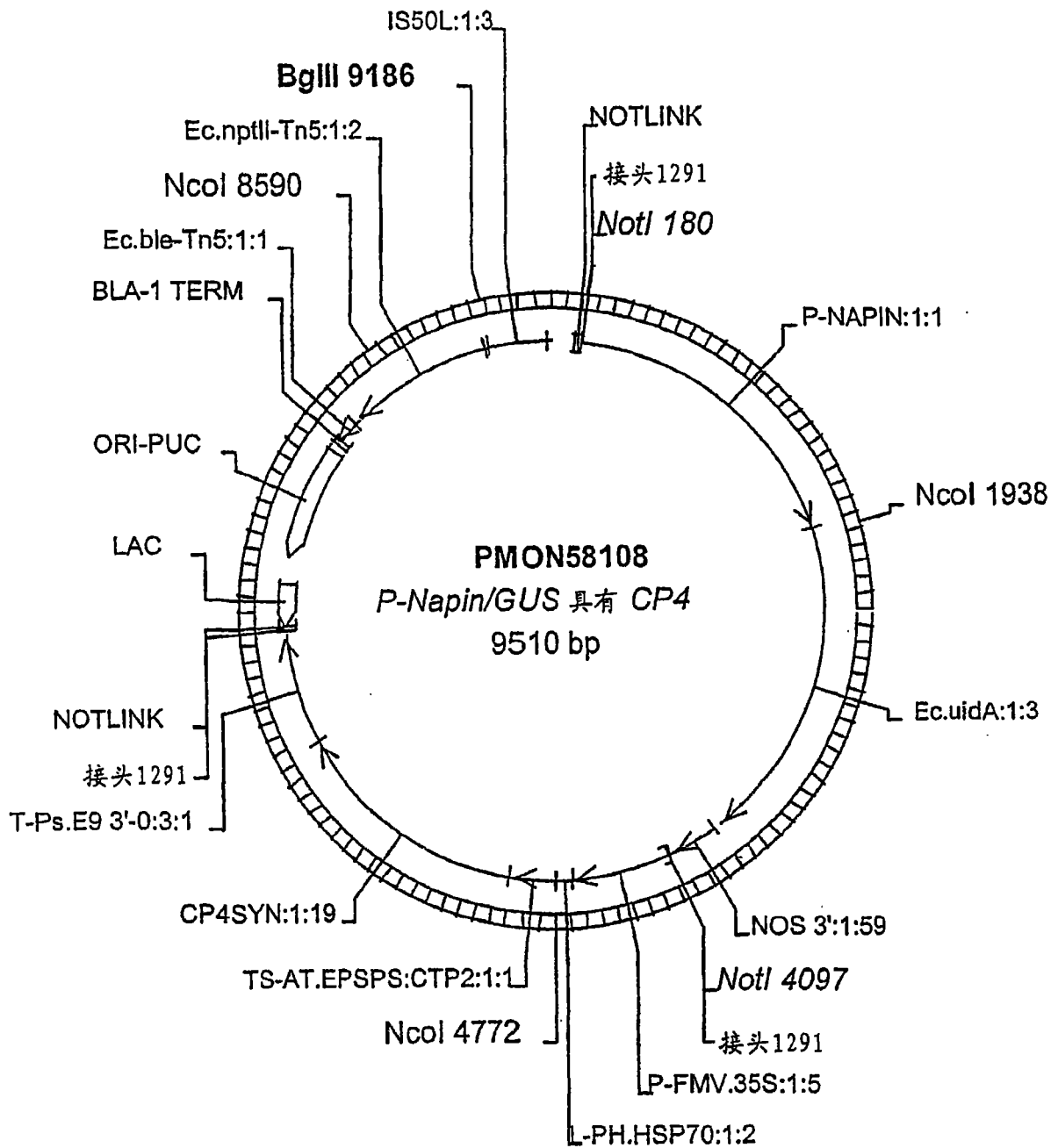


图 10

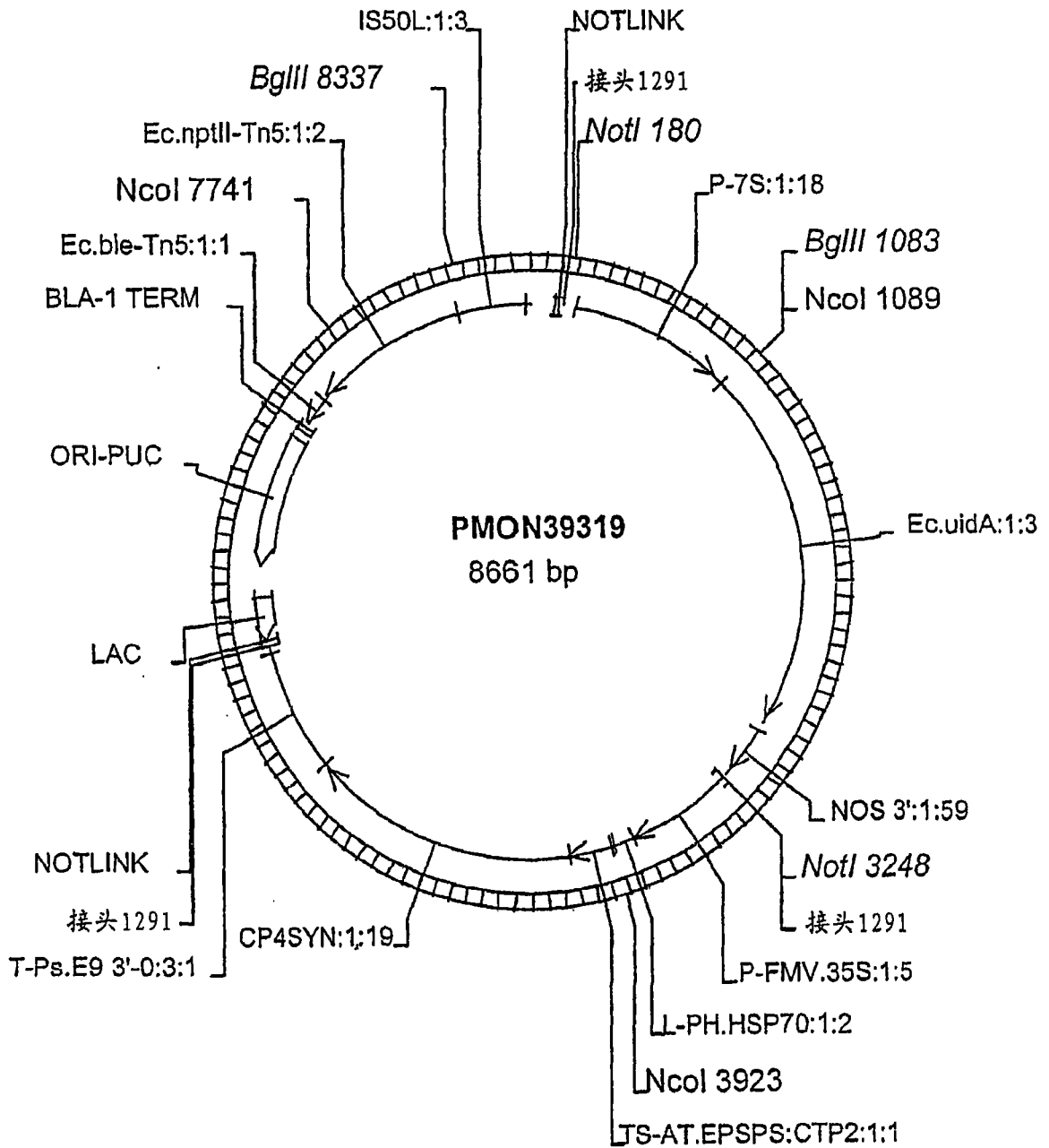


