



(12) 发明专利申请

(10) 申请公布号 CN 102361976 A

(43) 申请公布日 2012. 02. 22

(21) 申请号 201080010816. 2

(72) 发明人 J·L·罗西尚

(22) 申请日 2010. 03. 02

(74) 专利代理机构 上海专利商标事务有限公司
31100

(30) 优先权数据

61/157, 759 2009. 03. 05 US

代理人 杨帆

(85) PCT申请进入国家阶段日

2011. 09. 02

(51) Int. Cl.

C12N 15/00 (2006. 01)

(86) PCT申请的申请数据

PCT/US2010/025872 2010. 03. 02

(87) PCT申请的公布数据

W02010/101884 EN 2010. 09. 10

(71) 申请人 罗门哈斯公司

地址 美国宾夕法尼亚州

权利要求书 2 页 说明书 21 页

序列表 9 页 附图 20 页

(54) 发明名称

在植物中控制关键乙烯激素信号传导途径蛋白的目标周转率以调节乙烯敏感性

(57) 摘要

一种在植物细胞中可控地表达乙烯应答的基因表达系统,其包括激活盒,该激活盒包括识别效应元件的 DNA 结合域;蜕皮激素受体配体结合域;和激活域;以及靶标盒,该靶标盒包括诱导性启动子,其包含操作性相连的效应元件和对激活域起反应的最小启动子。该诱导性启动子控制编码所选调控蛋白的核酸序列的表达,从而改变植物中某种信号蛋白的乙烯敏感性。在植物细胞中,在诱导组合物存在时激活盒和靶标盒的组件之间的相互作用能提高所选调控蛋白的表达,进而降低信号蛋白在植物中的表达和蓄积,从而降低植物细胞的乙烯敏感性。可通过施加诱导组合物的时机、浓度和持续时间来控制该调控蛋白表达的增加(特别是在存在乙烯时)。提供在诱导组合物存在时能控制乙烯敏感性的转基因植物的细胞、组织、器官和完整植株。可控制这种转基因植物和组织的乙烯敏感性和/或乙烯生产以操纵熟化、开花、枯萎及植物的其他乙烯敏感性功能。

1. 一种可控地抑制乙烯诱导性蛋白在植物细胞中蓄积的基因表达系统,其包括:

激活盒,激活盒包括在与其操作性相连的合适启动子控制下的以下组件:(a) 识别选定效应元件的 DNA- 结合域 (DBD);(b) 蜕皮激素受体配体结合域 (EcRLBD);和 (c) 在诱导组合物存在下活化的激活域 (AD);和

靶标盒,其包括 (d) 诱导型启动子,该启动子包含操作性相连的效应元件和最小启动子,(a) 的 DBD 结合所述效应元件,最小启动子对 (c) 的 AD 起反应,所述诱导型启动子控制 (e) 编码表达时能降低 EIN2 或 EIN3 基因产物表达的所选调控蛋白的核酸序列的表达;

在植物细胞中,所述激活盒和靶标盒操作性相连,以便通过与诱导组合物相互作用可控地提高所选调控蛋白的表达,以抑制乙烯存在下 EIN2 或 EIN3 基因产物在植物中的蓄积,对所述 EIN2 或 EIN3 蛋白表达蓄积的抑制作用可受施加诱导组合物的时机、浓度和持续时间的控制。

2. 如权利要求 1 所述的系统,其特征在于,所述激活盒和靶标盒位于同一质粒上。

3. 如权利要求 1 所述的系统,其特征在于,所述激活盒和靶标盒分别位于单独的质粒上。

4. 如权利要求 1 所述的系统,其特征在于,所述激活盒的启动子选自:组成型启动子、植物细胞特异性启动子、植物组织特异性启动子、植物器官特异性启动子、诱导型启动子、发育特异性启动子和细胞分化特异性启动子。

5. 如权利要求 4 所述的系统,其特征在于,所述启动子是组成型启动子。

6. 如权利要求 1 所述的系统,其特征在于,所述 DBD 选自:GAL4 DBD、LexA DBD、转录因子 DBD、H 类核受体成员 DBD、甾体 / 甲状腺激素核受体超家族成员 DBD、细菌 LacZ DBD、EcR DBD、cI 启动子 DBD。

7. 如权利要求 1 所述的系统,其特征在于,所述蜕皮激素受体 EcRLBD 包括无脊椎动物蜕皮激素受体或其突变体的全部或其一部分。

8. 如权利要求 1 所述的系统,其特征在于,所述激活域选自:VP16、糖皮质激素激活域和 Gal4 激活域。

9. 如权利要求 1 所述的系统,其特征在于,所述诱导型启动子包含操作性相连的最小启动子序列以及一个或多个拷贝的对应于激活盒中 DNA 结合域的效应元件。

10. 如权利要求 1 所述的系统,其特征在于,编码所选蛋白质的所述核酸序列选自 EIN2 调控基因 ETP1 和 ETP2,结合 EIN3 的 F 盒蛋白,EBF1 和 EBF2,以及它们的突变体。

11. 如权利要求 1 所述的系统,其包括:

激活盒,包括与其操作性相连的组成型 G10-90 启动子控制下的以下组件:(a) 识别包含 5 个拷贝的 GAL4 效应元件的效应元件的 GAL4 DBD;(b) 蜕皮激素受体 LBD;和 (c) 在诱导组合物存在时活化的 VP16 AD;和

靶标盒,其包括 (d) 诱导型启动子,该启动子包含操作性相连的 5 个拷贝的 GAL 4 效应元件和对 VP16 AD 的活化起反应的最小 35S 启动子,所述诱导型启动子控制 (e) 编码 ETP1 或 ETP2 蛋白的核酸序列的表达;

在植物细胞中,所述激活盒和靶标盒操作性相连,以便通过与诱导组合物相互作用可控地提高 ETP1 或 ETP2 的表达,以抑制乙烯存在下 EIN2 基因产物在植物中的蓄积,对所述 EIN2 蛋白表达蓄积的抑制作用可受施加诱导组合物的时机、浓度和持续时间的控制。

12. 一种组合物,其包含稳定表达权利要求 1 或 11 所述的基因表达系统的转基因植物细胞、转基因植物组织或器官、或者转基因植物。

13. 一种制备转基因植物的方法,该方法包括:

(a) 用权利要求 1 或 11 所述的基因表达系统转化植物的至少一个细胞;

(b) 由转化的植物细胞形成植物;和

(c) 选择包含转化的植物细胞的植物,所述植物在乙烯存在下接触诱导组合物时能够表现出乙烯不敏感性,所述调节受施加所述诱导组合物的时机、浓度和持续时间的控制。

14. 一种调节植物乙烯敏感性的方法,该方法包括:

将诱导组合物施加于转基因植物的细胞,所述植物包括稳定表达权利要求 1 或 11 所述基因表达系统的细胞,通过施加所述诱导组合物的时机、浓度或持续时间过度表达所述调控蛋白,以降低或抑制乙烯存在下植物细胞中乙烯诱导性信号蛋白的蓄积。

15. 如权利要求 14 所述的方法,其特征在于,所述诱导组合物是二酰肼化合物。

16. 如权利要求 14 所述的方法,其特征在于,所述受控的乙烯敏感性选自下组:枯萎、果实熟化、应激反应、发芽、病原体抗性、叶片脱落、花蕾脱落、芽脱落、荚壳脱落、果实脱落、开花以及对干旱、热、群体密度和盐度的反应。

在植物中控制关键乙烯激素信号传导途径蛋白的目标周转率以调节乙烯敏感性

[0001] 发明背景

[0002] 植物激素乙烯是调节植物生命周期中的许多生理学过程的信号传导分子,这些生理学过程包括发芽期间、花和果实发育过程中,以及植物对各种环境应激源的反应,例如干旱、热、盐度过量以及疾病的反应(参见例如,Chen等,2005, *Annals of Botany*, 95: 901-915; Czarny等,2006 *Biotechnol. Adv.*, 24: 410-419)。乙烯生物合成途径和信号传导/调节途径及网络是公知的。例如,参见图1和2, Wang等,“乙烯生物合成和信号传导网络”(Ethylene Biosynthesis and Signaling Networks),《植物细胞》(*The Plant Cell*), 2002(美国植物生物学家协会(American Society of Plant Biologists)编著),第S131-S151页。

[0003] 在植物中,乙烯信号传导途径中的几个关键步骤受到高度调节。就EIN2(乙烯不敏感2)和EIN3(乙烯不敏感3)蛋白而言,它们的表达受乙烯诱导,其导致乙烯应答增加。此外,ETP1(EIN2靶向蛋白1)和ETP2(EIN2靶向蛋白2)或EBF1(结合EIN3的F盒蛋白1)和EBF2(结合EIN3的F盒蛋白2)分别靶向EIN2和EIN3蛋白的周转。在不存在乙烯激素的情况下,这些关键应答信号传导蛋白的周转帮助将植物维持在抑制或“关闭”状态。在转基因拟南芥(*Arabidopsis*)中组成型过表达ETP1或ETP2(Qiao等, *Genes Dev.*, 23: 000-000, 2008年2月12日在线公开)或EBF1或EBF2(Guo和Ecker, 2003 *Cell*, 115: 667-677)分别导致部分乙烯不敏感和EIN2或EIN3蛋白的蓄积减少。

[0004] 商业上,调节包括水果、蔬菜和花卉等植物中的乙烯效应的常规途径涉及将化学物质施用于植物、水果、花卉或蔬菜,例如1-甲基环丙烯(1-MCP;农新有限公司(AgroFresh Inc.))。1-MCP是可用作植物生长调节剂,防止乙烯与其在植物组织中受体结合的化合物。其应用增加了植物对乙烯的不敏感性。乙烯敏感性的短暂消除可能增加植物对应激的抵抗力或耐受、延迟成熟、枯萎或开花,以及进行其他有商业价值的植物生长操纵。

[0005] 近来,提出了用修饰的乙烯效应受体或参与植物中乙烯效应的其他蛋白质遗传转化植物细胞的建议,例如参见美国专利6294716;美国专利申请公开2006/0200875、2005/0066389、2005/0060772和2004/0128719等。这些系统针对乙烯途径中的各种突变基因的表达。这些系统通常采用各种建议的启动子来启动蛋白质的表达,包括组成型启动子和组织特异性启动子。

[0006] 虽然已经提出将化学调节的基因表达系统用于植物物种(M. Padidim 2003 *Curr. Opin Plant Biol.*, 6(2): 169-77),但许多这些系统仅仅是实验性的,或者已经报道存在一些缺陷。这些缺陷包括:使用毒性或挥发性诱导剂、诱导水平低、在植物中易位/移动差、“关闭”基因表达的能力慢、或者对植物非毒性诱导剂的特异性不够,等等。这些基因表达系统未能在所有植物中普遍使用,操作性组件及其组合件的选择常常充满挑战。

[0007] 在现有技术的例子中,乙烯途径基因常常总是在植物的所有组织和部件上或者总是在植物的特定组织上表达。然而,组织特异性启动子或低水平组成型启动子可能有漏洞或者由不希望的诱导剂诱导。这些常规启动子不能严格调节植物中的激素表达。乙烯途径

基因的表达时机、表达持续时间和表达水平是正常生理学功能的关键。随意诱导乙烯不敏感性并且保持预定的时间范围在现有技术中尚未成功实现。

[0008] 本领域仍然需要能够时间可控地调节乙烯敏感性的组合物和方法。这对于存在乙烯时直接诱导其表达的基因产物特别重要。需要将这种组合物和方法安全和有效地应用于农作物和食品,以及其它植物。

[0009] 附图简要说明

[0010] 图 1 是称为 p201 的质粒的一个例子的示意图,该质粒同时具有基因表达系统 (G10-90p-VGE-NosT-5xGAL-M35S-ETP1-35ST) 的激活盒和靶标盒,盒中各组件如实施例 1 和 SEQ ID NO :1 所示,括号中是以核酸位置表示的切割位点。该质粒的盒部分在 SEQ ID NO :1 中报道。序列表或图中未提供市售可得的质粒主链,这些主链可以容易地被其他质粒主链代替。

[0011] 图 2 是称为 p202 的质粒的一个例子的示意图,该质粒包含激活盒和靶向盒 (G10-90p-GVE-NosT-5xGAL-M35S-ETP2-35ST),利用 GVE 受体-介导的 ETP2 的诱导性表达来表达 ETP2。该质粒的盒部分在 SEQ ID NO :2 中报道。序列表或图中未提供市售可得的质粒主链,这些主链可以容易地被其他质粒主链代替。

[0012] 图 3A 是称为 p1004 的质粒的一个例子的示意图 [G10-90p-VGE-NosT-5xGAL-M35S-ETP1 和 MMVp-def-rbcS-E9t],其组件如 SEQ ID NO :3 和图 3B-3I 所示,即 G10-90 组成型启动子,VP16 激活域和 GAL4DNA 结合域,以及与 NOS 终止子序列相连的野生型蜕皮激素受体配体结合域,由五个拷贝的 GAL4 效应原件和最小 35S 启动子组成的诱导型启动子,ETP1 基因,35S 终止子序列,MMV 启动子,P-DEF 标记基因和 rbcS-E9 终止子。

[0013] 图 3B-3I 是显示 p1004 的组件的序列图。

[0014] 图 4A 是称为 p1005 的质粒的一个例子的示意图 [G10-90p-GVE-NosT-5xGAL-M35S-ETP2 和 MMVp-def-rbcS-E9t],其组件如 SEQ ID NO :4 和图 4B-4I 所示,即 G10-90 组成型启动子,GAL4 DNA 结合域,VP16 激活域,以及与 NOS 终止子序列相连的野生型蜕皮激素受体配体结合域,由五个拷贝的 GAL4 效应原件和最小 35S 启动子组成的诱导型启动子 (SEQ ID NO :4 的核苷酸 2493-2548),ETP2 基因,35S 终止子序列,MMV 启动子,P-DEF 标记基因和 rbcS-E9 终止子。

[0015] 图 4B-4I 是其显示 p1005 的组件的序列图。

发明内容

[0016] 本文所述组合物和方法通过提供能够以时序、定性和 / 或定量方式可靠而安全地控制乙烯敏感性调节的转基因植物、植物细胞、组织、器官、果实或花而满足了本领域的需要。这些组合物和方法能够严格调节基因表达和激素表达,可安全应用于农作物和食物以及其他有商业价值的植物。

[0017] 在一个方面,提供一种可控地抑制某些乙烯诱导性蛋白在植物细胞中蓄积的基因表达系统。该系统包括激活盒和靶标盒,激活盒和靶标盒可以位于一个或多个质粒上。激活盒包括合适的启动子、DNA-结合域 (DBD)、蜕皮激素受体配体结合域 (EcRLBD);和激活域 (AD)。靶标盒包括化学诱导的启动子 (可以是化学诱导性组织特异性启动子),该启动子包括操作性相连的效应元件和最小启动子,DBD 结合该效应元件,最小启动子对 AD 起反应。该

化学诱导性启动子控制着编码所选的调控蛋白或其片段的靶核酸序列的表达,表达时操作降低 EIN2 或 EIN3 基因产物的表达。植物细胞中两个盒中各组件与诱导组合物的相互作用能可控地增加调控蛋白的表达,并抑制存在乙烯时 EIN2 或 EIN3 基因产物在植物中的蓄积。可通过调节施加诱导组合物的时机、浓度和持续时间来控制对 EIN2 或 EIN3 基因产物蓄积的抑制作用。诱导组合物可被植物细胞吸收和易位到植物细胞内,在细胞内与激活域相互作用以启动靶标盒的化学诱导型启动子。因此,该系统允许以可控和选择性方式调节植物细胞的乙烯敏感性,通过使调控蛋白的表达水平足够高,以克服植物对乙烯的正常反应,即增加乙烯诱导性蛋白,如 EIN2 或 EIN3 的表达,进而在植物中导致乙烯信号传导途径下游的激活。

[0018] 在另一方面,提供了一种瞬时或稳定地表达上述基因表达系统的植物细胞。

[0019] 在另一方面,提供了一种瞬时或稳定地表达上述基因表达系统的植物组织或器官。

[0020] 在另一方面,提供了一种瞬时或稳定地表达上述基因表达系统的转基因植物。

[0021] 在另一方面,一种制备这种转基因植物或其部分的方法,该方法包括用本文所述的基因表达系统转化至少一个植物细胞;由转化的植物细胞产生植物细胞、组织、器官或完整植株;和当植物细胞、组织、器官或完整植株在乙烯存在下接触诱导组合物时选择显示能抑制或降低 EIN2 或 EIN3 的蓄积的植物细胞、组织、器官或完整植株。如上所述,可通过施加诱导组合物的时机、浓度和持续时间来控制对 EIN2/EIN3 蛋白蓄积的调控,导致调节性蛋白表达增加到足够高的水平。诱导组合物可以被植物细胞、组织、器官或完整植株吸收和易位到这些植物细胞、组织、器官或完整植株内。

[0022] 在又一方面,一种控制植物中由乙烯诱导蛋白如 EIN2 或 EIN3 表达导致的乙烯敏感性的方法,该方法包括将诱导组合物施加于转基因植物或其部分的细胞,所述植物包含能够稳定或瞬时表达本文所述基因表达系统的细胞。诱导组合物可以被植物细胞吸收和易位到植物细胞内。在诱导组合物存在下,植物细胞对乙烯的反应,即 EIN2 或 EIN3 的正常增加,以及乙烯信号传导途径下游的进一步激活在所选时间上受到抑制或降低;在所选时间过后通过去除植物中的诱导剂提高植物细胞对乙烯的反应。可通过施加诱导组合物的时机、浓度和持续时间,以及高水平表达调控蛋白使该植物对乙烯不敏感的能力,来控制 EIN2/EIN3 蛋白表达的调控。在一个实施方式中,诱导组合物的时机、浓度或持续时间允许过度表达调控蛋白,以降低或抑制乙烯存在下乙烯诱导性信号蛋白在植物细胞中的蓄积。

[0023] 下面的发明详述中将进一步描述这些方法和组合物的其他方面和优点。

[0024] 发明详述

[0025] 本文所述组合物和方法解决了本领域对于可控调节植物中乙烯敏感性的组合物和方法的需要。更具体说,本文所述的组合物和方法能够基于选定量化学诱导剂的使用和 EIN2/EIN3 调控蛋白(分别对应于 ETP1 和 ETP2 或 EBF1 和 EBF2)的高水平表达,有意改变乙烯诱导蛋白如 EIN2 或 EIN3 的表达水平,以使该植物对乙烯不敏感。这种操控植物激素调节的能力提供了农作物的生长和成熟的农业优势以及以下所述的其他优点。

[0026] 本发明人已确定,实现实践中乙烯不敏感性的关键在于获得足够水平的靶向关键乙烯诱导性信号蛋白周转的调控蛋白,以克服乙烯诱导的信号蛋白的量。使植物对乙烯不敏感能够为植物提供某些益处,例如本文所述的应激耐受性。然而,在植物正常生长和发育

的某些时间点,需要乙烯敏感性。因此,本文所述的组合物和方法可用于精确控制植物对乙烯不敏感的时机和水平,以便能够使该植物恢复到对乙烯敏感的状态,从而保证该植物进一步的正常生长和发育。

[0027] I. 基因表达系统

[0028] 利用基因表达或调节系统在植物细胞中稳定或瞬时表达。该系统的组件包括至少两个基因表达盒,每个盒能够在植物细胞中表达。

[0029] 在一个实施方式中,第一基因表达盒称为激活盒,其包括可以在与其操作性相连的合适启动子控制下在植物细胞中表达的多核苷酸,该多核苷酸编码以下组件:(a) DNA-结合域(DBD),DBD识别与准备调节其表达的基因,即编码靶向关键乙烯诱导性信号蛋白如EIN2和EIN3的周转的调控蛋白,如ETP1/ETP2或EBF1/EBF2的基因相关的效应元件;(b)配体结合域(LBD),LBD包括蜕皮激素受体配体结合域(EcRLBD)或其功能片段;和(c)激活或反式激活域(AD),AD在适用于植物的诱导组合物存在下激活。在一个实施方式中,激活盒中的组件采用以下5'到3'的顺序存在:LBD位于DBD下游,而DBD在AD下游。在另一实施方式中,激活盒中的组件采用以下5'到3'的顺序存在:LBD位于AD下游,而AD在DBD下游。在另一实施方式中,激活盒中的组件采用以下5'到3'的顺序存在:DBD位于LBD下游,而LBD在AD下游。在另一实施方式中,激活盒中的组件采用以下5'到3'的顺序存在:DBD位于AD下游,而AD在LBD下游。在另一实施方式中,激活盒中的组件采用以下5'到3'的顺序存在:AD位于LBD下游,而LBD在DBD下游。在另一实施方式中,激活盒中的组件采用以下5'到3'的顺序存在:AD位于DBD下游,而DBD在LBD下游。激活盒还包括优选位于盒3'末端的终止子。下面将描述这些组件的具体特征。

[0030] 第二基因表达盒即靶标盒,其包括编码以下组件的多核苷酸。一个组件是化学诱导的启动子,其包括操作性相连的效应元件(RE)和最小启动子,激活盒编码的蛋白质的DBD结合RE,而最小启动子对激活盒的AD起反应。另一个组件是靶核酸序列,该序列编码靶向关键乙烯诱导性信号蛋白如EIN2和EIN3的周转的调控蛋白,如ETP1/ETP2或EBF1/EBF2,或其功能片段。在一个实施方式中,核酸序列是正义取向。诱导型启动子控制选定调控蛋白的编码序列的表达。

[0031] 在另一实施方式中,激活盒和/或靶标盒还包含终止子序列,例如在编码蛋白质序列的核酸序列下游,以及任意的可选标记物。这种标记物是众所周知的,用于选择在抗生素或其他化学物质的存在下摄取基因的细胞。下面将更详细地描述这些任选组件。

[0032] 植物细胞中与诱导组合物协同作用时,基因表达系统发挥作用,使得激活盒和靶标盒的各组件能够调节选定调控蛋白的表达。选定调控蛋白的调控或调节能够选择性地调节植物细胞的乙烯敏感性。例如,一种调控方式包括通过降低调控蛋白的表达提高植物细胞的乙烯敏感性。在另一实施方式中,通过将调控蛋白的表达水平提高到足够高,以克服植物在存在乙烯时的竞争性正常反应,即正常操作提高信号蛋白EIN2或EIN3的基因产物的表达和蓄积,从而降低植物细胞的乙烯敏感性。通过基因表达系统中各组分与诱导组合物相互作用(特别是存在乙烯时)来控制植物细胞中调节乙烯途径的调控蛋白的表达。诱导组合物可以被植物细胞吸收和/或易位到植物细胞内。

[0033] 提到调节信号蛋白周转的调控蛋白的表达水平时,术语“足够高的表达水平”指能改变细胞或植物对10ppm乙烯或10 μ M ACC(1-氨基环丙基-1-羧酸)的反应的表达水平。

在一个实施方式中,调控蛋白的表达水平足够高,以克服乙烯或 ACC 存在时通常导致的信号蛋白表达增加。在另一实施方式中,调控蛋白的表达水平足够高,以降低乙烯或 ACC 存在时信号蛋白的表达水平。在另一实施方式中,调控蛋白的表达水平足够高,以克服应激植物中应激诱导的信号蛋白介导的乙烯生产,从而使植物对乙烯不敏感并克服应激对植物造成的不良影响。

[0034] 在该系统的一个实施方式中,第一盒和第二盒位于同一个质粒上,例如图 1 和 2 所示。在另一实施方式中,第一和第二盒位于不同质粒上。

[0035] 在基因表达系统的另一个实施方式中,第一基因表达盒可包含 DBD 和第一 LBD;第二盒可包含 AD 和第二不同的 LBD;第三盒可包含编码效应元件的多核苷酸、启动子和需要调节表达的靶调控基因,第一多肽的 DBD 结合该效应元件,第二盒的 AD 激活该启动子。在该系统中,AD 和 DBD 操作性连接于两个不同的蛋白质,在诱导组合物的存在下可激活靶基因的表达。在一个实施方式中,第一 LBD 可以是 EcR LBD,第二 LBD 可以是来自类视黄醇 X 受体的 LBD。在另一实施方式中,第二 LBD 可以是 EcR LBD,而第一 LBD 可以是来自类视黄醇 X 受体的 LBD。这种结构在美国专利 7091038 或 2005 年 12 月 1 日公开的美国专利公开 US 2005/0266457 中描述。