图 11

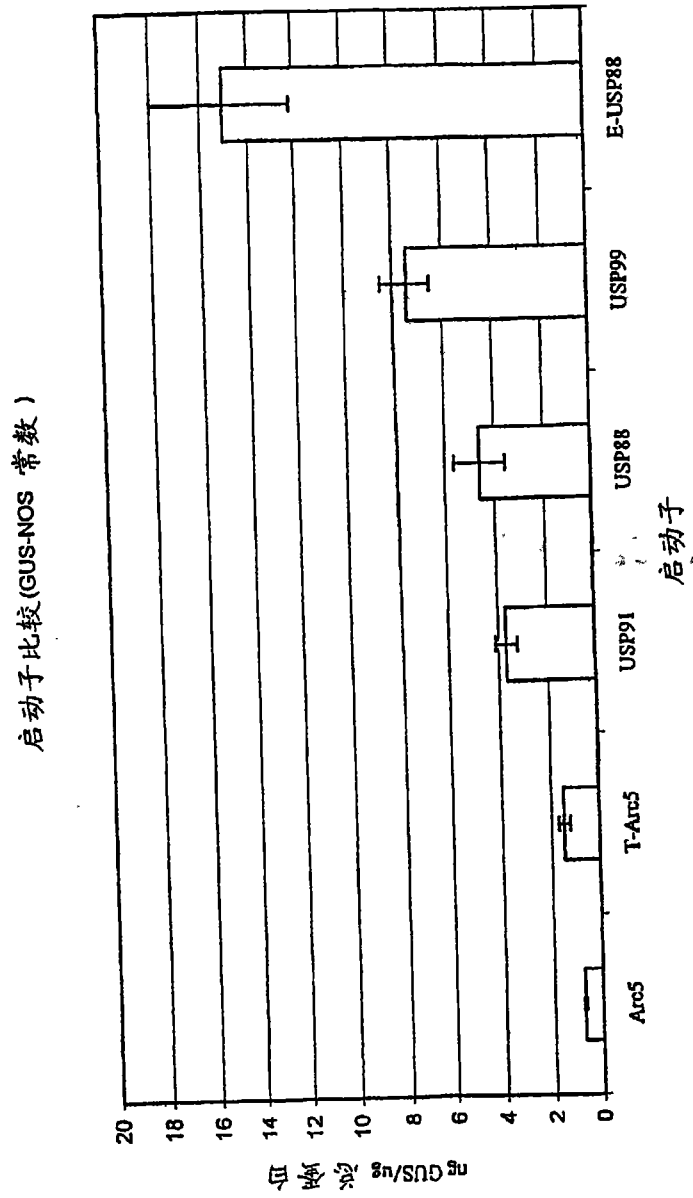


图 12