[0036] 术语“操作性连接”或“可操作地连接”表示单一核酸片段上核酸序列的结合,使得一个核酸序列的功能受到另一个核酸序列的影响。例如,当启动子能够影响编码序列的表达(即编码序列在启动子的转录控制下)时,启动子操作性连接于编码序列。编码序列以有义或反义取向操作性连接于调控序列。

[0037] 本文所用术语“表达”表示源自核酸或多核苷酸的有义(mRNA)或反义 RNA 的转录和稳定累积。在一个实施方式中,表达表示 mRNA 翻译形成蛋白质或多肽。在另一实施方式中,基因表达系统的组件可以在植物细胞中瞬时表达。在另一实施方式中,基因表达系统的组件可以通过整合入植物细胞的染色体内而稳定表达。瞬时或稳定整合表达的选择可以由产生和使用本文所述基因表达系统领域中的技术人员进行选择。

[0038] 术语“盒”、“表达盒”和“基因表达盒”表示能够在特定限制位点或者通过同源重组插入核酸或多核苷酸的 DNA 区段。该 DNA 区段包括编码感兴趣多肽的多核苷酸,盒和限制位点被设计成能够确保盒插入适当读码框以合适方向进行转录和翻译。这些载体或质粒可任选地包含编码感兴趣多肽的多核苷酸并具有除有利于特定宿主细胞转化的多核苷酸之外的元件。在某些实施方式中,这些盒还包含使得编码感兴趣多肽的多核苷酸在宿主细胞中表达提高的元件。这些元件包括但不限于:启动子、最小启动子、增强子、效应元件、终止子序列、聚腺苷酸化序列等。

[0039] 本文使用的所有其他术语采用本领域的常规涵义,除非另有说明。参见例如,美国专利 7,091,038 中的术语定义。

[0040] A. 激活盒的启动子

[0041] 在所述系统的一个实施方式中,激活盒的启动子是能够控制转化的植物细胞内 DBD、LBD 和 AD 序列的表达的核酸序列(DNA 或 RNA)。通常,这三种激活盒的主要组件位于选定启动子序列的 3' 端。启动子序列包括称为增强子的近端或较远端的上游元件。“增强子”是能够刺激启动子活性的 DNA 序列,可以是启动子的先天元件或是插入以增强启动子的水平或特异性的异源元件。本文中有用的启动子可完全源自天然基因,或者由源自天然存

在的不同启动子的不同元件组成,或者甚至包含合成的 DNA 区段。

[0042] 在一个实施方式中,激活盒的启动子是组成型启动子,例如导致基因大多数时候在大多数细胞类型中表达的启动子,使得用该盒转化的植物细胞不断产生激活盒组件。例如,适用于所述激活盒的某些组成型启动子包括但不限于:示例性的 G10-90 启动子、花椰菜花叶病毒 35S 启动子、木薯花叶病毒启动子、玄参花叶病毒启动子、杆状 DNA 病毒启动子、紫茉莉花叶病毒启动子、核酮糖-1,5-二磷酸羧化/加氧酶启动子(Rubisco 启动子)、肌动蛋白启动子或泛素启动子。

[0043] 在其他实施方式中,可采用介导基因在不同组织或细胞类型中表达的启动子(“组织特异性”、“细胞特异性”或“植物器官特异性启动子”)。理想情况下,这种启动子是植物组织和植物细胞固有或有功能的启动子,或者是植物组织和植物细胞固有或有功能启动子的突变形式。然而,也可使用在植物细胞中起作用的诸如哺乳动物、无脊椎动物等其他来源的其他组织和细胞的启动子。其他实施方式采用在发育的不同阶段表达组分的启动子(“发育特异性启动子”或“细胞分化特异性启动子”),或者响应不同环境或生理学条件的启动子。组织特异性启动子的详细列表可参见 Gallie,美国专利申请公开 2005/0066389,其描述了源自以下基因的种子特异性启动子:来自玉米的 MACI (Sheridan,1996Genetics 142:1009-1020);来自玉米的 Cat3(GenBank No. L05934, Abler 1993Plant Mol. Biol. 22:10131-10138);来自拟南芥属(Arabidopsis)的 viviparous-1(Genbank No. U93215);来自拟南芥属的 atmyc1(Urao,1996 Plant Mol. Biol. 32:571-576;Conceicao 1994 Plant 5:493-505);来自甘蓝型油菜(Brassica napus)的 napA 和 BnCysP1(GenBank No. J02798, Josefsson,1987 JBL26:12196-12201, Wan 等,2002 Plant J 30:1-10);来自甘蓝型油菜(Brassica napus)的 napin 基因家族(Sjodahl,1995Planta 197:264-271)。果实特异性启动子包括来自 CYP78A9 基因的启动子(Ito 和 Meyerowitz,2000Plant Cell 12:1541-1550)。其他组织特异性启动子包括 Reiser,1995Cell 83:735-742, GenBank No. U39944;Ray,1994Proc. Natl. Acad. Sci. USA 91:5761-5765 中描述的胚珠特异性 BEL1 基因,以及卵和中央细胞特异性 FIE1 启动子。萼片和花瓣特异性启动子包括:拟南芥属花同源异形基因 APETALA1(AP1)(Gustafson Brown,1994Cell 76:131-143;Mandel, 1992Nature 360:273-277),AP2 的相关启动子(例如参见 Drews,1991Cell 65:991-1002;Bowman,1991Plant Cell 3:749-758)。另一种有用的启动子是控制拟南芥的罕见花器官(ufo)基因表达的启动子(Bossinger,1996Development 122:1093-1102)。其他组织特异性启动子包括:玉米花粉特异性启动子(Guerrero,1990Mol. Gen. Genet. 224:161-168);也可参见 Wakeley,1998Plant Mol. Biol. 37:187-192;Ficker,1998Mol. Gen. Genet. 257:132-142;Kulikauskas,1997Plant Mol. Biol. 34:809-814;Treacy,1997Plant Mol. Biol. 34:603-611 描述的启动子)。有用的启动子包括来自 FUL 基因的启动子(Mandel 和 Yanofsky,1995Plant Cell,7:1763-1771)以及来自 SHP1 和 SHP2 基因的启动子(Flanagan 等,1996Plant J 10:343-353;Savidge 等,1995Plant Cell 7(6):721-733)。启动子可源自 TA29 基因(Goldberg 等,1995Philos Trans. R. Soc. Lond. B. Biol. Sci. 350:5-17)。

[0044] 其他合适的启动子包括来自以下基因的启动子:编码甘蓝型油菜的 2S 储存蛋白的基因(Dasgupta,1993Gene 133:301-302);来自拟南芥属的 2s 种子储存蛋白基因家族;编码甘蓝型油菜的油质蛋白 20kD 的基因,GenBank 编号 M63985;编码大豆油质蛋白 A 的

基因, Genbank 编号 U09118, 和编码大豆油质蛋白 B 的基因, Genbank 编号 U09119 ;编码拟南芥属油质蛋白的基因, Genbank 编号 Z17657 ;编码玉米油质蛋白 18kD 的基因, GenBank 编号 J05212 和 Lee, 1994Plant Mol. Biol. 26 :1981-1987 ;编码大豆低分子量富硫蛋白的基因 (Choi, 1995Mol Gen, Genet. 246 :266-268)。来自番茄的组织特异性 E8 启动子和来自 ATHB-8、AtPIN1、AtP5K1 或 TED3 基因的启动子 (Baima 等, 2001Plant Physiol. 126 :643-655, Galawailer 等, 1998Science 282 :2226-2230 ;Elge 等, 2001Plant J. 26 :561-571 ;Igarashi 等, 1998Plant Mol. Biol. 36 :917-927) 也是有用的。

[0045] 也可使用果实熟化、叶片枯萎和脱落期间以及在较小程度上花的枯萎和脱落期间有活性的番茄启动子 (Blume, 1997Plant J. 12 :731-746)。其他示例性的启动子包括 :编码雌蕊特异性碱性内切壳多糖酶的马铃薯 (*Solanum tuberosum* L.) SK2 基因中的雌蕊特异性启动子 (Ficker, 1997Plant Mol. Biol. 35, 425 :425-431) ;转基因苜蓿滋养和花柄尖的表皮组织中有活性的豌豆阿拉斯加栽培品种 (*Pisum sativum* cv. Alaska) 的 Blec4 基因。可使用在滋养型组织中有活性的各种启动子, 包括控制土豆的主要储存蛋白 patatin 的启动子 (例如, Kim, 1994Plant Mol. Biol. 26 :603-615 ;和 Martin, 1997Plant J. 11 :53-62) 和毛根农杆菌的 ORF 13 启动子 (Hansen, 1997Mol. Gen. Genet. 254 :337-343)。其它有用的滋养型组织特异性启动子包括编码主要的芋头 (*Colocasia esculenta* L. Schott) 球茎蛋白家族的球蛋白 tarin 的基因的 tarin 启动子 (Bezerra, 1995Plant Mol. Biol. 28 :137-144) ;仙茅甜蛋白 (curculin) 启动子 (de Castro, 1992Plant Cell 4 :1549-1559) 和烟草根部特异性基因 TobRB7 的启动子 (Yamamoto, 1991Plant Cell 3 :371-382)。叶特异性启动子包括 :二磷酸核酮糖羧化酶 (RBCS) 启动子、番茄 RBCS1、RBCS2 和 RBCS3A 基因 (Meier, 1997FEBS Lett. 415 :91-95)。几乎只在叶片和叶鞘的叶肉细胞中高水平表达的二磷酸核酮糖羧化酶启动子 (Matsuoka, 1994Plant J. 6 :311-319), 集光叶绿素 a/b 结合蛋白基因启动子 (Shiina, 1997Plant Physiol. 115 :477-483 ;Casal, 1998Plant Physiol. 116 :1533-1538), 拟南芥 myb- 相关基因启动子 (Atmyb5 ;Li, 1996FEBS Lett. 379 :117-121), 和 Atmyb5 启动子 (Busk, 1997Plant J. 11 :1285-1295) 都是有用的启动子。

[0046] 有用的滋养型组织特异性启动子包括 :分生组织 (根尖和苗端) 启动子, 如 “SHOOTMERISTEMLESS” 和 “SCARECROW” 启动子 (Di Laurenzio, 1996Cell 86 :423-433 ;和 Long, 1996Nature 379 :66-69)。另一种有用的启动子控制 3-羟基-3-甲基戊二酰辅酶 A 还原酶 HMG2 基因的表达 (例如参见 Enjuto, 1995Plant Cell. 7 :517-527)。有用的还有来自玉米和其他物种的具有分生组织特异性表达的 kn1 相关基因, 例如参见 Granger, 1996Plant Mol. Biol. 31 :373-378 ;Kerstetter, 1994Plant Cell 6 :1877-1887 ;Hake, 1995Philos. Trans. R. Soc. Lond. B. Biol. Sci. 350 :45-51, 例如拟南芥 KNAT1 或 KNAT2 启动子 (例如参见 Lincoln, 1994Plant Cell 6 :1859-1876)。

[0047] 在激活盒的某些实施方式中, 启动子可以是诱导型或调节型的, 例如在细胞接触试剂、生物分子、化学物质、配体、光或一些其他刺激物或用这些物质处理细胞后可引起核酸序列的表达。这种诱导型启动子的非限制性列表包括 :PR1-a 启动子、原核阻抑子-操纵子系统和高等真核转录激活系统, 例如美国专利 7091038 中详述。这种启动子包括 :来自大肠杆菌的四环素 (“Tet”) 和乳糖 (“Lac”) 阻抑子-操纵子系统。其他诱导型启动子包括 :玉米干旱诱导型启动子 ;马铃薯的寒冷、干旱和高盐诱导型启动子, 拟南芥属枯萎诱

导型启动子, SAG12, 以及 LEC1、LEC2、FUS3、AtSERK1 和 AGL15 的胚胎发生相关启动子, 都是本领域技术人员已知的。接触植物激素如植物生长激素或细胞激肽类后可诱导的其他植物启动子是也有用的, 例如在美国专利申请公开 US2005/0066389 和美国专利 6, 294, 716 中所述。

[0048] 在激活盒中, 能够启动序列 DBD、LBD 和 AD 的表达的任何启动子基本都是合适的, 包括但不限于: 病毒启动子、细菌启动子、植物启动子、合成启动子、组成型启动子、组织特异性启动子、发育特异性启动子、诱导型启动子、光调节的启动子; 发病或疾病相关启动子、花椰菜花叶病毒 19S、花椰菜花叶病毒 35S、CMV35S 最小、木薯叶脉花叶病毒 (CsVMV)、玄参花叶病毒、杆状 DNA 病毒、紫茉莉花叶病毒、叶绿素 a/b 结合蛋白、核酮糖 1-二磷酸羧化酶、苗特异性、根特异性、壳多糖酶、应激诱导型、水稻东格鲁杆状病毒、植物超-启动子、马铃薯亮氨酸氨基肽酶、硝酸还原酶、醇脱氢酶、蔗糖合酶、甘露碱合酶、胆脂碱合酶、真蛭碱合酶、泛素、玉米醇溶蛋白、肌动蛋白和花色苷启动子。在本发明优选的实施方式中, 启动子选自下组: 花椰菜花叶病毒 35S 启动子、木薯叶脉花叶病毒启动子和花椰菜花叶病毒 35S 最小启动子、玄参花叶病毒启动子、杆状 DNA 病毒启动子、紫茉莉花叶病毒、泛素 (Ubc) 启动子和肌动蛋白启动子。

[0049] B. DBD

[0050] 本文所用术语“DNA 结合域”包括 DNA 结合蛋白的最小多肽序列, 最长至整个 DNA 结合蛋白长度, 只要该 DNA 结合域能够与特定效应元件结合。在有或没有配体的存在下, DNA 结合域结合 RE 的 DNA 序列以启动或抑制受 RE 调控的下游基因的转录。在基因表达单元的某些实施方式中, DBD 位于激活盒中, 而效应元件位于靶标盒中。

[0051] DNA 结合域可以是具有已知效应元件的任何 DNA 结合域, 包括合成和嵌合的 DNA 结合域, 或其类似物、或组合、或修饰形式。在某些实施方式中, DBD 是 GAL4DBD、LexA DBD、转录因子 DBD、H 类核受体成员 DBD、甾体 / 甲状腺激素核受体超家族成员 DBD 或细菌 LacZ DBD。更优选地, DBD 是昆虫蜕皮激素受体 DBD, GAL4 DBD (参见本文所述实施例的质粒中所示的序列), 或 LexA DBD。这种 DBD 的序列公开可得, 在诸如美国专利 7091038 或美国专利申请公开 2005/0266457 中描述。在其它实施方式中, 盒中的 DBD 包括但不限于: 由 cI 启动子或 lac 启动子获得的 DNA 结合域, 这也是公开可得的序列。

[0052] C. 蜕皮激素 LBD 和任选的第二 LBD

[0053] 在基因表达系统的某些实施方式中, 蜕皮激素受体 (EcR) LBD 包括无脊椎动物蜕皮激素受体或其突变体的全部或其一部分。EcR 是核甾体受体超家族成员, 表现为特征 DNA 和配体结合域, 以及激活域 (Koelle 等, 1991, Cell, 67 :5977; 也可参见美国专利 6245531 (Stanford)。蜕皮激素受体对许多甾体化合物有响应, 例如松甾酮 A 和莫瑞甾酮 A (muristerone A) 和非甾体化合物。EcR 具有 5 个模块域, A/B (反式激活), C (DNA 结合, 异源二聚化), D (铰链, 异源二聚化), E (配体结合, 异源二聚化和反式激活) 和 F (反式激活) 结构域。这些结构域中的一些, 例如 A/B、C 和 E 在与其他蛋白融合后保留功能。可以用作本文所述基因表达系统中的 LBD 的 EcR 的合适部分包括 D、E 和 F。例如, 野生型 EcR LBD 的序列可参见 SEQ ID NO :1 中的核苷酸 990-1997。

[0054] 优选地, EcR 是鳞翅类 EcR、双翅类 EcR、节肢类 EcR、同翅类 EcR 和半翅类 EcR。更优选地, 使用的 EcR 是云杉卷叶蛾 (*Choristoneura fumiferana*) EcR (" CfEcR ")、黄粉虫甲

(*Tenebrio molitor*)EcR(" TmEcR")、烟草天蛾 (*Manduca sexta*)EcR(" MsEcR")、绿棉铃虫 (*Heliothies virescens*)EcR(" HvEcR")、家蚕蛾 (*Bombyx mori*)EcR(" BmEcR")、黑腹果蝇 (*Drosophila melanogaster*)EcR(" DmEcR")、埃及伊蚊 (*Aedes aegypti*)EcR(" AaEcR")、绿蝇 (*Lucilia capitata*)EcR(" LcEcR")、地中海果蝇 (*Ceratitis capitata*)EcR(" CcEcR")、东亚飞蝗 (*Locusta migratoria*)EcR(" LmEcR")、紫萁瘤蚜 (*Myzus persicae*)EcR(" MpEcR")、招潮蟹 (*Uca pugilator*)EcR(" UpEcR")、美洲钝眼蝉 (*Amblyomma americanum*)EcR(" AmaEcR")、银叶粉虱 (*Bamecia argentifoli*)EcR(" BaEcR")、或黑尾叶蝉 (*Nephotetix cincticeps*)EcR(" NcEcR")等。甚至更优选地, LBD 来自云杉卷叶蛾 EcR(" CfEcR")或黑腹果蝇 EcR(" DmEcR")。

[0055] 各种野生型或突变型 EcR 的序列公开可得, 例如美国专利 7091038 ; 国际专利申请 WO 97/38117 以及美国专利 6333318、6265173 和 5880333 中所描述的。虽然下述实施例采用了野生型 EcR 序列, 但预计本领域技术人员可选择突变序列进行类似应用。例如, 在一个实施方式中, 突变型蜕皮激素受体包含上文引用的美国专利申请公开号 2005/0266457 记载的突变, 例如包含至少一个突变的 H 类核受体配体结合域。在一个实施方式中, 在等同于 (例如) SEQ ID NO :2 的 1374-1376 的核苷酸位置含有将 Thr 密码子 ACA 改变成 Val 密码子 GTG 的突变的蜕皮激素 LBD 是有用的 EcR LBD。突变体 EcR LBD 称为 T52V, 其编码全长 Cf EcR 中氨基酸位置 335 处 Thr 到 Val 的突变。本文所述基因表达系统中也可使用该文献所述的其他 EcR LBD。在另一实施方式中, 突变体蜕皮激素受体 LBD 是美国专利 6, 245, 531 (Stanford) 所述的 LBD, 或者是截短的 EcRLBD 序列或缺失突变体, 或本领域已知的其它突变序列。在另一实施方式中, LBD 由常规杂交条件下与已知 EcR LBD 或突变序列杂交的多核苷酸编码, 所述杂交条件包括例如在小于 500mM 盐和至少 37°C 的杂交步骤中, 和在 2XSSPE 和至少 63°C 的洗涤步骤中。

[0056] 在某些实施方式中, 基因表达系统采用第二 LBD。第二 LBD 不是蜕皮激素受体多肽, 但可以是第二核受体的配体结合域。这种第二结合域包括但不限于: 脊椎动物类视黄醇 X 受体配体结合域、无脊椎动物类视黄醇 X 受体配体结合域、超气门蛋白 (ultraspiracle protein) 配体结合域, 以及包含两个多肽片段的嵌合配体结合域, 其中第一多肽片段来自脊椎动物类视黄醇 X 受体配体结合域、无脊椎动物类视黄醇 X 受体配体结合域或超气门蛋白配体结合域, 第二多肽片段来自不同的脊椎动物类视黄醇 X 受体配体结合域、无脊椎动物类视黄醇 X 受体配体结合域或超气门蛋白配体结合域。例如参见美国专利申请公开 US 2005/0266457 中描述的结合域。这种 LBD 是本领域技术人员熟知的并且已在文献中详细记载。

[0057] 本领域技术人员根据本文的教导无需过多实验即能够选择一种或多种合适的 EcR LBD 序列, 并用其代替下文中列举的序列。

[0058] D. 激活域

[0059] 可用于基因表达系统的激活或反式激活域 (缩写为“AD”) 可以是任何 H 类核受体成员 AD、甾体 / 甲状腺激素核受体 AD、合成或嵌合 AD、聚谷胺酰胺 AD、碱性或酸性氨基酸 AD、VP16AD、GAL4AD、NF- κ B AD、BP64AD、B42 酸性激活域 (B42AD)、p65 反式激活域 (p65AD)、糖皮质激素激活域或其类似物、组合或修饰形式。在具体实施方式中, AD 是合成或嵌合 AD, 或者由 EcR、糖皮质激素受体、VP16、GAL4、NF- κ B 或 B42 酸性激活域 AD 获得。优选地, AD

是 EcRAD、VP16AD、B42AD 或 p65AD。这种活化域的序列在诸如美国专利 7,091,038 或本文所述的其他文献中公开可得。示例性的 VP16AD 在本文实施例所述的质粒中描述。这些结构域是本领域技术人员熟知的并且已在文献中详细记载。

[0060] E. 靶标盒的“诱导型”启动子系统

[0061] 在某些实施方式中,靶标盒的启动子是多组件启动子序列。其包括与对应于激活盒中的 DNA 结合域的一个或多个拷贝的效应元件操作性相连的最小启动子。

[0062] 如本文所述,最小启动子包括核心启动子(即,介导转录启动的序列)和 5' 非翻译区(5' UTR)而没有增强子序列。因此,在基因表达系统的实施方式中,最小启动子可以是源自部分 A 所述激活盒中使用的任何启动子的最小启动子。在靶标盒的某些实施方式中,理想的最小启动子包括:花椰菜花叶病毒 35S 最小启动子;合成的 E1b 最小启动子(SEQ ID NO :8 ;参见美国专利 7091038)和合成的 TATA 最小启动子(TATATA ;参见美国专利申请公开 US2005/0228016)。本领域技术人员可以容易地从文献中详述的各种启动子选择本文所述基因表达系统中使用的最小启动子。

[0063] 35S 最小启动子的序列在以下实施例所述的质粒中描述。

[0064] 靶标盒中诱导型启动子的其他部分包括:位于最小启动子 5' 或 3' 位置的效应元件("RE")。一个 RE 的两侧可具有两种不同或者相同的最小启动子以表达两种不同的蛋白质。在一个实施方式中,RE 操作性连接于最小启动子。效应元件是一种或多种顺式作用的 DNA 元件,它能够赋予启动子以通过与激活盒的 DNA-结合域相互作用介导的反应性。这种 DNA 元件可以是回文序列(完美或不完美)或者由可变数量的核苷酸分隔的序列基序或半位点(half sites)组成。半位点可以类似或者相同,并排列成正向或反转重复序列或者形成单个半位点或相邻半位点串联形成的多聚体。天然蜕皮激素受体的效应元件的 DNA 序列的例子在 Cherbas L. 等,1991 Genes Dev. 5,120-131 ;D' Avino PP. 等,1995 Mol. Cell. Endocrinol,113 :19 和 Antoniewski C. 等,1994 Mol. Cell Biol. 14,4465-4474 以及其他出版物中描述。RE 可以是对应于激活盒中的 DNA 结合域的任何效应元件,或其类似物、组合或修饰形式。靶标盒中可采用单个 RE 或多个 RE,多个 RE 可以是相同 RE 的多个拷贝或是两种或更多种不同的 RE。RE 可以被针对其他 DNA 结合蛋白域,例如酵母的 GAL-4 蛋白(参见 Sadowski 等,1988 Nature,335 :563-564) 或大肠杆菌的 LexA 蛋白(参见 Brent 和 Ptashne 1985, Cell,43 :729-736) 的效应元件所修饰或取代,或根据特异性相互作用设计、修饰和选择的蛋白质的靶向相互作用特异的合成效应元件(例如参见 Kim 等,1997 Proc. Natl. Acad. Sci., USA,94 :3616-3620) 以适合嵌合受体。在一个具体实施方式中,RE 是 GAL4 的 RE("GAL4RE"),优选两个或更多个拷贝。以下实施例显示采用 5 个拷贝的 GAL4RE(Le, 5XGAL4)。然而,其他合适的 RE 包括但不限于:LexA、H 类核受体 RE、甾体/甲状腺激素核受体 RE、或识别合成 DNA 结构域的合成 RE。在其它实施方式中,RE 是蜕皮激素效应元件(EcRE)、或包含多核苷酸序列的 LexA RE(操纵子,"op")。所有这些 RE 已在文献中详细记载,本领域技术人员可以容易地根据本说明书内容进行选择。

[0065] 在靶标盒中,"诱导型启动子"操作性连接和控制调控性核酸序列或基因的表达。所述基因的高水平表达能降低乙烯诱导性信号蛋白如 EIN2 或 EIN3 的蓄积,从而调节乙烯灵敏度,如下所述。靶标盒的诱导型启动子由化学诱导组合物或者诱导剂诱导,当所述化学诱导组合物或者诱导剂接触激活盒的配体结合域时激活最小启动子的效应元件。

[0066] F. 编码选定调控蛋白的核酸序列

[0067] 可用于此系统的核酸序列编码通过调节乙烯诱导性信号蛋白如 EIN2 或 EIN3 的周转来调整植物的乙烯敏感性或乙烯生产的选定的调控蛋白。在某些实施方式中,这种核酸序列包括结合 EIN3 的 F 盒蛋白,EBF1 和 EBF2。在另一实施方式中,选定的调控蛋白是 F 盒蛋白 ETP1 和 ETP2(Qiao 等,2008,如上所述)。

[0068] 这种 ETP1 核酸序列的例子参见 SEQ ID NO:1 的核苷酸 2557 至 3804(也可参见 GENBANK 登录号 NM_112874)。这种 ETP2 核酸序列的例子参见 SEQ ID NO:2 的核苷酸 2551 至 3717(也可参见 GENBANK 登录号 NM_112777)。GENBANK 登录号 NM_128106 公开了 EBF1 核酸序列的例子。GENBANK 登录号 NM_122444 公开了 EBF2 核酸序列的例子。这些序列通过引用纳入本文。

[0069] 除了使用野生型或天然来源的植物调控基因外,基因表达系统也可采用某些核酸序列,在这些基因序列和编码蛋白中包含有用的突变。

[0070] 在本文所述的本发明的一个实施方式中,将控制特定植物的所选信号蛋白 EIN2 或 EIN3 周转的野生型调控基因用于本文所述的基因表达系统和方法,以便通过在乙烯存在下抑制信号蛋白蓄积来控制或调节该植物的乙烯敏感性。

[0071] 在另一实施方式中,采用介导植物细胞乙烯敏感性或乙烯产生的野生型蛋白的突变体。在又一些实施方式中,编码能够通过乙烯存在下抑制信号蛋白蓄积来调节一种类型的植物细胞的乙烯敏感性或乙烯产生的蛋白的基因的野生型或突变体用于另一种类型的植物细胞中,这种使用例如对于消除潜在的 RNA 沉默是理想的。

[0072] G. 任选的组件

[0073] 基因表达系统各个盒中的任选组件包括终止控制区。在本发明的某些实施方式中,激活盒或靶标盒中可使用这种终止子或聚腺苷酸化序列。该区域源自优选宿主天然具有的各种基因。在本发明的一个实施方式中,终止控制区包括或源自合成的聚腺苷酸化信号、胆脂碱合酶(nos)、花椰菜花叶病毒(CaMV)、真蛭碱合酶(ocs)、土壤杆菌、病毒和植物终止子序列等。

[0074] 选择性标记物包括:能够基于其作用进行选择抗生素或化学抗性基因,即抗生素耐药性、除草剂抗性、比色标记物、酶、荧光标记物等。本领域已知并使用的可选择标记基因的例子包括:提供氨苄青霉素、链霉素、庆大霉素、卡那霉素、潮霉素、放线酰胺素(PDF1 基因)、双丙氨磷除草剂、草甘磷除草剂、磺酰胺、甘露糖等抗性的基因;和可用作表型标记物的基因,即花色素苷调节基因、异戊烯转移酶基因、GUS 和荧光素酶。

[0075] 根据本文所述的某些实施方式转化到植物细胞内的盒或包含盒的质粒中可以存在其他调控序列,例如在质粒中用作间隔序列的核苷酸序列,以及其他小调控序列、酶切位点等。

[0076] 根据本文内容,本领域技术人员可以容易地从公知并且已在文献中详细记载的各种序列中选择合适的终止序列、选择性标记物和其他常规质粒调控序列。

[0077] H. 用于基因表达系统的诱导组合物 / 诱导剂

[0078] 在植物中表达该基因表达系统时,通过调节所选调控蛋白的表达,如过度表达该蛋白选择性,基于调控蛋白对所选信号蛋白的表达和蓄积的控制来调节植物细胞的乙烯敏感性。过度表达该调控蛋白引起信号蛋白的表达减少。在乙烯存在下,过度表达该调控蛋

白至足够高的水平,以克服乙烯对信号蛋白的正常影响(即表达升高)。可通过诱导组合物施加于植物的时机、持续时间和用量来控制调控蛋白的过度表达程度。在一个实施方式中,诱导组合物是与植物细胞相接触的化学物质。在另一实施方式中,诱导组合物是被植物细胞吸收的化学物质。在又一实施方式中,诱导组合物是易位到植物细胞内的化学物质。诱导组合物也可以是对激活盒的 EcR LBD 高度特异的配体。诱导组合物或配体与激活盒的 LBD 的结合可诱导靶标盒的诱导型启动子以及编码选定蛋白的核酸序列的表达。因此,该系统通过降低或抑制信号蛋白的表达在植物中调节乙烯敏感性。诱导组合物还表现为对植物细胞、组织或器官的毒性低。诱导组合物还能够从植物快速去除以“关闭”乙烯敏感性的调节,并有效控制这种调节,如下文更详细所述。

[0079] 其中,有效的诱导组合物是优先结合蜕皮激素配体结合域的配体。在某些实施方式中,这些配体包括二乙酰肼化合物,包括市售虫酰肼(陶氏农用科技公司(Dow AgroSciences))、甲氧虫酰肼(陶氏农用科技公司)、氯虫酰肼(陶氏农用科技公司)和环虫酰肼(日本化药株式会社(Nippon Kayaku))(参见国际专利公开 WO 96/027673 和美国专利号 5,530,028)。其他有用的诱导剂是非甾体配体,包括美国专利 6,258,603 所述的二苯甲酰肼衍生物。其他有用的诱导剂是美国专利申请公开 US 2005/0228016 中详细描述的四氢喹啉衍生物。许多其他合适的化合物如 1-芳酰基-4-(芳基氨基)-1,2,3,4-四氢喹啉(THQ)在 Kumar 等, *J. Biol. Chem.* 2004, 279(26):27211-8; Hormann 等, *J. Comput Aided Mol. Res* 2003 17(2-4):135-53; Tice 等, *Bioorg Med Chem Lett* 2003, 13(11):1883-6; 和 Tice 等, 2003 *Bioorg Med Chem Lett*. 2003, 13(3):475-8 中列出。

[0080] 因此,基因表达系统由化学诱导组合物或诱导剂诱导或“打开”,当诱导组合物或诱导剂接触激活盒的配体结合域时可激活最小启动子的效应元件,从而打开调控核酸序列或基因的表达,进而抑制信号蛋白 EIN2 或 EIN3 的表达,使该植物对乙烯不敏感。此种基因表达系统也提供了一种手段,即外源性表达调控基因至足够高的水平,以克服乙烯对信号蛋白表达的影响。

[0081] II. 转基因植物、植物细胞、组织或器官

[0082] 如上所述,基因表达系统设计成能够整合到植物、植物细胞或植物的其他组织或器官内。任选地,这种整合可以是瞬时的。然而,在本发明的某些实施方式中,希望稳定整合到植物染色体内。

[0083] 在一个实施方式中,转基因植物细胞设计成能够表达上述基因表达系统,暂时且可逆地控制乙烯敏感性。在一个实施方式中,这种植物细胞是用基因表达系统的激活盒和靶标盒转染或转化的细胞。在一个实施方式中,激活盒和靶标盒位于相同质粒上,并将该质粒转染或转化入植物细胞。在另一实施方式中,激活盒和靶标盒位于不同质粒上,两质粒单独或一起转染或转化入相同的植物细胞。或者,两个单独的质粒各自转染或转化入相同植物的不同细胞中。在又一个实施方式中,两个质粒各自转染或转化入不同植物,具有单一质粒的各植物有性杂交形成含有两种质粒的杂交体,从而提供功能诱导性系统。

[0084] 转染包括将外源或异源 RNA 或 DNA 引入细胞内,以实现表型改变。转化表示将核酸片段转移和整合到植物细胞的染色体 DNA 内,得到遗传稳定的遗传特性。因此,包含转化的核酸片段的植物细胞称为“转基因”或“重组”或“转化”生物体。因此,初始转化或转染的植物细胞的后代也具有瞬时或稳定整合入其染色体的盒。

[0085] A. 转化

[0086] 植物细胞的转化包括产生仅包含激活盒、仅包含靶标盒或包含两种盒的甾体或质粒。参见附图 1、2、3A-3I 和 4A-4I 的示例。合适的载体和植物组合是本领域技术人员容易明白的,可参见例如 Maliga 等,1994《植物分子生物学方法:实验室手册》(Methods in Plant Molecular Biology:A Laboratory Manual),纽约冷泉港(Cold Spring Harbor, NY)。

[0087] 例如,合适的“载体”是用于将核酸克隆和 / 或转移到植物细胞内的任何部件。载体可以是连接另一 DNA 区段的复制子,以实现所连接区段的复制。“复制子”是用作 DNA 自主复制单位的任何遗传元件(例如质粒、噬菌体、粘粒、染色体、病毒),即能够在其自身控制下进行复制。适合用基因表达系统转化植物细胞的载体包括将核酸引入细胞的病毒和非病毒载体。可使用本领域已知的大量载体来操纵核酸,将效应元件和启动子引入基因等。合适的载体包括例如:质粒和修饰的植物病毒,包括例如噬菌体如 λ 衍生物,或质粒如 pBR322 或 pUC 质粒衍生物,或 Bluescript 载体。将合适的 DNA 片段连接到具有互补粘性末端的选定载体内或通过将其核苷酸序列(接头)连接到 DNA 末端内以便对合适的插入位点进行酶促修饰的常规方式是已知的。可用于转化植物细胞的任何病毒或非病毒载体都是有用的。非病毒载体包括质粒、脂质体、带电脂质(细胞转染剂(cytfectin))、DNA-蛋白质复合物和生物聚合物。除了基因表达系统的盒之外,载体还可包含一个或多个调控区,和 / 或用于选择、测量和监测核酸转移结果(转移到哪个组织、表达持续时间等)的选择性标记物。

[0088] 术语“质粒”表示通常为环状双链 DNA 分子形式的额外染色体元件。

[0089] 这些元件可以是自主复制序列、基因组整合序列、噬菌体或核苷酸序列,源自任何来源的线性、环状或超螺旋的单链或双链 DNA 或 RNA,其中大量核苷酸序列已连接或重组入能够将启动子片段和选定基因产物的 DNA 序列与合适的 31' 未翻译序列引入细胞的独特结构。

[0090] 载体或质粒可以通过本领域已知的方式引入所需的植物细胞,例如转染、电穿孔、显微注射、转导、细胞融合、DEAE 右旋糖苷、磷酸钙沉淀、脂质转染(溶酶体融合)、利用基因枪、或 DNA 载体转运子(参见例如, Wu 等,1992, J. Biol. Chem. 267 :963-967 ;Wu 和 Wu, 1988, J. Biol. Chem. 263 :14621-14624 ;和 Hartmut 等,美国专利 5354844)。或者,利用阳离子脂质有利于负电荷核酸的包封,也有利于与负电荷细胞膜的融合(Feigner 和 Ringold, 1989 Science 337 :387 388)。用于转移核酸的尤其有用的脂质化合物和组合物如美国专利 6172048、6107286 和 5459127 所述。也可利用其他分子促进核酸转染,例如阳离子寡肽或阳离子聚合物(例如,美国专利 5856435),或源自 DNA 结合蛋白的肽(例如,美国专利 6200956)。也可将载体以裸 DNA 质粒的形式引入(参见美国专利 5693622、5589466 和 5580859)。也可利用受体-介导的 DNA 递送方法来实现基因表达盒转化入植物细胞。植物转化例如可以利用 Horsch 等,1988《叶盘转化:植物分子生物学手册》(Leaf Disc Transformation:Plant Molecular Biology Manual)所述的土壤杆菌介导的叶盘转化法或本领域已知的其他方法来实现。

[0091] B. 繁殖与筛选

[0092] 因此,用上述基因表达系统(在单一质粒中,或者作为多个转化质粒,各自包含不同的盒)转化植物中的至少一个细胞之后,产生转基因植物、植物组织或植物器官的方法还包括在选定植物的典型条件下由转化的植物细胞或植物繁殖植物、或如上所述的植物杂

交体。然后,筛选植物以选择包含或显示转化植物细胞表型性状的植物(细胞、组织、器官)。例如,随后进行所得植物或其细胞、组织和器官的筛选以确定所述植物是否包含所需的基因表达盒的整合核酸序列,这也是本领域技术人员已知的。例如,利用质粒中的一种或多种报道基因或标记物来选择将引入的 DNA 稳定整合到其染色体中的细胞。在下面的实施例中,采用了卡那霉素、放线酰胺素或双丙氨磷。

[0093] 成功整合表达系统的植物(组织或器官)在接触上述诱导组合物时,快速出现乙烯不敏感性。任何植物(包括植物细胞、组织或器官)易于进行这种转化,因而可通过常规方式培育重组植物。尤其适合用基因表达系统进行转化、因而易于调节其乙烯敏感性的植物包括:双子叶植物、单子叶植物、装饰性植物、开花植物以及用于人或动物消费的植物或植物部分。非限制性地,这种植物包括:水稻、玉米、小麦、大麦、高粱、粟米、柳枝稷、芒、牧草、燕麦、番茄、马铃薯、香蕉、猕猴桃、鳄梨、甜瓜、芒果、甘蔗、甜菜、烟草、木瓜、桃、草莓、悬钩子、黑莓、蓝莓、莴苣、卷心菜、花椰菜、洋葱、椰菜、抱子甘蓝、棉花、芸苔、葡萄、大豆、油菜、芦笋、豆类、胡萝卜、黄瓜、甜瓜、秋葵、防风草、花生、胡椒、菠萝、南瓜、甜马铃薯、黑麦、香瓜、豌豆、西葫芦、向日葵、菠菜、苹果、樱桃、酸果蔓、葡萄柚、柠檬、酸橙、油桃、橙、桃、梨、蜜柑与柚子杂交品(tangelos)、橘子、百合、康乃馨、菊、矮牵牛花、玫瑰、天竺葵、紫罗兰、剑兰、兰花、紫丁香、山楂、枫香树(sweetgum)、槭树、猩猩木、刺槐、岑树、白杨、菩提树和拟南芥。

[0094] 植物组织和器官包括但不限于:营养组织,例如根、茎或叶,或繁殖组织,例如果实、胚珠、胚、胚乳、珠被、种子、种皮、花粉、花瓣、萼片、雌蕊、花、花药或任何其他胚组织。

[0095] III. 控制乙烯敏感性的方法

[0096] 可对这类转基因植物、细胞、组织、花、种子或器官进行控制乙烯敏感性的方法,该方法在合适时间施加有效量的诱导组合物并持续足够时间,以抑制信号蛋白如 E1N2 或 E1N3 的蓄积,特别是在植物细胞接触乙烯时。诱导组合物可以与转基因植物、植物细胞、植物组织或植物器官的细胞相接触、被其吸收和/或易位到这些细胞内。应用技术包括但不限于:用诱导剂浸渍、喷雾、洒粉、灌透、滴加或灌溉植物或者与植物相接触的土壤或介质。

[0097] 在诱导组合物的存在下,可以通过提高选定的调控蛋白或周转蛋白的表达调节植物细胞对乙烯的反应。在选定蛋白是 EPT1 的下述实施例中,施加诱导剂能提高 EPT1 的表达,进而降低或抑制 E1N2 的表达并抑制其在细胞中蓄积,从而降低该植物对乙烯的敏感性。这种敏感性的降低持续一段时间,该持续时间包括诱导剂施加于植物(细胞、组织或器官)以及停止主动施加后植物继续代谢剩余的诱导剂的时间。另外,在存在乙烯时,可通过过度表达调控蛋白,至抵消乙烯对乙烯诱导性信号蛋白的诱导和表达的直接竞争作用所需的程度,来实现此种降低。一段时间后由于植物清除了诱导剂,植物细胞对乙烯的反应回复至野生型,在这种情况下增加。在植物细胞中“控制”、“调节”或“调控”影响信号蛋白表达和调节乙烯敏感性或乙烯生产的调控蛋白的表达可以多种方式实现。在一种实施方式的方法中,通过将诱导组合物应用于植物的时机选择来控制调控蛋白表达的调节(包括该表达的量化大小)。在另一实施方式中,利用应用于植物的诱导组合物的浓度来控制蛋白表达(包括该表达量化大小),因而控制乙烯敏感性。在另一个实施方式中,通过诱导组合物应用于植物的持续时间来控制蛋白表达的调节作用(包括该表达的量化大小)。植物生长期间可以改变施加诱导组合物的这三种参数中的任一种、两种或全部三种参数以获得所需结

果。

[0098] 作为一个通过时机选择进行控制的例子,可以在植物生长周期的选定时间施加诱导组合物以调节乙烯敏感性,例如在植物发芽、果实熟化或开花中的一个阶段之前或之后或者根据环境条件(例如在植物接触应激因子如病原体或干旱之前或之后)施加。在另一实施方式中,在植物生长周期过程中多次应用诱导组合物。在其它实施方式中,在选定时间停止应用诱导组合物以控制乙烯敏感性。理想的施加时间可根据需处理植物的类型、诱导组合物的效能、及其对植物可能产生的植物毒性作用进行选择 and 改变。

[0099] 作为通过诱导组合物浓度进行控制的一个例子,根据诱导组合物的种类、植物类型、施加时机(即在施加诱导组合物时该植物是否接触或接触过乙烯)、植株大小或株龄(如幼苗或是成熟植株),以及施加环境(如野外或组织培养物、花盆、或者其它实验室或培养容器),以所选浓度将诱导组合物施加于植物。在一个实施方式中,诱导组合物的施用浓度为每株植物(例如在组织培养物中)至少 $0.01 \mu\text{M}$ 。在另一实施方式中,诱导组合物的施用浓度为每株植物至少 $0.1 \mu\text{M}$ 。在另一实施方式中,诱导组合物的施用浓度为每株植物至少 $1 \mu\text{M}$ 。在另一实施方式中,诱导组合物的施用浓度为每株植物至少 $10 \mu\text{M}$ 。在另一实施方式中,诱导组合物的施用浓度为每株植物至少 $20 \mu\text{M}$ 。在另一实施方式中,诱导组合物的施用浓度为每株植物至少 $50 \mu\text{M}$ 。在另一实施方式中,诱导组合物的施用浓度为每株植物至少 $100 \mu\text{M}$ 。在另一实施方式中,诱导组合物的施用浓度为每株植物至少 $200 \mu\text{M}$ 。在另一实施方式中,诱导组合物的施用浓度为每株植物至少 $500 \mu\text{M}$ 。在另一实施方式中,诱导组合物的施用浓度为每株植物至少 $700 \mu\text{M}$ 。在另一实施方式中,诱导组合物的施用浓度为每株植物至少 1mM 。在另一实施方式中,诱导组合物的施用浓度为每株植物至少 3mM 。在另一实施方式中,诱导组合物的施用浓度为每株植物至少 5mM 。在其它实施方式中,该浓度选自 $0.01 \mu\text{M}$ 至至少 5mM 之间的任何小数浓度。

[0100] 第三种控制介导植物乙烯敏感性或乙烯产生的蛋白质的调节方法包括改变施加诱导组合物的持续时间。例如,将诱导组合物应用于植物的持续时间从至少 10 分钟到至少 2 周或者更长时间,取决于所需的作用、诱导剂的效能以及不良植物毒性作用的可能性。如果需要,诱导组合物的施加时间可以达到几天。在一些实施方式中,诱导组合物的施加时间可以达到数周。在一个实施方式中,通常将上述浓度施加于植物至少 10 分钟。在另一实施方式中,不存在乙烯时,通常将上述浓度施加于植物至少 20 分钟以降低信号蛋白的蓄积。在另一实施方式中,当植物细胞接触通常会增加信号蛋白表达的乙烯时,施加诱导组合物至少 30 分钟。在另一实施方式中,施加诱导组合物至少 1 小时。在另一实施方式中,施加诱导组合物至少 2 小时。在另一实施方式中,施加诱导组合物至少 5 小时。在另一实施方式中,施加诱导组合物至少 12 小时。在另一实施方式中,施加诱导组合物至少 1 天。在另一实施方式中,施加诱导组合物至少 2 天。在一些实施方式中,可能需要或想要施加诱导组合物至少或至多 2 周。在一个实施方式中,当植物中不存在乙烯时,如在植物出现应激之前或预期植物会出现应激情况时,本领域技术人员选择施加时机和浓度,以抑制信号蛋白的蓄积。在一个实施方式中,当植物中存在乙烯时,例如在植物出现应激时,本领域技术人员选择施加时机和浓度,以抑制信号蛋白的蓄积。

[0101] 例如,一种方法包括在植物未接触乙烯时,每植株施加大于 $10 \mu\text{M}$ 持续约 10 至 240 分钟,以降低或抑制信号蛋白的蓄积。例如,一种方法包括每植株施加大于 $50 \mu\text{M}$ 持续约 5

小时至 2 天,以克服乙烯对植物细胞中信号蛋白表达和蓄积的竞争性影响。例如,如实施例 7 所述,该实施例显示了诱导剂的浓度如何调节植物所具有的乙烯敏感性的程度。类似于通过施加时机进行的控制,诱导组合物的浓度可用于响应植物在不同生长阶段变化的需求或者响应变化的环境条件。

[0102] 前述各种诱导剂的组成和物理化学性质可能影响获得所需生物学效应所需要的施加时间和浓度。鉴于本文提供的教导,施加的浓度、时机和持续时间以及诱导组合物本身对有经验的培育员而言无需过多实验即可进行选择。

[0103] 通常,从停止施加诱导剂到植物对诱导剂的反应逆转之间的时间约为 2 天或更长,取决于植物大小、施加方法和诱导剂的用量。

[0104] 以下实施例详细描述了如何增加植物中 ETP1、ETP2 和的表达,和 / 或降低 EIN2 的表达可使该植物对乙烯的敏感性降低。然而,本领域技术人员不难理解,可以类似方式在植物中提高其它调控蛋白如 EBF1、EBF2 的表达,和 / 或降低另一乙烯诱导性信号蛋白如 EIN3 的表达以使该植物对乙烯的敏感性降低。一段时间后由于植物中诱导剂的消除,植物细胞对乙烯的反应被逆转、在这种情况下增加。这就使得植物能够在去除诱导剂后继续正常发育、成熟、熟化。