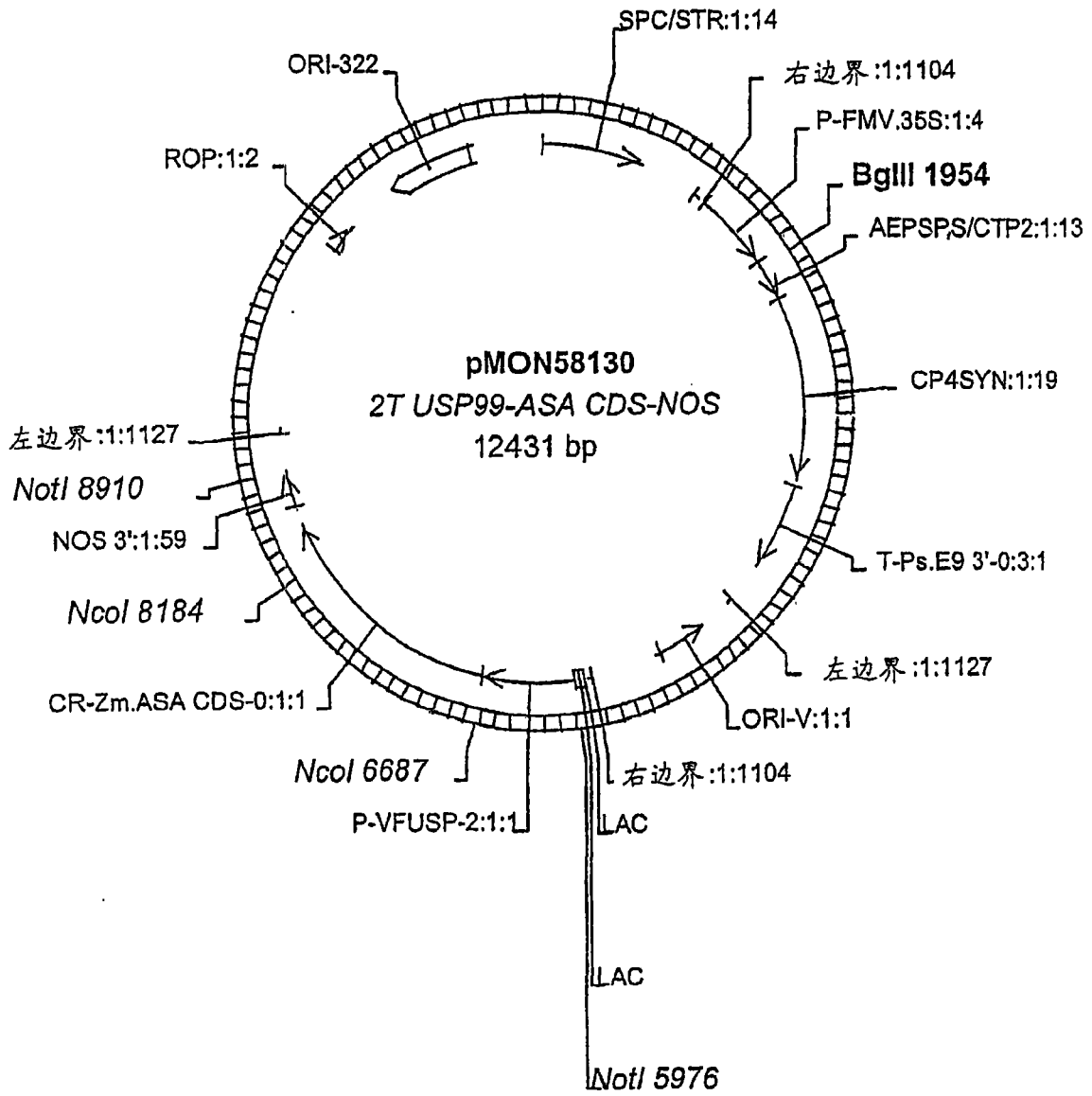


图 13

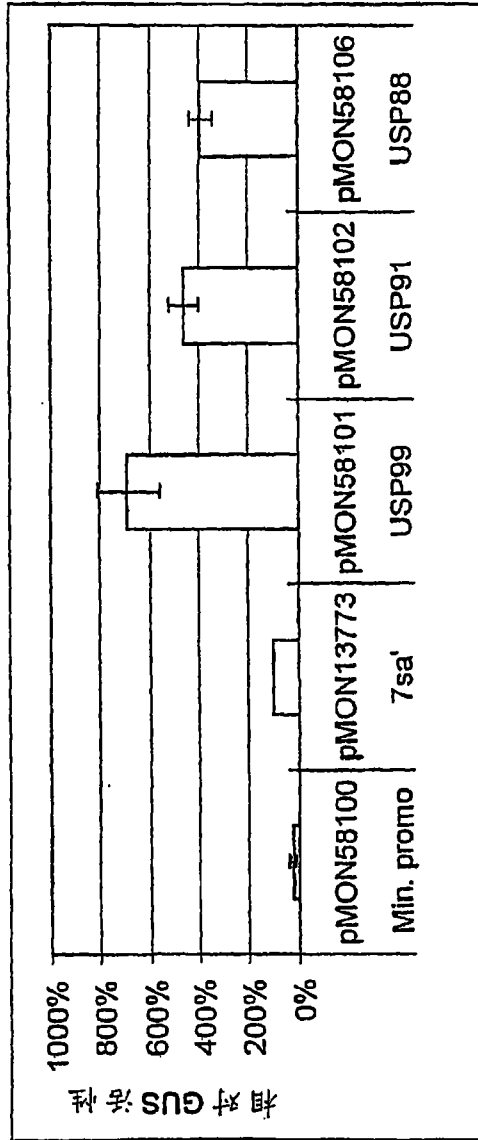