[0105] 将诱导剂应用于用本文所述的基因表达系统稳定转化的植物能够控制对乙烯敏感的一种或多种植物生长特性,例如枯萎、果实熟化、发芽、病原体抗性、叶片脱落、花蕾脱落、芽脱落、荚壳脱落、果实脱落和开花,以及植物对应激,例如干旱、热、群体密度和盐度等条件导致的应激的反应。

[0106] 更具体说,本文所述方法也可用作提高植物对疾病的抗性和 / 或耐受性的方法,该方法包括提高植物对乙烯的不敏感性。或者,该方法可通过提高植物对乙烯的不敏感性延迟植物、组织或器官的熟化或开花,例如用于储存或运输的目的。在另一个实施方式中,采用本文所述转化植物与在适当时机施加诱导组合物的方法能够处理处于不良生长条件,例如干旱或过热条件的植物,该方法包括施加诱导组合物以降低乙烯敏感性并使植物更易于忍受环境条件。基于本说明书的内容,植物繁殖与生长领域的技术人员可能容易地选择转化植物和诱导上述核酸序列表达的方法能够提供效益的情况。

[0107] 因此,利用本文所述方法,在植物生长期间可以改变施加诱导剂的时机、持续时间或浓度以控制乙烯敏感性,因而以相当准确地控制植物的生长特性。

[0108] IV. 实施例

[0109] 下述实施例介绍上述基因表达系统的应用,所述基因表达系统包括激活盒,该激活盒包含在组成型 G10-90 启动子控制下并与其操作性相连的 (a) GAL4DBD,该 DBD 识别包含五个拷贝的 GAL4 效应原件的效应原件;(b) 包含结构域 D、E 和 F 的蜕皮激素受体 LBD;和 (c) 在诱导组合物存在时活化的 VP16AD。靶标盒包括诱导型启动子,诱导型启动子包括操作性相连的位于最小 35S 启动子下游 5 个拷贝的 GAL4 效应元件,该效应元件对 VP16AD 起反应,诱导型启动子控制 (e) 编码 ETP1 蛋白的核酸序列的表达。根据该实施方式,在植物细胞中,活化盒和靶标盒各组件可调节 ETP1 蛋白的表达并选择性地降低植物细胞的乙烯敏感性。这种蛋白表达受到诱导组合物相互作用的控制,诱导组合物能增加选定调控蛋白的表达,进而降低 EIN2 的表达,并降低植物细胞的乙烯敏感性。此种蛋白表达调节作用可以通过施加诱导组合物的时机、浓度和持续时间进行控制。

[0110] 更具体说,示例性的基因表达系统包含位于单一质粒 p201 上的激活盒和靶标盒。该质粒的示意图如图 1 所示。另一种示例性的质粒 p202 如图 2 所示。图 1-2 所示质粒各自的基因表达盒组件的核酸序列进一步分别如 SEQ ID NO:1 和 2 所示。实施例所述的其它质粒 p1004 和 p1005 的基因表达盒组件的核苷酸序列可参见图 3A-3I 和 4A-4I,以及 SEQ ID NO:3 和 4。

[0111] 下面的实施例阐述了上述组合物和方法的某些实施方式。这些实施例并不限制权利要求书和说明书的范围。

[0112] 实施例 1:质粒

[0113] 单独克隆的基因盒组件的激活盒中包括:

[0114] G10-90 组成型启动子,

[0115] VP16 激活域,

[0116] GAL4 DNA 结合域,和

[0117] 与 NOS 终止子序列相连的蜕皮激素受体配体结合域。

[0118] 类似地,单独克隆靶标盒组件,靶标盒包括:

[0119] 由 5 个拷贝的 GAL4 效应元件和最小 35S 启动子构成的诱导型启动子,以及

[0120] ETP1 基因 (GENBANK 登录号 NM_112874) 或 ETP2 基因 (GENBANK 登录号 NM_112777) 和

[0121] 35S 终止子序列。

[0122] 所有各个序列参照下述 SEQ ID NO:1 的核苷酸编号标示。

[0123] 在组装激活盒的过程中,以下组件以两种不同顺序融合形成两种不同的激活盒:

[0124] VP16 AD 到 GAL4 LBD 到 DBD 的 EcR(DEF) (缩写为" VGE") 和

[0125] GAL4 LBD 到 VP16 AD 到 DBD 的 EcR(DEF) (缩写为 GVE)。

[0126] 应理解,虽然下述质粒对上述组件包括 EcR LBD,组件顺序如 VGE 或 GVE 有特别选择,但基因的选择还是要服从靶标盒的诱导型启动子和质粒主链,本领域技术人员无需过多实验即可选择所有这些组件,并且不难操作所述质粒。以下的具体质粒仅为示范性。

[0127] 随后,用 pBlueScript II SK⁻ 主链 (司查塔基公司 (Stratagene)) 制备质粒 DNA。SK⁻ 多克隆位点区域被包含 8bp 切割酶的识别位点的新多克隆位点取代。这种酶切位点中的一些在图 1、2、3A-3I 和 4A-4I 的示例性质粒中示出。

[0128] 制备了六个示例性的大肠杆菌 (E. coli) 质粒并测序以验证核苷酸序列。这些质粒包含独特的酶切位点,用于添加/删除/交换各个组件,如图 1、2、3A-3I 和 4A-4I 所示。因此,每个完整的构建物可转移至包括植物转化二元载体在内的所选择的任何其他载体中。

[0129] 将 SK⁻ 质粒中的构建物转移至二元质粒 pBIN19 (美国典型培养物保藏中心编号 37327)。由于 pBIN19 已经具有新霉素植物选择性标记基因、LB、RB 和 nptII 选择性标记物,所以附图和 / 或列表未示出主链序列,而是仅仅显示了感兴趣的基因表达序列,即基因表达系统的主要组件,例如克隆到左右边界之间的 Ec 受体和诱导型 ETP1 或 ETP2。

[0130] 选择以下土壤杆菌 (Agrobacterium) 二元质粒用于产生转基因植物:

[0131] p201[G10-90p-VGE-NosT-5xGAL-M35S-ETP1] 用于获得包含 G10-90 启动子驱动的 VGE 受体和诱导型 ETP1 的转基因植物。p201 的表达盒组件如 SEQ ID NO:1 所示,即 G10-90 组成型启动子 (SEQ ID NO:1 的核苷酸 1-243)、VP16 激活域 (SEQ ID NO:1 的核苷

酸 249-529)、GAL4DNA 结合域 (SEQ ID NO :1 的核苷酸 534-985)、和与 NOS 终止子序列 (SEQ ID NO :1 的核苷酸 2070-2364) 相连的蜕皮激素受体配体结合域 (SEQ ID NO :1 的核苷酸 990-1997), 由 5 个拷贝的 GAL4 效应元件 (SEQ ID NO :1 的核苷酸 2391-2492) 和最小 35S 启动子 (SEQ ID NO :1 的核苷酸 2499-2554) 组成的诱导型启动子、ETP1 基因 (SEQ ID NO :1 的核苷酸 2557-3804) 和 35S 终止子序列 (SEQ ID NO :1 的核苷酸 3828-4038)。

[0132] p202[G10-90p-GVE-NosT-5xGAL-M35S-ETP2] 是用于获得包含 G10-90 启动子驱动的 GVE 受体和诱导型 ETP2 的转基因植物的质粒。参见图 2 和 SEQ ID NO :2。p202 的表达盒组件是 G10-90 组成型启动子 (SEQ ID NO :2 的核苷酸 1-243)、GAL4 DNA 结合域 (SEQ ID NO :2 的核苷酸 276-716)、VP16 激活域 (SEQ ID NO :2 的核苷酸 717-974)、和与 NOS 终止子序列 (SEQ ID NO :2 的核苷酸 2064-2358) 相连的上述突变型 T52V 蜕皮激素受体配体结合域 (SEQ ID NO :2 的核苷酸 984-1991), 由 5 个拷贝的 GAL4 效应元件 (SEQ ID NO :2 的核苷酸 2385-2486)、最小 35S 启动子 (SEQ ID NO :2 的核苷酸 2493-2548) 组成的诱导型启动子、ETP2 基因 (SEQ ID NO :2 的核苷酸 2551-3717) 和 35S 终止子序列 (SEQ ID NO :2 的核苷酸 3741-3951)。

[0133] p1004[G10-90p-VGE-NosT-5xGAL-M35S-ETP1 和 MMVp-def-rbcS-E9t] 和其组件可参见 SEQ ID NO :3, 组件的示意图和序列可参见图 3A-3I。这些组件是 :G10-90 组成型启动子 (SEQ ID NO :3 的核苷酸 1-243)、VP16 激活域 (SEQ ID NO :3 的核苷酸 249-529)、GAL4 DNA 结合域 (SEQ ID NO :3 的核苷酸 534-985)、和与 NOS 终止子序列 (SEQ ID NO :3 的核苷酸 2070-2364) 相连的蜕皮激素受体配体结合域 (SEQ ID NO :3 的核苷酸 990-1997), 由 5 个拷贝的 GAL4 效应元件 (SEQ ID NO :3 的核苷酸 2391-2492) 和最小 35S 启动子 (SEQ ID NO :3 的核苷酸 2499-2554) 组成的诱导型启动子、ETP1 基因 (SEQ ID NO :3 的核苷酸 2557-3804)、35S 终止子序列 (SEQ ID NO :3 的核苷酸 3828-4038)、MMV 启动子 (SEQ ID NO :3 的核苷酸 6265-5629)、P-DEF 标记基因 (SEQ ID NO :3 的核苷酸 5567-4746) 和 rbcS-E9 终止子 (SEQ ID NO :3 的核苷酸 4719-4075)。

[0134] p1005[G10-90p-GVE-NosT-5xGAL-M35S-ETP2 和 MMVp-def-rbcS-E9t] 和其组件可参见 SEQ ID NO :4, 组件的示意图和序列可参见图 4A-4I。这些组件是 :G10-90 组成型启动子 (SEQ ID NO :4 的核苷酸 1-243)、GAL4 DNA 结合域 (SEQ ID NO :4 的核苷酸 276-716)、VP16 激活域 (SEQ ID NO :4 的核苷酸 717-974)、和与 NOS 终止子序列 (SEQ ID NO :4 的核苷酸 2064-2358) 相连的上述蜕皮激素受体配体结合域 (SEQ ID NO :4 的核苷酸 984-1991), 由 5 个拷贝的 GAL4 效应元件 (SEQ ID NO :4 的核苷酸 2385-2486) 和最小 35S 启动子 (SEQ ID NO :4 的核苷酸 2493-2548) 组成的诱导型启动子、ETP2 基因 (SEQ ID NO :4 的核苷酸 2551-3717)、35S 终止子序列 (SEQ ID NO :4 的核苷酸 3741-3951)、MMV 启动子 (SEQ ID NO :4 的核苷酸 6178-5542)、P-DEF 标记基因 (SEQ ID NO :4 的核苷酸 5483-4662) 和 rbcS-E9 终止子 (SEQ ID NO :4 的核苷酸 4632-3988)。

[0135] 实施例 2 :转基因拟南芥属植株的制备

[0136] 采用标准花浸渍方法, 用实施例 1 的土壤杆菌质粒转化拟南芥属植株。收获种子并接种到含卡那霉素的培养基上。选择转化植株能够在卡那霉素上生长的能力并通过 PCR 筛选以验证 ETP1 或 ETP2 基因的存在。阳性转化体自交产生 T1 种子。种子在含卡那霉素的培养基上生长以鉴定转基因纯合品系。采用纯合植株进行诱导乙烯不敏感性的试验。

[0137] 实施例 3 :调节乙烯敏感性对拟南芥属植株生长的影响

[0138] 进行三重反应分析(根据上文引用的 Guzman 和 Ecker, 1990 进行如下所述的改良)以确定实施例 2 转化的拟南芥属植株中乙烯敏感性的调节。测定野生型 *ein2-5*(乙烯不敏感型突变体对照)和 p1004 或 p1005 拟南芥属转化体幼苗。对拟南芥种子表面除菌并浸泡在 20 μ M 诱导剂中, 4 $^{\circ}$ C 保持避光 4 天。将种子接种到含 1% 蔗糖和 20 μ M 诱导剂, 含和不含 20 μ M ACC(乙烯前体)的 0.5X MS 上, 21 $^{\circ}$ C 避光培养 4-8 天。在一些实验中, 向培养基中加入 5 μ M AgNO₃(诱导乙烯不敏感的乙烯抑制剂)作为对照。最后一天记录对乙烯的反应。根据与乙烯前体 ACC 存在时培养的未诱导的转基因植株相比发芽长度增加和/或根生长改变判断, 含有激活盒和靶标盒用以在诱导剂存在下表达 ETP1 或 ETP2 的转基因植株显示出对乙烯不敏感。据预测, 该实施例的结果表明, 用基因表达系统对其有反应的选定诱导组合物处理时, 基因表达系统和用该系统转化的植株能够成功调节乙烯敏感性。根据相同条件下未诱导的转基因植株相比根长度增加, 预计含有激活盒和靶标盒用以在诱导剂存在下表达 ETP1 或 ETP2 的转基因植株对乙烯不敏感。

[0139] 实施例 4 :产生转基因番茄植株和调节乙烯敏感性对植株生长的影响

[0140] 采用标准方法, 用实施例 1 的土壤杆菌质粒转化番茄子叶片。用卡那霉素或放线酰胺素选择推定的转化体并用 PCR 试验进行验证。阳性转化体自交两次, 获得转基因纯合品系。使用纯合植物进行测定, 以类似于实施例 3 的方式诱导乙烯敏感性。

[0141] 采用三重反应试验(Guzman 和 Ecker, 1990, 同上, 如下所述进行改良)来确定转化的番茄植株中乙烯敏感性的调节。来自纯合的独立 p1004 和 p1005 品系的种子在含 20 μ M ACC(乙烯前体)和 1% 蔗糖的 0.5x MS 培养基上避光发芽。在 ACC 上发芽抑制番茄幼苗生长。来自相同品系的种子在 20 μ M ACC 加 20 μ M 诱导剂的存在下发芽。在 21 $^{\circ}$ C 避光培育幼苗 4-8 天。最后一天记录对乙烯的反应。

[0142] 根据与乙烯前体 ACC 存在时培养的未诱导的转基因植株相比发芽长度增加和/或根生长改变判断, 含有激活盒和靶标盒用以在诱导剂存在下表达 ETP1 或 ETP2 的转基因植株显示出对乙烯不敏感。该实施例进一步表明, 用基因表达系统对其有反应的选定诱导组合物处理时, 基因表达系统和用该系统转化的植株能够成功调节乙烯敏感性。

[0143] 实施例 5 :产生转基因玉米植株和调节乙烯敏感性对玉米植株生长的影响

[0144] 采用实施例 1 的质粒通过微粒轰击转化玉米植株。用卡那霉素或双丙氨磷选择推定的转化体并用 PCR 试验进行验证。阳性转化体与近交 B73 回交以增加活力。测定如上所述制备的转基因玉米植株中乙烯敏感性的调节。通过使植株暴露于 ACC(乙烯前体)并测定乙烯诱导型基因的诱导变化, 在分子水平测定乙烯敏感性的调节。从在温室中培育的转基因 T0 玉米植株切割茎鞘组织用于体外生物试验。用水或 20 μ M 诱导剂处理切下的组织 2 天以诱导 ETP1 或 ETP2 表达。诱导 2 天后, 用 0、1 或 10 μ M ACC 处理该组织 1 天以产生乙烯, 然后收获并用于制备 RNA。

[0145] 用定量 PCR 在应用生物系统公司(Applied Biosystems)7900HT 快速实时 PCR 系统(ABI)上测定乙烯诱导性基因(ACC 氧化酶)的诱导, 利用 TaqMan 测定试剂盒(ABI)根据厂商推荐方案进行逆转录(RT)和 PCR。使用玉米 18s 作为内标以均一化各样品的表达。

[0146] 引物和探针的序列如下:

[0147] 18s

[0148] 正向引物 :CGTCCCTGCCCTTTGTACAC SEQ ID NO :9

[0149] 反向引物 :ACACTTCACCGGACCATTCAA SEQ ID NO :10

[0150] 探针 :CCGCCCGTCGCTCCTACCG SEQ ID NO :11

[0151] ACC 氧化酶 (aco) :

[0152] 正向引物 :GTTGTAGAAGGACGCGATGGA SEQ ID NO :5

[0153] 反向引物 :CAGGTACAAGAGCGTCATGCA SEQ ID NO :6

[0154] 探针 :TCCTGTTCCCCTGGGCTGC SEQ ID NO :7

[0155] 为了测定基因表达,标准化各个玉米品系中 $0 \mu\text{M}$ 诱导剂加上 $0 \mu\text{M}$ ACC 对照处理中 ACC 氧化酶的表达。

[0156] 根据 ACC 氧化酶表达降低判断,含有激活盒和靶标盒用以在诱导化合物存在下表达 ETP1 或 ETP2 的转基因植株对乙烯不敏感。采用类似方法,本领域技术人员还可测定编码 EIN2 或 EIN3 乙烯诱导性基因的诱导水平降低。

[0157] 实施例 6 :ACC 浓度对乙烯不敏感性的影响

[0158] 来自一种 p1004 品系和一种 p1005 品系的纯合种子在包含 $20 \mu\text{M}$ 诱导剂和 $5 \mu\text{M}$ AgNO_3 或不同水平的乙烯前体 ACC (即 $1.0\text{--}20 \mu\text{M}$) 的培养基上发芽 4 天并生长。避光生长 10 天后测定胚轴和根的长度。由于乙烯诱导 EIN2 而 ETP1 和 ETP2 的诱导靶向 EIN2 蛋白的周转,所以预计在较高 ACC 浓度下实现的不敏感水平可能会降低,正如胚轴和根延长较少所示。这一实施例表明,诱导 ETP1 或 ETP2 能够诱导乙烯不敏感性,植株的不敏感水平可通过 ETP1 或 ETP2 的诱导水平进行调节。

[0159] 实施例 7 :植株中乙烯不敏感性的程度与诱导剂浓度的关系

[0160] 来自一种含 p1004 品系和一种含 p1005 品系的纯合种子在包含 $20 \mu\text{M}$ ACC 和各种水平的诱导剂 (即 $0.5\text{--}20 \mu\text{M}$) 的培养基上发芽。避光生长 10 天后测定胚轴和根的长度。由于乙烯诱导 EIN2 而 ETP1 和 ETP2 的诱导靶向 EIN2 蛋白的周转,所以预计在较低诱导剂浓度下实现的乙烯不敏感水平可能会降低,正如胚轴和根延长较少所示。这一实施例表明,诱导 ETP1 或 ETP2 能够诱导乙烯不敏感性,植株的不敏感水平可通过 ETP1 或 ETP2 的诱导水平进行调节。

[0161] 实施例 8 :在植株中瞬时诱导乙烯不敏感性

[0162] 为了验证当不再提供诱导剂的情况下诱导的乙烯不敏感性植株能够恢复到乙烯敏感,使来自一种 p1004 品系或 p1005 品系的纯合种子在包含 $20 \mu\text{M}$ ACC 和 $10\text{--}20 \mu\text{M}$ 诱导剂的培养基上发芽 2、4、6、8 或 14 天。避光培育 14 天后,测定胚轴和根以评价乙烯不敏感性。在 ACC 存在时,预计乙烯不敏感性幼苗的胚轴和根较长。在没有诱导剂的情况下,预计幼苗会变为对乙烯敏感并显示避光生长迟缓。因此,逆转幼苗 (reversed seedling) 的根和胚轴生长是介于敏感和不敏感幼苗间的中间状态。从诱导幼苗中去除诱导剂时,预计它们会恢复到与非诱导对照幼苗相比对乙烯敏感的状态。虽然上述实施例显示使用了 ETP1 或 ETP2 的核苷酸序列,但说明书清楚地向本领域技术人员提供了以相同方式调节其它调控基因,如 EBF1 和 EBF2 的表达的能力。类似地,乙烯能诱导乙烯信号传导途径蛋白 EIN3, EBF1 和 EBF2 靶向该蛋白的周转。因此,利用本文所述的方法和组合物在植株中诱导和调节 EBF1 和 EBF2 应高度选择性和临时性地产生乙烯不敏感状态,去除诱导剂后这种状态可恢复至乙烯敏感状态。

[0163] 上述实施方式的各种改良和变化包括在本说明书中,并且是本领域技术人员显而易见的。据信这种对本文所述组合物和方法的改良和改变包括在本文所附权利要求书的范围内。

[0164] 上文列出或引用的所有文献,包括专利、专利申请和出版物以及非专利出版物以及附图和 / 或序列列表的全部内容都包括在此作为参考,只要它们与本说明书的明确涵义不发生矛盾。然而,任何参考文献的引用不应解释为承认这些参考文献构成本申请的“现有技术”。

[0001]

序 列 表

- <110> 罗门哈斯公司(Rohm and Haas Company)
J. L. 罗西尚(Rosichan, Jeffrey L.)
- <120> 在植物中控制关键乙烯激素信号传导途径蛋白的目标周转率以调节乙烯敏感性
- <130> A02042PCT
- <150> US 61/157,759
<151> 2009-03-05
- <160> 11
- <170> PatentIn 3.5 版
- <210> 1
<211> 4038
<212> DNA
<213> 人工
- <220>
<223> 合成, 基于拟南芥属(*Arabidopsis*)和云杉卷叶蛾(*Choristoneura fumiferana*)
- <400> 1
atagtttaaa ctgaaggcgg gaaacgacaa tctgatccaa gctcaagcta agcttgcattg 60
cctgcaggat atcgtggatc caagcittgcc acgtgccgcc acgtgccgcc acgtgccgcc 120
acgtgcctct agaggatcca tctccactga cgttaaggat gacgcacaat cccactatcc 180
ttcgcaagac ccttctctata tataaggaag ttcatattcat ttggagagga cacgctggga 240
tccccaccat ggcccccccg accgatgtca gcttggggga cgaactccac ttagacggcg 300
aggacgtggc gatggcgcat gccgacgcgc tagacgattt cgatctggac atgttggggg 360
acggggattc cccaggctcg ggatttacc cccacgactc cgccccctac ggcgctctgg 420
atatggccga cttcgagttt gagcagatgt ttaccgatgc ccttgaatt gacgagtacg 480
gtgggaaagt tctaggctacc tccagaagaa tatcaggcgg ggaattcggc gggatgaagc 540
tactgtcttc tatcgaacaa gcatgcgata ttgcccact taaaaagctc aagtgtctca 600
aagaaaaacc gaagtgcgcc aagtgtctga agaacaactg ggagtgctgc tactctccca 660
aaaccaaag gctcccctg actagggcac atctgacaga agtggaatca aggctagaaa 720
gactggaaca gctatttcta ctgattttc ctgagaaga ccttgacatg atttgaaaa 780
tggattcttt acaggatata aaagcattgt taacaggatt atttgtacaa gataatgtga 840
ataaagatgc cgtcacagat agattggctt cagtggagac tgatatgcct ctaacattga 900
gacagcatag aataagtgcg acatcatcat cgaagagag tagtaacaaa ggtcaaagac 960
agttgactgt atcgggaggc ggtgggatcc ggctgagtg cgtagtacc gagactcagt 1020
gcgcatgaa gcgaaagag aagaaagcac agaaggagaa ggacaaactg cctgtcagca 1080
cgacgacggg ggacgaccac atgccgccca ttatgcagtg tgaacctcca cctcctgaag 1140
cagcaaggat tcacgaagtg gtcccaaggt tctctccga caagctgtt gtgacaaaacc 1200
ggcagaaaaa catccccag ttgacagcca accagcagtt ccttatcgcc aggctcatct 1260
ggtaccagga cgggtacgag cagccttctg atgaagattt gaagaggatt acgcagacgt 1320
ggcagcaagc ggacgatgaa aacgaagagt cggacactcc cttccgccag atcacagaga 1380
tgactatcct caggttccaa cttatcgtgg agttccgcaa gggattgcca gggttcgcca 1440
agatctcgca gctgatcaa attacgctgc ttaaggctt ctcaagtgag gtaatgatgc 1500
tccgagtcgc gcgacgatac gatgcggcct ccgacagtg tctgttcgcg aacaaccaag 1560
cgtacactcg cgacaactac cgcaaggctg gcatggccta cgtcatcgag gatctactgc 1620
acttctgccc gtgcatgtac tctatggcgt tggacaacat ccattacgag ctgctcacgg 1680
ctgtctcat ctttctgac cggccagggt tggagcagc gcaactggtg gaagagatcc 1740
agcggtaacta cctgaatacg ctccgcatct atatcctgaa ccagctgagc gggctcggcg 1800
gttctgctcg catatacggc aagatcctct caatcctctc tgagctacgc acgctcggca 1860
tgcaaaactc caacatgtgc atctccctca agctcaagaa cagaaagctg ccgctttcc 1920
tcgaggagat ctgggatgtg gcggacatgt cgcacacca accgccct atcctcgagt 1980
ccccacgaa tctctagccc ctgcgcgcac gcacaccga tgcgcgctcc ggccgctg 2040
ctctgagaat tcgatatcaa gcttctagac cgggctgca gagatctac cgttaagctt 2100
aattcccgat cgttcaaaca ttggcaata aagtttctta agatgaaac ctgttccgg 2160
tcttgcatg attatcatat aatttctgtt gaattacgtt aagcatgtaa taattaacat 2220
glaatgcatg acgttattta tgagatgggt ttttatgatt agagtcccgc aattatacat 2280
ttaatacgcg atagaaaaca aaatatagcg cgcaactag gataaattat cgcgcgctg 2340
gtcatctatg ttactagatc ggggactagt aaggccggcc gcttggatcc gctcggagga 2400
cagtactccg ctccgaggac aglactccgc tcggaggaca gtactccgct cgaggacagt 2460
actccgctcg gaggacagta ctccgatccg tcagatctgc aagacccttc ctctataata 2520
ggaagttcat ttcatttga gaggacagc tgaacatga cgataccgga tctctgtaac 2580