图 14

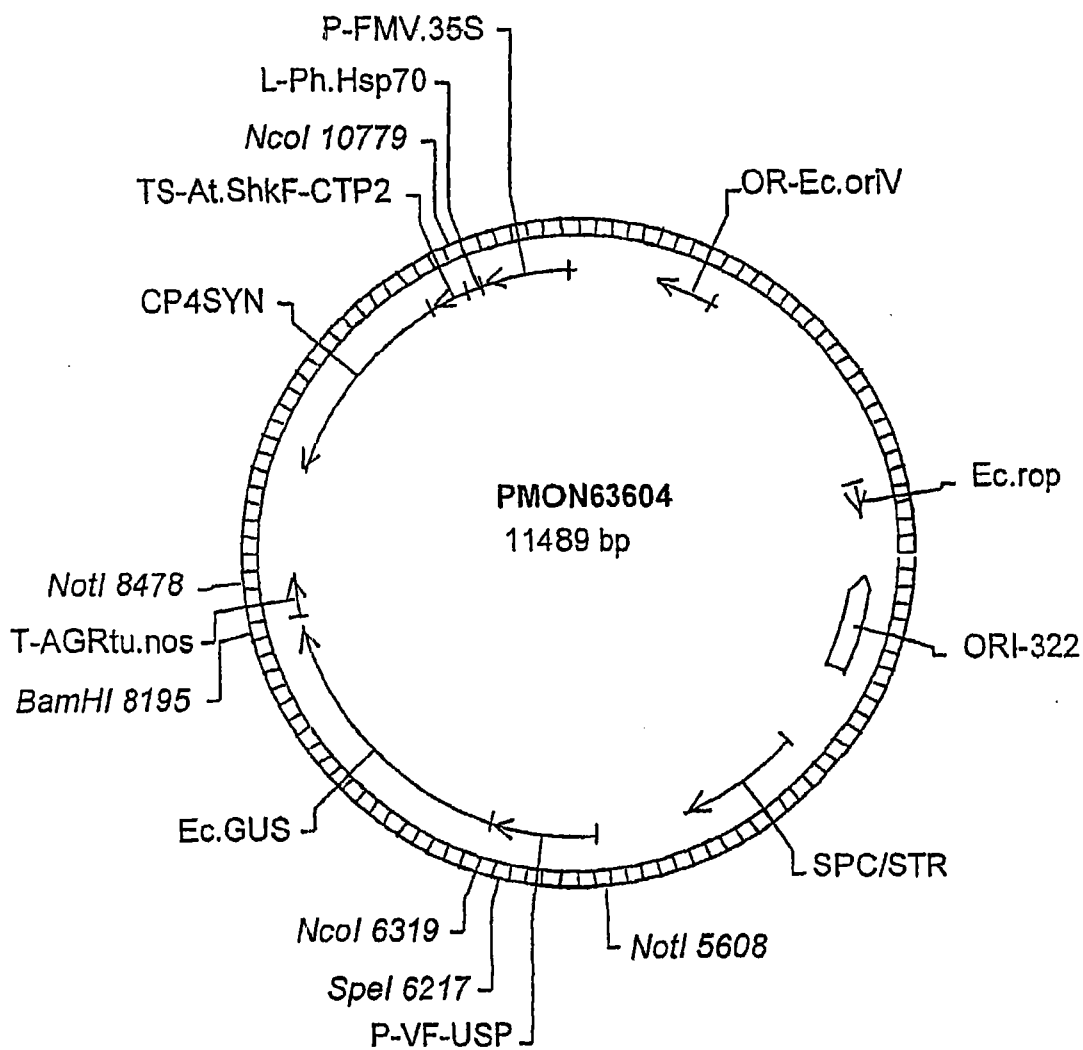


图 15

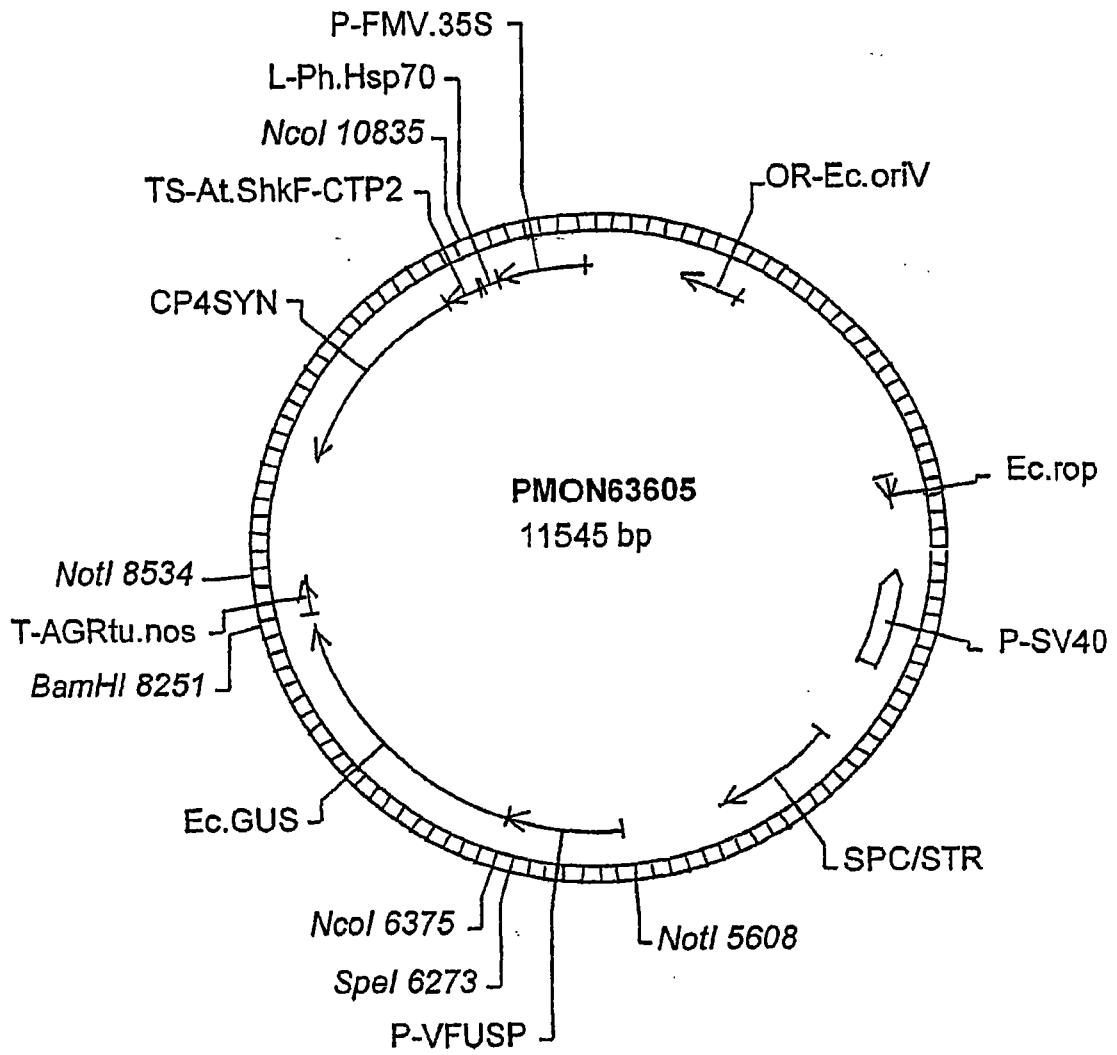


图 16

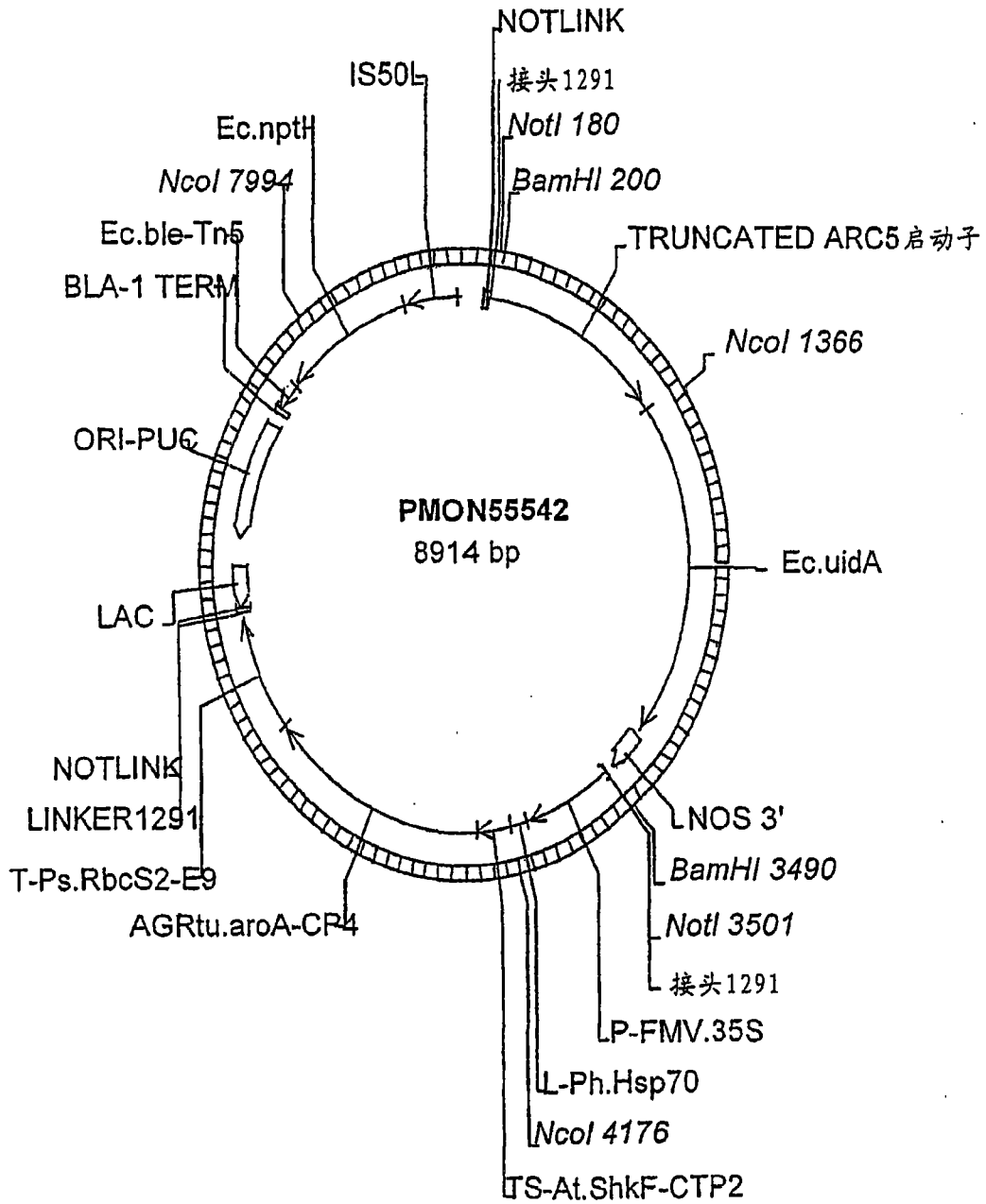