[0002]

gatttggctg	atgagatact	ctgtcgcgtt	ccggcgagga	atctgaaacg	gttacgatct	2640
accagcaaac	gatggaaccg	tttattcaaa	gatgatagga	gattcgcaag	agagcacatg	2700
cataaagccc	caaaggagta	tctacctctc	atgittgaaa	gcgagtacag	gatctgtccc	2760
gtgacgatca	atctccaagg	agatgttcct	tctgtagtgt	taaagagaga	gcttagccta	2820
ccagatccgg	attattcaca	tcaattcgat	ataggctgag	tctttcactg	cgacggccta	2880
ttgggatgca	accacgtagg	caagaatccc	cgatacggat	ctaaaaatcg	ggtttggaac	2940
ccgctgactg	gtcaaaccag	gtggatcgaa	gccggctatc	gttggaagga	atacgaagtc	3000
agatttgttc	tcggatactg	ctaccagcaa	gacgagaaca	attcctgcag	taaaaaaatc	3060
tacaaaattt	tgtgttttta	tcctaattggc	caagatacag	aaatctacga	gcttaactac	3120
tctgataggt	ggacaaggac	gattcctgat	ggtgatctca	ctccaggctg	gaccttgata	3180
tactcagagc	agaccgtgtc	tatgaatgga	aatctttact	tgtttgcttc	ggagaaatca	3240
aaaccccatc	ttggcgtgtc	cttgctcaga	tttgatttct	caacagagaa	atcatctcta	3300
tgtgtgactc	tccctatca	gcgtccaagg	tatgaaattt	tgagtatttc	cgccgttaga	3360
ggaggagaga	atctttctct	gttgttgacg	ctcgattttg	aatctaagac	tgagatatgg	3420
gtgacgaata	agattgatga	caccaccacc	aaaggagcag	cagtccttg	gaccaaggte	3480
ctagcatttg	atttaagccc	tgatcttcaa	ttattttcgg	aggaggtaaa	ttttttgctt	3540
gacgaggata	agaaagtcgc	tgtgtgtgtg	gagagatggt	tggaaccgca	agagcaccac	3600
agggtaccagt	gcaggagaga	gtacaagatc	accgacaaga	tatacattct	cggggaggat	3660
aataaagtcg	atgaagtagg	ttctggagag	ggagaggcta	cagattcact	tgaaggaatt	3720
tcgcaagtta	ttctcaatta	cgctccaagt	ttggtcctaa	tcgagcaagc	cggaggaggc	3780
aaaacaaaaa	gaggtgacga	ctaagcggcc	gctagggcat	gtctagaagt	ccgcaaaaat	3840
caccagtcct	tctctacaaa	tctatctctc	tctatttttc	tccagaataa	tgtgtgagta	3900
gttcccagat	aaggaatta	gggttcttat	agggtttcgc	tcatgtgttg	agcatataag	3960
aaacccttag	tatgtatttg	tatttgtaaa	atacttctat	caataaaatt	tctaattcct	4020
aaaacaaaaa	tccagtga					4038

<210> 2

<211> 3951

<212> DNA

<213> 人工

<220>

<223> 合成, 基于拟南芥属(Arabidopsis)和云杉卷叶蛾(Choristoneura fumiferana)

<400> 2

atagtttaaa	ctgaaggcgg	gaaacgacaa	tctgatccaa	gctcaagcta	agcttgcattg	60
cctgcaggat	atcgtggatc	caagcttgcc	acgtgcccgc	acgtgcccgc	acgtgcccgc	120
acgtgacctc	agaggatcca	tctccactga	cgtaagggat	gacgcacaat	cccactatcc	180
ttcgcaagac	ccttcctcta	tataagggaag	ttcatttctat	ttggagagga	cacgctggga	240
tccccaccat	ggatccgcca	ccatgctagc	ccaccatgaa	gctactgtct	tctatcgaac	300
aagcatgcca	tatttgccga	cttaaaaaagc	tcaagtgtct	caaagaaaaa	ccgaagtgcg	360
ccaagtgtct	gaagaacaac	tgggagtgtc	gctactctcc	caaaacccaa	aggtctccgc	420
tgactagggc	acatctgaca	gaagtggaa	caaggctaga	aagactggaa	cagctatttc	480
tactgatttt	tctcagagaa	gaccttgaca	tgattttgaa	aatggattct	ttacaggata	540
taaaagcatt	gttaacagga	ttatttgtac	aagataatgt	gaataaagat	gccgtcacag	600
atagattggc	ttcagtgag	actgatatgc	ctctaaccatt	gagacagcat	agaataagtg	660
cgacatcatc	atcggaagag	agtagtaaca	aaggtcaaag	acagttgact	gtatccatgg	720
cccccccgac	cgatgtcagc	ctgggggacg	aactccactt	agacggcgag	gacgtggcga	780
tggcgcgatc	cgacgcgcta	gacgatttcg	atctggacat	gttgggggac	ggggattccc	840
caggctccggg	atttaccccc	caogactccg	ccccctacgg	cgcctctggat	atggccgact	900
tcagatttga	gcagatgttt	accgatgcc	ttggaattga	cgagtacggt	gggaagcttc	960
taggtacctc	tagaagaata	tcgtggcctg	agtgcgtagt	acccgagact	cagtgcgcca	1020
tgaagcggaa	agagaagaaa	gcacagaagg	agaaggacaa	actgcctgtc	agcacgacga	1080
cggtggacga	ccacatgccg	cccattatgc	agtgtgaacc	tccacctcct	gaagcagcaa	1140
ggattcaaga	agtggtccca	aggtttctct	ccgacaagct	ggtggagaca	aaccggcaga	1200
aaaacatccc	ccagttgaca	gccaaccagc	agttccttat	cgccaggctc	atctggtacc	1260
aggacgggta	cgagcagcct	tctgatgaag	atttgaagag	gattacgcag	acgtggcagc	1320
aagcggacga	tgaaacgaa	gagtcggaca	ctcccctccg	ccagatcaca	gagatgacta	1380
tcctcacggg	ccaacttata	gtggagttcg	cgaagggatt	gccagggttc	gccaagatct	1440
cgcagcctga	tcaaatctacg	ctgcttaagg	cttgctcaag	tgaggtaatg	atgctccgag	1500
tcgctgcgac	atcagatgcg	gcctccgaca	ggttctgtgt	cgcaacaac	caagcgtaca	1560
ctcgcgacaa	ctaccgcaag	gctggcatgg	ctctacgtcat	cgaggatcta	ctgcaactct	1620
gccggtgcat	gtactctatg	gcgttggaca	acatccatta	cgcgctgctc	acggctgtcg	1680
tcatcttttc	tgaccggcca	gggttggagc	agccgcaact	ggtggaagag	atccagcggg	1740
actacctgaa	tacgctccgc	atctatatcc	tgaaccagct	gagcgggtcg	gcgcgttcgt	1800
ccgtcatata	cggaagatc	ctctcaatcc	tctctgagct	acgcacgctc	ggcatgcaaa	1860
actccaacat	gtgcactccc	ctcaagctca	agaacagaaa	gctgccgect	ttctcagagg	1920
agatctggga	tgtggcggac	atgtcgcaca	cccaaccgcc	gcctatcctc	gagtccecca	1980

[0003]

cgaatctcta	gcccctgcgc	gcacgcacgc	ccgatgccgc	gtccggccgc	gctgctctga	2040
gaattcgata	tcaagcttct	agaccggggc	tcagagatc	tacgcgtaa	gcttaattcc	2100
cgatcgttca	aacatttggc	aataaagt	cttaagattg	aatcctgttg	ccggtcttgc	2160
gatgattatc	atataat	tgttgaatta	cgtaagcat	gtaataatta	acatgtaatg	2220
catgacgtta	tttatgagat	gggtttttat	gattagagtc	ccgcaattat	acatttaata	2280
cgcgatagaa	aacaaaaat	agcgcgc	ctaggataaa	ttatcgcgc	cggtgtcatc	2340
tatgttacta	gatcggggac	tagtaaggcc	ggccgcttgg	atccgctcgg	aggacagtac	2400
iccgcctgga	ggacagtact	ccgctcggag	gacagtactc	cgctcggagga	cagtactccg	2460
ctcggaggac	agtactccga	tccgtcagat	ctgcaagacc	cttccctctat	ataaggaagt	2520
tcatttcatt	tggagaggac	acgctgaacc	atgaagacaa	tacaggagca	gcttccaaat	2580
gacttggtag	aggagatact	ctgtcgcggt	ccggcaacat	ctctgagacg	tttacgatcg	2640
acttgc	catggaaccg	tttattcaaa	ggtgatcggga	tattagcaag	taagcatttt	2700
gaaaaatccg	caaaacagtt	tagatctcta	tcgttaagga	atgattacag	gatttttccg	2760
attagcttca	atctccatgg	aaatagcca	tccttagagc	ttaaaagtga	gctaatacgt	2820
ctctattcta	agaattcagc	tgtctcattc	gaaatatctc	gagtcattca	ctgtgagggga	2880
ttgttgttgt	gctcctccca	attggacgaa	tcctagagtcg	tggtttggaa	tcctttaacc	2940
ggtgaaacca	ggtggatcag	aaccggcgat	tttcgccaaa	aaggccgtag	ctttgatgtc	3000
gggtactact	accaaaaaga	caagagatcc	tgatcaaga	gctacaaact	ctttgtgctat	3060
tatcgtggtg	ccaaatattt	tgaaatctac	gattttgact	ctgattcatg	gaggattctt	3120
gatgatatta	tcgctccacg	ggggagtatt	ggatactcgg	aacttagcgt	gtctctgaaa	3180
ggaaatactt	actggttcgc	taaagggtga	acagaagagc	ggccccgcac	catactattg	3240
ctcaaatitg	atTTTTtatac	agagaaatct	gtacctgtgc	ttcttcccta	tcagagtcgt	3300
cgctctttcc	aagctagtag	cttttctgtt	gttagagaag	ataaactttc	tggtttattg	3360
cagctagatc	aaagtccaa	gactgagata	tggttgacaa	atgtgattga	tgagaccacc	3420
aaaggagcag	tttcttggac	caaggcttca	gcattggatt	tgagccctca	cttccagatt	3480
gggaatgatg	gaagtttctt	cttaggcgag	gataagaaag	tcgtcatggt	ctgtgagaaa	3540
ttgattgatg	agaacaaggt	caaagacatg	gtctacattg	ttggggagga	taatgttgtc	3600
acagaagtgg	gatttggagt	agatgaaatg	gatggalgtc	gggcagttat	cttaattat	3660
gttccaaagt	tggttcaaat	cgagcgcgct	ggaggcaaca	ggaaaagagg	gcactaagcg	3720
cccgcctagg	catgtctaga	agtcgcgaaa	aatcaccagt	ctctctctac	aaatctatct	3780
ctctctat	ttctccagaa	taatgtgtga	gtagttccca	gataagggaa	ttagggttct	3840
tatagggttt	cgctcatgtg	ttgagcatat	aagaaacct	tagtatgtat	ttgtatttgt	3900
aaaatacttc	tatcaataaa	atttctaatt	cctaaaacca	aatccagtg	a	3951

- <210> 3
- <211> 8716
- <212> DNA
- <213> 人工

- <220>
- <223> 合成, 基于拟南芥属(Arabidopsis)和云杉卷叶蛾(Choristoneura fumiferana)

<400> 3						
atagtttaaa	ctgaaggcgg	gaaacgacaa	tctgatccaa	gctcaagcta	agcttgcgatg	60
cctgcaggat	atcgtggatc	caagcttgcc	acgtgccgcc	acgtgccgcc	acgtgccgcc	120
acgtgcctct	agaggatcca	ctccactga	cgtaaaggat	gacgcacaat	cccactatcc	180
ttcgcaagac	cttctctcta	tataagggaag	ttcatttcat	ttggagagga	cacgctggga	240
tccccacat	ggcccccccg	accgatgtca	gcctggggga	cgaactccac	ttagacggcg	300
aggacgtggc	gatggcgcat	gccgacgcgc	tagacgattt	cgatctggac	atgttggggg	360
acggggattc	cccaggtccg	ggattttacc	cccacgactc	cgccccctac	ggcgcctctg	420
atatggccga	cttcgagttt	gagcagatgt	ttaccgatgc	ccttgggaat	gacgagtagc	480
gtgggaagct	tctaggtacc	tccagaagaa	tatcaggcgg	ggaattcggc	gggatgaagc	540
tactgtcttc	tatcgaacaa	gcatgcgata	tttccgact	taaaaagctc	aagtgtctca	600
aagaaaaacc	gaagtgcgcc	aagtgtctga	agaacaactg	ggagtgtcgc	tactctccca	660
aaaccaaaag	gtctccgctg	actagggcac	atctgacaga	agtggaatca	aggctagaaa	720
gactggaaaca	gctatttcta	ctgatttttc	ctcgagaaga	ccttgacatg	atTTTgaaaa	780
tggattcttt	acaggatata	aaagcattgt	taacaggatt	atTTgtacaa	gataatgtga	840
ataaagatgc	cgtcacagat	agattggctt	cagtggagac	tgatatgcct	ctaacattga	900
gacagcatag	aataagtgcg	acatcatcat	cggaaagagag	tagtaacaaa	ggtcaaagac	960
agttgactgt	atcgggaggc	ggtgggatcc	ggcctgagtg	cgtagtacc	gagactcagt	1020
gcgcatgaa	gcggaagag	aagaaagcac	agaaggagaa	ggacaaactg	cctgtcagca	1080
cgacgacggt	ggacgaccac	atgccgccca	ttatgcagtg	tgaacctcca	cctcctgaag	1140
cagcaaggat	tcacgaagtg	gtcccaaggt	ttctctccga	caagctgttg	gtgacaaacc	1200
ggcagaaaaa	catccccag	ttgacagcca	accagcagtt	ccttatcgcc	aggctcatct	1260
ggtaccagga	cgggtacgag	cagccttctg	atgaagattt	gaagaggatt	acgcagacgt	1320
ggcagcaagc	ggacgatgaa	aacgaagagt	cggacactcc	cttccgccag	atcacagaga	1380
tgactatcct	cacggtccaa	cttatcgtgg	agttcgcgaa	gggattgccca	gggttcgcca	1440
agatctcgca	gcctgatcaa	attacgctgc	ttaaggcttg	ctcaagtgag	gtaatgatgc	1500

[0004]

tccgagtcgc	gcgacgatac	gatgcggcct	ccgacagtgt	tctgttcgcg	aacaaaccaag	1560
cgtacactcg	cgacaactac	cgcaaggctg	gcatggccta	cgatcatcgag	gatctactgc	1620
acttctgccc	gtgcatgtac	tctatggcgt	tggacaacat	ccattacgcg	ctgctcacgg	1680
ctgtcgtcat	cttttctgac	cggccagggt	tggagcagcc	gcaactgggt	gaagagatcc	1740
agcggtacta	cctgaatacg	ctccgcatct	atatectgaa	ccagctgagc	gggtcggcgc	1800
gttcgtccgt	catatacggc	aagatcctct	caatcctctc	tgagctacgc	acgctcggca	1860
tgcaaaactc	caacatgtgc	atctccctca	agctcaagaa	cagaaagctg	ccgcctttcc	1920
tcgaggagat	ctgggatgtg	gcggacatgt	cgcacaccca	accgccgcct	atcctcgagt	1980
ccccacgaa	tctctagccc	ctgcgcgcac	gcatcgccga	tgccgcgtcc	ggccgcgctg	2040
ctctgagaat	tcgatatcaa	gcttctagac	ccgggctgca	gagatctacg	cgtaagctt	2100
aattcccgat	cgttcaaaaca	tttggcaata	aagtttctta	agattgaaat	ctgttgcgg	2160
tcttgcgatg	attatcatat	aatttctgtt	gaattacgtt	aagcatgtaa	taattaacat	2220
gtaatgcatg	acgttatit	tgagatgggt	ttttatgatt	agagtccgcg	aattatacat	2280
ttaatacgcg	atagaaaaca	aaatatagcg	cgaaaactag	gataaattat	cgcgcgcggt	2340
gcatctatg	ttactagatc	ggggactagt	aaggccggcc	gcttggatcc	gctcggagga	2400
cagttactccg	ctcggaggac	agtactccgc	tcggaggaca	gtactccgct	cgaggacagt	2460
actccgctcg	gaggacagta	ctccgatccg	tcagatctgc	aagacccttc	ctctataata	2520
ggaagtccat	ttcatttggg	gaggcacacgc	tgaacctga	cgataccgga	tctctgtaac	2580
gatttggctg	atgagatact	ctgtcgcggt	ccggcgagga	atctgaaacg	gttacgatct	2640
accagcaaac	gatggaaccg	tttatcaaaa	gatgatagga	gattcgcgag	agagcacatg	2700
cataaagccc	caaaggagta	tctacctctc	atgttgacaa	gcgagtacag	gatctgtccg	2760
gtgagcatca	atctccaagg	agatgttctt	tctgtagtgt	taaagagaga	gcttagccta	2820
ccagatccgg	attatcaca	tcaatcogat	ataggtcgag	tctttcactg	cgacggctta	2880
ttggtatgca	accacgtagg	caagaatccc	cgataccgat	ctaaaatcgt	ggtttggaac	2940
ccgctgactg	gtcaaaaccg	gtggatcgaa	gccggctatc	gttggaaagga	atacgaagtc	3000
agatttgttc	tcggatactg	ctaccagcaa	gacgagaaca	atcctgcag	taaaaaaatc	3060
tacaaaat	tgtgttttta	tcctaattgc	caagatacag	aaatctacga	gcttaactac	3120
tctgataggt	ggacaaggac	gattcctgat	ggtgatctca	ctccaggctg	gaccttgata	3180
tactcagagc	agaccgtgtc	tatgaatgga	aatctttact	tgtttgcttc	ggagaaatca	3240
aaaccccact	ttggcgtgtc	cttgcacaga	tttgatttct	caacagagaa	atcatctcta	3300
tggttgactc	tccctatca	gcgtccaagg	tatgaaaatt	tgagtatttc	cgccgttaga	3360
ggaggagaga	atctttctct	gttgttcgag	ctcgattttg	aatctaagac	tgagatatgg	3420
gtgacgaata	agattgatga	caccaccacc	aaaggagcag	cagtctcttg	gaccaaggtc	3480
ctagcatttg	atllaaagccc	tgatcttcaa	ttatlltcgg	aggaggtaaa	ttttttgctt	3540
gacgaggata	agaaaagtcgc	tgtgtgttgt	gagagatggt	tggaaccgca	agagcaccac	3600
aggttaccagt	gcaggagaga	gtacaagatc	accgacaaga	tatacatctc	cggggagatg	3660
aataaagtcg	atgaagttag	ttclggagag	ggagaggcta	cagattcact	tgaaggaatt	3720
tcgcaagtta	ttctcaatta	cgctccaagt	ttggttcaaa	tcgagcaagc	cgaggagggc	3780
aaaacaaaaa	gaggtgacga	ctaagcggcc	gctagggcat	gtctagaagt	ccgcaaaaaa	3840
caccagctctc	tctctacaaa	tctatctctc	tctatllttc	lccagaataa	tgtgtgagta	3900
gttcccagat	aagggaatta	gggttcttat	agggtttcgc	tcagtgtgtg	agcatataag	3960
aaacccttag	tatgtatitg	tatttgtaaa	atacttctat	caataaaatt	tctaattcct	4020
aaaaccaaaa	tccagtgact	gcaggcatgc	aagcttatcg	ataccgtcga	cgattgatgc	4080
atgttgtcaa	tcaattggca	agtcataaaa	tgcattaaaa	aatatlltca	tactcaacta	4140
caaattccatg	agtataacta	taattataaa	gcaatgatta	gaatctgaca	aggattctgg	4200
aaaattacat	aaaggaaagt	tcataaatgt	ctaaaacaca	agaggacata	ctgtattca	4260
gtaacatttg	cagcttttct	aggctctgaaa	atataattgt	tgccctagtga	ataagcataa	4320
tggtacaact	acaagtgttt	tactcctcat	attaactctg	gtcattagag	gccacgattt	4380
gacacatttt	tactcaaaac	aaaatgtttg	catatctctt	ataatllcaa	attcaacaca	4440
caacaaataa	gagaaaaaac	aaataatatt	aatlltgaaa	tgaacaaaag	gaccataatca	4500
ttcattaact	cttctccatc	catttccatt	tcacagttcg	atagcgaaaa	ccgaataaaa	4560
aacacagtaa	attacaagca	caacaaatgg	tacaagaaaa	acagttttcc	caatgccata	4620
atactcaaac	tcagtaggat	tctgtgtgtg	gcgcaatgaa	actgatgcat	tgaacttgac	4680
gaacgttgtc	gaaaccgatg	atacgaacga	aagctctaga	ggatcaattc	gagctcttag	4740
gtcgacccac	glltgccaaa	accaactcct	gctctccttt	tttgcgtgct	ttctactctt	4800
tcagggtctg	gcaatccagt	ttttctctcg	tacttctttt	ctagggcctc	tagctcttca	4860
cgaatgctat	caagaacttg	atccgtcatt	ctgtcaagaa	agagaactcc	ctccagatgg	4920
tcgtattcgt	gctgaaagat	tcgtgcaggt	aaacgtgata	gactgattga	aaatctttca	4980
ccagtaatat	cccttgcac	aatcttgaca	gattgtggtc	gaacaacttc	agcatagatc	5040
cccgggaagg	agaggcatcc	ttcatcaaac	ggtactaatt	tatcggaata	ttctttgatt	5100
ttcggattta	caaggacaat	ttctttctct	ctccaggctc	ctccagctgg	atlaaacacc	5160
atgagttgaa	cattgagacc	tacttgtggt	gctgagagcc	caatgccatc	cgttttgtac	5220
ataacatcaa	acatagcact	aaccaagttc	tttaattctc	cgtaaaaaat	atcaatctct	5280
ttgttcttag	cccgtagtat	aggatccgga	tactcaacaa	tcttcaagg	cgcttcaaat	5340
tgaacatcag	tagctgaagc	tactttatcg	tctttacgag	agacgcgttt	tacttctgag	5400
cggaaccgaag	atgtcagagg	actggctcgg	ttcacagtag	agcagaactg	gacctggatg	5460
ttgagccgac	cataaccggc	agagagagta	gtagctcggc	gagataaaac	cggtaggagt	5520
atgagagaga	gtgggtggagc	ttggagggaag	cagttacaga	cggtcccat	ggtggaagta	5580