图 17

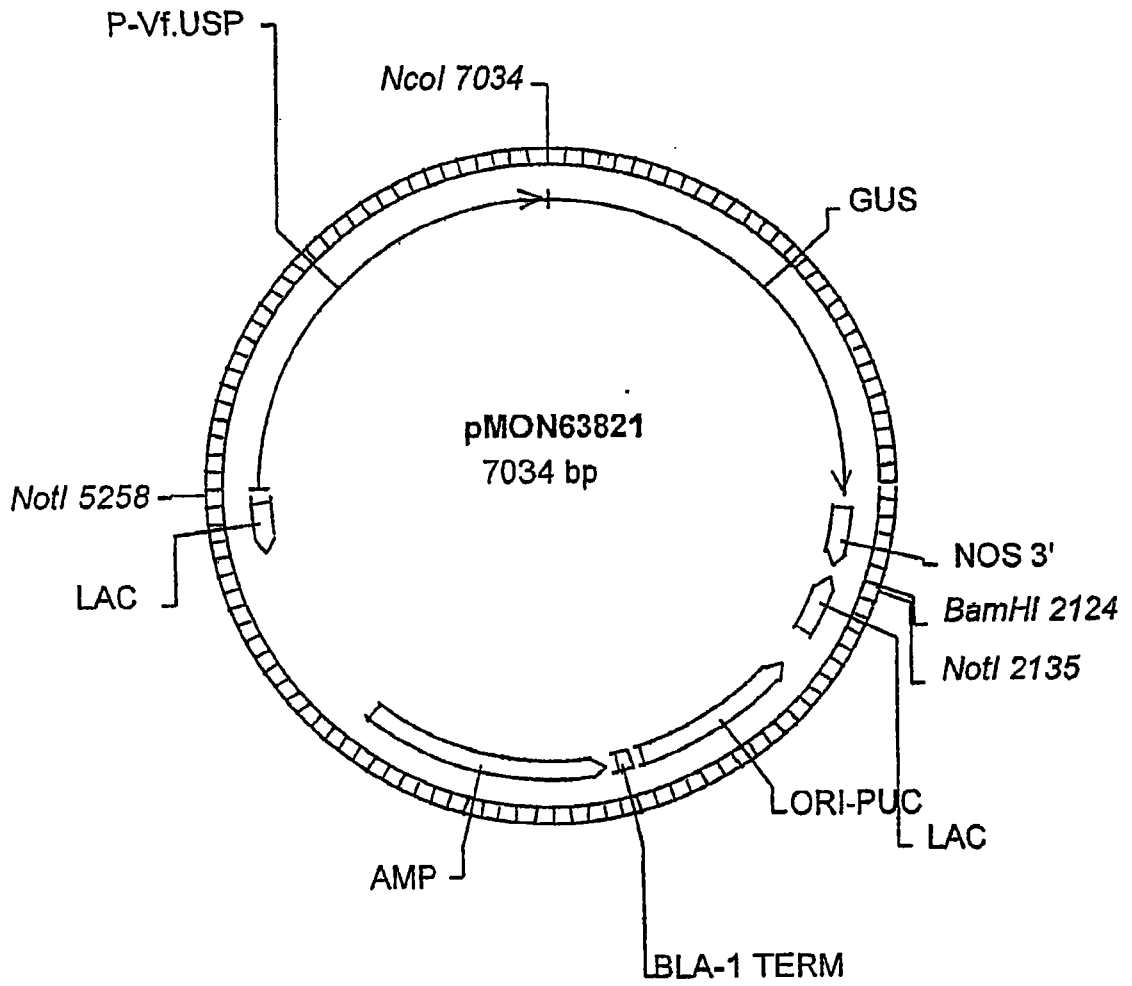


图 18

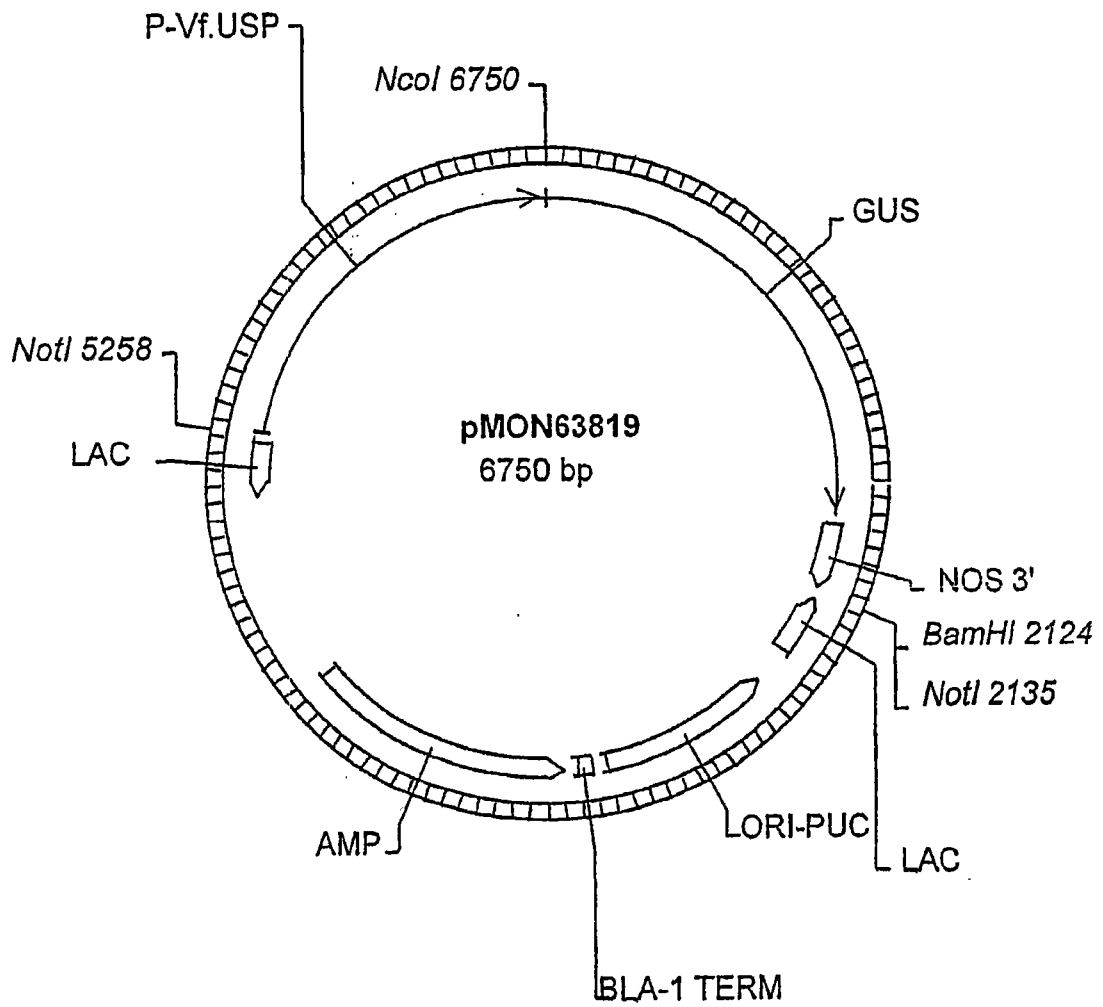


图 19

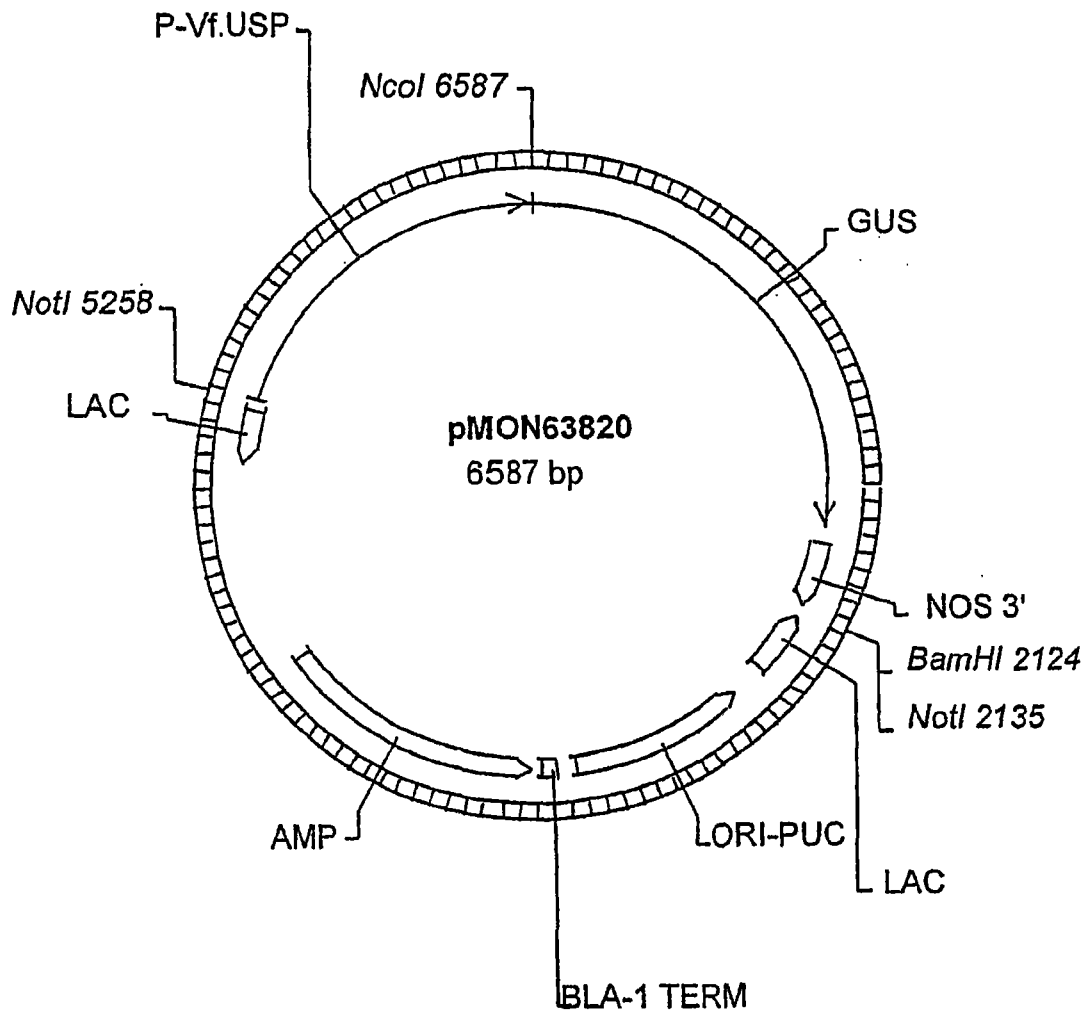