[0005]

```

tttgaagaa aattaaat aaaaagatcc gctcaggat ccaagcttag atgagagatt 5640
tcgattccga ttttgattc gattccgatt ttgatttcca ttgatctctt ccttctgatt 5700
tgtgttcctt atataaggaa attcttgtgg gatttagact catggcttac gtcatttctt 5760
tcgtcctgtt gctcactgat tgagctgtga gtggaggac cactggaaga tgcttcacta 5820
atthtcttag tggaggacc ggcttcacat gcttcacaca agtggctgtc gggcatcatc 5880
ttttttagct ttgacaaag caatgtttta gtggaggctc ccactcttat ctccaacatt 5940
attatcttat ctccaagga cgataagatg ttgatgtctg tggacgaagt tgggattaga 6000
cgtcatggct tacgtcattt ccttcgtcct gttgctcact gatfagctg tgagtggagg 6060
gaccactgga agatgcttca ctaatthtct tagtggaggg accggcttct catgcttcac 6120
acaagtggct gtcgggcac atctthttha gctthtgaac aagcaatggt ttagtggggg 6180
ctcccactct tatcttaaac attattatct tatcttcaaa ggacgataag atgttgatgt 6240
ctgtggacga agttgacgaa ttgacactg caggcatgca agcttggcgt aatcatggct 6300
atagctgttt cctgtgtgaa attgttatcc gctcacaaat ccacacaaca tacgagccgg 6360
aagcataaag tgtaaagcct ggggtgccta atgagtgagc taactacat taattgctt 6420
gcgctcactg ccgcttctc agtcgggaaa cctgtcgtgc cagctgcatt aatgaatcgg 6480
ccaacgcgcg gggagaggcg gtttgcgtat ttggccaaag acaaaaggcc gatattcaac 6540
cgattgaggg agggaaggta aatattgacg gaaattatc attaaagggt aattatcacc 6600
gtcaccgact tgagccattt gggaattaga gccagcaaaa tcaccagtag caccattacc 6660
attagcaagg ccggaacagt caccaatgaa accatcgata gcagcaccgt aatcagtagc 6720
gacagaatca agtttgcctt tagcgtcaga ctgtagcgcg thtcatcgg cattttcgg 6780
catagccccc ttattagcgt ttgccatctt ttcataatca aaatcaccgg aaccagagcc 6840
accaccggaa ccgctcctc cagagccgcc accctcagaa ccgccacct cagagccacc 6900
accctcagag ccgccaccag aaccaccacc agagccgcgc ccagcattga caggaggccc 6960
gatctagtaa catagatgac accgcgcgcg ataatttctc ctagtthtgc cgctatatt 7020
tgtthtctat cgcgtattaa atgtataatt gcgggactct aatcataaaa acccatctca 7080
taaataacgt catgcattac atgttaatta ttacatgctt aacgtaattc aacagaat 7140
atatgataat catgcaaga ccggcaacag gattcaatct taagaaactt tattgcaaa 7200
tgtttgaacg atcggggatc atccgggtct gtggcgggaa ctccacgaaa atatccgaac 7260
gcagcaagat atcgcggtgc atctcggctt gccttgggca gtcgccgccg acgccgttga 7320
tgtggacgcc gggcccgatc atattgtcgc tcaggatcgt ggcgttgtgc ttgtcggccg 7380
ttgctgtcgt aatgatatcg gcaccttca ccgctgttc ccagagatc ccgtggcga 7440
agaactccag catgagatcc ccgcgttga ggatcatcca gccggcgtcc cggaaaacga 7500
ttccgaagcc caaccttca tagaaggcgg ccgttgaatc gaaatctcgt gatggcagg 7560
tgggcgtcgc ttggtcggtc atttcaaac ccagatccc gtcagaaga actcgtcaag 7620
aaggcgatag aaggcgatgc gctgcgaatc gggagcggcg ataccgtaa gcacgaggaa 7680
gcggtcagcc cattcgcgc caagctctc agcaatatca ccggtagcca acgctatgtc 7740
ctgalagcgg tccgccacac ccagccggcc acagtcgat aatccagaaa agcggccatt 7800
ttccaccatg atattcggca agcaggcatc gccatgggtc acgacgagat catgccgtc 7860
gggcatgcgc gcccttgacc tggcgaacag ttcggctggc gcgagccctt gatgctctt 7920
gtccagatca tctgatcga caagaccggc ttcatecca gtacgtgctc gctcgtatgc 7980
atgtttcgtc ttgttgcga atgggcagg atccggatca agcgtatgca gccgccgcat 8040
tgcatcagcc atgatggata ctttctcggc aggagcaagg tgagatgaca ggagatcctg 8100
ccccggcact tcgccaata gcagccagtc cctccccgtc tcagtgacaa cgtcagcac 8160
agctgcgcaa ggaacgcccg tctgtggccag ccacgatagc ccgctgcctc cgtcctgcag 8220
ttcattcagg gcaccggaca ggtcggcttt gacaaaaaga accgggcgcc cctgcgtga 8280
cagccggaac accggcgcat cagagcagcc gattgtctgt tgtgccagt catagccgaa 8340
tagcctctcc acccaaggcg ccggagaacc tgcgtgcaat ccactttgtt caatcatgcg 8400
aaacgatcca gatccgggtc agattatttg gattgagagt gaatatgaga ctctaattgg 8460
ataccgaggg gaatttatgg aacgtcagtg gagcatttt gacaagaaat atttgctagc 8520
tgatagtgac cttaggcgac ttttgaacgc gcaataatgg tttctgacgt atgtgcttag 8580
ctcattaaac tccagaaacc cgcggctgag tggctcctc aacgttgcgg ttctgtcagt 8640
tccaaacgta aaacggcttg tcccgcgtca tggcgggggg tcataacgtg actcccttaa 8700
ttctccgctc atgac 8716

```

<210> 4

<211> 8629

<212> DNA

<213> 人工

<220>

<223> 合成, 基于拟南芥属(*Arabidopsis*)和云杉卷叶蛾(*Choristoneura fumiferana*)

<400> 4

```

atagtttaaa ctgaaggcgg gaaacgacaa tctgatccaa gctcaagctc agcttgcagt 60
cctgcaggat atcgtggatc caagcttccc acgtgccgcc acgtgccgcc acgtgccgcc 120
acgtgcctct agaggatcca tctccactga cgtaaaggat gacgcacaat cccactatcc 180
ttcgcaagac ctttctctta tataagggaag ttcatthtcat ttggagagga cacgctggga 240
tcccaccat ggatccgcca ccagcttagc ccaccatgaa gctactgtct tctatcgaac 300

```

[0006]

aagcatgcga	tatttgccga	cttaaaaaagc	tcaagtgctc	caaagaaaa	ccgaagtgcg	360
ccaagtgtct	gaagaacaac	tgggagtgct	gctactctcc	caaaccaaa	aggtctccgc	420
tgactagggc	acatctgaca	gaagtggaat	caaggctaga	aagactggaa	cagctatttc	480
tactgatttt	tcctcgagaa	gaccttgaca	tgattttgaa	aatggattct	ttacaggata	540
taaaagcatt	gttaacagga	ttatttgtag	aagataatgt	gaataaagat	gccgtcacag	600
atagattggc	ttcagtgagg	actgatatgc	ctctaaccatt	gagacagcat	agaataagtg	660
cgacatcatic	atcggaagag	agtagtaaca	aaggtcaaaag	acagttgact	gtatccatgg	720
ccccccgac	cgatgtcagc	ctgggggacg	aactccactt	agacggcgag	gacgtggcga	780
tggcgcagtc	cgacgcgcta	gacgatttcg	atctggacat	gttgggggac	ggggattccc	840
caggtccggg	atttacccc	cacgactccg	ccccctacgg	cgctctggat	atggccgact	900
tcgagtttga	gcagatgttt	accgatgccc	ttggaattga	cgagtacggg	gggaagttc	960
taggtacctc	tagaagaata	tcgtggcctg	agtgcgtagt	acccgagact	cagtgcccca	1020
tgaagcggaa	agagaagaaa	gcacagaagg	agaaggacaa	actgcctgtc	agcacgacga	1080
cggtggacga	ccacatgccg	cccattatgc	agtgtgaacc	tccacctcct	gaagcagcaa	1140
gaattcacga	agtggtocca	aggtttctct	ccgacaagct	gttggagaca	aaccggcaga	1200
aaaacatccc	ccagttgaca	gccaacccagc	agttccctat	cgccaggctc	atctgtacc	1260
aggacgggta	cgagcagcct	tctgatgaag	atltgaagag	gattacgcag	acgtggcagc	1320
aagcggacga	tgaaaacgaa	gagtcggaca	ctcccctccg	ccgatcaca	gagatgacta	1380
tcctcacggt	ccaacttatc	gtggagttcg	cgaagggatt	gccagggttc	gccaagatct	1440
cgagcctga	tcaaattacg	ctgcttaagg	cttgctcaag	tgaggtaatg	atgctccgag	1500
tcgctcgacg	atacagatgc	gcctccgaca	gtgttctggt	cgcgaacaac	caagcgtaca	1560
ctcgcgacaa	ctaccgcaag	gctggcatgg	cctacgtcat	cgaggatcta	ctgcacttct	1620
gccggtgcat	gtactctatg	gcgttggaca	acatccatta	cgcgctgctc	acggctgtcg	1680
tcactctttc	tgaccggcca	gggttggagc	agccgcaact	gggtggaagag	atccagcggg	1740
actacctgaa	taagctccgc	atctatatcc	tgaaccagct	gagcgggtcg	gcgcgttctg	1800
ccgctatata	cggcaagatc	ctctcaatcc	tctctgagct	acgcacgctc	ggcatgcaaa	1860
actccaacat	gtgcactctc	ctcaagctca	agaacagaaa	gctgccgctt	ttcctcgagg	1920
agatctggga	tgtggcggac	atgtcgcaca	cccaaccgcc	gcctatctc	gagtccccca	1980
cgaatctcta	gccccctgcg	gcacgcacgc	ccgatgccgc	gtccggccgc	gctgctctga	2040
gaattcgata	tcaagcttct	agaccggggc	tgcagagatc	taagcgttaa	gcttaattcc	2100
cgatcgttca	aacatttggc	aataaaagttt	cttaagattg	aatcctgttg	ccggcttgc	2160
gatgattatc	atataatttc	tgttgaatta	cgttaaagat	gtaataatta	acatgtaatg	2220
catgacgtta	tttatgagat	gggtttttat	gattagagtc	ccgcaattat	acatttaata	2280
cgcgatagaa	aacaaaalal	agcgcgcaaa	ctaggataaa	ttatcgcgcg	cggtgtctac	2340
tatgttacta	gatcggggac	tagtaaggcc	ggccgcttgg	atccgctcgg	aggacagtac	2400
tcgctcggga	ggacagtagt	ccgctcggag	gacagtactc	cgctcagga	cagtactccg	2460
ctcggaggac	agtactccga	lccglcagat	ctgcaagacc	cttctctat	ataaggaagt	2520
tcatttcatt	tggagaggac	acgctgaacc	atgaagacaa	tacaggagca	gcttccaaat	2580
gacttggtag	aggagatact	ctgtccgctt	ccggcaacat	ctctgagacg	tttacgatcg	2640
acttgcaaa	catggaaccg	tttatlcaaa	ggtgatcggg	tattagcaag	taagcatttt	2700
gaaaatccc	caaaaacagtt	tagatctcta	tcgttaaagga	atgattacag	gatttttccg	2760
attagcttea	atctccatgg	aaatagtcca	tctctagagc	ttaaaagtga	gctaatcgat	2820
cctcattcta	agaattcagc	tgctccattc	gaaatatctc	gagtcattca	ctglgagggg	2880
ttgttgttgt	gctcctccca	attggacgaa	tctagagtcg	tggtttggaa	tcctttaacc	2940
gggtgaacca	gggtgatcag	aaccggcgat	tttcgcaaa	aaggccgtag	ctttgatgtc	3000
gggtactact	accaaaaga	caagagatcc	tggatcaaga	gctacaactc	cttgtgctat	3060
tatcgtggta	ccaaatattt	tgaatctac	gattttgact	ctgattcatg	gaggattctt	3120
gatgatatta	tcgctccacg	ggggagtatt	ggatactcgg	aacttagcgt	gtctctgaaa	3180
ggaaatactt	actggttcgc	taaagggtga	acagaagagc	ggccccgcac	catatcattg	3240
ctcaaatttg	atltttatac	agagaaatct	gtacctgtgc	ttcttcccta	tcagagtcgt	3300
cgctctttcc	aagctagtag	ccittctggt	gttagagaag	ataaacttct	tgtgttattg	3360
cagctagatc	aaagttccaa	gactgagata	tgggtgacaa	atgtgattga	tgagaccacc	3420
aaaggagcag	tttcttggac	caaggtctta	gcattggatt	tgagccctca	tcttcagatt	3480
gggaatgatg	gaagtttctt	cctaggcgag	gataagaaa	tcgtcatggt	ctgtgagaaa	3540
ttgattgatg	agaacaaggt	caaagacatg	gtctacattg	ttggggagga	taatgttgtc	3600
acagaagtgg	gatttggagt	agatgaaatg	gatggatgtc	gggcagttat	tcttaattat	3660
gttccaagtt	tggttcaaat	cgagcgagct	ggaggcaaca	ggaaaagagg	gcactaagcg	3720
gccgctaggg	catgtctaga	agtcgcgaaa	aatcaccagt	ctctctctac	aaatctatct	3780
ctctctatct	ttctccagaa	taatgtgtga	gtagttccca	gataagggaa	ttagggttct	3840
tatagggttt	cgctcatgtg	ttgagcata	aagaaacctt	tagtatgtat	ttgtatttgt	3900
aaaaacttct	tatcaataaa	atlttcaatt	cttaaaacca	aaatccagtg	actgcaggca	3960
tgcaagctta	tcgataaccg	cgacgattga	tgcattgtgt	caatcaattg	gcaagtcata	4020
aaatgcatta	aaaaatattt	tcatactcaa	ctacaaatcc	atgagtataa	ctataattat	4080
aaagcaatga	ttagaatctg	acaaggattc	tggaaaatta	cataaaggaa	agttcataaa	4140
tgtctaaaa	acaagaggac	atacttgtat	tcagtaacat	ttgcagcttt	tctaggtctg	4200
aaaatatatt	tgttgcttag	tgaataagca	taatgttaca	actacaagtg	ttttactcct	4260
catattaact	tcggtcatta	gaggccacga	tttgacacat	ttttactcaa	aacaaaatgt	4320
ttgcatatct	cttataattt	caaattcaac	acacaacaaa	taagagaaaa	aacaaataat	4380

[0007]

attaatttga	gaatgaacaa	aaggaccata	tcattcatta	actcttctcc	atccatttcc	4440
atttcacagt	tcgatagcga	aaaccgaata	aaaaacacag	taaattacaa	gcacaacaaa	4500
tggtacaaga	aaaacagttt	tcccaatgcc	ataaactca	aactcagtag	gattctgttg	4560
tgtgogcaat	gaaactgatg	cattgaactt	gacgaacgtt	gtcgaaaccg	atgatacгаа	4620
cgaagctct	agaggatcaa	ttcgagctct	taggtcgacc	cacgtttgcc	aaaaccaact	4680
ctctctctcc	ttttttgtcg	tgcttctact	ctttcagggc	ttggcaatcc	agttttttct	4740
tcgtacttct	tttctagggc	ctctagctct	tcacgaatgc	tatcaagaac	ttgatccgtc	4800
attctgtcaa	agaagagaac	tcctccaga	tggtcgtatt	cggtcgtaaa	gattcgtgca	4860
ggtaaacgtg	atagactgat	tgaaaatctt	tcaccagtaa	tatcccttgc	atcaatcttg	4920
acagattgtg	gtcgaacaac	ttcagcatag	atccccggga	aggagaggca	tccttcata	4980
aacgttacta	atttatcgga	atatttcttg	attttcggat	ttacaaggac	aatttctttt	5040
ccttctccag	gctctccagc	tggtataaac	accatgagtt	gaacattgag	acctacttgt	5100
gggtcgtgaga	gccccaatgcc	atccgttttg	tacataacat	caaacatagc	atcaaccaag	5160
ttctttaaat	tctcgtcaaa	aatatcaatc	ctcttgttct	tagcccgtag	tataggatcc	5220
ggatactcaa	caatcttcaa	aggcgtctca	aattgaacat	cagtagctga	agctacttta	5280
tcgtctttac	gcgagacgcg	ctttacttct	gcgcgaccg	aagatgtcag	aggactggtc	5340
cggttcacag	tagagcagaa	cgtgaccgtg	gatttgacc	gaccataacc	ggcagagaga	5400
gtagtagctc	ggcggagataa	aaccggtagg	agtatgcgag	agagtgggtg	agcttggagg	5460
aagcagttac	agacggctcc	catggtgaa	gtatttgaaa	gaaaattaaa	aataaaaaga	5520
tcgctcagag	gatccaagct	tagatgagag	atttcgattc	cgattttgat	ttcgattccg	5580
attttgattt	cgattgatct	cttctctctg	atttgtgttc	cttatataag	gaaattcttg	5640
tggtgattaga	cgtcatgget	tacgtcattt	ccttcgtcct	gttgctcact	gattgagctg	5700
tgagtggagg	gaccactgga	agatgcttca	ctaattttct	tagtggaggg	accggcttca	5760
catgcttcac	acaagtggct	gtcgggcatc	atctttttta	gcttttgaca	aagcaatggt	5820
ttagtgggtg	ctcccactct	tatcttcaac	attattatct	tatcttcaaa	ggacgataag	5880
atgttgatgt	ctgtggacga	agttgggatt	agacgtcatg	gcttacgtca	ttcctctgt	5940
cctgttgctc	actgattgag	ctgtgagtg	agggaccact	ggaagatgct	tcactaattt	6000
tcttagtgga	gggaccggct	tctcatgctt	cacacaagtg	gctgtcgggc	atcatctttt	6060
ttagcttttg	acaaaagcaat	gttttagtgg	gggctcccac	tcttatcttc	aacattatta	6120
tcttatcttc	aaaggacgat	aagatgttga	tgtctgtgga	cgaagttagc	gaatttcgac	6180
ctgcaggcat	cgtaactatg	gtcatagctg	gtcatagctg	ttcctgtgtg	gaaattgtta	6240
tcgctcaca	attccacaca	acatacgagc	cggaagcata	aagtgtaaag	cctgggggtgc	6300
ctaattgagtg	agctaactca	cattaattgc	gttgcgctca	ctgcccgtt	tcagtcggg	6360
aaacctgtcg	lgccagctgc	attaatgaat	cggccaacgc	gcggggagag	gcggtttgcc	6420
tattgggcca	aagacaaaag	ggcgacattc	aaccgattga	gggagggaag	gtaaatatg	6480
acggaaatta	ttcattaaag	gtgaattatc	accgtaccgc	acttgagcca	tttgggaatt	6540
agagccagca	aaatcaccag	lagcaccatt	accattagca	aggccggaaa	cgtcaccaat	6600
gaaaccatcg	atagcagcac	cgtaatcagt	agcgacagaa	tcaagtttgc	cttagcgtc	6660
agactgtagc	gcgttttcat	cggeattttc	ggtcatagcc	cccttattag	cgtttgccat	6720
cttttcataa	tcaaaatcac	cggaaaccaga	gccaccaccg	gaaccgcctc	cctcagagcc	6780
gccaccctca	gaaccgccac	cctcagagcc	accaccctca	gagccgccac	cagaaccacc	6840
accagagccg	ccgccagcat	tgacaggagg	cccgatctag	taacatagat	gacaccgcgc	6900
gcgataatth	atcctagttt	gcgcgctata	ttttgtttc	tatcgclat	taaatgtata	6960
attggcggac	tctaatacata	aaaaccatc	tcataaataa	cgctatgcat	tacatgttaa	7020
ttattacatg	cttaacgtaa	ttcaacagaa	attatatgat	aatcatcgca	agaccggcaa	7080
ggatttcaa	tcttaagaaa	ctttattgcc	aaatgttga	acgatcgggg	atcatccggg	7140
tctgtggcgg	gaactccacg	aaaatatccg	aacgcagcaa	gatatcgcg	tgcatctcgg	7200
tcttgccctg	gcagtcgccg	ccgacccgct	tgatgtggac	gcccggcccg	atcatattgt	7260
cgctcaggat	cgtggcgttg	tgcttgcctg	ccgttgctgt	cgtaatgata	tcggcacctt	7320
cgaccgcctg	ttccgcagag	atcccgtggg	cgaagaactc	cagcatgaga	tccccgcct	7380
ggaggatcat	ccagccggcg	tcccggaaaa	cgattccgaa	gcccaccctt	tcatagaagg	7440
ggcgggtgga	atcgaaatct	cgtgatggca	ggttggcgct	cgcttggctg	gtcatttcga	7500
accccagagt	cccgtcaga	agaactcgtc	aagaaggcga	tagaaggcga	tgccctgcga	7560
atcgggagcg	gcgataccgt	aaagcacgag	gaagcggtca	gcccattcgc	cgccaagctc	7620
ttcagcaata	tcacgggtag	ccaacgctat	gtcctgatag	cggtccgcca	caccagccg	7680
ggcacagtcg	atgaatccag	aaaagcggcc	attttcacc	atgatattcg	gcaagcagcc	7740
atcgccatgg	gtcacgacga	gatcatcgcc	gtcgggcatg	cgcgccttga	gcctggcgaa	7800
cagttcggct	ggcgcgagcc	cctgatgctc	ttcgtccaga	tcacctgat	cgacaagacc	7860
ggcttccatc	cgagtacgtg	ctcgtcgtat	gcgatgtttc	gcttgggtgt	cgaatgggca	7920
ggtagccgga	tcaagcgtat	gcagcccgcc	cattgcata	gccatgatgg	atactttctc	7980
ggcagagca	agttgagatg	acaggagatc	ctgccccggc	acttcgcca	atagcagcca	8040
gtcccttccc	gcttcagtga	caacgtcgag	cacagctgcg	caaggaacgc	ccgtcgtggc	8100
cagccacgat	agccgcgctg	cctcgtcctg	cagttcattc	agggcaccgg	acaggtcggg	8160
cttgacaaaa	agaaccgggc	gcccctgcgc	tgacagccgg	aacacggcgg	catcagagca	8220
gccgattgtc	tgttgtgccc	agtcatalgc	gaatagcctc	tccaccaag	cgccgggaga	8280
acctgcgtgc	aatccatctt	gttcaatcat	gcgaaacgat	ccagatccgg	tcagattat	8340
ttggattgag	agtgaaatg	agacttcaat	tgataaccga	ggggaattta	tggaacgtca	8400
gtggagcatt	tttgacaaga	aatatttgc	agctgatagt	gaccttaggc	gacttttgaa	8460

[0008]

cgcgcaataa tggtttctga cgtatgtgct tagctcatta aactccagaa acccgcggt 8520
gagtggtcct ttcaacgttg cggttctgtc agttccaaac gtaaacggc ttgtcccgcg 8580
tcatcggcgg gggtcataac gtgactcct taattctccg ctcatgatc 8629