图 20

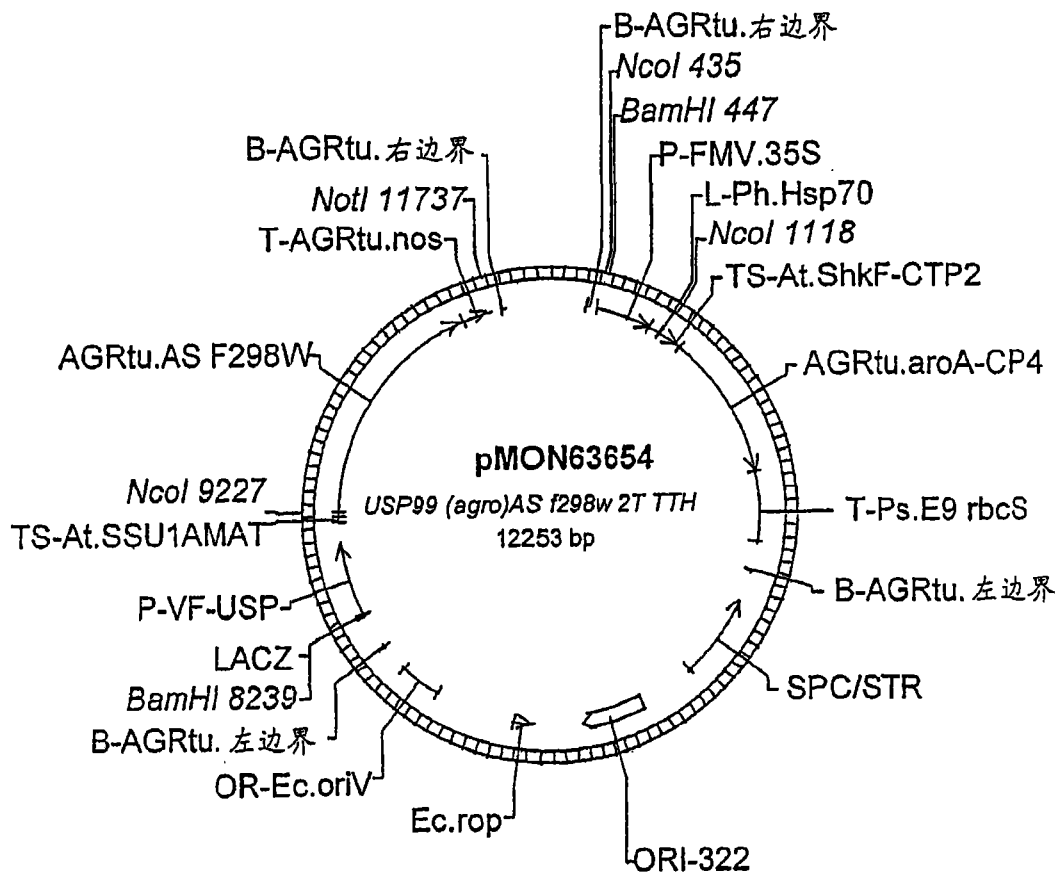


图 21