<210> 5
<211> 21
<212> DNA
<213> 人工

<220>
<223> 玉米

<400> 5
gttgtagaag gacgcgatgg a 21

<210> 6
<211> 21
<212> DNA
<213> 人工

<220>
<223> 玉米

<400> 6
caggtacaag agcgtcatgc a 21

<210> 7
<211> 20
<212> DNA
<213> 人工

<220>
<223> 玉米

<400> 7
tctgttccc gctgggctgc 20

<210> 8
<211> 24
<212> DNA
<213> 人工

<220>
<223> 分离自腺病毒

<400> 8
tatataatgg atccccgggt accg 24

<210> 9
<211> 20
<212> DNA
<213> 人工

<220>
<223> 玉米

<400> 9
cgtccctgcc ctttgtacac 20

<210> 10
<211> 21
<212> DNA
<213> 人工

[0009]

<220>		
<223>	玉米	
<400>	10	
acacttcacc	ggaccattca a	21
<210>	11	
<211>	19	
<212>	DNA	
<213>	人工	
<220>		
<223>	玉米	
<400>	11	
ccgcccgtcg	ctcctaccg	19

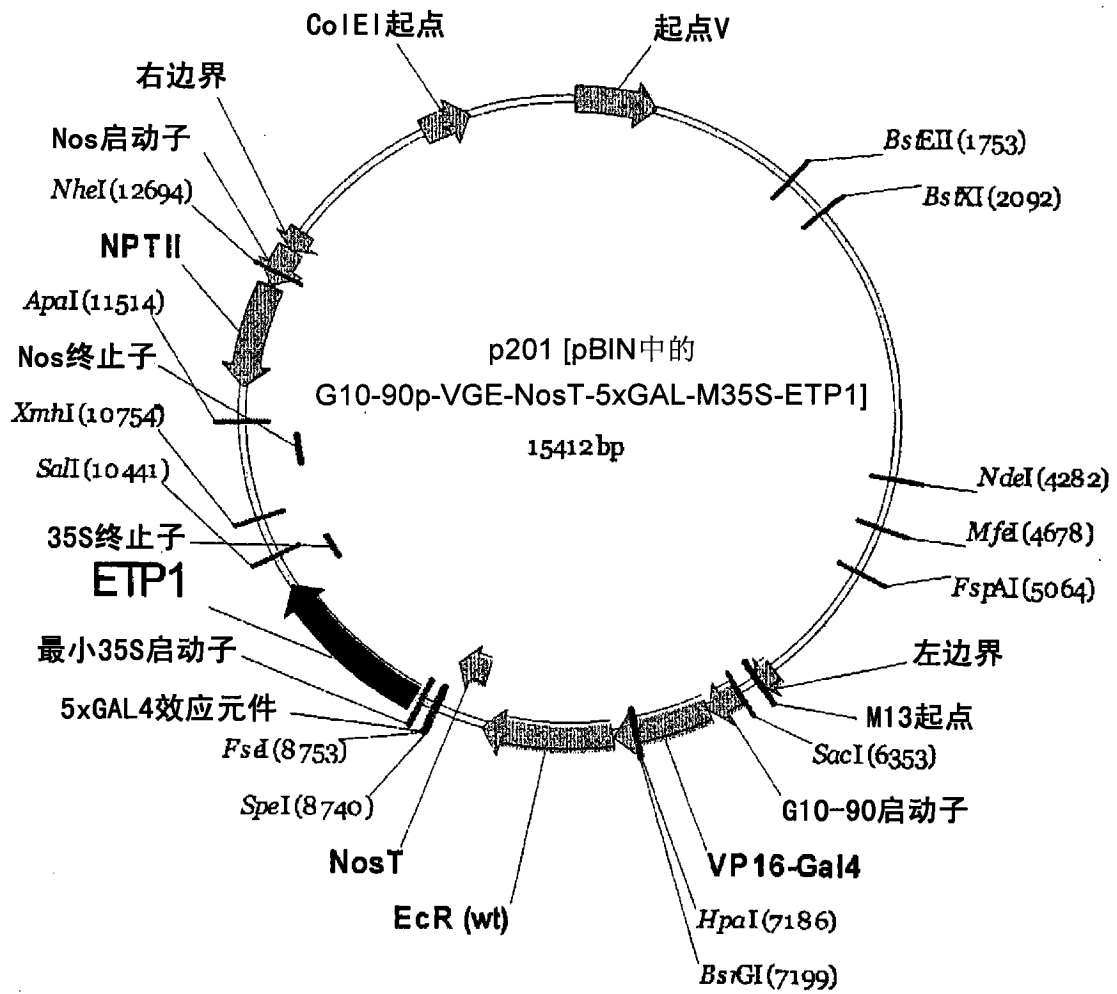


图 1

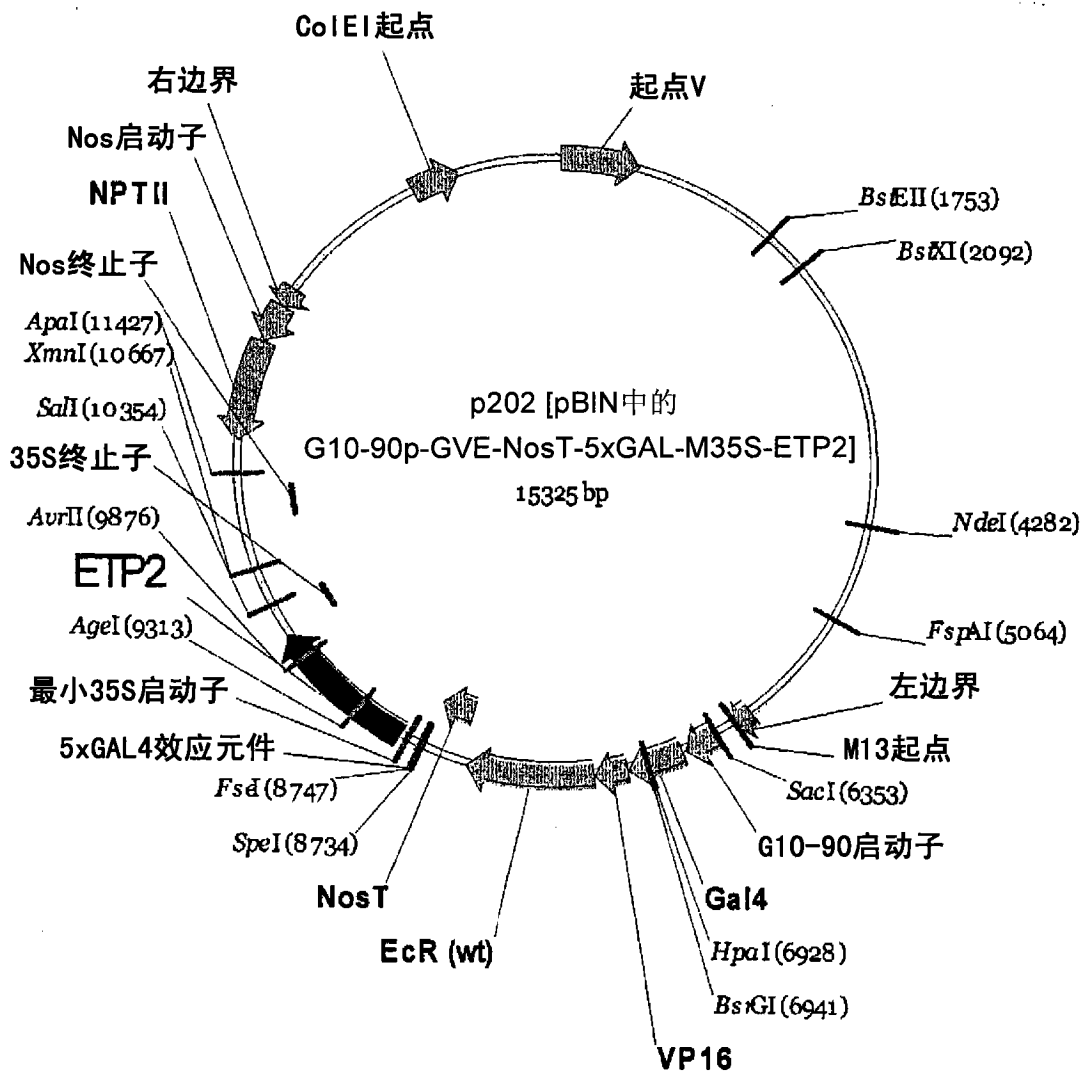


图 2

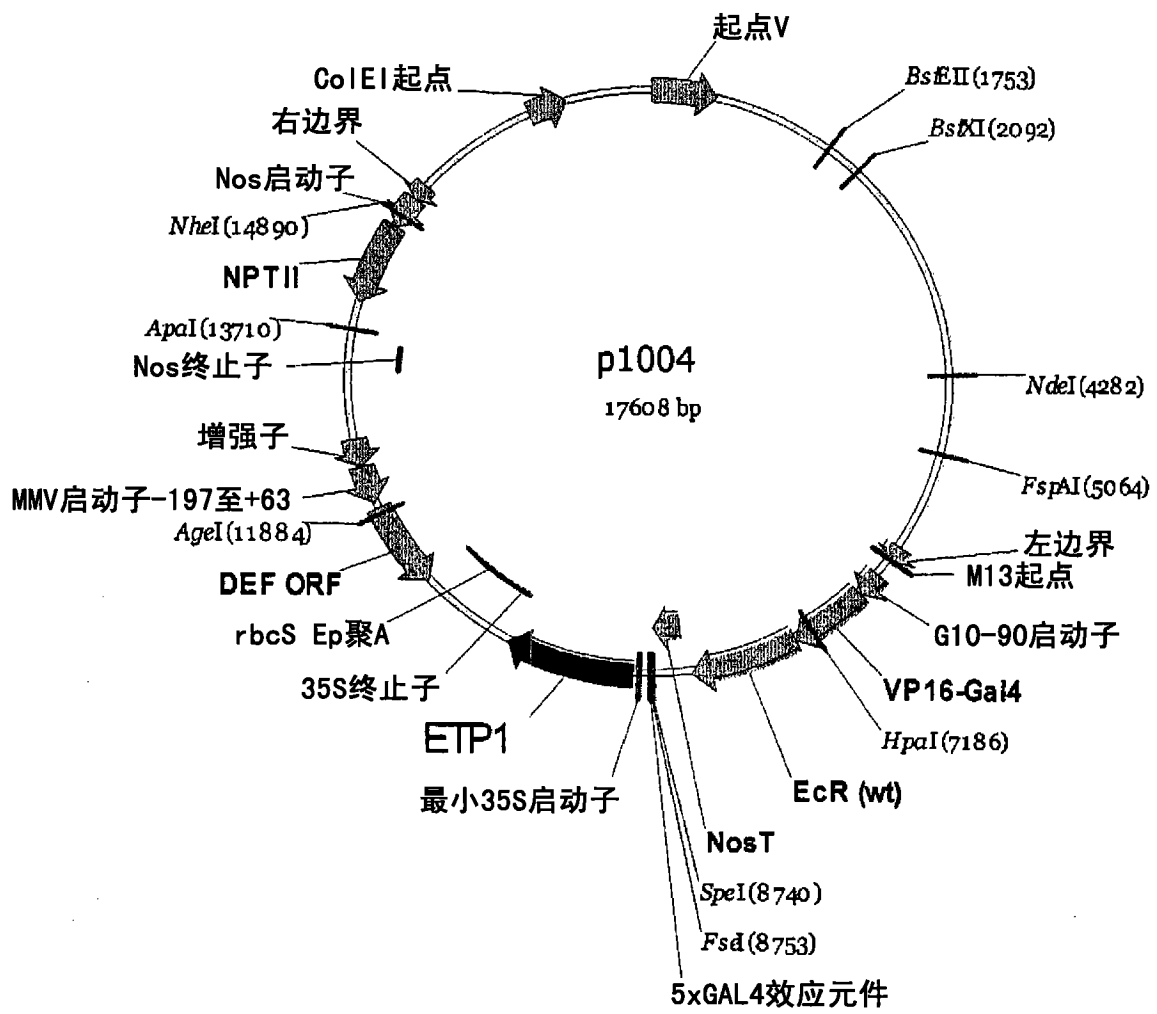


图 3A

G10-90启动子

6301 ~~~~~
 ATAGTT TAAACTGAAG GCGGGAACG
 ~~~~~  
 6401 ACAATCTGAT CCAAGCTCAA GCTAAGCTTG CATGCCTGCA GGATATCGTG GATCCAAGCT TGCCACGTGC CGCCACGTGC CGCCACGTGC  
 ~~~~~  
 6501 CTCTAGAGGA TCCATCTCCA CTGACGTAAG GGATGACGCA CAATCCCACT ATCCTTCGCA AGACCCCTCC TCTATATAAG GAAGTTCATT TCATTTGGAG
 ~~~~~  
 G10-90启动子  
 VP16-Gal4

G10-90启动子

6601 AGGACACGCT GGGATCCCCA CCTAGGCCCC CCGACCGAT GTCAGCCTGG GGGACGAACT CCACTTAGAC GCGGAGGACG TGGCGATGGC GCATGCCGAC  
 ~~~~~  
 6701 GCGCTAGACG ATTTCGATCT GGACATGTTG GGGGACGGGG ATTCCCACAG TCCGGGATTT ACCCCCACAG ACCTCCGCCC CTACGGCGCT CTGGATATGG
 ~~~~~  
 6801 CCGACTTCGA GTTTGAGCAG ATGTTTACCG ATGCCCTTGG AATTGACGAG TACGGTGGGA AGCTTCTAGG TACCTCCAGA AGAATATCAG GCGGGGAATT  
 ~~~~~  
 6901 CGGGGGATG AAGCTACTGT CTTCTATCGA ACAAGCATGC GAFATTTGCC GACTTAAAAA GCTCAAGTGC TCCAAAGAAA AACCGAAGTG CGCCAAGTGT
 ~~~~~  
 7001 CTGAAGAACA ACTGGGAGTG TCGCTACTCT CCCAAARACCA AAAGGCTCC GCTGACTAGG GCACATCTGA CAGAAGTGA ATCAAGGCTA GAAGACTGG  
 ~~~~~  
 7101 AACAGCTATT TCTACTGATT TTTCCCTCGAG AAGACCTTGA CATGATTTTG AAAATGGATT CTTTACAGGA TATAAAGCA TTGTTAACAG GATTATTTGT
 ~~~~~  
 7201 ACAAGATAAT GTGAATAAAG ATGCCGTCAC AGATAGATTG GCTTCAGTGG AGACTGATAT GCCTCTAACA TTGAGACAGC ATAGAATAAG TCGGACATCA  
 ~~~~~  
 7301 TCATCCGAAG AGAGTAGTAA CAAAGGTCNA AGACAGTTGA CTGTATCGGG AGGCCGTGGG ATCCGGCCIG AGTCCGTAGT ACCCGAGACT CAGTCCGCCA
 ~~~~~  
 ECR (wt)

图 3B

```

7401 TGAAGCGGAA AGAGAAGAAA GCACAGAAGG AGAAGACAA ACTGCCTGTC AGCAGGACGA CCGTGGACGA CCACATGCCG CCCATTATGC AGTGTGAACC
Ecr (wt)
7501 TCCACCTCCT GAAGCAGCAA GGATTCACGA AGTGGTCCCA AGGTTTCTCT CCGACAAGCT GTTGGTGACA AACCGGCAGA AAAACATCCC CCAGTTGACA
Ecr (wt)
7601 GCCAACCCAGC AGTTCCTTAT CGCCAGGCTC ATCTGGTACC AGGACGGGTA CGAGCAGCCT TCTGATGAG AATTGAAGAG GATTACGCAG ACCTGGGCAGC
Ecr (wt)
7701 AAGCGGACGA TGA AAAACGAA GAGTCGGACA CTCCCTCCG CCAGATCACA GAGATGACTA TCCTCACGGT CCAACTTATC GTGGAGTTCC CGAAGGGATT
Ecr (wt)
7801 GCCAGGGTTC GCCAAGATCT CGCAGCCTGA TCAAATTACG CTGCTTAAGG CTTGCTCAAG TGAGGTAATG ATGCTCCGAG TCGCGCGAGG ATACGGATGG
Ecr (wt)
7901 GCCTCCGACA GTGTTCTGTT CGCGAACAA CAAAGCGTACA CTGCGGACAA CTACCGCAG GCTGGCATGG CCTACGTCAT CGAGGATCTA CTGCACCTCT
Ecr (wt)
8001 GCCGGTGCAT GTA CTCTATG GCGTTGGACA ACATCCATTA CGCGCTGCTC ACGGCTGTCC TCATCTTTTC TGACCGGCCA GGGTTGGAGC AGCCGCAACT
Ecr (wt)
8101 GGTGGAAGAG ATCCAGCGGT ACTACCTGAA TAGCTCCGC ATCTATATCC TGAACCCAGT GAGCGGGTCC GCGCGTTCCG CCGTCATATA CGGCAAGATC
Ecr (wt)
8201 CTCTCAATCC TCCTTGAGCT ACGCAGCTC GGCATGCARA ACTCCAACAT GTGCATCTCC CTC AAGCTCA AGAACAGAAA GCTGCGCGCT TTCTCCGAGG
Ecr (wt)
8301 AGATCTGGGA TGTGGGGGAC ATGTGGCACA CCCAACCGCC GCCTATCTCTC GAGTCCCCCA CGAATCTCTA GCCCCTGGCC GCACGCATCC CCGATGCCGC
NosF
8401 GTCCGGCCGC GCTGCTCTGA GAATTCGATA TCAAGCTTCT AGACCCGGGC TGCAGAGATC TACCGGTTAA GCTTAATTC CGATCGTTCA AACATTTGGC
NosF
8501 AATAAGTTT CTTAAGATTG AATCCTGTTG CCGGTCCTTC GATGATATC ATATAATTC TGTGAATTA CGTTAAGCAT GTAATAATTA ACAATGTAATG
NosF

```



```

ETP1
9701 AAGGTATGAA ATTTGAGTA TTTCGCCCGT TAGAGGAGGA GAGAACTCTT CTCTGTTGTT GCAGCTCGAT TTTGAATCTA AGACTGAGAT ATGGGTGACG
ETP1
9801 AATAAGATTG ATGACACCAC CACCAAAGGA GCAGCAGTCT CTGGACCAA GGTCTTAGCA TTGATTAA GCCCTGATCT TCAATTATTT TCGGAGGAGG
ETP1
9901 TAAATTTTTT GCTTGACGAG GATAAGAAAG TCGCTGTGTG TTGTGAGAGA TGGTTGGAAC CGCAAGAGCA CCACAGGTAC CAGTGCAGGA GAGAGTACAA
ETP1
10001 GATCACCGAC AAGATATACA TTCTCGGGGA GGATAATAAA GTCCATGAAG TAGGTTCTGG AGAGGGAGAG GCTACAGATT CACTTGAAGG AATTTCCGCAA
ETP1
10101 GTTATTCICA ATTACGCTCC AAGTTTGGTC CAAATCGAGC AAGCCGGAGG AGGCAAAACA AAAAGAGGTG ACGACTAAGC GGCCGCTAGG GCATGICTAG
35S终止子
10201 AAGTCCGCAA AATCACCAG TCTCTCTCTA CAAATCTATC TCTCTCTATT TTCTCCAGA ATPATGTGTG AGTAGTCCC AGATAAGGGA ATTAGGGTTC
35S终止子
10301 TTATAGGGTT TCGCTCATGT GTTGGCATA TAAGAAACCC TTAGTATGTA TTTGTATTTG TAAATFACTT CTATCAATAA AATTTCTAAT TCCTAAAACC
35S终止子
10401 AAAATCCAGT GACTGCAGGC ATGCCAGCTT ATCGATACCG TCGACGATTG ATGCATGTTG TCAATCAATT GGCAAGTCAT AAAATGCATT AAAAAATATT
rbcSEp聚A
10501 TTCATACTCA ACTACAAATC CATGACTATA ACTATAATTA TAAAGCAATG ATTAGAATCT GACAAGGATT CTGGAAAT ACATAAAGGA AAGTTCATAA
rbcSEp聚A
10601 ATGTCTAAA CACAAGAGGA CATACTGTA TTCAGTACA TTTCAGCTT TTCTAGGCTT GAAATATAT TTGTTGCCTA GTGAATAAGC ATATGGTAC
rbcSEp聚A
10701 AACTACAAGT GTTTTACTCC TCATATTAAC TTGGGTCAAT ATTGACACA TTTTACTCA AAACAAAATG TTGTCATATC TCTTATAATT

```

图 3E

**rbcSEp 聚A**  
 10801 TCAAAATCAA CACACAACAA ATAAGAGAAA AAACAATAA TATTAATTG AGAATGAACA AAAGGACCAT ATCATTATT AACTTTTCTC CATCCATTTC  
 ~~~~~  
rbcSEp 聚A
 10901 CATTTCACAG TTCGATAGCG AAAACCGAAT AAAAACAACA GTAATATTACA AGCACAACAA ATGGTACAAG AAAAACAGTT TTCCCAATGC CATAATACTC
 ~~~~~  
**rbcSEp 聚A**  
 11001 AAACTCAGTA GGATTCYGGT GTGTGGCAA TGAACCTGAT GCATTGAAC TGCAGAACGT TGTCGAAACC GATGATACGA ACGAAAGCTC TAGAGGATCA  
 ~~~~~  
rbcSEp 聚A
 11101 ATTCGAGCTC TTAGGTCGAC CCACGTTTGC CAAAACCAAC TCCTGCTCTC CTTTTTTGTC GTGCTTCTAC TCTTTCAGGG CTGGCAATC CAGTTTTTTC
 ~~~~~  
 DEF ORF  
 11201 TTCGACTTC TTTTCTAGGG CCTCTAGCTC TTCACGAATG CTATCAAGAA CTTCATCCGT CATTCTGTCA AAGAAGAGAA CTCCCTCCAG ATGGTCGTAT  
 ~~~~~  
 DEF ORF
 11301 TCGTCTGAA AGATTCGTGC AGGTAACCGT GATAGACTGA TTGAAAATCT TTCACCAGTA ATATCCCTTG CATCAATCTT GACAGATTGT GGTCGAACAA
 ~~~~~  
 DEF ORF  
 11401 CTTCCAGATA GATCCCCGG AAGGAGAGGC ATCCTTCATC AAACGGTACT AATTATCGG AAVATTCTT GATTTTCGGA TTACAAGGA CAATTTCTTT  
 ~~~~~  
 DEF ORF
 11501 TCCTTCTCCA GGCTCTCCAG CTGGATTAAA CACCATGAGT TGAACATTGA GACCTACTTG TGGTCTGAG AGCCCAATGC CATCCGTTTT GTACATAACA
 ~~~~~  
 DEF ORF  
 11601 TCAAACATAG CATCAACCA GTTCTTTAAA TTCTCGTCAA AAATATCAAT CCTCTTGTTC TTAGCCCGTA GTATAGGATC CGGATACTCA ACAATCTTCA  
 ~~~~~  
 DEF ORF
 11701 AAGGGTCTC AAATTGAACA TCAGTAGCTG AAGTACTTTT ATCGTCTTTA CCGGAGACGC GCTTACTTTC TCGCGGACC GAAGATGTCA GAGGACTGGT
 ~~~~~  
 DEF ORF  
 11801 CCGGTTTACA GTAGAGCAGA ACGTGACCGT GGATTTGAGC CGACCATAC CCGCAGAGAG AGTAGTAGT CCGCGAGATA AAACCGGTAG GAGTATGCGA  
 ~~~~~

Age1

图 3F

```

DEF ORF
11901 GAGAGTGGTG GAGCTTGGAG GAAGCAGTTA CAGACGGCTC CCATGGTGGG AGTATTTGAA AGRARAATTAA AAATAAAAAG ATCCGGCTCGA GGATCCCRAGC
~::~:
DEF ORF
12001 TTAGATGAGA GATTTCGATT CCGATTTTGA TTTGATTC CCATGATTC TCGATTGATC TCCTCCTCT GATTTGTGTT CCTATATATA GGAATTTCTT
~::~:
MMV启动子-197至+63
12101 GTGGGATTAG ACGTCATGGC TTACGTCATT TCCTTCGTCC TGTGCTCAC TGATTGAGCT GTGAGTGGAG GGACCACCTGG AAGATGCTTC ACTAATTTTC
~::~:
MMV启动子-197至+63
12201 TTAGTGGAGG GACCGGCTTC ACATGCTTCA CACAAGTGGC TGTCGGGCAT CATCTTTTT AGCTTTTGAC AAAGCAATGT TTTAGTGGTG GCTCCCCACTC
~::~:
MMV启动子-197至+63
12301 TTATCTTCAA CAAATATATC TTATCTTCAA AGGACGATAA GAATGTGATG TCTGTGGAGC AAGTTGGGAT TAGACGTCAAT GGCTTACGTC ATTTCCCTTCG
~::~:
MMV启动子-197至+63
12401 TCCTGTTGCT CACTGATTGA GCTGTGAGTG GAGGACCAC TGGAGATGC TTCACTAATT TTCTTAGTGG AGGGACCAGGC TTCTCATGCT TCACACAAGT
~::~:
增强子
12501 GGCTGTCCGG CAAATCTTT TTTAGCTTTT GACAAAGCAA TGTTTAGTG GGGGCTCCA CTTTATCTT CAACATTAAT ATCTTATCTT CAAAGGACGA
~::~:
增强子
12601 TAAGATGTTG ATGTCGTGG ACGAAGTTGA CGAATTTGA CCTGCAGGCA TGCAAGCTTG GCGTAATCAT GGTCATAGCT GTTTCCTGTG TGAAATGTT
~::~:
增强子
12701 ATCCGCTCAC AATCCACAC AACATACGAG CCGGAAGCAT AAAGTGTAAG GCCTGGGGTG CCTAATGAGT GAGCTAATC ACAATTAATT CGTTGGGCTC
12801 ACTGCCCGCT TTCAGTCCG GAACCTGTC GTGCCAGCTG CATTAATGAA TCGGCCAACG CCGGGGAGA GCGGTTTGC GTATTGGGC AAAGACAAA
12901 GGGCGACATT CAACCGATTG AGGAGGAA GGTAAATATT GACGGAATTT ATTCAATTA AATTCAATTA CACCGTCACC GACTTGAGCC ATTTGGGAAT
13001 TAGAGCCAGC AAAATCACCA GTAGCACCAT TACCATTAGC AAGGCCGAA AGTCAACCA TGRAACCATC GATAGCAGCA CCGTAATCAG TAGCGACAGA
13101 ATCRAAGTTG CCTTTAGCGT CAGACTGTAG CCGGTTTTCA TCGGCATTTT CCGTCATAGC CCCCTTATTA GCGTTTTGCA TCTTTTCATA ATCAAATCA
13201 CCGGAACCCAG AGCCACCACC GGAACCGCT CCTCAGAGC CGCACCCCTC AGAACCGCA CCTCAGAGC CACCACCTC AGAGCCGCCA CCAGAACCCAC
13301 CACCAGAGCC GCGCCAGCA TTGACAGGAG GCGCCGATCTA GTAACATAGA TGACACCGCG CCGCATTAAT TATCCTAGTT TCGCGGCTAT ATTTGTGTTT

```

图 3G

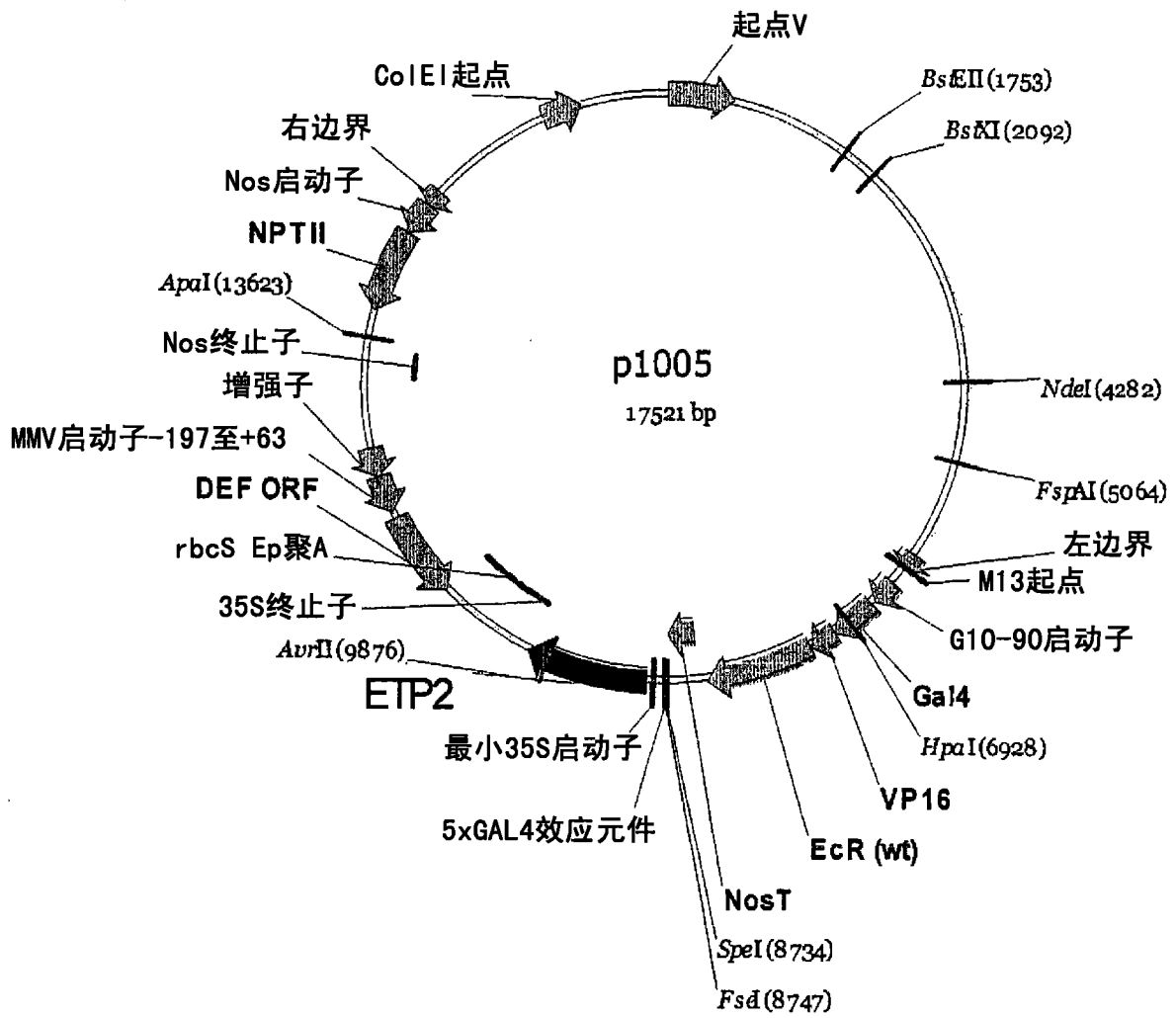


图 4A

```

          ~~~~~
          ATAGTT TAAACTGAAG GCGGGAAACG
          ~~~~~
G10-90启动子
          ~~~~~
6301  ACAATCTGAT CCAAGCTCAA GCTARGCTTG CATGCCCTGCA GGATACCTGTG GATCCAAGCT TGCCACGTTGC CGCCACGTTGC CGCCACGTTGC
          ~~~~~
G10-90启动子
          ~~~~~
6401  CTCTAGAGGA TCCATCTCCA CTGACGTPAAG GGATGACGCA CAATCCCCTACT ATCCTTCGCA AGACCCCTCC TCTATATAAG GAAAGTTCATT TCATTTGGAG
          ~~~~~
G10-90启动子
          ~~~~~
6501  AGGACACGCT GGGATCCCCA CCAATGGATCC GCCACCATGC TAGCCACCCA TGAAGCTACT GTCTTCTATC GAACAAGCAT GCGATATTG CCGACTTAAA
          ~~~~~
G10-90启动子
          ~~~~~
6601  AAGCTCAAGT GCTCCAAAGA AAMACCGAAG TCGGCCAAGT GTCTGAAGAA CAACTGGGAG TGTCCCTACT CTCCCAAAAC CAAAAGGTCT CCGCTGACTA
          ~~~~~
G10-90启动子
          ~~~~~
6701  GGGCACATCT GACAGAAGTG GAATCAAGGC TAGAAGACT GGAACAGCTA TTTTCTACTGA TTTTCTCCTCG AGAAGACCTT GACATGATTT TGAARAATGGA
          ~~~~~
          ~~~~~
          HpaI
          ~~~~~
6901  TTCTTTTACAG GATATAAAG CATTGTTAAC AGGATTATTT GTACAAGATA AFGTGAATAA AGATGCCGTC ACAGATAGAT TGGCTTCAGT GGAGACTGAT
          ~~~~~
          ~~~~~
          Gal4
          ~~~~~
7001  ATGCCCTTAA CATTGAGACA GCATAGAATA AGTGGACAT CATCATCGGA AGAGAGTAGT AACAAAGGTC AAAGACAGTT GACTGTATCC ATGGCCCCCC
          ~~~~~
VP16
          ~~~~~
7101  CGACCGATGT CAGCCTGGG GACGAACTCC ACTTAGACGG CGAGGACGTG GCGATGGCGC ATGCCGAGCC GCTAGACGAT TTCGATCTGG ACAFTTTGGG
          ~~~~~
VP16
          ~~~~~
7201  GGACGGGGAT TCCCCAGGTC CGGGATTTAC CCCCCAGAC TCCGCCCTCT ACGGGCTCT GGATATGGCC GACTTCGAGT TTGAGCAGAT GTTTACCGAT
          ~~~~~

```

图 4B


```

ETP2
9601 GCACCATATC ATTGCTCAAA TTGATTTTT ATACAGAGAA ATCTGTACCT GTGCTTCTTC CCTATCAGAG TCGTCGTCTT TTCCAAGCTA GTAGCCTTTC
ETP2
9701 TGTTCGTAGA GAAGATAAAC TTTCGTGTT ATTGCAGCTA GATCAAGT CCAAGACTGA GATATGGTG ACAATGTGA TTGATGAGAC CACCAARAGGA
AvrII
9801 GCAGTTTCTT GGACCAAGGT CTTAGCATTG GATTGAGCC CTCATCTTCA GATTGGAAAT GAUGGAAGTT TCTTCCTAGG CGAGGATAAG AAAGTCGTCA
ETP2
9901 TGTTCGTAGA GAAATTGATT GATGAGAACA AGGTCAAGA CATGGTCTAC AFTGTTGGG AGGATAATGT TGTACAGAA GTGGGATTTG GAGTAGATGA
ETP2
10001 AATGGATGA TGTCGGGCAG TTATCTTAA TTATGTCCA AGTTGGTTC AAATCGAGG AGCTGGAGC AACAGGAAA GAGGCACATA AGCGGCCGCT
35S终止子
10101 AGGGCATGTC TAGAAGTCCG CAAAATCAC CAGTCCTCT CTACAAATCT ATCTCTCTCT ATTTTCTCC AGAATAATGT GTGAGTAGTT CCCAGATAAG
35S终止子
10201 GGAATTAGG TCTTTATAGG GTTTCGCTCA TGTGTGACC ATATAAGAAA CCTTAGTAT GTATTTGTAT TTGTAAAATA CTCTATCAA TAAAATTTCT
35S终止子
10301 AATTCCPAA ACCAAATCC AGTGACTGCA GGCAATGCAAG CTTATCGATA CCGTCGACGA TTGATGCATG TTGTCAATCA ATTGCCAAGT CATAAAATGC
rbcSEp聚A
10401 ATTAATAAAT ATTTTCAATC TCAACTACAA ATCCATGAGT ATAACTATAA TTATAAAGCA ATGATTAGAA TCTGACAAGG ATTCGTGAAA ATTACATAAA
rbcSEp聚A
10501 GGAAGTTCA TAAATGCTA AAACACAAGA GGACATACTT GTATTCAGTA ACAATTCAG CTTTTCTAGG TCTGAAAATA TATTTGTTC CTAGTCAATA
rbcSEp聚A
10601 AGCATAATGG TACAACACTACA AGTGTTTTAC TCCTCAATT AACTTCGGTC ATTAGAGCC ACGATTTGAC ACATTTTAC TCAAAACAAA ATGTTTTCAT

```

图 4E

```

10701 ATCTCTTTATA ATTTCAAATT CAACACACAA CAAATPAGAG AAAAAACAAA TAATATTAAT TTGAGAATGA ACAAAGGAC CATATCATTC ATTAACTCCTT
~::~:
~::~:
~::~:
10801 CTCCATCCAT TTCCATTTCA CAGTTCCGATA GCGAAAACCG AATAAAAAAC ACAGTAAAT ACAAGCACAA CAAATGGTAC AAAAAACA GTTTTCCCAA
~::~:
~::~:
~::~:
10901 TGCCATAATA CTCAAACTCA GTAGGATTCT GGTGTGTCCG CAATGAAACT GATGCATTGA ACTTGACGAA CGTTGTGCGA ACCGATGATA CGAACGGAAG
~::~:
~::~:
~::~:
11001 CTCTAGAGGA TCAATTCGAG CTCTTAGGTC GACCCACGTT TGCCAAAACC AACTCCTGCT CTCCTTTTTT GTCCGTGCTTC TACTCTTTCA GGGCTTGCCA
~::~:
~::~:
~::~:
.....
11101 ATCCAGTTTT TTCTTCGTAC TTCTTTTCTA GGGCCTCTAG CTCTTCAGGA ATGCTATCAA GAACTTGATC CGTCATTCG TCAAAGAAGA GAACTCCTC
~::~:
~::~:
~::~:
DEF ORF
11201 CAGATGGTCG TATTCGTGCT GAAAGATTCC TGCAGGTAAA CGTGATAGAC TGATTGAAAA TCTTTCACCA GTAATATCCC TTGCATCAAT CTTGACAGAT
~::~:
~::~:
~::~:
DEF ORF
11301 TGTGGTCGAA CAACTTCAGC ATAGATCCCC GGAAGGAGA GGCATCCTTC ATCAAACGGI ACTAATTAT CGGAATATTT CTTGATTTTC GGATTTACAA
~::~:
~::~:
~::~:
DEF ORF
11401 GGACAATTTC TTTTCTTCTI CCAGGCTCTC CAGCTGGATT AAACACCATG AGTTGAACAT TGAGACCTAC TTGTGGTGTG GAGAGCCCAA TGCCATCCGT
~::~:
~::~:
~::~:
DEF ORF
11501 TTTGTACATA ACATCAAACA TAGCATCAAC CAAGTTCITTT AAATTCCTGT CAAAAATATC AATCCTCTTG TTCITAGCCC GTAGTATAGG ATCCGGATAC
~::~:
~::~:
~::~:
DEF ORF
11601 TCAACAATCT TCBAAGGCGT CTCBAATTGA ACATCAGTAG CTGAAGCTAC TTTATCGTCT TTACGGGAGA CGCGCTTTAC TTCTGGCGG ACCGAAGATG
~::~:
~::~:
~::~:
DEF ORF
11701 TCAGAGGACT GGTCCGGTTC ACAGTAGAGC AGAACGTGAC AGCCGACCAT AACCCGCAGA GAGAGTAGTA GCTCGGGGAG ATAAAACCGG
~::~:
~::~:
~::~:
DEF ORF
11801 TAGGAGTATG CGAGAGAGTG GTGGAGCTTG GAGGAAGCAG TTPACAGACGG CTCCCATGCT GGAAGTATTT GAAAGAAAAT TAAAAATAAA AAGATCCGCT
~::~:
~::~:
~::~:

```

图 4F

```

DEF ORF
11901 CGAGGATCCA AGCTTAGATG AGAGATTTCG ATTCCGATTT TGATTTCCGAT TCCGATTTTG ATTTCCGATTG ATCTCTTCCCT TCTGATTTGT GTTCCTTTATA
~::~:
MMV启动子-197至+63
12001 TAAGGAAATT CTTGTGGGAT TAGACGTCAAT GGCTTACGTC ATTTCTTCCG ATTTCTTCCG TCCGTGTGCT CACTGATGA GCTGTGAGTG GAGGACCAC TGAAGAATGC
~::~:
MMV启动子-197至+63
12101 TTCACTAATT TTCTTAGTGG AGGACCAGG TFCACATGCT TCACACAAGT GGCCTGCGGG CATCATCTTT TTAGCTTTT GACAAAGCAA TGTTTTACTG
~::~:
MMV启动子-197至+63
12201 GTGGCTCCCA CTCTTATCTT CAACATTATT ATCTTATCTT CAAGGACCGA TAAGATGTTG ATGCTGTGG ACGAAGTTGG GATTAGACGT CATGGCTTAC
~::~:
MMV启动子-197至+63
12301 GTCATTTCCCT TCGTCTGTT GCTCACTGAT TGAGCTGTA GTGGAGGAC CACTGGAAGA TGCITTCACTA ATTTTCTTAG TGGAGGGACC GGCTTCTCAT
~::~:
增强子
12401 GCTTCACACA AGTGGCTGTC GGGCATCATC TTTTTTAGCT TTTGACAAAG CAATGTTTTA GTGGGGGCTC CCACCTTTAT CTCAACATTT ATTATCTTAT
~::~:
增强子
12501 CTTCAAAGGA CGATAAGATG TTGATGTCTG TGGACGAAGT TGACGAATTT CGACCTGCAG GCATGCAAGC TTGGCGTAAT CATGGTCATA GCTGTTCCT
~::~:
增强子
12601 GTGTGAAATT GTTATCCGCT CACAATCCA CACAACATAC GAGCCGGAAG CATAAAGTGT AAAGCCTGGG GTGCCTAATG AGTCAGCTAA CTCACATTAA
12701 TTGGTTGCG CTCACTGCC GCTTCCAGT CCGGARACCT GTCGTGCCAG CTGCATTAAT GAATCGGCCA ACGCGGGGG AGAGGGGTTT TCGGTATTGG
12801 GCCAAGACA AAAGGGGAC ATTCACACCGA TTGAGGGAGG GAAGGTAAT ATTGACGGAA ATTATTCATT AAAGGTGAAT TATCACCGTC ACCGACTTGA
12901 GCCATTTGG AATTAGAGCC AGCAAAATCA CCAGTAGCAC CATTACCATT AGCAAGGCCG GAAACCTCAC CAATGAAACC ATCGATAGCA GCACCGTAAT
13001 CAGTAGCGAC AGAATCAAGT TTGCTTTAG CGTCAGACTG TAGCGGTTT TCATCGGCAT TTTCCGTTA AGCCCCCTTA TTAGCGTTTG CCACTTTTTC
13101 ATRATCAAAA TCACCGGAAC ACCGGACCG CCTCCCTCAG AGCCGCCACC CTCAGAACCG CCACCCCTCAG AGCCACCACC CTCAGAGCCG
13201 CCACCAGAAC CACCACCAGA GCGCCGCCA GCATTTGACAG GAGGCCCGAT CTAGTAACAT AGATGACACC GCGCGGATA ATTTATCCTA GTTTGCGCGC
~::~:
Nos终止子
13301 TATATTTTGT TTCTATCGC GTATTAATG TATAATGCG GGACTCTAAT CATAAAAACC CATCTCATAA ATAACGTCAT GCATTACATG TFAATTAATA
~::~:
Nos终止子
13401 CATGCTTAAC GTAATTCAAC AGAAATATA TGATAATCAT CGCAAGACCG GCAACAGGAT TCAATCTTAA GAACTTTAT TGCCAAATGT TTGACGATC
~::~:

```

图 4G

Nos 终止子
 13501 GGGGATCATC CCGGTCTGTG GCGGAACTC CACGAAAATA TCCGAACGCA GCAAGATATC GCGGTGCATC TCGGTCTTGC CTGGGCAGTC GCGGCGGACG
 ~~~~~  
 ApaI  
 ~~~~~  
 13601 CCGTTGATGT GGACGCCGGG CCCGATCATA TTGTGCTCA GGATCGTGGC GTTGTGCTTG TCGGCCGTTG CTGTGCTAAT GATATCGCA CCTTCGACCG
 13701 CCTGTTCGGC AGAGATCCCG TGGGCCAAGA ACTCCAGCAT GAGATCCCG CGCTGGAGGA TCATCCAGCC GGCCTCCCGG AAAACGATTC CGAAGCCCAA
 13801 CCTTTCATAG AAGCGGGCGG TGGAAATCGAA ATCTCTGTAT GGCAGGTTGG GCGTCCGTTG GTCGGTCATT TCGAACCACA GAGTCCCGCT CAGAAGAACT
 ~~~~~  
**NPT II**  
 13901 CGTCAAGAAG GCGATAGAAG GCGATCGGCT GCGATCGGG AGGGCGGATA CCGTAAAGCA CGAGGAAGCG GTCAGCCCAT TCGCCGCCAA GCTCTTCAGC  
 ~~~~~  
NPT II
 14001 AATATCAGG GTAGCCAAG CTATGTCTTG ATAGGCTCC GCCACACCA GCCGGCCACA GTCGATGAAT CCAGAAAAGC GGCCATTTTC CACCAATGATA
 ~~~~~  
**NPT II**  
 14101 TTCGGCAAGC AGGCATCGCC ATGGTCAAG ACGAGATCAT CGCCGTCCGG CATGGGGCC TTGAGCCTGG CGAACAGTTC GGCTGGCGG AGCCCTTGAT  
 ~~~~~  
NPT II
 14201 GCTCTTCGTC CAGATCATCC TGATCGACAA GACCGGCTTC CATCCGAGTA CGTGTCTGCT CGATGGGATG TTTCGCTTGG TGGTCAATG GGCAGGTAGC
 ~~~~~  
**NPT II**  
 14301 CCGATCAAGC GTATGCAGCC GCCGATTGC ATCAGCCATG ATGGATACTT TCTCGGCAGG AGCAAGTGA GATCACAGGA GATCCTGCC CGGCACCTCG  
 ~~~~~  
NPT II
 14401 CCCAATAGCA GCCAGTCCCT TCCCGCTTCA GTGACAAGT CGAGCACAGC TGGCAAGGA AGCCCGTCCG TGGCCAGCCA CGATAGCCGC GCTGCCCTCGT
 ~~~~~  
**NPT II**  
 14501 CCTGCAGTTC ATTCAGGCA CCGACAGGT CCGTCTTGAC AAAAAGAAC GGGCGCCCT GCGGTGACAG CCGGAACAGC GCGGCATCAG AGCAGCCGAT  
 ~~~~~  
NPT II
 14601 TGCTGTGTGT GOCAGTCAT AGCGAATAG CCTCTCCACC CAAGCGGCGG GAGAACCCTGC GTGCAATCCA TCTGTTCAA TCATGCGAAA CGATCCAGAT
 ~~~~~  
**NPT II**  
**Nos**  
 14701 CCGGTGCAGA TTATTGGAT TGAGAGTAA TATGAGACTC TAATTGGATA CCGAGGGGAA TTTATGGAC GTCAGTGGAG CATTTTIGAC AAGAAATATT  
 ~~~~~

图 4H
52

Nos/启动子

14801 TGCTAGCTGA TAGTGACCTT AGGCGACTTT TGAACCGCA ATAAATGGTTT CTGACGTAAG TGCTTAGCTC ATTAACCTCC AGAAACCCGC GGCTGAGTGG
 ~~~~~  
 14901 CTCCTTCAAC GTTGGGGTTC TGTCAGTTCC AAACGTAAAA CGGCTTGTTCC CGGCTCATCG GCGGGGGTCA TAACGTGACT CCTTAATTC FCCGGCTCATG  
 ~~~~~  
 15001 ATC
 ~~~~~  
 . . .

图 4I